ISSN 2472-1972

Identifying Geographic Disparities in Diabetes Prevalence Among Adults and Children Using Emergency Claims Data

David C. Lee,[1,2] Mary Pat Gallagher,[3] Anjali Gopalan,[4] Marcela Osorio,[1] Andrew J. Vinson,[1] Stephen P. Wall,[1] Joseph E. Ravenell,[2] Mary Ann Sevick,[2] and Brian Elbel[2,5]

[1]Ronald O. Perelman Department of Emergency Medicine, New York University School of Medicine, New York, New York 10016; [2]Department of Population Health, New York University School of Medicine, New York, New York 10016; [3]Division of Endocrinology, Department of Pediatrics, New York University School of Medicine, New York, New York 10016; [4]Division of Research, Kaiser Permanente Northern California, Oakland, California 94612; and [5]Wagner Graduate School of Public Service, New York University, New York, New York 10012

Geographic surveillance can identify hotspots of disease and reveal associations between health and the environment. Our study used emergency department surveillance to investigate geographic disparities in type 1 and type 2 diabetes prevalence among adults and children. Using all-payer emergency claims data from 2009 to 2013, we identified unique New York City residents with diabetes and geocoded their location using home addresses. Geospatial analysis was performed to estimate diabetes prevalence by New York City Census tract. We also used multivariable regression to identify neighborhood-level factors associated with higher diabetes prevalence. We estimated type 1 and type 2 diabetes prevalence at 0.23% and 10.5%, respectively, among adults and 0.20% and 0.11%, respectively, among children in New York City. Pediatric type 1 diabetes was associated with higher income ($P = 0.001$), whereas adult type 2 diabetes was associated with lower income ($P < 0.001$). Areas with a higher proportion of nearby restaurants categorized as fast food had a higher prevalence of all types of diabetes ($P < 0.001$) except for pediatric type 2 diabetes. Type 2 diabetes among children was only higher in neighborhoods with higher proportions of African American residents ($P < 0.001$). Our findings identify geographic disparities in diabetes prevalence that may require special attention to address the specific needs of adults and children living in these areas. Our results suggest that the food environment may be associated with higher type 1 diabetes prevalence. However, our analysis did not find a robust association with the food environment and pediatric type 2 diabetes, which was predominantly focused in African American neighborhoods.

Freeform/Key Words: clinical science and care, epidemiology, information technology, nutrition and diet, socioeconomic aspects

As diabetes rates continue to rise toward epidemic levels, better surveillance is needed to identify specific communities with the highest burden of disease [1, 2]. Emergency department data have been used to track several epidemics, including pandemic influenza and outbreaks of other infectious diseases [3, 4]. Geographic analysis of these data has enhanced these surveillance methods by revealing critical hotspots of disease, identifying environmental associations, and directing local interventions [5, 6]. Through recent research, we have shown that geospatial analysis (or the application of statistical analytic techniques on spatially related data) and emergency claims data can be used for surveillance of outbreaks or

Abbreviations: ICD-9, International Classification of Diseases, Ninth Revision; SPARCS, Statewide Planning and Research Cooperative System.

Received 2 January 2018
Accepted 29 March 2018
First Published Online 17 April 2018

pandemic illnesses but can be equally important in identifying geographic patterns of chronic diseases like diabetes [7].

The use of emergency claims data provides access to a large population sample with data already collected though several national and state-level databases [8]. The Centers for Disease Control and Prevention estimates that one in five adults visits an emergency department in a given year [9]. In New York City, the 2009 Child Community Health Survey found that 30% of children had visited an emergency department at least once [10]. By using unique identifiers to track individuals across hospitals, several years of emergency claims data can provide health records for more than half of the New York City population [7, 11].

In this study, we investigated geographic disparities in type 1 and type 2 diabetes prevalence among adults and children in New York City using emergency claims data. Whereas adult type 2 diabetes is prevalent, adult type 1 diabetes and pediatric diabetes are relatively rare [2, 12]. Therefore, a large population sample is needed to accurately estimate the prevalence of diabetes at a local level [13]. Several years of emergency claims data can provide this large sample, and prior studies have developed algorithms to differentiate patients with type 1 vs type 2 diabetes using claims data [14, 15]. Here, we perform a geographically detailed estimation of diabetes prevalence in New York City and identify associations of this prevalence with neighborhood-level factors.

1. Research Design and Methods

A. Study Design and Setting

Using cross-sectional emergency claims data, we identified unique New York City residents with a history of type 1 or type 2 diabetes by diagnosis codes and geocoded their location by home addresses. Geospatial analysis was used to estimate type 1 and type 2 diabetes prevalence among adults and children by New York City Census tract. Multivariable regression was then performed to identify demographic, socioeconomic, and food environment factors associated with higher diabetes prevalence.

B. Data Sources

The Statewide Planning and Research Cooperative System (SPARCS) is an all-payer, New York State claims database that includes all inpatient hospitalizations and emergency department visits [16]. Using our previously validated method of emergency department surveillance, we estimated neighborhood-level adult diabetes prevalence with similar accuracy to traditional health surveys [7]. In this study, we use this method to separately analyze the prevalence of type 1 and type 2 diabetes among adults and children using SPARCS data from 2009 to 2013.

To identify sociodemographic factors associated with geographic disparities in type 1 and type 2 diabetes prevalence, we obtained population characteristics from the 2009 to 2013 American Community Survey 5-year estimates. We used Census tract level data for the proportion of residents that were elderly, female, African American, or Hispanic to provide demographic characteristics. We also included median household income, high school graduation rate, and employment rate as socioeconomic factors.

To characterize the restaurant food environment from 2009 to 2013, we used restaurant inspection data from the New York City Department of Health and Mental Hygiene. Fast food "swamps" were identified by the proportion of nearby restaurants categorized as fast food [17]. After excluding nonrestaurants and collapsing observations to unique restaurants by name and location, we identified fast food restaurants by a venue marked as fast food or service marked as take-out or counter service only.

For the retail food environment, we used inspection data from the New York State Department of Agriculture and Markets for 2009 to 2013. Retail food "swamps" were identified as the proportion of retail food stores categorized as bodegas or small convenience stores, which often have poorer food choices than larger grocery stores [18]. Consistent with previously published definitions, we used a cutoff of 2000 ft^2 to identify these stores [19]. For both food environment measures, we counted restaurants and food stores within a 1-mile radius of each Census tract's centroid based on prior studies [20]. A sensitivity analysis was also performed to test half-mile and 2-mile distances [21].

C. Participants

Our study included adults (≥18 years) and children (≤17 years) who visited a New York State emergency department at least once from 2009 to 2013. We included patients with a home address that matched to Census tracts. We excluded patients from correctional facilities and nursing homes to capture a noninstitutionalized population. Emergency departments included all 911-receiving, general acute care New York City hospitals.

D. Main Outcomes

Our main outcome was type 1 and type 2 diabetes prevalence among adults and children by Census tract. To identify cases among unique individuals, identifiers from SPARCS were used to match the same individual across different hospitals and multiple emergency department visits. Individuals with diabetes were identified if they ever had a primary or secondary International Classification of Diseases, Ninth Revision (ICD-9) diagnosis code with the prefix 250.

Consistent with algorithms developed in prior studies, individuals were identified as having type 1 diabetes if they had 50% or more of their diabetes diagnosis codes listed as type 1 (last digit of ICD-9 code of 1 or 3) or were <10 years old. Based on pediatric studies, this algorithm was 95.6% sensitive and 92.4% specific [14]. Individuals were identified as having type 2 diabetes if they had <50% of their diabetes diagnosis codes listed as type 1 and were at least 10 years old. This algorithm was 88.4% sensitive and 88.5% specific among children [14]. Several studies have shown only a modest benefit of more complex algorithms that included data from laboratory values or prescription records not available in claims data [14, 15].

E. Statistical Analysis

Given the low prevalence rates of adult type 1 diabetes and pediatric type 1 and type 2 diabetes, we used spatial empirical Bayes methods to calculate spatially smoothed rates [22]. This method uses estimation that borrows strength from a prior distribution based on local observations. In our analysis, we used first-order Queen contiguity, which means that only adjoining tracts were used to correct variance instability for prevalence estimates within each Census tract.

To perform our multivariable analysis of factors associated with higher type 1 or type 2 diabetes prevalence among adults and children, we used the proportion of elderly, female, African American, and Hispanic residents; median household income; high school graduation rates; and employment rates by Census tract from the American Community Survey. We also included measures of the food environment described above for the proportion of fast food restaurants and bodega retail stores. Prior to inclusion in our model, factors were tested to ensure the absence of multicollinearity.

A probit fractional regression model with robust standard errors was used because our main outcome of diabetes prevalence followed β probability distributions but also had occasional zeros. We also used an adjusted P value of <0.0125 using Bonferroni correction to account for the four separate regression analyses.

We then used a margins analysis to estimate the relative effect that each variable had on diabetes prevalence. For each type of diabetes, a ratio of prevalence was calculated between the highest and lowest values for each population characteristic while keeping all other factors at their means. Thus, these ratios compared diabetes prevalence across the range of values for each variable.

The lower and upper bounds for each variable ranged from 0% to 100% for all variables with the following exceptions: elderly, upper bound 48.1%; female, range 26.6% to 71.9%; African

American, upper bound 99.7%; Hispanic, upper bound 97.3%; income, range $10,114 to $250,001; graduation rate, lower bound 29.4%; and employment rate, lower bound 51.8%. Thus, our subsequent margins analysis comparatively evaluated the estimated diabetes prevalence at the lower vs upper bounds of each variable, keeping all other variables at their means.

Consistent with published tract level multivariable analyses, we excluded tracts where the estimated population error was greater than half of the total number of residents estimated in each Census tract [23]. This exclusion was performed to reduce the influence of areas where the Census did not survey enough residents to provide confident estimates [24]. Of 2167 Census tracts, this exclusion applied to 40 tracts with zero population (mostly parks and airports) and 27 tracts with significant sampling error.

An additional three Census tracts with substantial error in estimates for the proportion of elderly residents and employment rate were also excluded. In addition, to account for Census tracts with an inadequate number of individuals identified in emergency department surveillance, we excluded Census tracts that did not have at least 100 individuals to estimate diabetes prevalence. This additional exclusion applied to only three tracts among adults and 25 tracts among children.

Statistical analyses were performed in Stata 14.2 (StataCorp, College Station, TX). Geographic analysis was performed using ArcGIS Desktop 10.3.1 (ESRI, Redlands, CA) and GeoDa 1.8 (Center for Spatial Data Science, University of Chicago, Chicago, IL). Our study protocol was approved by the Institutional Review Board at the New York University School of Medicine.

2. Results

A. Study Population

We identified 5.0 million unique adults and 1.6 million unique children with a New York City address who visited an emergency department at least once from 2009 to 2013. This represents a substantial majority of the estimated 6.5 million adults and 1.8 million children in New York City based on Census data. Using diagnosis codes, we identified 11,561 adults with type 1 and 528,862 adults with type 2 diabetes. In our sample, adult type 1 diabetes prevalence was 0.23%, and adult type 2 diabetes prevalence was 10.5%. We also identified 3333 children with type 1 and 1794 children with type 2 diabetes. In our sample, pediatric type 1 diabetes prevalence was 0.20% for ages 0 to 17, and pediatric type 2 diabetes prevalence was 0.11% for ages 10 to 17 (Table 1).

Adults and children with type 1 diabetes were generally younger than those with type 2 diabetes. There was a substantially higher proportion of children with type 2 diabetes who were African American (42%) compared with the proportion of children with type 1 diabetes who were African American (31%). Among adults, there was a substantially higher proportion of adults with type 2 diabetes insured by Medicare (44%) compared with the proportion of adults with type 1 diabetes insured by Medicare (22%).

B. Multivariable and Margins Analysis

We used spatial empirical Bayes estimation to smooth rates of diabetes by New York City Census tract (Fig. 1 and 2). In our multivariable analysis (Fig. 3 and Supplemental Table 1), we found that type 1 diabetes prevalence among adults was 0.91 times lower in the mostly African American neighborhoods ($P = 0.007$) and 0.77 times lower in the mostly Hispanic neighborhoods ($P < 0.001$). Among children, type 1 diabetes prevalence was 0.79 times lower in the mostly African American neighborhoods ($P = 0.001$). Higher type 1 diabetes prevalence among children was also associated with higher-income neighborhoods ($P = 0.001$). For type 2 diabetes, higher adult prevalence was identified in lower-income neighborhoods ($P < 0.001$) and in neighborhoods with a higher proportion of elderly residents ($P < 0.001$). Among

Table 1. Characteristics of Adults and Children with Type 1 vs Type 2 Diabetes in New York City Based on Emergency Department Data From 2009 to 2013

	Adults		Children	
Population Characteristics	**Type 1 Diabetes**	**Type 2 Diabetes**	**Type 1 Diabetes**	**Type 2 Diabetes**
Prevalence, %	0.23	10.5	0.20	0.11
Individuals, n	11,561	528,862	3333	1794
Age groups, %				
0–5 y	0	0	24	0
6–12 y	0	0	43	21
13–17 y	0	0	33	79
18–44 y	51	13	0	0
45–64 y	30	42	0	0
≥65 y	19	45	0	0
Sex				
Male	51	46	49	42
Female	49	54	51	58
Race/ethnicity				
White	43	39	34	25
African American	30	32	31	42
Hispanic	23	23	31	30
Asian	4	6	4	3
Insurance				
Private	28	19	35	30
Medicare	22	44	0	0
Medicaid	35	26	57	58
Uninsured	15	11	8	12

children, type 2 diabetes was 1.92 higher in the mostly African American neighborhoods ($P < 0.001$), which was the only factor associated with higher pediatric type 2 diabetes prevalence.

As for the food environment, we found that the prevalence of all types of diabetes was higher in fast food swamps ($P < 0.001$), except for pediatric type 2 diabetes. In areas with the highest proportion of fast food restaurants (Supplemental Fig. 1), there was a 1.55 times higher adult type 1 diabetes prevalence, a 2.52 times higher adult type 2 diabetes prevalence, and a 2.03 times higher pediatric type 1 diabetes prevalence compared with areas without fast food. In a univariate analysis, pediatric type 2 diabetes was 2.91 times higher in the food swamps with the highest proportion of fast food restaurants. However, in line with the multivariable approach used in all other results, this association was no longer statistically significant after adjustment for all other factors included in our multivariable analysis. In addition, changing the distance of restaurants counted to a half-mile or 2 miles did not change the statistical significance of results reported for fast food swamps.

We did not find any influence of retail food swamps or the proportion of bodegas or small convenience retail food stores (Supplemental Fig. 2) on the prevalence of type 1 or type 2 diabetes among adults or children using a distance of 1 mile. In our sensitivity analyses, we did not find that changing the distance of retail food stores counted within a half-mile changed the statistical significance of results reported above. However, using 2 miles to count retail food stores resulted in retail food swamps being associated with a lower prevalence of adult type 1 diabetes ($P = 0.004$) but higher prevalence of pediatric type 1 diabetes ($P = 0.002$).

3. Discussion

In our study, we used emergency claims data to identify geographic disparities in type 1 and type 2 diabetes prevalence among adults and children in New York City. Identifying specific communities with a higher burden of diabetes is critical for exploring associated environmental factors and health behaviors [25]. Without these geographically detailed surveillance

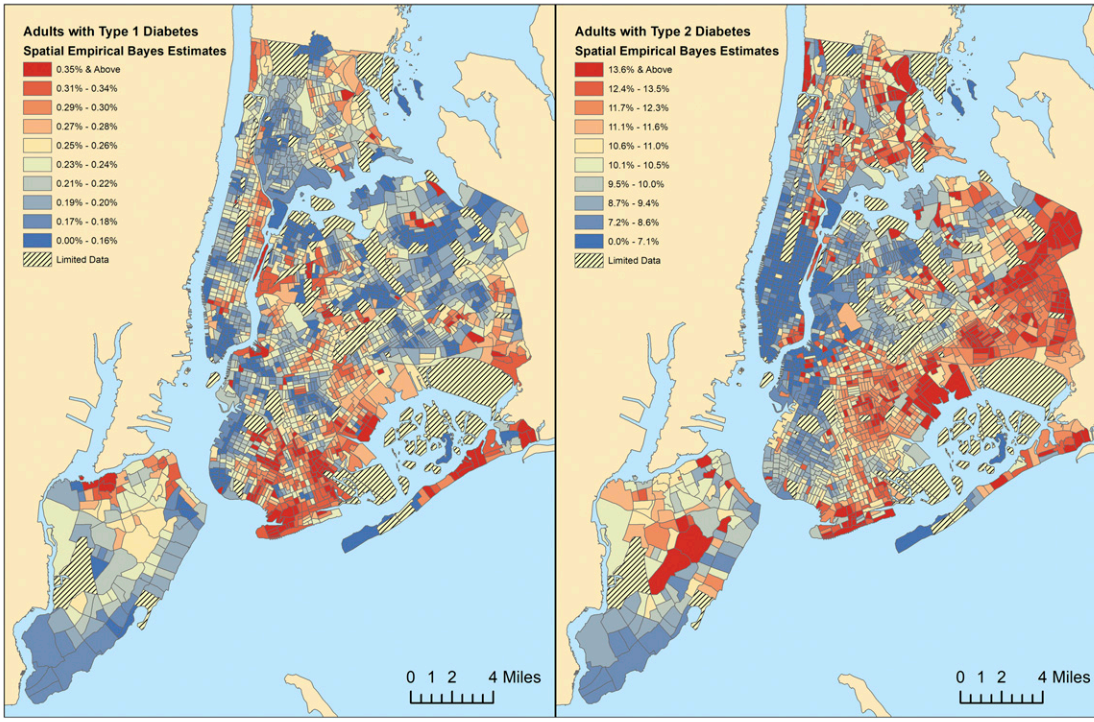

Figure 1. Spatial empirical Bayes estimates for prevalence of type 1 vs type 2 diabetes among adults by Census tract. Spatially smoothed rates of (a) type 1 vs (b) type 2 diabetes prevalence among adults.

methods, these areas would go undetected by traditional health surveys, which lack the sample size to estimate diabetes prevalence at a local level [26].

Using emergency department surveillance, we can capture health data for an overwhelming majority of New York City adults and children with 5 years of claims data. Because we only analyzed emergency patients, there is inherent sampling bias that skews observations toward those more likely to visit an emergency department for care (*e.g.*, patients on Medicaid) [23]. However, this bias is balanced by our approach of analyzing individuals using small geographic areas. Local neighborhoods tend to be more demographically and socioeconomically homogeneous, and these over-represented subpopulations are attributed only to the specific Census tracts where they live using geocoded home addresses [11].

Furthermore, our citywide estimates of diabetes prevalence are similar to those based on available health survey data and national studies. In 2013, the New York City Community Health Survey estimated overall adult diabetes prevalence at 10.7%, which is the same result as adding our estimates of adult type 1 and type 2 diabetes at 0.23% and 10.5% [27]. Using National Health and Nutrition Examination Survey data, a recent study estimated the prevalence of adult type 1 diabetes to be ~0.27% [12]. The similarity with our estimates is striking and consistent with prior research that has demonstrated that an ICD-9 diagnosis code for diabetes in emergency department records can be 99% sensitive and 95% specific for identifying individuals with diabetes [28].

In 2009, the SEARCH for Diabetes in Youth Study, the most comprehensive epidemiological study of diabetes among children, estimated a pediatric prevalence of type 1 diabetes at 0.19% and type 2 diabetes at 0.05% [2]. Using emergency department surveillance in New York City, we estimated a pediatric prevalence of type 1 diabetes at 0.20% and type 2 diabetes at 0.11%. Our estimate for pediatric type 2 diabetes was much higher, but this may be due to the lower specificity of ICD-9 codes for pediatric type 2 diabetes or the cultural diversity of New York City (Supplemental Fig. 3) [14, 29].

In our analysis of demographic and socioeconomic factors associated with higher diabetes prevalence, many of our findings are consistent with prior research [30]. We found that

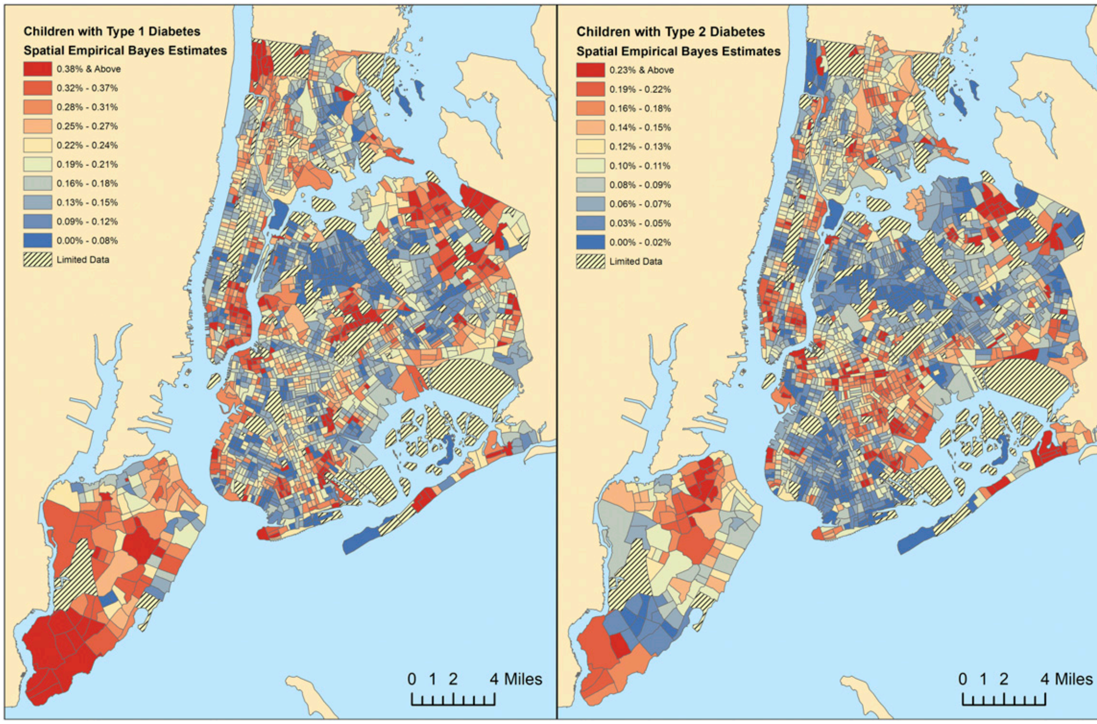

Figure 2. Spatial empirical Bayes estimates for prevalence of type 1 vs type 2 diabetes among children by Census tract. Spatially smoothed rates of (a) type 1 vs (b) type 2 diabetes prevalence among children.

pediatric type 1 diabetes prevalence was higher in areas of higher income [31, 32]. We also found that type 1 diabetes prevalence was generally lower among minorities, whereas type 2 diabetes among children was predominantly concentrated in African American neighborhoods [12, 33]. In fact, after multivariable adjustment, the only predictor of higher pediatric type 2 diabetes was living in a neighborhood with a higher proportion of African American residents [34].

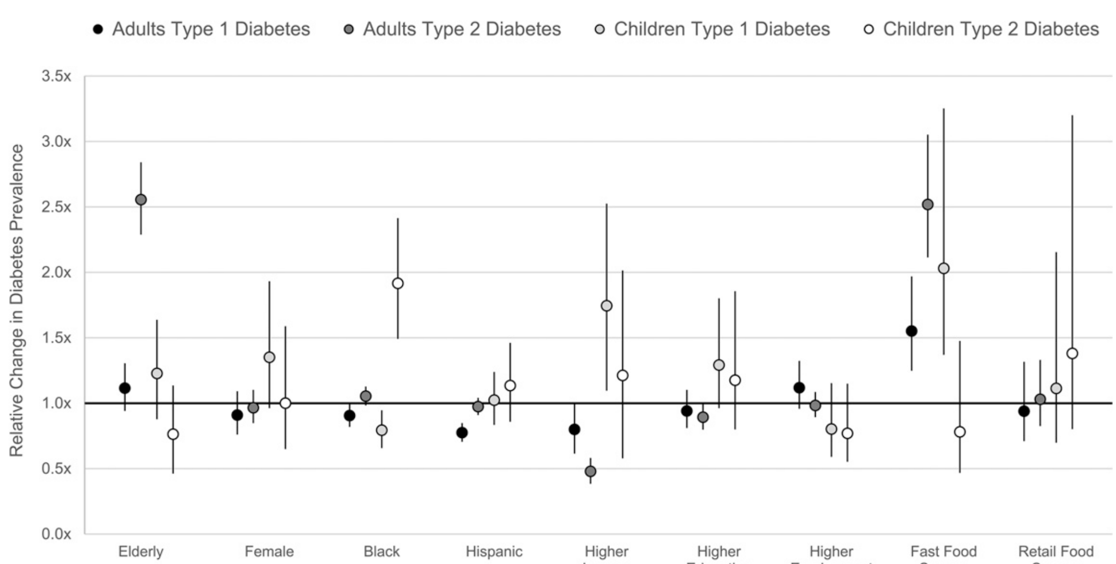

Figure 3. Relative change in the prevalence of type 1 and 2 diabetes among adults and children between highest and lowest values for a given population characteristic while adjusting for all other demographic, socioeconomic, and environmental factors.

On the other hand, we did not find a higher prevalence of adult type 2 diabetes in African American neighborhoods. In a univariate analysis, adult type 2 diabetes prevalence was 1.37 times higher in African American neighborhoods. However, this result was not statistically significant when other factors were included. In our multivariate analysis, older age and lower income were the only demographic or socioeconomic factors associated with higher adult type 2 diabetes prevalence.

In terms of the food environment, we did not find a consistent effect of the retail food environment as measured by the proportion of bodegas or small convenience retail food stores. This result may suggest that the retail food environment does not have a strong association with local diabetes prevalence or may be due to how we measured the retail food environment. Although a half-mile has previously been used in retail food environments studies, there is evidence that some people do not shop at a nearby store and may be willing to travel further for groceries [20].

We were surprised to find that there was a significantly higher prevalence of type 1 diabetes among adults and children in areas identified as fast food swamps [35]. In a visual univariate comparison of diabetes prevalence and food environment maps, one finds that not all clusters of higher type 1 diabetes prevalence were located in food swamps. However, many of these areas, for instance among children, were Census tracts where other sociodemographic factors (fewer minorities or higher income) from our model did predict higher pediatric type 1 diabetes prevalence.

There is increasing belief that certain environmental influences contribute sharply to the increased prevalence of type 1 diabetes [36]. There is some emerging literature suggesting that pregnant women in adverse food environments may have a higher likelihood of gestational diabetes, which may affect health outcomes in their offspring [17, 37]. Given the high prevalence of obesity among children with type 1 diabetes, the association we identified between the food environment and type 1 diabetes merits further research [38].

As expected, we also found that areas identified as fast food swamps had a higher prevalence of type 2 diabetes among adults [21, 39]. However, we did not anticipate that the association between fast food swamps and pediatric type 2 diabetes prevalence would not be significant after adjustment in a multivariable model. Instead, the only factor that we found associated with higher type 2 diabetes prevalence among children was the proportion of African American residents. Although this finding may be a type II statistical error (incorrectly retaining a null hypothesis), we see a clear and strong univariate relationship between pediatric type 2 diabetes prevalence and an adverse food environment that disappears once the proportion of African American residents is included in a multivariable model. Overall, our results suggest that the physical food environment may not play as strong a role in characterizing the risk of type 2 diabetes among children and that other factors, such as genetics, health behaviors, environmental exposures, or family influences, may play more important roles [40, 41].

A. Limitations

Our data sources may have contained errors that contributed to inaccuracies in our estimates. Our study is observational and may have been confounded by omitted variables. Associations identified cannot be considered as evidence of causation. Our use of emergency claims data means that our study population is biased toward individuals more likely to use an emergency department for care. However, over-represented subpopulations are appropriately attributed to the neighborhoods where they live using geocoded addresses. Finally, our study setting was New York City, a unique urban environment; thus, our findings may not be generalizable to other regions.

In summary, in this study we have shown that emergency claims data are a valuable tool in health surveillance, specifically regarding rates of type 1 and type 2 diabetes, for an overwhelming majority of adults and children in New York City. Our estimates of diabetes prevalence are similar to those found in health survey data and national studies, and our analysis of demographic and socioeconomic factors associated with higher diabetes prevalence is consistent with prior research. However, in our study living in a majority African American community was the only predictor of higher pediatric type 2 diabetes prevalence. Furthermore, we found that there was a higher prevalence of type 1 diabetes among adults and children in areas identified as fast food swamps. As pediatric and adult diabetes rates continue to rise, our data suggest that linkages beyond genetics, such as environmental exposures, should be considered for type 1 diabetes, whereas a more thorough investigation of genetics, health behaviors, and cultural influences should be considered for type 2 diabetes.

Future studies should seek to further validate these methods of estimating type 1 and type 2 diabetes prevalence among adults and children using alternative data sources. In addition, detailed neighborhood-level studies of diabetic hotspots should investigate reasons (beyond already measured sociodemographic and environmental factors) why these communities have a disproportionately higher burden of diabetes. Examples of these neighborhood-level factors could include cultural dietary patterns specific to a local community, engrained patterns of belief around how one develops diabetes that may be incorrect, or other unidentified influences that cause certain neighborhoods to face a much higher burden of diabetes.

Acknowledgments

Financial Support: This study was funded by National Institute of Diabetes and Digestive and Kidney Disease Grant K23DK110316 (to D.C.L.) to study environmental factors associated with diabetes among adults, National Institute of Diabetes and Digestive and Kidney Disease Grant R01DK097347 (to B.E.) to study the impact of the food environment on childhood obesity, and Juvenile Diabetes Research Foundation Grant 1-INO-2017-449-A-N (to D.C.L.) to study environmental factors associated diabetes among children.

Author Contributions: D.C.L., M.P.G., and B.E. participated in study conception and design. D.C.L. and A.J.V. acquired the data. D.C.L., M.P.G., A.G., M.O., A.J.V., S.P.W., J.E.R., M.A.S., and B.E. analyzed and interpreted the data. D.C.L. drafted the manuscript. M.P.G., A.G., M.O., A.J.V., S.P.W., J.E.R., M.A.S., and B.E. critically revised the manuscript for intellectual content. D.C.L. is the guarantor of this work and had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Correspondence: David C. Lee, MD, 462 First Avenue, Room A345, New York, New York 10016. E-mail: david.lee@nyumc.org.

Disclosure Summary: The authors have nothing to disclose.

References and Notes

1. Selvin E, Parrinello CM, Sacks DB, Coresh J. Trends in prevalence and control of diabetes in the United States, 1988-1994 and 1999-2010. *Ann Intern Med*. 2014;**160**(8):517–525.
2. Dabelea D, Mayer-Davis EJ, Saydah S, Imperatore G, Linder B, Divers J, Bell R, Badaru A, Talton JW, Crume T, Liese AD, Merchant AT, Lawrence JM, Reynolds K, Dolan L, Liu LL, Hamman RF; SEARCH for Diabetes in Youth Study. Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA*. 2014;**311**(17):1778–1786.
3. Hiller KM, Stoneking L, Min A, Rhodes SM. Syndromic surveillance for influenza in the emergency department: a systematic review. *PLoS One*. 2013;**8**(9):e73832.
4. Balter S, Weiss D, Hanson H, Reddy V, Das D, Heffernan R. Three years of emergency department gastrointestinal syndromic surveillance in New York City: what have we found? *MMWR Suppl*. 2005; **54**:175–180.
5. Curtis AJ, Lee WA. Spatial patterns of diabetes related health problems for vulnerable populations in Los Angeles. *Int J Health Geogr*. 2010;**9**(1):43.
6. Stewart JE, Battersby SE, Lopez-De Fede A, 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. *Int J Health Geogr*. 2011;**10**(1):18.
7. Lee DC, Long JA, Wall SP, Carr BG, Satchell SN, Braithwaite RS, Elbel B. Determining chronic disease prevalence in local populations using emergency department surveillance. *Am J Public Health*. 2015; **105**(9):e67–e74.

8. Livingood WC, Razaila L, Reuter E, Filipowicz R, Butterfield RC, Lukens-Bull K, Edwards L, Palacio C, Wood DL. Using multiple sources of data to assess the prevalence of diabetes at the subcounty level, Duval County, Florida, 2007. *Prev Chronic Dis*. 2010;**7**(5):A108.
9. Centers for Disease Control and Prevention. *Health, United States, 2013: With Special Feature on Prescription Drugs*. Hyattsville, MD: National Center for Health Statistics; 2014.
10. New York City Department of Health and Mental Hygiene. Epiquery: Child Community Health Survey. Available at: https://a816-healthpsi.nyc.gov/epiquery/Child/CCHSIndex.html. Accessed 31 July 2016.
11. Lee DC, Yi SS, Fong HF, Athens JK, Ravenell JE, Sevick MA, Wall SP, Elbel B. Identifying local hot spots of pediatric chronic diseases using emergency department surveillance. *Acad Pediatr*. 2017;**17**(3):267–274.
12. Menke A, Orchard TJ, Imperatore G, Bullard KM, Mayer-Davis E, Cowie CC. The prevalence of type 1 diabetes in the United States. *Epidemiology*. 2013;**24**(5):773–774.
13. Naing L, Winn T, Rusli BN. Practical issues in calculating the sample size for prevalence studies. *Arch Orofac Sci*. 2006;**1**:9–14.
14. Zhong VW, Pfaff ER, Beavers DP, Thomas J, Jaacks LM, Bowlby DA, Carey TS, Lawrence JM, Dabelea D, Hamman RF, Pihoker C, Saydah SH, Mayer-Davis EJ; Search for Diabetes in Youth Study Group. Use of administrative and electronic health record data for development of automated algorithms for childhood diabetes case ascertainment and type classification: the SEARCH for Diabetes in Youth Study. *Pediatr Diabetes*. 2014;**15**(8):573–584.
15. Lawrence JM, Black MH, Zhang JL, Slezak JM, Takhar HS, Koebnick C, Mayer-Davis EJ, Zhong VW, Dabelea D, Hamman RF, Reynolds K. Validation of pediatric diabetes case identification approaches for diagnosed cases by using information in the electronic health records of a large integrated managed health care organization. *Am J Epidemiol*. 2014;**179**(1):27–38.
16. Quan JM. SPARCS: the New York State health care data system. *J Clin Comput*. 1980;**8**(6):255–263.
17. Kahr MK, Suter MA, Ballas J, Ramin SM, Monga M, Lee W, Hu M, Shope CD, Chesnokova A, Krannich L, Griffin EN, Mastrobattista J, Dildy GA, Strehlow SL, Ramphul R, Hamilton WJ, Aagaard KM. Geospatial analysis of food environment demonstrates associations with gestational diabetes. *Am J Obstet Gynecol*. 2016;**214**(1):110.e1–110.e9.
18. Gordon C, Ghai N, Purciel M, Talwalkar A, Goodman A. *Eating Well in Harlem: How Available is Healthy Food?* New York, NY: New York City Department of Health and Mental Hygiene; 2007.
19. Sturm R, Hattori A. Diet and obesity in Los Angeles County 2007-2012: is there a measurable effect of the 2008 "Fast-Food Ban"? *Soc Sci Med*. 2015;**133**:205–211.
20. Hirsch JA, Hillier A. Exploring the role of the food environment on food shopping patterns in Philadelphia, PA, USA: a semiquantitative comparison of two matched neighborhood groups. *Int J Environ Res Public Health*. 2013;**10**(1):295–313.
21. Mezuk B, Li X, Cederin K, Rice K, Sundquist J, Sundquist K. Beyond access: characteristics of the food environment and risk of diabetes. *Am J Epidemiol*. 2016;**183**(12):1129–1137.
22. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. *Geogr Anal*. 2006;**38**(1):5–22.
23. Lee DC, Doran KM, Polsky D, Cordova E, Carr BG. Geographic variation in the demand for emergency care: a local population-level analysis. *Healthc (Amst)*. 2016;**4**(2):98–103.
24. Spielman SE, Folch DC. Reducing uncertainty in the American community survey through data-driven regionalization. *PLoS One*. 2015;**10**(2):e0115626.
25. 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**(4):434–439.
26. 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.
27. New York City Department of Health and Mental Hygiene. Epiquery: Community Health Survey. Available at: https://a816-healthpsi.nyc.gov/epiquery/CHS/CHSXIndex.html. Accessed 31 July 2016.
28. Zgibor JC, Orchard TJ, Saul M, Piatt G, Ruppert K, Stewart A, Siminerio LM. Developing and validating a diabetes database in a large health system. *Diabetes Res Clin Pract*. 2007;**75**(3):313–319.
29. Mayer-Davis EJ, Bell RA, Dabelea D, D'Agostino R Jr, Imperatore G, Lawrence JM, Liu L, Marcovina S; SEARCH for Diabetes in Youth Study Group. The many faces of diabetes in American youth: type 1 and type 2 diabetes in five race and ethnic populations: the SEARCH for Diabetes in Youth Study. *Diabetes Care*. 2009;**32**(Suppl 2):S99–S101.
30. Spanakis EK, Golden SH. Race/ethnic difference in diabetes and diabetic complications. *Curr Diab Rep*. 2013;**13**(6):814–823.

31. Liese AD, Puett RC, Lamichhane AP, Nichols MD, Dabelea D, Lawson AB, Porter DE, Hibbert JD, D'Agostino RB Jr, Mayer-Davis EJ. Neighborhood level risk factors for type 1 diabetes in youth: the SEARCH case-control study. *Int J Health Geogr*. 2012;**11**(1):1.
32. Puett RC, Lamichhane AP, D Nichols M, Lawson AB, A Standiford D, Liu L, Dabelea D, Liese AD. Neighborhood context and incidence of type 1 diabetes: the SEARCH for Diabetes in Youth study. *Health Place*. 2012;**18**(4):911–916.
33. Gaskin DJ, Thorpe RJ Jr, McGinty EE, Bower K, Rohde C, Young JH, LaVeist TA, Dubay L. Disparities in diabetes: the nexus of race, poverty, and place. *Am J Public Health*. 2014;**104**(11):2147–2155.
34. Mayer-Davis EJ, Beyer J, Bell RA, Dabelea D, D'Agostino R Jr, Imperatore G, Lawrence JM, Liese AD, Liu L, Marcovina S, Rodriguez B; SEARCH for Diabetes in Youth Study Group. Diabetes in African American youth: prevalence, incidence, and clinical characteristics: the SEARCH for Diabetes in Youth Study. *Diabetes Care*. 2009;**32**(Suppl 2):S112–S122.
35. Rovner AJ, Nansel TR. Are children with type 1 diabetes consuming a healthful diet?: a review of the current evidence and strategies for dietary change. *Diabetes Educ*. 2009;**35**(1):97–107.
36. Knip M, Veijola R, Virtanen SM, Hyöty H, Vaarala O, Akerblom HK. Environmental triggers and determinants of type 1 diabetes. *Diabetes*. 2005;**54**(Suppl 2):S125–S136.
37. Tam WH, Ma RCW, Ozaki R, Li AM, Chan MHM, Yuen LY, Lao TTH, Yang X, Ho CS, Tutino GE, Chan JCN. In utero exposure to maternal hyperglycemia increases childhood cardiometabolic risk in offspring. *Diabetes Care*. 2017;**40**(5):679–686.
38. Liu LL, Lawrence JM, Davis C, Liese AD, Pettitt DJ, Pihoker C, Dabelea D, Hamman R, Waitzfelder B, Kahn HS; SEARCH for Diabetes in Youth Study Group. Prevalence of overweight and obesity in youth with diabetes in USA: the SEARCH for Diabetes in Youth study. *Pediatr Diabetes*. 2010;**11**(1):4–11.
39. Christine PJ, Auchincloss AH, Bertoni AG, Carnethon MR, Sánchez BN, Moore K, Adar SD, Horwich TB, Watson KE, Diez Roux AV. Longitudinal associations between neighborhood physical and social environments and incident type 2 diabetes mellitus: the Multi-Ethnic Study of Atherosclerosis (MESA). *JAMA Intern Med*. 2015;**175**(8):1311–1320.
40. Thompson SJ, Auslander WF, White NH. Comparison of single-mother and two-parent families on metabolic control of children with diabetes. *Diabetes Care*. 2001;**24**(2):234–238.
41. Auslander WF, Thompson S, Dreitzer D, White NH, Santiago JV. Disparity in glycemic control and adherence between African-American and Caucasian youths with diabetes: family and community contexts. *Diabetes Care*. 1997;**20**(10):1569–1575.