



Neighborhood Socioeconomic Environment and Risk of Type 2 Diabetes: Associations and Mediation Through Food Environment Pathways in Three Independent Study Samples

Diabetes Care 2022;45:798–810 | <https://doi.org/10.2337/dc21-1693>

Lorna E. Thorpe,¹ Samrachana Adhikari,¹ Priscilla Lopez,¹ Rania Kanchi,¹ Leslie A. McClure,² Annemarie G. Hirsch,³ Carrie R. Howell,⁴ Aowen Zhu,⁵ Farrokh Alemi,⁶ Pasquale Rummo,¹ Elizabeth L. Ogburn,⁷ Yasemin Algur,² Cara M. Nordberg,³ Melissa N. Poulsen,³ Leann Long,⁸ April P. Carson,⁵ Shanika A. DeSilva,² Melissa Meeker,² Brian S. Schwartz,^{3,9} David C. Lee,^{1,10} Karen R. Siegel,¹¹ Giuseppina Imperatore,¹¹ and Brian Elbel^{1,12}

OBJECTIVE

We examined whether relative availability of fast-food restaurants and supermarkets mediates the association between worse neighborhood socioeconomic conditions and risk of developing type 2 diabetes (T2D).

RESEARCH DESIGN AND METHODS

As part of the Diabetes Location, Environmental Attributes, and Disparities Network, three academic institutions used harmonized environmental data sources and analytic methods in three distinct study samples: 1) the Veterans Administration Diabetes Risk (VADR) cohort, a national administrative cohort of 4.1 million diabetes-free veterans developed using electronic health records (EHRs); 2) Reasons for Geographic and Racial Differences in Stroke (REGARDS), a longitudinal, epidemiologic cohort with Stroke Belt region oversampling ($N = 11,208$); and 3) Geisinger/Johns Hopkins University (G/JHU), an EHR-based, nested case-control study of 15,888 patients with new-onset T2D and of matched control participants in Pennsylvania. A census tract-level measure of neighborhood socioeconomic environment (NSEE) was developed as a community type-specific z-score sum. Baseline food-environment mediators included percentages of 1) fast-food restaurants and 2) food retail establishments that are supermarkets. Natural direct and indirect mediating effects were modeled; results were stratified across four community types: higher-density urban, lower-density urban, suburban/small town, and rural.

RESULTS

Across studies, worse NSEE was associated with higher T2D risk. In VADR, relative availability of fast-food restaurants and supermarkets was positively and negatively associated with T2D, respectively, whereas associations in REGARDS and G/JHU geographies were mixed. Mediation results suggested that little to none of the NSEE–diabetes associations were mediated through food-environment pathways.

CONCLUSIONS

Worse neighborhood socioeconomic conditions were associated with higher T2D risk, yet associations are likely not mediated through food-environment pathways.

¹Department of Population Health, New York University Grossman School of Medicine, New York, NY

²Department of Epidemiology and Biostatistics, Drexel University Dornsife School of Public Health, Philadelphia, PA

³Department of Population Health Sciences, Geisinger, Danville, PA

⁴Division of Preventive Medicine, University of Alabama at Birmingham School of Medicine, Birmingham, AL

⁵Department of Epidemiology, University of Alabama at Birmingham School of Public Health, Birmingham, AL

⁶Department of Health Administration and Policy, George Mason University, Fairfax, VA

⁷Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

⁸Department of Biostatistics, University of Alabama at Birmingham School of Public Health, Birmingham, AL

⁹Department of Environmental Health and Engineering, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

¹⁰Department of Emergency Medicine, New York University Grossman School of Medicine, New York, NY

¹¹Division of Diabetes Translation, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA

¹²New York University Wagner Graduate School of Public Service, New York, NY

Corresponding author: Lorna E. Thorpe, lorna.thorpe@nyulangone.org

Received 12 August 2021 and accepted 5 January 2022

This article contains supplementary material online at <https://doi.org/10.2337/figshare.18092759>

© 2022 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

Type 2 diabetes (T2D) is a significant cause of morbidity and mortality among adults in the United States. In 2018, the Centers for Disease Control and Prevention (CDC) estimated that 13% of U.S. adults aged ≥ 18 years had diabetes, corresponding to slightly more than 34 million adults (1). In adults, $>90\%$ of diabetes is represented by T2D. Although both prevalence and incidence have increased nationally since the late 1990s, the distribution of and risk for T2D have not been equal in all communities. County-level analyses have highlighted substantial geographic disparities between neighboring counties, with prevalence ranging from 1.5% to 33.0% in 2016 and incidence ranging from 1.2 to 46.2 per 1,000 persons (1–3). These surveillance findings suggest that heterogeneity in neighborhood-level socioeconomic conditions, the built environment, and other neighborhood-level factors may influence geographic variation in T2D risk (4–6).

Studies have consistently identified a robust association between worse neighborhood socioeconomic conditions and risk of T2D (7–11). Mechanisms through which neighborhood socioeconomic status influences T2D risk are not well elucidated but are generally attributed to variation in community characteristics that are, in themselves, socioeconomically patterned. For example, although increased access to supermarkets and reduced exposure to fast-food restaurants have each been associated with decreased T2D risk in longitudinal analyses, the extent to which these contribute to socioeconomic patterns in diabetes risk is unclear (9,12,13). To date, to our knowledge, one study has examined the longitudinal contribution of neighborhood socioeconomic status on availability of healthy food sources in neighborhoods (11). The authors found that residents living in neighborhoods of stable and high socioeconomic status had more options overall, including both fast-food and non-fast-food restaurants, but there was no difference in the number of supermarkets relative to residents living in neighborhoods with low socioeconomic conditions (11). Other studies have explained how neighborhoods with lower socioeconomic status are more likely to have a greater proportion of fast-food and other unhealthy food sources than neighborhoods of higher socioeconomic

status (14,15). However, it appears no studies have formally measured the mediating influence of food environments as a pathway through which neighborhood socioeconomic conditions influence T2D risk.

Two aspects of prior studies that complicate the synthesis of findings on neighborhood attributes relative to T2D risk are the varying definitions of socioeconomic and food environments across studies and that attributes may function differently across regional and community contexts (16). Research has highlighted the importance of considering community type (e.g., urban versus rural) when measuring the effect of neighborhood environments on individual health outcomes (17). These issues pose challenges for 1) building a consistent body of literature across studies, populations, and geographies; 2) harmonizing analytic approaches across different study populations and designs; and, ultimately, 3) designing evidence-based policies to mitigate negative effects of community socioeconomic and food environment attributes. In this study, we aimed to measure the extent to which the relative availability of food-outlet types mediates an established inverse association between neighborhood socioeconomic conditions and T2D occurrence, using data from the Diabetes Location, Environmental Attributes, and Disparities (LEAD) Network (18). As part of this network, three independent study teams used the same environmental data sources and applied harmonized analytic methods to three distinct study samples to collectively evaluate these associations across four strata of community types, ranging from rural to higher density urban areas (18).

RESEARCH DESIGN AND METHODS

Diabetes LEAD Network Partners

The Diabetes LEAD Network (hereafter, “the Network”) is a 5-year (2017–2022) research collaboration of four academic centers that is funded by the Centers for Disease Control and Prevention: one coordinating center (Drexel University) and three study sites: Geisinger and Johns Hopkins University (G/JHU), New York University Grossman School of Medicine (NYU), and the University of Alabama at Birmingham (18). The overarching goal of the Network is to iden-

tify modifiable community determinants of T2D and cardiometabolic conditions using electronic health records (EHRs), administrative claims, and survey data from across the United States. For this analysis, Network partners collaborated to characterize the relationship between neighborhood socioeconomic environment (NSEE) and new T2D diagnoses, and to measure the extent to which attributes of the food environment mediated the NSEE–T2D association. Our objective was to discern what proportion of the total effect of NSEE on T2D risk could potentially be mitigated by modifying food environments. Researchers at Drexel University, the data coordinating center for this study, led discussions across the Network to develop a set of harmonized community factors, health outcomes, and analysis plans; the three study sites applied the harmonized analytic plan to each of their site cohorts and geographies.

NSEE Measurement Development

The primary exposure for these analyses was NSEE. For this study, we developed an index to measure NSEE at the census-tract level. We defined NSEE as a z-score sum of six U.S. census tract–derived variables (the percentage of each of the following: persons with less than a high school education, persons unemployed, households earning $< \$30,000$ /year, households in poverty, households on public assistance, and households with no car) that were selected based on the work of Messer et al. (19) and adapted by Xiao et al. (20). The summed z-scores were scaled to range from 0 to 100, with higher NSEE z-score sums indicating greater socioeconomic disadvantage compared with lower z-score sums. The scaled z-score sum of NSEE was computed using 2000 and 2010 Census data and then put into quartiles by LEAD community type separately for each of the three study samples.

Community Type

Community types were derived by the LEAD Network using a novel modification of the rural-urban commuting area codes from the U.S. Department of Agriculture, developed by the Network and described elsewhere (18,21). Briefly, after collapsing the original 10 rural-urban commuting-area categories into three, we

divided census tracts within urbanized areas into two categories based on land area. This resulted in four LEAD Network-derived community-type categories reflecting distinct categories along a rural-urban continuum: higher density urban, lower density urban, suburban/small town, and rural (18).

Food-Environment Mediator Metrics

We used food-establishment data from the Retail Environment and Cardiovascular Disease (RECV) study (22), which classified health-related neighborhood amenities using the National Establishment Time Series database (1990–2014). The RECV team re-geocoded the National Establishment Time Series data to improve locational accuracy and then assigned establishments to subcategories using Standard Industrial Classification codes, employee and sales information, and chain names obtained from Technomic/Restaurants and Institutions and TDLinx. Details on classification methods are described elsewhere (23).

Food-environment mediator metrics included two relative measures of food availability: 1) the percentage of fast-food restaurants out of the total of food service establishments, and 2) the percentage of supermarkets out of the total of food retail establishments. Fast-food restaurants were defined as quick-service restaurants offering low-preparation-time foods for takeaway or cafeteria-style service (i.e., no-wait service). The supermarkets category was composed of three mutually exclusive food store subcategories: supermarkets, supercenters, and medium-sized grocers. We selected relative food-environment measures because absolute count or density (count per unit area) measures do not account for the presence of other food outlets and because absolute counts of food outlets often mirror population density rather than quality of the food environment, particularly in higher-density urban settings (24,25).

Relative food variables were operationalized as 5-year averages (ending the year prior to study entry) of fast-food or supermarket food outlets relative to all restaurants and food retail establishments, respectively, calculated using a street-network buffer around the population-weighted centroid of the census tracts of study participants' home

addresses in ArcGIS. We selected buffer type and size on the basis of data from the U.S. Department of Agriculture's National Household Food Acquisition and Purchase Survey (26), which calculated average driving distances between individuals' residential addresses and their primary food store in rural and nonmetropolitan community types. Extending these categories to our community-type measure, we derived four community type-tailored (e.g., census-based) buffer sizes (surrounding the population-weighted centroid of the census tract): a 1-mile walking buffer in higher-density urban communities, a 2-mile driving buffer in lower-density urban communities, a 6-mile driving buffer in suburban/small town communities, and a 10-mile driving buffer in rural communities. We separately considered the influence of multiple buffer sizes on observed associations between the food-environment metrics and diet, which are published elsewhere (17).

Community-Level Covariates

Census tract-level covariates included continuous percentages of the population that were non-Hispanic (NH) Black and Hispanic, derived from the decennial Census, and a land-use environment metric, calculated from a multiple-group confirmatory factor analysis based on seven components of the built environment: average block length, average block size, intersection density, street connectivity, establishment density, percentage of developed land, and household density, derived from ESRI 2009 Vintage Street data and computed via ArcGIS Pro 2.3 (18,22,27). Tracts with a higher land-use environment factor score were considered more walkable than tracts with a lower score (22).

Individual-Level Covariates

Each site incorporated individual-level covariates in their models, including age (continuous) and sex (male or female). Age transformations were considered for analyses as necessary, depending on associations between age and T2D risk in the respective study sample for each site. Other individual-level covariates included site-specific measures of race/ethnicity, smoking status, and individual-level socioeconomic status (defined in the next section). Race/ethnicity was

adjusted for as a confounder in models as a social construct that can cause, or be correlated with, many social factors affecting health.

Site-Specific Analytic Plans

NYU Grossman School of Medicine

The Veterans Administration Diabetes Risk (VADR) cohort was developed using the Veterans Affairs EHR VINCI system, described elsewhere (28). This dynamic retrospective cohort included 4.1 million patients who are U.S. veterans with at least two primary care service visits at least 30 days apart prior to cohort entry (between 1 January 2008 and 31 December 2018). Patients were excluded if they had evidence of T2D prior to cohort entry date and censored upon first evidence of T2D diagnosis, death, or loss to follow-up (loss to follow-up was defined as no encounter with the Veterans Affairs system for 2 years). New diagnoses of T2D were defined using T2D encounter diagnoses, medication orders, and laboratory test results (Table 1). Person-time was calculated by subtracting the cohort entry date from the date of censoring. Baseline address was identified as the address closest to the cohort entry date. If an address was not on file before cohort entry date, the first address after cohort entry date was accepted if it was within 2 years. Baseline address was geocoded and spatially joined to the corresponding census tract. NSEE, food environment, and neighborhood confounders were linked to cohort patients by residential census tract. For patients who entered the cohort before 2010, NSEE information based on the 2000 decennial Census data was used; NSEE information, based on 2006–2010 American Community Survey data, was used for patients who entered the cohort during or after 2010. Food-environment mediators were assigned on the basis of the year prior to cohort entry.

To assess mediation by food-environment pathways, the NYU researchers first examined exposure-outcome and mediator-outcome associations using piecewise exponential (PWE) survival models (29). Assuming a constant hazard function within intervals over time, they used generalized linear mixed-effects regression models with a Poisson link

Table 1—Definition of T2D outcomes across the three studies in the Diabetes LEAD Network

VA EHR (NYU)	<p>Incident T2D: meets at least one of the following criteria during the study period:</p> <ol style="list-style-type: none"> 1. At least two separate inpatient or outpatient encounters with T2D ICD-9/10 code 2. Any Rx of T2D medication (excluding metformin or acarbose alone) 3. At least one encounter with a T2D ICD-9/10 code and two elevated ($\geq 6.5\%$) glycosylated hemoglobin levels according to laboratory test results 4. One encounter with T2D ICD-9/10 code and two or more $HbA_{1c} \geq 6.5\%$ (on different days)
REGARDS (UAB)	<p>Prevalent T2D: meets at least one of the following criteria at baseline:</p> <ol style="list-style-type: none"> 1. Fasting glucose level ≥ 126 mg/dL 2. Random glucose level ≥ 200 mg/dL 3. Use of oral or injectable T2D medications or insulin <p>Incident T2D: meets at least one of the following criteria at second visit, among participants without prevalent T2D at baseline:</p> <ol style="list-style-type: none"> 1. Fasting glucose level ≥ 126 mg/dL 2. Random glucose level ≥ 200 mg/dL 3. Use of oral or injectable T2D medications or insulin
G/JHU EHR (G-JHU)	<p>New-onset T2D: meets criterion 1 and at least one other criterion during the study period:</p> <ol style="list-style-type: none"> 1. At least 2 years of contact with the health system prior to the first T2D criterion being met 2. At least two separate encounter dates with T2D ICD-9/10 code or Geisinger specific electronic diagnosis grouper codes 3. At least one T2D medication order after age 10 years, other than metformin or acarbose, if female 4. At least one encounter with T2D ICD-9/10 code and an abnormal laboratory result

ICD, *International Classification of Diseases*; Rx, medication prescription; VA, Veterans Affairs.

function and an offset of the logarithm of time-at-risk during each interval to estimate hazard ratios. PWE survival models were selected instead of frailty models because of their efficiency in estimating mediation effects for generalized linear mixed-effects regression models, using readily available statistical software, but their equivalence has been discussed extensively elsewhere (30,31). NYU researchers used 2-year intervals and county-level random effects. Network-harmonized community-level covariates were linked by census tract and added to the model. Hypothesized individual-level confounders were age, sex, race/ethnicity (NH White, NH Black, Hispanic, Asian, and Other/Unknown), and income or a disability indicator that served as a proxy for socioeconomic status (defined as disabled irrespective of income status, low income/nondisabled, or neither). Mediation models were conducted using the R packages “lme4” and “mediation” (32). Simulation-based 95% CIs were generated using 50 simulations each. Total effects, direct effects, and mediation effects were reported as differences in 1-year T2D incidence risk.

University of Alabama at Birmingham

The Reasons for Geographic and Racial Differences in Stroke (REGARDS) (33) Study is a longitudinal cohort that enrolled adults aged ≥ 45 years at baseline (2003–2007)

from the contiguous United States with oversampling in the Stroke Belt region (i.e., North Carolina, South Carolina, Georgia, Tennessee, Mississippi, Alabama, Louisiana, and Arkansas). For this analysis, 11,208 study participants without prevalent T2D at baseline and who completed the follow-up examination in 2013–2016 were included. The outcome was incident T2D defined as a fasting glucose level ≥ 126 mg/dL or random glucose measurement ≥ 200 mg/dL or use of T2D medication noted at the follow-up examination (Table 1). Follow-up time was measured as the time between the two in-home visits. NSEE 2000 was assigned to the census tract corresponding to participants’ geocoded residential addresses at baseline. Food-environment mediators were assigned using the measure prior to the year of cohort entry.

Poisson mixed models with robust variance estimation were used to model the relationship between NSEE and incident T2D accounting for correlation of participants within census tracts. Linear mixed models were used to investigate food environment as a function of NSEE accounting for participant correlation, and the R package “mediation” was used to assess the mediating effect of food environment on the NSEE–incident T2D relationship (32,34). All models were adjusted for individual-level covariates (age, sex, race [NH White or NH

Black], income [$< \$20,000$, $\$20,000$ – $\$34,999$, $\$35,000$ – $\$74,999$, $\geq \$75,000$, or refused], smoking status [current smoker or not a current smoker]) and LEAD Network–harmonized community-level covariates. Total effects, direct effects, and mediation effects were reported as differences in 1-year T2D incidence risk.

Geisinger/Johns Hopkins University

G/JHU conducted a nested case-control study as previously reported (35). Using Geisinger EHRs, individuals with new onset of T2D ($n = 15,888$) were identified using T2D encounter diagnoses, medication orders, and laboratory test results (Table 1). Control participants ($n = 79,435$, with 65,084 unique persons)—individuals who never met any of the T2D criteria used for cases—were randomly selected with replacement and were frequency-matched to cases (5:1) on age, sex, and year of encounter. At least previous two encounters were required on different days with a primary care provider to ensure we could detect T2D if present. To ensure T2D was of new onset, we required individuals to have at least one encounter with the health system without evidence of T2D ≥ 2 years prior to T2D onset date. NSEE 2000 was used for new case patients and matched control participants in the years 2008–2012 and NSEE 2010

for new case patients and matched control participants in years 2013–2016. Food-environment mediators were assigned for the year prior to onset or the matched encounter date.

Mediation analysis was conducted in R using the “medflex” package (36). This package allows for logistic regression models appropriate for the case-control study design as well as bootstrapping to account for clustering in census tracts (36). A total of 1,000 bootstrap samples were drawn to calculate 95% bootstrap CIs. All analyses were adjusted for age (centered within community type), sex, race/ethnicity (NH Black vs. NH White, Other), medical assistance status (ever or never; a surrogate for family socioeconomic status), and smoking (current or former/never). Age was evaluated for nonlinearity within each LEAD community type. Models were adjusted for linear age in the higher-density urban and lower-density urban community types, and for linear and quadratic age in the rural and suburban/small town community types. We also adjusted for harmonized community-level (census tract-based) variables used by the three study designs. Total effects, direct effects, and mediation effects were reported as odds ratios.

Sensitivity Analyses

Three network-wide sensitivity analyses were conducted. To further evaluate whether differences in NSEE distribution between site samples influenced main effects model results, NSEE information was reset in quartiles by its national tract-level distribution (rather than site-specific distribution), and each site reran the NSEE–T2D models using the data separated in quartiles relative to the national distribution. To determine the robustness of the mediation results according to the choice of measure of neighborhood socioeconomic conditions, all sites also refitted their mediation models using quartiles of percentage below the federal poverty line in LEAD community types, rather than NSEE. To assess consistency of findings on the effect of the food-environment exposures on risk of T2D, results for the REGARDS and G/JHU geographies (census tracts) were rerun using patient data from the VADR cohort, which has patients in the same census tracts.

Site-specific sensitivity analyses were also conducted on the basis of the unique geographic, population, and analytic factors of each site. Examination of main effects of NSEE on new-onset T2D in the VADR sample was repeated using a shared frailty model to confirm consistency of effect estimates, and the mediation model was repeated, adjusting for smoking status, which was not included in the original mediation model, because there were much missing data on this covariate. The REGARDS data were refitted to include a measure characterizing the study’s sampling regions (Stroke Belt vs. non-Stroke Belt) in the mediation model to rule out potential unmeasured confounding of regional effects on NSEE, food environment, and T2D (33). The researchers at the G/JHU site reexamined their results for the subset of participants aged ≥ 20 years ($n = 15,584$ case patients; $n = 77,915$ control patients [person-visits]) to rule out effect-measure modification of the estimates by age, considering differing risk levels and biological development of T2D among young adults compared with older adults.

RESULTS

Demographic characteristics and geographic coverage of participants in each of the three study samples are summarized in Table 2. Briefly, NYU’s VADR cohort included >4.1 million veterans in 71,835 census tracts (98% of 73,056 tracts across the United States, per the 2010 Census). The REGARDS cohort included 11,208 participants in 7,502 census tracts, approximately half of which are located in the southeastern United States. G/JHU’s nested case-control study sample included 15,888 case patients and 79,435 control participants (control participants sampled with replacement) in 785 census tracts in Pennsylvania. The VADR cohort and G/JHU study periods were similar: 2008–2018 and 2008–2016, with an average duration of follow-up of 5.0 and 11.0 years, respectively (time from first clinical encounter to year of T2D onset or matched encounter year); the REGARDS study period was 2003–2016, with an average follow-up of 9.5 years (the time between the two in-home visits). REGARDS cohort participants were, on average, slightly older than VADR or

G/JHU participants (63.0 vs. 59.4 and 54.9 years, respectively). Approximately half of G/JHU and REGARDS study participants were women, whereas VADR participants were predominantly men (only 7.8% were women), reflecting veteran population characteristics. Among study samples, the REGARDS cohort was racially more diverse (32.8% NH Black) than the VADR and G/JHU study samples (16.0% and 1.8%, respectively).

The distribution of NSEE values differed across study sites. NSEE values across tracts represented in the VADR cohort had a wider distribution of values, including a long tail of higher values (indicating worse NSEE) than the other two cohorts, whereas tracts represented in the G/JHU cohort had the smallest range (Supplementary Figure 1). In general, central tendency values varied modestly within community-specific quartiles for each site (Table 2). The relative availability of fast food compared with other restaurant types was approximately similar across LEAD community types among VADR and REGARDS cohorts (range: 26–35%), with a slightly higher mean proportion of restaurants being fast-food in lower density urban and suburban/small-town community types (31–35%) than in higher-density urban or rural settings (26–31%). G/JHU’s case patients and control participants had nearly identical levels of fast-food relative availability within each LEAD community type (18–27%), averaging $\sim 10\%$ lower availability compared with the other study samples. Supermarket relative availability was nearly the same across all three study samples (9–14%).

NSEE–T2D Occurrence Association

We first examined the relationship between NSEE quartiles and risk of T2D in the data from all study sites, adjusting for individual- and community-level confounders (Fig. 1). In most community types, greater socioeconomic disadvantage, measured as higher quartiles of NSEE, was associated with greater T2D occurrence, with one exception. In the higher-density urban community type, a null or nonsignificant inverse association was seen for the most advantaged NSEE quartile across study sites. In the VADR cohort, T2D incidence increased significantly from better to worse NSEE quartiles in the suburban/and small town

Table 2—Baseline characteristics of three samples examining the mediating effects of food environment on the association between neighborhood socioeconomic environment and T2D occurrence: Diabetes LEAD Network

	VADR cohort	REGARDS cohort	G/JHU T2D case-control study population	
			Case patients	Control participants
Counties, <i>n</i>	3,108	1,349	37	
Census tracts, <i>n</i>	71,835	7,502	785	
Study design	Cohort	Cohort	Nested case-control	
Study period	2008–2018	2003–2016	2008–2016	
Follow-up, median (IQR) (years)	5.0 (2.4, 9.0)	9.5 (8.7, 9.9)	11.2 (7.5, 14.1)	11.2 (7.5, 14.1)
Participants, <i>n</i> [†]	4,100,650	11,208	15,888	79,435
New T2D cases, <i>n</i> (%)	539,369 (13.15)	1,409 (12.6)	N/A	N/A
Demographics				
Age, mean (SD) (years)	59.4 (17.2)	63.0 (8.5)	54.9 (15.1)	54.9 (15.3)
Age-groups, <i>n</i> (%) (years)				
10–17	NA	NA	230 (1.5)	1,184 (1.5)
18–34	486,304 (11.9)	NA	1,361 (8.6)	6,771 (8.5)
35–49	637,553 (15.6)	649 (5.8)	4,038 (25.4)	20,189 (25.4)
50–64	1,332,938 (32.5)	5,884 (52.5)	6,202 (39.0)	31,007 (39.0)
65–79	1,087,570 (26.5)	4,321 (38.6)	3,331 (21.0)	16,654 (21.0)
≥80	556,227 (13.6)	354 (3.2)	726 (4.6)	3,630 (4.6)
Sex, women, <i>n</i> (%)	321,013 (7.8)	6,256 (55.8)	7,798 (49.1)	38,988 (49.1)
Race/ethnicity, <i>n</i> (%)[‡]				
Non-Hispanic White	2,783,756 (76.3)	7,534 (67.2)	15,112 (95.1)	76,971 (96.9)
Non-Hispanic Black	584,655 (16.0)	3,674 (32.8)	293 (1.8)	905 (1.1)
Hispanic	189,177 (5.2)	N/A	369 (2.3)	1,094 (1.4)
Asian	34,838 (1)	N/A	63 (0.4)	267 (0.34)
Other/unknown	56,804 (1.6)	N/A	51 (0.32)	198 (0.25)
Health/clinical status				
Smoking status, <i>n</i> (%) [§]				
Current smoker	610,506 (40.3)	1,245 (11.1)	3,272 (20.6)	14,831 (18.7)
Not smoker (former and never smoker)	902,901 (59.7)	9,926 (88.9)	12,223 (76.9)	63,241 (79.7)
BMI, mean (SD) (kg/m ²)	28.7 (5.3)	28.6 (5.6)	36.2 (8.4)	29.3 (6.3)
Hypertension, <i>n</i> (%) ^{¶,#}	2,216,777 (54.1)	5,532 (49.4)	10,228 (64.4)	35,942 (45.3)
HbA _{1c} , mean (SD)**	5.7 (0.6)	N/A	7.3 (1.9)	5.6 (0.4)
Neighborhood characteristics by LEAD community type				
Higher-density urban, <i>n</i>	478,668	1,810	1,039	4,121
NSEE, median (IQR)	20.7 (13.5, 31.4)	24.8 (16.5, 34.0)	23.6 (19.1, 29.3)	23.6 (18.8, 28.3)
NSEE quartiles, <i>n</i> (%)				
1	126,567 (26.5)	415 (22.9)	243 (23.4)	1,082 (26.3)
2	135,038 (28.2)	509 (28.1)	310 (29.8)	1,142 (27.7)
3	115,778 (24.2)	486 (26.9)	288 (27.7)	1,040 (25.2)
4	100,868 (21.1)	400 (22.1)	189 (19.1)	857 (20.8)
Fast-food relative density, mean (SD) (%)	0.26 (0.14)	0.27 (0.16)	0.18 (0.07)	0.19 (0.07)
Supermarket relative density, mean (SD) (%)	0.09 (0.08)	0.11 (0.08)	0.10 (0.06)	0.10 (0.06)
LUE, median (IQR)	−0.07 (−0.59, 0.50)	0.13 (−0.50, 0.64)	0.24 (−0.34, 0.66)	0.24 (−0.38, 0.70)
Non-Hispanic Black, mean (SD) (%)	26.1 (32.5)	64.4 (32.9)	6.2 (7.8)	5.1 (5.9)
Hispanic, mean (SD) (%)	20.2 (23.1)	8.4 (11.4)	5.8 (7.3)	5.8 (8.2)
Lower-density urban, <i>n</i>	1,509,042	4,524	1,890	8,665
NSEE, median (IQR)	10.9 (7.2, 16.6)	14.5 (8.4, 25.3)	17.1 (13.4, 22.0)	16.4 (12.8, 21.7)
NSEE quartiles, <i>n</i> (%)				
1	320,370 (21.2)	948 (21.0)	84 (4.4)	673 (7.8)
2	411,960 (27.3)	1,094 (24.2)	410 (21.7)	2,144 (24.7)
3	430,590 (28.6)	1,205 (26.6)	547 (28.9)	2,486 (28.7)
4	345,444 (22.9)	1,277 (28.2)	849 (44.9)	3,362 (38.8)
Fast-food relative density, mean (SD)	0.31 (0.13)	0.34 (0.15)	0.26 (0.10)	0.27 (0.10)
Supermarket relative density, mean (SD)	0.10 (0.08)	0.13 (0.09)	0.11 (0.05)	0.11 (0.05)
LUE, median (IQR)	0.09 (−0.58, 0.68)	0.00 (−0.63, 0.65)	0.31 (−0.47, 1.02)	0.15 (−0.57, 0.93)
Non-Hispanic Black, mean (SD) (%)	15.1 (22.6)	43.4 (35.9)	2.6 (3.3)	2.5 (2.9)
Hispanic, mean (SD) (%)	12.3 (16.7)	5.0 (7.9)	3.0 (5.3)	2.7 (4.4)

Continued on p. 804

Table 2—Continued

	VADR cohort	REGARDS cohort	G/JHU T2D case-control study population	
			Case patients	Control participants
Suburban/small town, <i>n</i>	919,281	2,224	5,009	24,886
NSEE, median (IQR)	11.5 (7.7, 16.7)	11.3 (6.9, 17.8)	16.1 (10.6, 21.7)	15.0 (9.7, 20.4)
NSEE quartiles, <i>n</i> (%)				
1	209,164 (22.8)	482 (21.7)	779 (15.6)	4,817 (19.4)
2	262,205 (28.5)	493 (22.2)	906 (18.1)	5,140 (20.7)
3	263,587 (28.7)	560 (25.2)	1,537 (30.7)	7,509 (30.2)
4	184,071 (20.0)	689 (31.0)	1,787 (35.7)	7,420 (29.8)
Fast-food relative density, mean (SD)	0.32 (0.10)	0.35 (0.11)	0.26 (0.07)	0.26 (0.07)
Supermarket relative density, mean (SD)	0.11 (0.05)	0.13 (0.06)	0.12 (0.05)	0.12 (0.05)
LUE, median (IQR)	−0.06 (−0.70, 0.54)	0.02 (−0.64, 0.57)	0.71 (−0.54, 1.85)	0.26 (−0.68, 1.82)
Non-Hispanic Black, mean (SD) (%)	9.3 (15.3)	28.0 (30.3)	1.9 (4.4)	1.8 (4.3)
Hispanic, mean (SD) (%)	8.2 (13.4)	3.3 (5.3)	1.5 (2.6)	1.5 (2.5)
Rural, <i>n</i>	1,193,659	2,650	7,950	41,763
NSEE, median (IQR)	17.9 (13.5, 23.1)	22.6 (16.2, 30.0)	16.2 (13.6, 19.1)	16.0 (13.2, 18.6)
NSEE quartiles, <i>n</i> (%)				
1	290,901 (24.4)	500 (18.9)	2,256 (28.4)	13,010 (31.2)
2	321,373 (26.9)	531 (20.0)	1,802 (22.7)	9,744 (23.3)
3	322,193 (27.0)	638 (24.1)	2,470 (31.1)	12,096 (29.0)
4	259,049 (21.7)	981 (37.0)	1,422 (17.9)	6,931 (16.6)
Fast-food relative density, mean (SD)	0.29 (0.15)	0.27 (0.20)	0.23 (0.10)	0.23 (0.10)
Supermarket relative density, mean (SD)	0.12 (0.08)	0.13 (0.12)	0.14 (0.06)	0.14 (0.06)
LUE, median (IQR)	−0.04 (−0.53, 0.59)	0.11 (−0.42, 0.79)	0.27 (−0.11, 0.67)	0.28 (−0.10, 0.67)
Non-Hispanic Black, mean (SD) (%)	6.9 (13.5)	31.5 (27.5)	2.2 (4.4)	2.2 (4.3)
Hispanic, mean (SD) (%)	4.9 (9.7)	2.6 (5.1)	2.2 (3.6)	2.0 (3.3)

HbA_{1c}, glycosylated hemoglobin; IQR, interquartile range; LUE, land-use environment; N/A, not applicable. Baseline clinical measures except hypertension were as follows: VADR: most recent measure prior or on cohort entry date; REGARDS: at enrollment; and G/JHU: BMI and HbA_{1c}: closest within 1 year prior to T2D diagnosis or matched encounter date. †For G/JHU, data represent person-visits, control participants sampled with replacement. There were 69,084 unique people in the control group. ‡Missing data: VADR: *n* = 451,420. §Missing data: VADR: *n* = 2,587,243; REGARDS: *n* = 37. ||Missing data: VADR: *n* = 183,068 patients; REGARDS: *n* = 32; G/JHU: *n* = 2,397 case patients and 24,491 control participants. ¶Hypertension definitions are as follows for the cohorts: VADR: at least one ICD-9 or ICD-10 code for hypertension ever or elevated blood pressure ($\geq 140/90$ mmHg) within 2 years of cohort entry date; REGARDS: systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg or self-reported use of medication to control blood pressure; and G/JHU: ICD-9 code 401 or ICD-10 code I10 prior to T2D onset. Patients with secondary hypertension (ICD-9 code 405.x or ICD-10 code I15.x) were excluded. #Missing data: REGARDS: *n* = 16. **Missing data: VADR: *n* = 2,423,788 patients; G/JHU: *n* = 8,391 case patients and 75,280 control participants.

and rural community types, although significant associations were also observed in higher- and lower-density urban community types, the pattern was less consistent. T2D risk was generally elevated in more disadvantaged NSEE quartiles across community types in the REGARDS study, but CIs were wide and associations were largely nonsignificant. In the G/JHU cohort, associations were strongest within the lower-density urban community type.

Food Environment—T2D Occurrence Association

In mediator-outcome analytic models (Supplementary Table 1A), there was a statistically significant increased risk of T2D incidence with greater exposure to fast-food restaurants in the VADR cohort across all community types. Similar directionality, though nonsignificant, was seen in the data from higher-density urban and lower-density urban areas in

the REGARDS study, but associations were null in suburban/small town and rural areas. In the G/JHU sample, we observed lower odds of T2D with higher fast-food metrics in all community types except rural, a finding that was confirmed when associations were analyzed using the subsample of the VADR cohort living in the same Pennsylvania census tracts as G/JHU participants (Supplementary Table 1A and B). Supermarkets were associated with a significant decrease in T2D risk in the VADR cohort across all community types. The results from the REGARDS and G/JHU samples were mixed but not statistically significant.

Food-Environment Mediation Analyses

Overall, mediation results across the three study samples suggested that the relationships of NSEE with T2D onset were not mediated through the relative

availability of fast-food restaurants (Table 3). A notable exception was for the lower-density urban community type in the VADR cohort, where the fast-food environment metric mediated 2–3% of the total effect between NSEE and T2D risk in the second, third, and fourth quartiles compared with the first quartile.

Sensitivity Analyses

To assess whether differences in NSEE distribution between site samples influenced main effects model results, the NSEE–T2DM model was reanalyzed using the reset NSEE quartiles by an all-US distribution; effect sizes for the NSEE–T2DM association followed the same overall pattern as our main effects models (Supplementary Fig. 2). All sites also reran the main mediation models, replacing the NSEE measure with quartiles of percentage of the resident population living below the federal poverty line, and

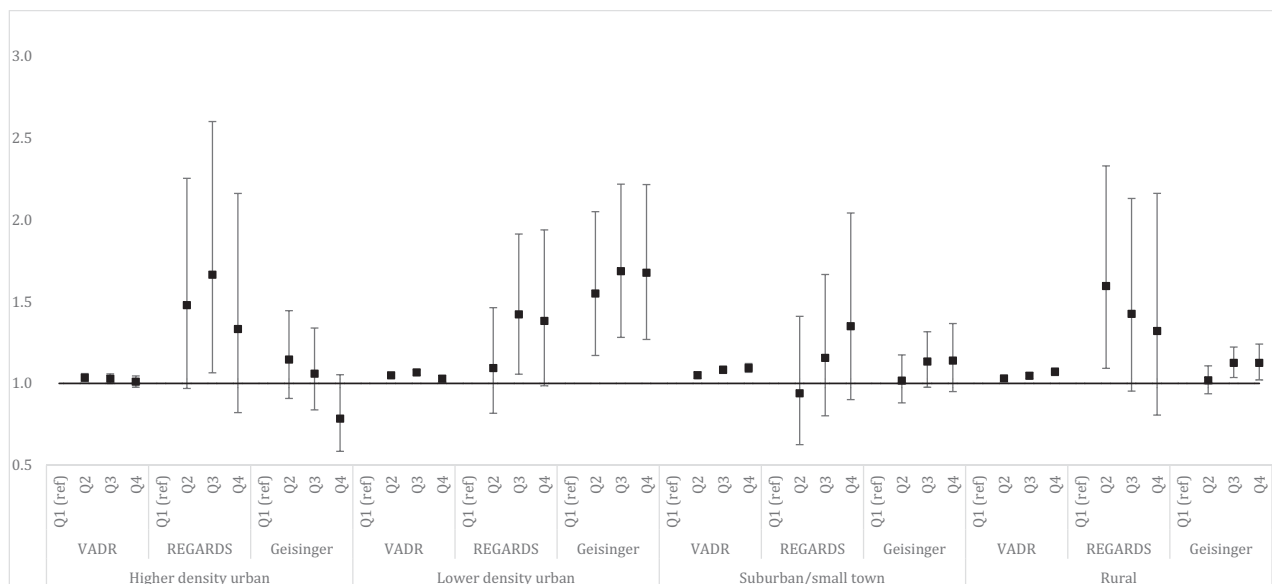


Figure 1—Associations of NSEE (in quartiles separately derived by Diabetes LEAD Network community type so that quartiles cannot be compared across LEAD community types) with risk of new-onset T2D. Effect is presented as hazard ratio for VADR, risk ratio for REGARDS, and odds ratio for G/JHU. VADR data were adjusted for baseline age, quadratic age, race/ethnicity, sex, income/disability flag, neighborhood land-use environment, and percentages of Hispanic and Black participants. REGARDS data were adjusted for age, race, sex, income, current smoking, neighborhood land-use environment, and percentages of Hispanic and Black participants. G/JHU data were adjusted for age, sex, race/ethnicity, medical assistance, smoking, neighborhood land-use environment, percentages of Hispanic and Black participants. To address nonlinearity, the higher-density urban model included a quadratic fast-food variable and the suburban/small town model included quadratic and cubic fast-food variables. Q, quartile.

results remained consistent for all study sites (Supplementary Table 2). When patient data from the VADR cohort were used to assess the effect of the food-environment exposure metrics on risk of T2D in the REGARDS and G/JHU census tracts, there were no substantial differences from what those sites observed using their own patient records (Supplementary Table 1B). In the VADR cohort, no major differences were seen with respect to main effects of NSEE on T2DM, using frailty survival models versus PWE survival models (Supplementary Fig. 2), and adjusting for smoking status did not change the magnitude or direction of effect estimates of the mediation model (Supplementary Table 3). For the REGARDS cohort, adding U.S. region also did not substantially change results by community type (Supplementary Table 4). When excluding participants younger than 20 years from the mediation analyses, there were no substantial changes in either associations or inferences in the G/JHU sample (Supplementary Table 5).

CONCLUSIONS

Through this collaboration, we investigated whether the effects of NSEE on T2D risk could be explained, in part, by

attributes of the food environment in three large, geographically distinct study samples comprising the Diabetes LEAD Network. To our knowledge, this is the first empirical study to examine these potential mediating pathways as a step toward understanding geographic and socioeconomic disparities in T2D risk. Overall, we observed that greater neighborhood socioeconomic disadvantage, indeed, was associated with greater T2D risk across community types, contributing to geographic disparities. We observed that relative availability of fast-food restaurants and supermarkets influenced T2D risk across most community types; however, the directionality of the food environment–T2D association varied geographically, a finding that suggests potential local variation in types of restaurants or food establishments themselves. Central to the goal of this article, however, we found little to no mediating influence of relative availability of food outlets on the relationship between NSEE and T2D risk in any community-type settings.

Consistent with previous findings (8,9), our results suggest that individuals who live in more disadvantaged neighborhoods have, on average, higher risk of T2D, although associations varied in magnitude by community type and across

study sites. These findings were invariant to how we defined NSEE distributions (i.e., site-specific or national distribution), although findings were less consistent in higher-density urban settings, where each site observed some null or inverse associations. The association between NSEE and T2DM risk in lower-density urban, suburban/small town, and rural settings may point to underlying features of NSEE that may be contributing to T2D risk in those settings, such as lack of access to public transportation or limited access to public social services, factors that may be less common in urban communities (37). Although previous studies in urban areas have generally reported positive associations between neighborhood disadvantage and T2D prevalence (38), few have investigated how neighborhood factors may affect T2DM risk either nationally or within strata of urbanicity.

In our study, the associations of the food-environment metrics with new-onset T2D varied across study samples. In the national VADR cohort, fast-food restaurant availability was associated with an increased T2D risk in all community types, whereas availability of supermarkets reduced T2D risk in suburban and rural communities only, a

Table 3—Mediation effect of neighborhood food environment in the association between the neighborhood social and economic environments and the risk of T2D

	Relative fast-food restaurants			Relative supermarkets		
	Total effect (95% CI)	Average direct effect (95% CI)	Average indirect effect (95% CI)	Total effect (95% CI)	Average direct effect (95% CI)	Average indirect effect (95% CI)
Higher-density urban						
VADR (risk difference, %)*						
Q2	0.1495 (0.0998–0.2344)	0.1488 (0.0992–0.2335)	0.0007 (0.0003–0.001)	0.1506 (0.0854–0.2235)	0.1506 (0.0854–0.2235)	0 (–0.0001 to 0.0001)
Q3	0.1895 (0.0978–0.2704)	0.1916 (0.1002 to 0.2722)	–0.0021 (–0.003 to –0.0009)	0.1812 (0.079–0.2596)	0.1805 (0.0777–0.2594)	0.0007 (–0.0001 to 0.0016)
Q4	0.1587 (0.0848–0.2444)	0.1625 (0.089–0.2475)	–0.0038 (–0.0055 to –0.0021)	0.1849 (0.1095–0.268)	0.1852 (0.1097–0.2685)	–0.0003 (–0.0007 to 0.0001)
REGARDS (risk difference, %) [†]						
Q2	4.0619 (–1.7143 to 9.0608)	4.0613 (–1.6172 to 9.0909)	0.0006 (–0.2214 to 0.2386)	4.1648 (–1.4812 to 9.1464)	4.1729 (–1.4911 to 9.1770)	–0.0081 (–0.1931 to 0.1641)
Q3	5.6835 (–0.5649 to 11.2082)	5.6291 (–0.7326 to 11.2343)	0.0543 (–0.1782 to 0.3682)	5.9538 (0.1267–11.4083)	5.9702 (0.1002–11.4280)	–0.0164 (–0.2448 to 0.1638)
Q4	2.7952 (–2.8888 to 8.4708)	2.8476 (–2.8835 to 8.5545)	–0.0524 (–0.3431 to 0.1710)	3.0214 (–2.8578 to 8.9549)	3.0134 (–2.8452 to 8.9836)	0.0080 (–0.1267 to 0.1897)
G/JHU (odds ratio) [‡]						
Q2	1.1901 (0.9271–1.5457)	1.1059 (0.8717–1.4184)	1.0761 (0.9752–1.1885)	1.1899 (0.9267–1.5457)	1.194 (0.9315–1.5582)	0.9966 (0.9685–1.0191)
Q3	1.111 (0.8827–1.4054)	1.0456 (0.8428–1.3048)	1.0626 (0.9602–1.1747)	1.1123 (0.8837–1.4073)	1.1169 (0.8896–1.4205)	0.9959 (0.9687–1.0160)
Q4	0.7745 (0.5496–1.1034)	0.7698 (0.5707–1.0538)	1.0061 (0.8961–1.1254)	0.7776 (0.5532–1.1068)	0.7845 (0.5622–1.1209)	0.9912 (0.9447–1.0285)
Lower-density urban						
VADR (risk difference, %)*						
Q2	0.1203 (0.0989–0.1412)	0.1169 (0.0951–0.1377)	0.0034 (0.0025–0.0043)	0.1309 (0.1123–0.1493)	0.1308 (0.1122–0.149)	0.0001 (–0.0003 to 0.0004)
Q3	0.1766 (0.1368–0.2181)	0.1704 (0.1307–0.2114)	0.0061 (0.0043–0.0077)	0.1842 (0.156–0.2049)	0.1843 (0.1559–0.2053)	0 (–0.0007 to 0.0008)
Q4	0.1153 (0.0822–0.1727)	0.1116 (0.0783–0.1689)	0.0037 (0.0031–0.0045)	0.1111 (0.0833–0.1428)	0.1111 (0.0832–0.1423)	0 (–0.0006 to 0.0006)
REGARDS (risk difference, %) [†]						
Q2	1.0367 (–2.3973 to 4.2590)	0.9713 (–2.4721 to 4.1819)	0.0654 (–0.0527 to 0.2345)	0.9263 (–2.5576 to 4.1751)	0.9269 (–2.5662 to 4.1417)	–0.0006 (–0.0692 to 0.0555)
Q3	4.2572 (0.5388–7.8820)	4.0912 (0.3541–7.7326)	0.1660 (–0.1387 to 0.4753)	4.2955 (0.3851–7.8567)	4.2933 (0.3758–7.8498)	0.0021 (–0.0728 to 0.0719)
Q4	3.6776 (–0.6152 to 7.6582)	3.5915 (–0.7722 to 7.6100)	0.0862 (–0.0788 to 0.3208)	3.8804 (–0.4388 to 8.1474)	3.8782 (–0.4052 to 8.1741)	0.0022 (–0.0837 to 0.1042)
G/JHU (odds ratio) [‡]						
Q2	1.5484 (1.1575–2.0631)	1.5112 (1.1393–2.0288)	1.0246 (0.9397–1.0995)	1.5482 (1.1579–2.0623)	1.5375 (1.1432–2.0774)	1.0069 (0.9462–1.0626)
Q3	1.6747 (1.2509–2.2326)	1.5923 (1.2244–2.0849)	1.0518 (0.9533–1.1476)	1.6747 (1.2512–2.2325)	1.6622 (1.2459–2.2318)	1.0075 (0.9456–1.0625)
Q4	1.7003 (1.2822–2.2672)	1.6605 (1.2847–2.2023)	1.024 (0.9301–1.1048)	1.7006 (1.2829–2.2676)	1.6958 (1.2775–2.2736)	1.0028 (0.9637–1.0393)
Suburban/small town						
VADR (risk difference, %)*						
Q2	0.13 (0.0922–0.1694)	0.1313 (0.0936–0.1707)	–0.0013 (–0.0016 to –0.0009)	0.1274 (0.0936–0.1624)	0.1292 (0.0955–0.1639)	–0.0018 (–0.0025 to –0.0013)
Q3	0.204 (0.166–0.2533)	0.2048 (0.1669–0.2544)	–0.0008 (–0.0012 to –0.0005)	0.2149 (0.1811–0.2673)	0.2183 (0.1841–0.2694)	–0.0034 (–0.005 to –0.002)
Q4	0.2571 (0.1911–0.3001)	0.26 (0.1935–0.3029)	–0.0029 (–0.0038 to –0.002)	0.2581 (0.1974–0.3139)	0.264 (0.2029–0.3196)	–0.0059 (–0.0087 to –0.0046)
REGARDS (risk difference, %) [†]						
Q2	–0.7554 (–5.6646 to 3.3976)	–0.7417 (–5.6261 to 3.4446)	–0.0138 (–0.2132 to 0.1658)	–0.7999 (–5.8290 to 3.9437)	–0.8269 (–5.7967 to 3.9241)	0.0270 (–0.1292 to 0.2227)
Q3	1.6791 (–3.1482 to 5.9948)	1.7063 (–3.0969 to 6.1044)	–0.0271 (–0.4053 to 0.3239)	1.6828 (–3.3518 to 6.4491)	1.8242 (–3.1168 to 6.5940)	–0.1414 (–0.6024 to 0.3008)
Q4	3.9079 (–1.7098 to 9.3131)	3.9331 (–1.6844 to 9.4345)	–0.0251 (–0.3906 to 0.2514)	3.9196 (–1.6493 to 10.1089)	4.0754 (–1.5386 to 10.2197)	–0.1558 (–0.7832 to 0.3676)
G/JHU (odds ratio) [‡]						
Q2	1.0574 (0.8821–1.2704)	1.0658 (0.8939–1.2740)	0.9921 (0.9673–1.0173)	1.0575 (0.8822–1.2706)	1.0559 (0.8795–1.2646)	1.0016 (0.9848–1.0235)
Q3	1.1479 (0.9698–1.3498)	1.1492 (0.9770–1.3429)	0.9989 (0.9727–1.0257)	1.1483 (0.9700–1.3506)	1.1521 (0.9701–1.3514)	0.9967 (0.9746–1.0255)
Q4	1.216 (1.0015–1.4611)	1.2132 (0.9990–1.4556)	1.0023 (0.9720–1.0353)	1.217 (1.0020–1.4630)	1.2248 (1.0068–1.4766)	0.9937 (0.9638–1.0231)

Continued on p. 807

Table 3—Continued

	Relative fast-food restaurants			Relative supermarkets		
	Total effect (95% CI)	Average direct effect (95% CI)	Average indirect effect (95% CI)	Total effect (95% CI)	Average direct effect (95% CI)	Average indirect effect (95% CI)
Rural						
VADR (risk difference, %)*						
Q2	0.0627 (0.0323–0.0925)	0.0672 (0.0366–0.0976)	–0.0045 (–0.0059 to –0.0034)	0.0686 (0.0176–0.1004)	0.0717 (0.0206–0.1038)	–0.003 (–0.0039 to –0.0021)
Q3	0.1229 (0.0685–0.1655)	0.1319 (0.0797–0.1745)	–0.0089 (–0.0115 to –0.0066)	0.116 (0.0783–0.159)	0.1191 (0.0806–0.1628)	–0.0031 (–0.0045 to –0.0019)
Q4	0.1662 (0.1307–0.2257)	0.1792 (0.1454–0.24)	–0.0131 (–0.017 to –0.0096)	0.184 (0.1282–0.2427)	0.1886 (0.1325–0.2477)	–0.0045 (–0.0059 to –0.0029)
REGARDS (risk difference, %) [†]						
Q2	5.4849 (0.7844–9.8848)	5.5412 (0.8414–9.9011)	–0.0563 (–0.2976 to 0.1198)	5.4866 (1.1072–10.1278)	5.4172 (1.0430–10.1214)	0.0694 (–0.0649 to 0.2977)
Q3	3.7945 (–0.6019–7.9959)	3.8548 (–0.5369–8.0826)	–0.0604 (–0.2941 to 0.1155)	3.8329 (–0.9781 to 8.1141)	3.6756 (–1.1995 to 7.9824)	0.1573 (–0.1043 to 0.5153)
Q4	2.8038 (–2.4782–8.1218)	2.8129 (–2.433–8.1101)	–0.0091 (–0.1954 to 0.1522)	2.8600 (–2.6076 to 7.9544)	2.6367 (–2.7676 to 7.8459)	0.2234 (–0.2248 to 0.8000)
G/JHU (odds ratio) [‡]						
Q2	1.0397 (0.9359–1.1534)	1.0417 (0.9376–1.1576)	0.998 (0.9880–1.0067)	1.0396 (0.9359–1.1533)	1.0403 (0.9367–1.1545)	0.9994 (0.9927–1.0056)
Q3	1.1541 (1.0560–1.2603)	1.1525 (1.0543–1.2606)	1.0014 (0.9917–1.0098)	1.1541 (1.0560–1.2603)	1.1576 (1.0591–1.2628)	0.9969 (0.9784–1.0171)
Q4	1.1735 (1.0562–1.3150)	1.182 (1.0625–1.3270)	0.9928 (0.9755–1.0098)	1.1736 (1.0562–1.3151)	1.1759 (1.0607–1.3155)	0.998 (0.9850–1.0107)

Q, quartile. *VADR: Model adjusted for baseline age, quadratic age, race/ethnicity, sex, income/disability flag, neighborhood land-use environment, and percentages of Hispanic and Black participants. †REGARDS: Model adjusted for age, race, sex, income, current smoking, neighborhood land-use environment, and percentages of Hispanic and Black participants. ‡G/JHU: Model adjusted for age, sex, race/ethnicity, medical assistance, smoking, neighborhood land-use environment, percentages of Hispanic and Black participants. To address nonlinearity, the higher-density urban model included a quadratic fast-food variable and suburban/small town included quadratic and cubic fast-food variable.

finding we have published and described in more detail elsewhere (39). Among the three samples, the VADR cohort had the widest distribution of food environments, including a long tail of tracts with much higher levels of fast-food availability. Counterintuitively, we found fast-food availability was protective against T2D in most community types in the G/JHU sample, a finding we replicated in a subsample of patients in the VADR cohort living in the same census tracts. Although this finding could suggest true geographic variation in the influence of food environments on T2D risk (due either to actual variation in the types of fast-food restaurants present in this area of Pennsylvania or in how residents use restaurants), it may also reflect unmeasured confounding differentially influencing these observed associations, because the neighborhood food environment has been associated with protective community features, such as physical activity assets (40) in the G/JHU study geography. This speaks to the challenges of multilevel causal inference and the implicit assumptions of population-average associations in multilevel models (41). To date, only a few studies have examined the association between neighborhood food environment and incident T2D using longitudinal data mostly from regional or metropolitan data sets. In one study, using data from the Jackson Heart Study in a single urban area, researchers found that higher density of less-healthy food stores was associated with a higher risk of T2D incidence (hazard ratio 1.34 [95% CI 1.12–1.61]), consistent with VADR findings (13). Findings from the Multi-Ethnic Study of Atherosclerosis, representing a number of urban settings, also indicated that better neighborhood resources for physical activity and healthy food establishments reduced risk of T2D incidence (HR 0.62 [95% CI 0.43–0.88]) (12). A novel contribution of the present study was the ability to examine such relationships across diverse settings, including in understudied suburban/small town and rural settings, where we observed some novel patterns, particularly with respect to supermarket availability. Prior work indicates that supermarkets offer more diverse and healthful food inventories than do many other food outlets, such as corner stores or convenience stores, and, in rural

settings, are often the main source of healthy food (42,43).

Our mediation analysis results by community type indicated that the indirect influence of the food environment on the association between NSEE and T2D risk was largely negligible for both food-environment measures. We found weak mediation effects of the relative availability of fast food within lower-density urban community types for the VADR cohort only. In contrast to our study findings, authors of the majority of published studies exploring mediation pathways between neighborhood exposures and T2D or obesity have largely examined the influence of individual-level mediators (44–46). Exceptions include a very small number of studies examining the mediating pathways of neighborhood attributes on obesity (47,48). One of these studies applied causal mediation methods similar to what we did in the present study and identified a suppressive effect of air pollution on the association between greenness and adiposity in China (47), whereas another study examining the same association in Spain found no mediating impacts (48).

There are several limitations to the work described here. Most broadly, because data for these analyses come from three distinct studies, it is challenging to tease out how differences in study populations, study designs, data collection, measures of association, or length of follow-up may have influenced observed variation in results. In particular, each study adjusted for individual-level socioeconomic status differently, and individual-level measures of socioeconomic status or other social determinants of health were particularly limited in the EHR-based samples (i.e., the VADR and G/JHU cohorts). Despite these differences, we obtained fairly consistent mediation results across cohorts, including when rerunning VADR analyses using the geographic footprints of REGARDS and G/JHU. Indeed, similar mediation findings across study samples when using harmonized measures strengthen inference. Another potential limitation is that we stratified our measures and analyses by community type, because of contextual heterogeneity and the importance of taking into account differences in spatial scales, reducing the ability to directly compare results across community types.

However, the importance of locally tailoring implementation of policies and evidence-based programs warrants understanding the interrelationships between neighborhood-level determinants and T2D risk separately for each community type. We also did not assess the influence of time spent in food environments beyond home residences or how that influence varies across geography. Although the VADR cohort was large and racially and ethnically diverse, most patients in the cohort were male veterans, which limits generalizability and comparability with the other sites. Nonetheless, the absolute number of women in this cohort was still large relative to the other two data sets. In REGARDS, there is a potential risk of survivorship bias, because participants had to remain in the cohort long enough to attend the follow-up examination when incident T2D was assessed. Also, REGARDS included NH Black and White participants only. The G/JHU sample was limited to a predominantly NH White population, but the study sample was representative of the general population in the study region.

This study also has several important strengths. To our knowledge, this is the first large-scale, multicohort collaboration to examine the role of potentially modifiable community determinants such as the food environment in mediating the consistently observed association between NSEE and T2D risk. The harmonized definitions and analyses undergirding these results minimized the potential for varying measurement approaches to influence findings across studies and provided an opportunity to compare, identify consistencies and anomalies, and synthesize findings efficiently. The widespread geographic coverage provided by the three studies and tailored analyses across a rural to urban spectrum using empirically derived geographic buffers reduced the possibility of differential item functioning by community type (16). To our knowledge, previous studies of the food environment have not accounted for differing buffer sizes by community type based on empirical studies of resident behavior. Other strengths include the fact that all three studies incorporated temporality into their designs; two studies examined longitudinal associations with T2D

incidence, and the third used a nested case-control design with ascertainment of neighborhood exposures prior to T2D onset. The combination of large study sample sizes, wide geographic coverage, tailored analyses, and longitudinal outcome ascertainment all reduced potential uncertainties due to random error, selection bias, or temporality common in cross-sectional studies or studies with small sample sizes. Analytically, researchers at each site used multilevel causal mediation analyses of neighborhood exposures and mediators on individual-level outcomes. Measures of the food environment used in this study were based on highly curated data from the RECDV study (22). Although the main analytic approach we used in the present study used relative measures of the food environment, the Network also investigated absolute measures within the VADR cohort in a separate study, and results suggested that repeated sensitivity analyses with the other two samples were not indicated (39). Our choice to use relative availability as the primary food-environment measure for these analyses builds on prior research, including studies in which models with relative measures had better model fit than models with absolute measures (49,50). Indeed, in other analyses using these data, we also observed a high degree of collinearity between absolute measures of fast-food restaurant and supermarket availability (16).

The established pattern of higher T2D risk among residents living in disadvantaged neighborhoods contributes to geographic disparities in T2D across the United States. Although the contribution of biological pathways such as epigenetic embedding of socioeconomic deprivation and poverty cannot be discounted, such patterns suggest that identifying modifiable community-level mechanisms contributing to this association could yield meaningful impacts on reducing risk (51). Other multilevel studies are needed to understand how the respective and potentially interacting roles of individual- and community-level social determinants of health influence local use of food-environment resources. Within similar food-establishment types, geographic variation in food choices and consumption patterns also needs to be better understood; such research potentially can be facilitated

by analysis of large consumer-purchase databases. By carefully examining mediating pathways that contribute to this pattern across different rural, suburban/small town, and urban settings, however, the Network aimed to gain better understanding of plausible interventions or policies that may reduce population-specific determinants of T2D risk. Although the mediating influence of the food environment was limited, we did find that the relative availability of food outlets was independently associated with T2D risk in multiple community types, suggesting that tailored interventions restricting fast-food restaurants and targeting availability of supermarkets may be an effective strategy for population risk reduction in some community types in certain regions of the United States. Alternatively, policies that limit how unhealthy food is sold in such food establishments, such as sugar-sweetened beverage taxes, or that promote healthy options may ameliorate some of the direct harmful effects of food environments. However, it appears no study has yet to determine whether such taxes reduce risk of diabetes. Importantly, based on our findings, such interventions would not necessarily mitigate the role of poor NSEE on T2D risk. Ultimately, to guide effective tailoring of policies to local environments, natural experiment studies that exploit variation in local policies will provide important contributions to evidence regarding what works.

Funding. This research was conducted by the Diabetes LEAD Network, funded by the Centers for Disease Control and Prevention (CDC) cooperative agreements U01DP006293 (Drexel University), U01DP006296 (Geisinger–Johns Hopkins University), U01DP006299 (New York University School of Medicine), and U01DP006302 (University of Alabama at Birmingham), along with collaboration with the CDC Division of Diabetes Translation. The study was also supported by the National Institute of Diabetes and Digestive and Kidney Diseases (grant R01DK124400). The REGARDS Study was supported by cooperative agreement U01 NS041588, cofunded by the National Institute of Neurological Disorders and Stroke and the National Institute on Aging, National Institutes of Health, Department of Health and Human Services.

The findings and conclusions are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. Concept and design: L.E.T., A.G.H., L.A.M., P.R., and B.E. contributed to the study concept and design. L.E.T., S.A., R.K., P.L., P.R., D.C.L., and B.E. contributed to data acquisition, analysis, or interpretation. L.E.T., P.L., R.K., S.A., P.R., A.G.H., C.M.N., C.R.H., A.Z., and L.A.M. drafted the manuscript. B.E., D.C.L., B.S.S., E.L.O., Y.A., M.N.P., L.L., A.P.C., S.A.D., M.M., F.A., K.R.S., and G.R. critically revised the manuscript for important intellectual content. S.A., L.L., C.M.N., and A.Z. conducted the statistical analyses. L.E.T., B.E., L.A.M., A.G.H., B.S.S., and A.P.C. obtained funding. L.E.T. and R.K. are the guarantors of this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Centers for Disease Control and Prevention. National Diabetes Statistics Report 2020: Estimates of Diabetes and Its Burden in the United States. Atlanta, GA, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, 2020
- Dwyer-Lindgren L, Mackenbach JP, van Lenthe FJ, Flaxman AD, Mokdad AH. Diagnosed and undiagnosed diabetes prevalence by county in the U.S., 1999–2012. *Diabetes Care* 2016;39:1556–1562
- Hipp JA, Chalise N. Spatial analysis and correlates of county-level diabetes prevalence, 2009–2010. *Prev Chronic Dis* 2015;12:E08
- Barker LE, Kirtland KA, Gregg EW, Geiss LS, Thompson TJ. Geographic distribution of diagnosed diabetes in the U.S.: a diabetes belt. *Am J Prev Med* 2011;40:434–439
- Shrestha SST, Thompson TJ, Kirtland KA, et al. Changes in disparity in county-level diagnosed diabetes prevalence and incidence in the United States, between 2004 and 2012. *PLoS One* 2016;11:e0159876–e0159876
- Dendup T, Feng X, Clingan S, Astell-Burt T. Environmental risk factors for developing type 2 diabetes mellitus: a systematic review. *Int J Environ Res Public Health* 2018;15:78
- Piccolo RS, Duncan DT, Pearce N, McKinlay JB. The role of neighborhood characteristics in racial/ethnic disparities in type 2 diabetes: results from the Boston Area Community Health (BACH) Survey. *Soc Sci Med* 2015;130:79–90
- Piccolo RS, Subramanian SV, Pearce N, Florez JC, McKinlay JB. Relative contributions of socioeconomic, local environmental, psychosocial, lifestyle/behavioral, biophysiological, and ancestral factors to racial/ethnic disparities in type 2 diabetes. *Diabetes Care* 2016;39:1208–1217
- Bilal U, Auchincloss AH, Diez-Roux AV. Neighborhood environments and diabetes risk and control. *Curr Diab Rep* 2018;18:62
- Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic status and incidence of type 2 diabetes: results from the Black Women's Health Study. *Am J Epidemiol* 2010;171:564–570
- Richardson AS, Meyer KA, Howard AG, et al. Neighborhood socioeconomic status and food environment: a 20-year longitudinal latent class analysis among CARDIA participants. *Health Place* 2014;30:145–153
- Auchincloss AH, Diez Roux AV, Mujahid MS, Shen M, Bertoni AG, Carnethon MR. Neighborhood resources for physical activity and

healthy foods and incidence of type 2 diabetes mellitus: the Multi-Ethnic study of Atherosclerosis. *Arch Intern Med* 2009;169:1698–1704

- Gebreab SY, Hickson DA, Sims M, et al. Neighborhood social and physical environments and type 2 diabetes mellitus in African Americans: the Jackson Heart Study. *Health Place* 2017;43:128–137
- Ohri-Vachaspati P, DeWeese RS, Acciai F, et al. Healthy food access in low-income high-minority communities: a longitudinal assessment—2009–2017. *Int J Environ Res Public Health* 2019;16:2354
- Hilmers A, Hilmers DC, Dave J. Neighborhood disparities in access to healthy foods and their effects on environmental justice. *Am J Public Health* 2012;102:1644–1654
- Rummo PE, Algur Y, McAlexander T, et al. Comparing competing geospatial measures to capture the relationship between the neighborhood food environment and diet. *Ann Epidemiol* 2021;61:1–7
- McAlexander TP, Yasemin A, Schwartz BS, et al. Categorizing community type for epidemiologic evaluation of community factors and chronic disease across the United States. *Soc Sci Hum Open* 2022;5:100257
- Hirsch AG, Carson AP, Lee NL, et al. The Diabetes Location, Environmental Attributes, and Disparities Network: protocol for nested case control and cohort studies, rationale, and baseline characteristics. *JMIR Res Protoc* 2020;9:e21377
- Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health* 2006;83:1041–1062
- Xiao Q, Berrigan D, Powell-Wiley TM, Matthews CE. Ten-year change in neighborhood socioeconomic deprivation and rates of total, cardiovascular disease, and cancer mortality in older US adults. *Am J Epidemiol* 2018;187:2642–2650
- Rural-Urban Continuum Codes. U.S. Department of Agriculture. Accessed 17 January 2022. Available from <https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx>
- Drexel University Urban Health Collaborative. The Retail Environment and Cardiovascular Disease (RECVD) project. Accessed 17 January 2022. Available from <https://drexel.edu/uhc/research/projects/retail-environment-cardiovascular-disease/>
- Hirsch JA, Moore KA, Cahill J, et al. Business data categorization and refinement for application in longitudinal neighborhood health research: a methodology. *J Urban Health* 2021;98:271–284
- Rummo PE, Guilkey DK, Ng SW, et al. Does unmeasured confounding influence associations between the retail food environment and body mass index over time? The Coronary Artery Risk Development in Young Adults (CARDIA) study. *Int J Epidemiol* 2017;46:1456–1464
- Rummo PE, Guilkey DK, Ng SW, et al. Understanding bias in relationships between the food environment and diet quality: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *J Epidemiol Community Health* 2017;71:1185–1190
- Kaufman TK, Sheehan DM, Rundle A, et al. Measuring health-relevant businesses over 21 years: refining the National Establishment Time-

- Series (NETS), a dynamic longitudinal data set. *BMC Res Notes* 2015;8:507
27. Esri. ArcGIS Online. Available from <https://www.esri.com/en-us/landing-page/product/2019/arcgis-online/overview>
28. Avramovic S, Alemi F, Kanchi R, et al. US Veterans Administration Diabetes Risk (VADR) national cohort: cohort profile. *BMJ Open* 2020;10:e039489
29. Friedman M. Piecewise exponential models for survival data with covariates. *Ann Stat* 1982;10:101–113
30. Laird N, Olivier D. Covariance analysis of censored survival data using log-linear analysis techniques. *J Am Stat Assoc* 1981;76:231–240
31. Austin PC. A tutorial on multilevel survival analysis: methods, models and applications. *Int Stat Rev* 2017;85:185–203
32. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods* 2010;15:309–334
33. Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuro-epidemiology* 2005;25:135–143
34. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. *J Stat Softw* 2014;59:1–38
35. Schwartz BS, Pollak J, Poulsen MN, et al. Association of community types and features in a case-control analysis of new onset type 2 diabetes across a diverse geography in Pennsylvania. *BMJ Open* 2021;11:e043528
36. Steen J, Loeyts T, Moerkerke B, Vansteelandt S. Medflex: an R package for flexible mediation analysis using natural effect models. *J Stat Softw* 2017;76:1–46
37. Kolak M, Bhatt J, Park YH, Padrón NA, Molefe A. Quantification of neighborhood-level social determinants of health in the continental United States. *JAMA Netw Open* 2020;3:e1919928
38. Mirowsky JE, Devlin RB, Diaz-Sanchez D, et al. A novel approach for measuring residential socioeconomic factors associated with cardiovascular and metabolic health. *J Expo Sci Environ Epidemiol* 2017;27:281–289
39. Kanchi R, Lopez P, Rummo PE, et al. Longitudinal analysis of neighborhood food environment and diabetes risk in the Veterans Administration Diabetes Risk Cohort. *JAMA Netw Open* 2021;4:e2130789
40. Poulsen MN, Glass TA, Pollak J, et al. Associations of multidimensional socioeconomic and built environment factors with body mass index trajectories among youth in geographically heterogeneous communities. *Prev Med Rep* 2019;15:100939
41. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med* 2004;58:1929–1952
42. Cannuscio CC, Tappe K, Hillier A, Buttenheim A, Karpyn A, Glanz K. Urban food environments and residents' shopping behaviors. *Am J Prev Med* 2013;45:606–614
43. Chrisinger BW, Kallan MJ, Whiteman ED, Hillier A. Where do U.S. households purchase healthy foods? An analysis of food-at-home purchases across different types of retailers in a nationally representative dataset. *Prev Med* 2018;112:15–22
44. Saelee R, Gazmararian JA, Haardörfer R, Suglia SF. Associations between the Neighborhood Social Environment and Obesity Among Adolescents: Do Physical Activity, Screen Time, and Sleep Play a Role? *Health Place* 2020;64:102380
45. do Carmo AS, Rodrigues D, Nogueira H, et al. Influence of parental perceived environment on physical activity, TV viewing, active play and body mass index among Portuguese children: a mediation analysis. *Am J Hum Biol* 2020;32:e23400
46. Van Dyck D, Cerin E, Akram M, et al. Do physical activity and sedentary time mediate the association of the perceived environment with BMI? The IPEN Adult Study. *Health Place* 2020;64:102366
47. Huang B, Liu Y, Chen Y, Wei H, Dong G, Helbich M. Establishing associations between residential greenness and markers of adiposity among middle-aged and older Chinese adults through multilevel structural equation models. *Int J Hyg Environ Health* 2020;230:113606
48. O'Callaghan-Gordo C, Espinosa A, Valentin A, et al. Green spaces, excess weight and obesity in Spain. *Int J Hyg Environ Health* 2020;223:45–55
49. Clary CM, Ramos Y, Shareck M, Kestens Y. Should we use absolute or relative measures when assessing foodscape exposure in relation to fruit and vegetable intake? Evidence from a wide-scale Canadian study. *Prev Med* 2015;71:83–87
50. Wilkins E, Morris M, Radley D, Griffiths C. Methods of measuring associations between the Retail Food Environment and weight status: importance of classifications and metrics. *SSM Popul Health* 2019;8:100404
51. Oblak L, van der Zaag J, Higgins-Chen AT, Levine ME, Boks MP. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res Rev* 2021;69:101348