CROSS-SECTIONAL AND PROSPECTIVE ASSOCIATIONS BETWEEN NEIGHBORHOOD SOCIOECONOMIC STATUS AND DIABETES IN THE ATHEROSCLEROSIS RISK IN COMMUNITIES (ARIC) STUDY

by
Emily A. Knapp

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Health Science

Baltimore, Maryland
April 30, 2015

Abstract

Background

Diabetes is increasing in the United States, and affects almost a third of older adults. Previous research has focused on clinical and individual risk factors for diabetes, and there is little research on the contribution of the neighborhood environment to diabetes incidence and prevalence.

Methods

The Atherosclerosis Risk in Communities (ARIC) study enrolled 15,792 patients at four sites in the United States between 1987-89. Participants were followed for three additional study visits and annual telephone follow-up. Residential address at the first visit was linked to census tract. The primary exposure was a neighborhood socioeconomic summary score created from 6 census tract variables. The outcomes of interest were diabetes incidence and prevalence, measured by self-report and fasting plasma glucose. Multivariate logistic regression and Cox Proportional Hazards models stratified by race were used to evaluate the affect of neighborhood socioeconomic status on prevalence and incidence of diabetes.

Results

Among 11,032 white participants in the ARIC study, higher neighborhood socioeconomic status was associated with a statistically significant decrease in the prevalence of diabetes at the first ARIC study visit (OR for each unit increase in socioeconomic status 0.978, 95% CI 0.961-0.995) and with a statistically significant decrease in the incidence of diabetes over time (HR for each unit increase in socioeconomic status 0.972, 95% CI 0.962-0.982) after

adjusting for age, sex, income and education. Among black participants, neighborhood socioeconomic status was not associated with prevalence of diabetes at baseline or incidence of diabetes across the study period.

Conclusion

This study is one of the first evaluations of the affect of neighborhood socioeconomic status and diabetes. Further studies are needed to explore the differential effects of neighborhood socioeconomic status by race.

Acknowledgments

I would like to thank my advisor, Jessica Yeh, for her support for this project and her encouragement and mentoring during the last two years. Additionally, I am grateful to Thomas Glass, for his insightful comments on this thesis and for his guidance during this program.

I appreciate the guidance of the ARIC writing group, especially their help in refining the goals of this project: Drs. Priya Palta and Gerardo Heiss at University of North Carolina at Chapel Hill School of Public Health, Department of Epidemiology; and Usama Bilal and Elizabeth Selvin at the Johns Hopkins School of Public Health, Department of Epidemiology.

Finally, thank you to the staff and participants of the ARIC study for your important contributions, and to the ARIC Publications Committee for sharing data for this project.

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C).

Table of Contents

Introduction

Previous research suggests that neighborhood disadvantage is an important risk factor for many adverse health outcomes, often independent of individual socioeconomic status.[1-3] The Atherosclerosis Risk in Communities Study (ARIC) cohort has previously been used to demonstrate associations of neighborhood socioeconomic status with coronary heart disease, mortality, body mass index, metabolic syndrome, and other outcomes.[1,4-7] Since the first ARIC cohort visit (1987-1989), the prevalence of type II diabetes has increased dramatically in the United States.[8] Nearly thirty million people in the US are estimated to have diabetes, including a third of those aged 65 and older.[9] Cross-sectional studies suggest that the prevalence of diabetes is higher among those with low individual-level socioeconomic status, as measured by education, occupation, or income.[10,11] The few studies that have been published on the relationship between diabetes and neighborhood socioeconomic status (SES) suggest that neighborhood SES may be an important risk factor for diabetes, but inferences from these studies are limited by the methodological challenges common to studies of neighborhood effects.[10-14] Using data from the ARIC study, our aim was to evaluate the cross-sectional and prospective associations between neighborhood SES and incident diabetes.

Neighborhoods and Health

Place-based determinants of health have always been central to epidemiologic research, and interest in the neighborhood context has increased in recent years.[2,14,15] There are many pathways through which neighborhood characteristics may affect individual health outcomes: availability of health services, physical infrastructure, social support and influence of peer attitudes toward health-related behaviors, segregation and discrimination, employment opportunities, concentration of poverty, and crime.[11,15-17] For example, availability of fast food outlets, grocery stores, liquor stores, and public parks vary by neighborhood and contribute to the quality of the living environment.[10] Previous studies have demonstrated that those living in lower quality neighborhoods are more likely to smoke and consume unhealthy diets, and less likely to perform physical activity, regardless of the level of individual income.[16] These behaviors are associated with a higher risk of diabetes, and are likely intermediaries between the neighborhood environment and diabetes. [11] If neighborhood factors are important in the risk of developing diabetes, data such as those in ARIC can be used to inform neighborhood-level intervention and suggest geographical areas to focus preventive efforts. Efforts to address the increase in diabetes across the US at the individual level (i.e., behavior change) may not be successful if larger, structural barriers to health are not identified and mitigated.[11]

Type II Diabetes

Since 1980, the number of Americans with diagnosed diabetes has more than tripled, from 5.6 million Americans in 1980 to 20.9 million Americans in 2011.[18] Diabetes is most likely to effect older persons; the rate of diabetes for those 65 and older is over 20%.[19] Additionally, 1 out of 3 American adults—86 million people—have prediabetes.[19]

As with many health conditions, diabetes is more common among those with lower education and income.[20] Black and Hispanic people are also more likely to have diabetes.[21] Other risk factors for diabetes include: being overweight or obese, family history, gestational diabetes, high blood pressure (over 140/90), high cholesterol (HDL 35 or lower/ triglycerides 250 or higher), and physical inactivity.[22]

Beyond the burden of care for diabetes itself, diabetes is associated with an increased risk of cardiovascular disease, end-stage renal disease, lower extremity impairment, and visual impairment. Diabetes is the 7th leading cause of death in the US.[9] The American Diabetes Association estimates that diabetes cost the United States $176 billion in direct medical costs and an additional $69 billion in lost productivity in 2012. [23]

Neighborhood Socioeconomic Status and Risk of Diabetes

There is limited published literature on the relationship between the residential environment and type II diabetes. In general, these studies have several limitations. Most of these studies rely on census tract or other administrative units to define residential neighborhoods. This may mask meaningful variation within a census tract, or may offer only an approximate exposure if location of work, school, or other activities occurs in other areas and is also important for the risk of diabetes. Additionally, several of the studies that have been published utilize cross-sectional data and/or analysis, meaning that selection bias may be threatening the validity of the results. Some of the studies rely on self-reported diabetes outcomes, which could bias results if misclassification differed as a function of the neighborhood environment.

There is one randomized study of neighborhood effects in the health literature: the Moving to Opportunity (MTO) Study. Between 1994 and 1998, the department of Housing and Urban Development (HUD) randomly assigned 4498 women and children who lived in public housing in high poverty census tracts to receive housing vouchers redeemable for housing in low-poverty areas and counseling on moving, unrestricted housing vouchers, or to a control group with no vouchers. Diabetes, as measured by HgA1C of 6.5% or more, was lower in the group randomized to receive vouchers for a low-poverty area than the control group by 4.3 percentage points (95% CI: -7.82 to -0.80). The difference between the group randomized to receive unrestricted vouchers and the group randomized to move to a low-poverty area was not statistically significant. Moreover, adherence was low, suggesting that this is a conservative estimate.[24] This study provides the strongest evidence that neighborhood poverty does influence the risk of diabetes, in an experimental setting where selection bias is reduced.

One of the few longitudinal studies on this topic used data from the Black Women's Health Study to evaluate the relationship between the neighborhood environment and self-reported type II diabetes among African-American women.[25] A factor analysis was used to select six census tract variables that were determined to best describe neighborhood socioeconomic status: median household income, median housing value, percent of households receiving interest or rental income, percent of adults 25+ with college completed, percent of adults ages 16+ who have professional, managerial, executive jobs, and percent of families with children not headed by single female. Individual socioeconomic status was included in the final model to determine the independent effect of neighborhood socioeconomic status. Even when individual socioeconomic status was controlled for, neighborhood socioeconomic status was still statistically significantly related to incidence of type II diabetes [IRR comparing lowest quintile of NB SES to the highest: 1.65 (95% confidence interval: 1.46, 1.85)]. This study was limited by self-report of diabetes and the sample only included black women.[25]

Murray et al. explored the "neighborhood SES trajectories" of MESA participants instead of absolute neighborhood disadvantage. They found little change in neighborhood socioeconomic status over the 20-year study period for most participants, and increased diabetes among participants who lived in higher poverty census tracts.[26] In another study using MESA data, Auchincloss et al. found that residential environments that support physical activity and healthy diets were associated with lower incidence of type 2 diabetes. Associations between type 2 diabetes incidence and residential environment remained after adjustment for individual-level variables, including age, sex, family history of diabetes, socioeconomic characteristics, smoking status, and alcohol intake. Physical activity level, dietary factors, and BMI were also included as potential mediators. Associations were slightly reduced after adjustment for baseline BMI, providing support for conceptualizing BMI as a mediator in the relationship between the neighborhood environment and incidence of diabetes.[12]

The African-American Health Study, a population-based cohort of African-American residents of St. Louis, Missouri and surrounding suburbs born between 1936 and 1950, was used to investigate the relationship between the neighborhood environment and self-reported diabetes over three years of follow-up.[27] The exposure was defined as interviewer ("objective") observations of the "block-face", by rating the condition of the houses, noise, air quality, condition of the streets, and conditions of the yards and sidewalks on a four-point scale. The participants' ("subjective") perceptions of their neighborhoods were also collected. None of the objective or subjective measures of block face quality was statistically significantly associated with development of diabetes over the three-year follow-up models using propensity score matched analysis. Given that this study had a limited length of follow-up and a small, geographically limited sample, it is possible that there is not sufficient variation in neighborhood conditions to detect an effect. Additionally, the exposure variables were dichotomized, further decreasing the power of the study.

In a third study using MESA data, Auchincloss and colleagues explored the relationship between neighborhood deprivation and insulin resistance.[28] The neighborhood exposures of interest were defined as: poverty in the residential neighborhood, defined as 0.25 mile buffer around the residence of the participant; poverty in the 0.75 miles surrounding the neighborhood; and distance from the participants' neighborhood to the nearest census block group with per capita income of $33,000 or greater. Participants living in areas of greater poverty were more likely to have insulin resistance, even after adjusting for individual income and education. However, these results were not statistically significant after adjusting for race. Distance to wealthy areas was also associated with increased insulin resistance, even after adjustment for individual income, education, and race. The authors suggest that diet, BMI, and physical activity may be mediating this relationship, based on the attenuation of the effect estimates when these factors are included in the model.

A similar theoretical approach was taken by Cox et al. to assess the importance of "locality deprivation" in predicting type II diabetes in a Scottish population. [29] Deprivation in census output areas was compared to deprivation in the surrounding areas using the Carstairs Deprivation Index (composed of percent of residents in households with no car; the residents in households with 1 or more persons per room as a percent of all residents in households (overcrowding); the percent of residents in households with a head of household in social class IV and V; and unemployed male residents aged over 16 as a proportion of all residents aged over 16. The authors found that area deprivation was associated with increased risk of type II diabetes. Additionally, areas surrounded by less deprived areas have less type II diabetes than would be expected after adjusting for age, sex, and area deprivation, and areas surrounded by more deprived areas had higher rates of diabetes than expected. The authors interpret this as supporting a neo-material theory of the effect of social inequality on diabetes risk: living near more advantaged areas may provide greater access to resources such as employment, recreational facilities, and healthier food outlets.

Using cross-sectional data from the second visit of the Coronary Artery Risk Development in Young Adults (CARDIA) Study, Diez Roux, Jacobs and Kiefe evaluated the associations of neighborhood traits with insulin resistance syndrome (IRS) in younger adults (ages 28-40).[30] The primary exposure was a neighborhood summary score, created by summing z-scores of six indicators of neighborhood socioeconomic status from 1990 census data: median household income, median housing value, percent of households earning rental or interest income, percent completing high school, percent completing college, and percent employed in managerial or executive occupations. The outcome of this analysis was a summary score of IRS: BMI; fasting plasma HDL cholesterol, triglycerides, insulin, and glucose; and systolic blood pressure. Lower neighborhood summary score (i.e. lower neighborhood socioeconomic status) was associated with higher average IRS score (indicating more of the factors comprising the score are present) among white men and women and black women, but this association was not observed among black men. For white participants, this association persisted after adjustment for individual income and education. The differences in the associations by race and sex suggest that these characteristics may modify the relationship between neighborhood socioeconomic status and insulin resistance syndrome.

Liu et al. used a Neighborhood Physical and Social Environment Index to characterize neighborhoods using zip code as a surrogate. This index was made up of eight questions about the neighborhood social, physical, economic, and food environment. Using multilevel models to account for individual level covariates such as age, race, sex, and smoking, residents of Philadelphia who lived in neighborhoods with higher disadvantage were found to have higher odds of diabetes.[31]

Using NHANES and 2000 Census data, Gaskin et al. demonstrated that the income composition of the neighborhood of residence was associated with risk of diabetes. Overall, poor persons living in poor neighborhoods had a higher risk of diabetes than poor persons living in non-poor neighborhoods. Blacks living in poor neighborhoods, regardless of their individual income, were observed to have over twice the odds of diabetes as non-poor whites living in non-poor neighborhoods. [32] In a similar study using Colorado Behavioral Risk Factor Surveillance System (BRFSS) data, the prevalence of cardiovascular risk factors, including diabetes, was lower among those living in affluent communities, even after adjusting for individual SES.[33]

Using cross-sectional data from five regions in Germany, Muller et al. used unemployment as an indicator of the socioeconomic status of neighborhoods, along with the number of immigrants, married persons, and percent of the population under 17 and over 65 years of age.[34] Mixed effects logistic regression models demonstrated that more variation in rates of diabetes was observed at the regional level compared to the neighborhood level, and individual characteristics could account for some of the variation in diabetes risk between neighborhoods, but not between regions. The unemployment rate was observed to be a strong predictor of neighborhood and regional diabetes rates, even after adjusting for individual characteristics.

Objectives

Existing literature suggests that there is a relationship between the neighborhood environment and risk of precursors to type II diabetes as well as type II diabetes. We will evaluate the following hypotheses: 1) there is a cross-sectional relationship between neighborhood socioeconomic status and type II diabetes at visit 1; and 2) there is a relationship between neighborhood socioeconomic status and incident diabetes across 25 years of follow-up in the ARIC study. Similar to several studies cited above, we will use a summary of socioeconomic status of the census tract of residence to evaluate associations with diabetes in a cohort of older American adults. [25,30] This construct of general neighborhood socioeconomic status measures the income, wealth, education, and occupational composition of each census tract. This is one way to measure the extent of the human and material resources available to a community that could be deployed by residents to mitigate health risks. Such a general measure is useful because it allows us to compare neighborhoods using readily available administrative data, and could be refined by future research.

The conceptual framework used to guide these analyses is shown in Figure 1. First, existing literature suggests the role of race is complicated, and few studies explicitly modeled the role of race. Several studies suggest that the relationship between neighborhood socioeconomic status and diabetes differs by race. It is also often recognized that even after adjusting for individual factors, black and white Americans live in very different neighborhoods.[35] Therefore, we chose to stratify all models by race. Age and sex are risk factors for diabetes that may also be associated with neighborhood of residence. The majority of existing studies are in agreement that BMI, physical activity, and diet patterns are mediators on the pathway between the neighborhood environment and diabetes. We add alcohol use and smoking to this list of factors that are directly affected by neighborhood of residence and in turn affect risk of diabetes. For the purposes of this analysis, individual income and education are treated as confounders. We are interested in the effect of the neighborhood SES on diabetes, regardless of the income and education of individuals. While neighborhood socioeconomic status does likely influence individual socioeconomic status, we do not have a measure of neighborhood socioeconomic status at birth or during childhood. Thus, we cannot be sure whether the socioeconomic status of the neighborhood at baseline (i.e., during middle-age), preceded individual SES. Therefore, individual SES is not considered a mediator in the framework guiding this analysis. This will result in underestimates of the effect of neighborhood SES if individual income and education are truly mediators.

The aims of this study are to 1) investigate the cross-sectional association between neighborhood and individual socioeconomic status at baseline and type II diabetes with visit 1 and 2) investigate the prospective relationship between neighborhood-level socioeconomic status at baseline and incident diabetes during follow-up, after adjusting for individual-level socioeconomic status. We hypothesize that participants living in neighborhoods of lower socioeconomic status have higher risk of diabetes than those living in higher socioeconomic status neighborhoods, independent of individual socioeconomic status. We also hypothesize that the associations between neighborhood socioeconomic disadvantage and diabetes differ by race.

Methods

Atherosclerosis Risk in Communities (ARIC) Study

The Atherosclerosis Risk in Communities Study (ARIC) is a prospective study of atherosclerosis and its design has been described elsewhere.[36] Four communities in the US are included in the ARIC study: Forsyth county, North Carolina; Jackson, Mississippi; the northwestern suburbs of Minneapolis, Minnesota; and Washington county, Maryland. In each community, participants were selected by household sampling from administrative lists of persons of eligible age (e.g. lists of driver's licenses or state IDs). The ARIC study included home interviews, clinic examinations, annual telephone follow-up, and identification of clinical events. At four sites, 15,792 persons ages 45-64 were enrolled in the study. The baseline visit took place between November 1986 and December 1989. [36] The in-person clinical visits took place between 1990-92, 1993-95, and 1996-98, and annual telephone follow-up continued through 2012. Data from all four study visits and the annual telephone follow-up are used in this analysis.

Neighborhood Socioeconomic Summary Score

The neighborhood socioeconomic status summary score variable was previously created by Diez Roux et al. [1] During the first visit, ARIC participants were asked to report their residential addresses. These addresses were geocoded and connected to U.S. census tracts, which are statistical units of counties containing 1200—8000 residents that vary in geographic size.[37,38] Diez Roux and colleagues used a factor analysis of 1990 census data to identify variables to represent a construct of the neighborhood socioeconomic environment. The following six indicators of census tract ("neighborhood") socioeconomic status were selected: percentage of households receiving interest, dividend, or net rental income; percentage of adults over age 25 who completed high school; percentage of adults over age 25 who completed college; percentage of persons over age 16 in executive, managerial, or professional specialty occupations; log median household income; and log median value of household units.[1,4]

Diez Roux and colleagues created a z-score for each of these six census tract indicators by subtracting the mean across the population of all ARIC census tracts and dividing by the standard deviation. Higher z-scores for these variables indicate higher socioeconomic status. Then, they calculated the neighborhood summary score as the sum of the six z-scores. Higher neighborhood summary scores indicate higher socioeconomic status.

Figure 2 shows the frequency of neighborhood socioeconomic status summary score values by black and white race with a logistic lowess curve modeling the log odds of incident diabetes for each value of the neighborhood socioeconomic status. The logistic lowess curves, also stratified by race, suggest that the log odds of incident diabetes is approximately linearly related to neighborhood socioeconomic summary score in whites and blacks. Therefore, we choose to include the neighborhood socioeconomic summary score as a continuous variable. For descriptive purposes, race-specific tertiles of the socioeconomic score were created such that tertile 1 represents the lowest SES neighborhoods and tertile 3 represents the highest SES neighborhoods.

Outcome: Type II Diabetes

During the first four ARIC study visits (1987-9, 1990-2 1993-5, and 1996-8), a person could be defined as a diabetes case in one of three ways: self-report of a doctor's diagnosis of diabetes, self-report of medications for diabetes, or fasting plasma glucose greater than or equal to 126 mg/dL or non-fasting glucose of greater than or equal to 200mg/dL. During annual telephone follow-up after visit 4, diabetes was diagnosed through self-reported doctor diagnosis of diabetes or diabetes medication only.

Prevalent diabetes cases were defined by the presence of at least one of the following criteria at the baseline visit: self-reported previous diagnosis of diabetes, self-reported medications for diabetes control, or fasting blood glucose greater than or equal to 126mg/dL or non-fasting plasma glucose of 200mg/dL. Between visit two (1990-1992) visit four (1996-1998), incident diabetes was defined by the presence of at least one of the following criteria in individuals who were not prevalent for diabetes at baseline: self-reported medication use, self-report of a doctor's diagnosis, or fasting plasma glucose greater than or equal to 126 mg/dL or non-fasting glucose of greater than or equal to 200mg/dL. After visit four (1998-2012), incident diabetes was defined by self-reported medication use or self-reported doctor diagnosis of diabetes during annual telephone follow-up. Measured glucose is not used to define incident diabetes after 1998 because it was not available.

Other Baseline Covariates

Age at baseline is included in all models as a continuous variable. Sex is also included in all models. Three levels of self-reported education are considered: less than high school education, high school education and less than 4 years of college, or college education and above. Self-reported income at the first visit (1987-9) is included as a categorical variable with the following categories: under \$5,000; \$5000-7999; \$8000-11,999; \$12,000-15,999; \$16,000-24,999; \$25,000-34,999; \$35,000-49,999; and over \$50,000. BMI at baseline was calculated using study measured height and weight. Smoking, diet, physical activity, and alcohol use were self-reported. However, because the neighborhood environment affects these health behaviors, BMI, smoking, diet, and alcohol use are considered mediators in the framework guiding these analyses (Figure 1). Therefore, these factors are not included in the models.

Statistical Analysis

There were 15,689 participants with data available at baseline. Participants who were not of white or black race were excluded from the analysis (n= 48). Black participants in Washington County, MD (n=33) and Minneapolis, MN (n=22) were also excluded from the analysis, for consistency with other ARIC studies.[4-7] All models were stratified by black and white race due to *a priori* assumptions about the difference in the distribution of neighborhood summary score by race. Means and frequencies were used to describe the distribution of potential confounders and mediators by race-specific tertile of neighborhood socioeconomic status.

A series of multivariate logistic regression models were used to evaluate the cross-sectional association between prevalent diabetes and neighborhood socioeconomic status at baseline. First, only age and sex were adjusted. Model 2 was adjusted for age, sex, individual income, and individual education.

Kaplan Meier plots were created to show the cumulative hazard of incident diabetes across follow-up, overall and by race-specific tertiles of neighborhood socioeconomic status summary score. Cox proportional hazards models were used to evaluate the prospective association between neighborhood socioeconomic status summary score at baseline and incident diabetes. Neighborhood socioeconomic status summary score was treated as a fixed exposure. Follow-up time begins with date of first ARIC visit and ends with date of diabetes diagnosis, date of last ARIC visit or last annual follow-up if lost to follow-up, date of death, or date of most recent diabetes-free follow-up if administratively censored. All models accounted for clustering on census tract through the estimation of robust standard errors using a clustered sandwich estimator. Ignoring this clustering would lead to an underestimation of standard errors and perhaps incorrect inferences because observations in the same census tract are not independent of each other. The proportional hazards assumption was evaluated using Kaplan Meier plots of diabetes-free survival by tertile of neighborhood socioeconomic status summary score and Kaplan Meier plots on the natural log scale. A qualitative assessment of these graphs suggests that the proportional hazards assumption is not violated. A proportional hazards test using Schoenfeld residuals supported this assessment (P-value 0.1195).

Because those diagnosed with diabetes during the first four visits (through 1998) could be confirmed using a laboratory test, while those diagnosed after visit 4 could only be diagnosed through self-report, two sensitivity analyses were conducted. The first was conducted to evaluate whether the accuracy of self-report varies by neighborhood socioeconomic status. Participants who had fasting blood glucose levels of 126mg/dL or greater but did not report a doctor's diagnosis of diabetes were considered "unaware" of their diabetes status. The neighborhood SES summary score value for these participants was compared to that of participants who did not have a fasting blood glucose indicating diabetes and who did not report a doctor's diagnosis of diabetes using a two-sample t-test and using a mixed effects model clustering on census tract.

The second sensitivity analysis restricted the definition of incident diabetes to self-report only to evaluate the effect of using different definitions of diabetes throughout the follow-up period. All models were run again using only self-reported diabetes, since this method of ascertainment was consistent across the entire follow-up period (self-reported doctor diagnosis or self-reported diabetes medications; i.e., fasting plasma glucose diagnoses alone were not considered).

Three additional sensitivity analyses were considered among black persons in the study sample after the main analysis was completed. The neighborhood socioeconomic status summary score had previously been created and used in analyses using overall mean and standard deviations from the entire ARIC study population. As a sensitivity analysis, this summary score was re-created for the black participants as the sum of six race-specific z-score, which were centered on the mean and standard deviation for the black participants only. The same Cox models were run using this race-specific neighborhood socioeconomic status score as the exposure among black participants. Next, a sensitivity analysis considered the effects of the competing risk of death on our results for black participants. The final sensitivity analysis evaluated the impact of removing the 55 black participants from the primarily white study sites in Maryland and Minnesota. Our results were considered robust if inferences about the relationship between neighborhood socioeconomic score and incident diabetes were not changed.

All analyses were conducted using STATA version 13 (StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.).

Results

Baseline Characteristics

Table 1 shows baseline characteristics of ARIC participants by race and race-specific tertile of neighborhood socioeconomic status. Tertile 1 is the lowest socioeconomic status neighborhoods (most disadvantaged) and tertile 3 is the highest socioeconomic status neighborhoods (least disadvantaged). Tertile ranks were race-specific. Of the 15,689 participants with baseline data, the neighborhood socioeconomic summary score for address at visit 1 was available for 15,028 participants living in 356 unique census tracts. Neighborhood socioeconomic summary score values ranged from -8.979 to 14.497 among 11,032 white participants living in 289 census tracts and from -12.992 to 11.096 among 3,996 black participants living in 109 census tracts. Figure 3 shows the distribution of neighborhood summary scores by race and site.

On average, the neighborhood socioeconomic status summary score was lower among Black participants than white participants. Age at baseline is similar for all groups. Black participants are slightly more likely to be female, especially those residing in neighborhoods in the lower two tertiles of neighborhood SES. Across each tertile of neighborhood SES, Black participants had higher average BMI at visit 1 (for example, mean BMI is 27.6 among white participants in the lowest tertile and 30.0 among black participants in the lowest tertile). Within both races, lower individual income and lower level of education is more likely among those living in lower SES neighborhoods. At each level of neighborhood SES, black participants were less likely to be current or former smokers than white participants, with less smoking observed among those living in higher SES neighborhoods. White participants had higher consumption of alcohol in higher SES neighborhoods, and higher consumption of alcohol than black participants. Black participants were more likely to have a household income of less than the US median household income at every level of neighborhood socioeconomic status, and similarly more likely to have not attained a high school degree. Interestingly, at each tertile of neighborhood SES, blacks are more likely than whites to have earned a college degree.

Type II Diabetes

Table 2 shows diabetes outcomes by race and race-specific tertile of neighborhood socioeconomic status. There were 1863 prevalent cases of diabetes observed at baseline, and an additional 3782 cases of incident diabetes observed over 226,558 person years of follow-up. Black participants were more likely to have diabetes at baseline (p-value < 0.001) and had statistically significantly higher rates of incident diabetes across follow-up [23.9 cases compared to 14.6 cases per 100,000 person years in blacks (95% CI: 22.6 – 25.3) and whites (95% CI: 14.1 - 15.2) respectively]. For both black and white participants, those living in neighborhoods in the lowest tertile of neighborhood SES were most likely to have diabetes at baseline, with lower odds of diabetes at baseline seen in each higher tertile of neighborhood SES. Similarly, the highest rates of diabetes incidence were observed among the lowest tertiles of neighborhood SES in both black and white participants. Black participants had higher rates of incident diabetes at every tertile of neighborhood SES than white participants.

Cross-sectional results: logistic regression models

Results from cross-sectional models are shown separately for white and black participants in Tables 3 and 4. Among the 11,032 white participants, there were 997 cases of prevalent diabetes. There were 772 cases of prevalent diabetes among the 3,996 black participants. Among white participants, each unit increase in neighborhood socioeconomic status summary score is associated with a decrease in the odds of diabetes of 5.40 % (OR 0.946, 95% CI 0.932 - 0.961) when adjusting for age and sex. This means that white persons living in neighborhoods at the 75th percentile of the neighborhood socioeconomic status summary score have 22.8% lower odds of diabetes at visit 1 than those living in neighborhoods at the 25th percentile (OR 0.771, 95% CI 0.718, 0.828). This finding holds with a slight attenuation of the estimate when adjusting for individual income and education: the odds of diabetes for a one-unit change in neighborhood socioeconomic status summary score are 0.978 (95% CI 0.961 – 0.995). White persons living in neighborhoods at the 75th percentile of the neighborhood socioeconomic status summary score have 10.0% lower odds of diabetes at visit 1 than those living in neighborhoods at the 25th percentile (OR 0.900, 95% CI 0.829, 0.975). For Black participants, neighborhood socioeconomic status summary score is not associated with odds of diabetes at baseline when age and sex are accounted for (OR 0.982, 95% CI: 0.957 - 1.008), and the results are further attenuated when individual education and income are added to the model (OR 0.999, 95% CI 0.972 - 1.026).

Prospective results: Cox Proportional Hazards Model

The proportional hazards assumption was evaluated using Kaplan Meier plots of diabetes-free survival by tertile of neighborhood socioeconomic score and Kaplan Meier plots on the natural log scale (figures 4 and 5). A qualitative assessment of these graphs suggests that the proportional hazards assumption is not violated. A proportional hazards test using Schoenfeld residuals supported this assessment (P-value 0.1195), suggesting that the proportional hazard assumption is reasonable in these data. The cumulative hazard of incident diabetes stratified by race and race-specific tertiles of neighborhood socioeconomic status is shown in Figure 6. Among black participants, there were 1149 cases of incident diabetes over 48209 person years. Among white participants, there were 2459 cases of incident diabetes over 169,343.person years of follow-up.

Tables 5 and 6 show the prospective relationship between neighborhood socioeconomic status for white and black participants, respectively. Among white participants, higher neighborhood socioeconomic status, as measured by the neighborhood summary score, is associated with lower incidence of diabetes across the 25-year study period (Table 5). For each one unit increase in the neighborhood summary score (indicating higher neighborhood SES), the risk of developing diabetes decreases decreased by 4.45 % [HR 0.955, 95% CI: 0.947 - 0.964] when adjusting for age and sex. When adjusting for individual level socioeconomic indicators, the relationship between neighborhood socioeconomic summary score and incident diabetes is attenuated by 37.3% on the log hazard scale but is still statistically significant: a one unit increase in neighborhood socioeconomic status summary score is associated with a 2.8% reduction in the hazard of developing diabetes [HR 0.972, 95% CI: 0.962 - 0.982]. Alternatively, when adjusting for age and sex, white persons living in neighborhoods at the 75th percentile of the neighborhood socioeconomic status summary score have 19.3% lower hazard of diabetes during follow-up than those living in neighborhoods at the 25th percentile (HR 0.807, 95% CI 0.775, 0.842). When controlling for age, sex, and income and education, the hazard of diabetes over the study period is 12.5% lower among white participants living in neighborhoods with neighborhood socioeconomic status summary scores at the 75th percentile compared to the 25th percentile (HR 0.875, 95% CI 0.834, 0.917).

Among black participants, there is no statistically significant relationship between neighborhood socioeconomic status and incidence of diabetes when adjusting for sex and age alone (Table 6). For each unit increase in neighborhood socioeconomic score, there is 1.5% decrease in the hazard of diabetes across the study period (HR 0.985, 95% CI 0.970 - 1.001). When individual education and income are added to the model, there is no effect of neighborhood socioeconomic status on risk of incident diabetes (HR 0.999, 95% CI: 0.983 – 1.015).

Results of the Sensitivity Analyses

In the main models, participants could be classified as diabetic based on self-reported doctor's diagnosis, self-reported medication use, or measured plasma glucose. However, starting in 1998, measured glucose was not available. All diagnoses after this time were based on self-reported doctor's diagnosis or medication use. If accuracy of self-report varies by neighborhood socioeconomic status score, this could result in biased inferences.

Of the 1863 people who were diagnosed with diabetes at visit 1 using any of the definitions (self-report of a doctor's diagnosis of diabetes, self-report of medications for diabetes, or fasting plasma glucose), 635 had a fasting plasma glucose indicating diabetes but did not report a previous doctor diagnosis of diabetes or that they were taking diabetes medications. These participants are considered "unaware" of their diabetes status for the purpose of this sensitivity analysis. There were 14,424 participants aware of their diabetes status who had a non-missing neighborhood summary score. A t-test comparing the mean neighborhood SES summary score of those who were aware of their diabetes diagnosis compared to those unaware of their diabetes diagnosis suggested that those living in higher SES neighborhoods were more likely to be aware of their diabetes status (p-value < 0.001). However, when clustering by census tract was accounted for using a mixed effects model, the difference in neighborhood summary score between the groups was not statistically significant (p-value = 0.0829).

Secondly, cross-sectional and prospective models run with only self-reported diabetes (doctor diagnosis or medication use) as the outcome), i.e. removing those diagnosed with diabetes using measured glucose, did not result in different inferences compared to the main models for either white or black participants (Tables 7 and 8).

In the primary analysis, the neighborhood socioeconomic score was based on the distributions of all ARIC census tracts. However, there is concern that the assumption that community indicators work in the same way in predominantly white and black communities may not be valid. To test this, we conducted a sensitivity analysis by recreating the neighborhood summary score so that it was centered on the mean of primarily black communities, rather than all communities. Results showed among black participants only that our inferences were unchanged. Our results are not sensitive to our decision to use the entire distribution of community variables to scale our exposure variable (Table 9).

We re-estimated models to explore the possibility of bias due to competing risk of death prior to diabetes diagnosis using the cause-specific cumulative incidence method described by Fine and Gray.[39] Our inferences remained unchanged (Table 10) suggesting that it is unlikely our failure to account for competing risk of mortality within the study period introduced substantial bias.

Finally, our main cross-sectional and prospective models excluded 55 black participants who lived in the mostly white sites in Minnesota and Maryland. Because this may be excluding a subset of black participants who differ from the majority of black participants in our study sample, we re-estimated the cross-sectional and prospective models including black participants to include these 55 participants. Including these patients did not change the inferences observed in our main models (Tables 11 and 12).

Discussion

This study represents one of the first evaluations of the association between neighborhood socioeconomic status and diabetes using data collected prospectively. Using a longitudinal cohort of middle-aged Americans, neighborhood socioeconomic status was an important predictor of type II diabetes for white, but not black, participants in the ARIC cohort, both at baseline and over the follow-up period. Additionally, among white participants, lower neighborhood socioeconomic status was statistically significantly associated with higher incidence of diabetes independent of individual socioeconomic status, as measured by household income and education. Neighborhood socioeconomic status has been previously demonstrated to affect rates of obesity,[24,40] so it is not surprising that neighborhood socioeconomic status also affects the incidence of diabetes. Within the ARIC study population, for example, neighborhood SES has been linked to obesity and metabolic syndrome among white and black women, but not men.[6,7]

There are many pathways through which neighborhood socioeconomic status may be affecting diabetes risk. Individual risk factors for diabetes include overweight and obesity, hypertension, high cholesterol, and physical inactivity. These individual risk factors are in turn influenced by the neighborhood environment: the availability and types of food outlets, public parks and recreation facilities, and crime are just a few factors that may make healthy diets and recommended levels of physical activity difficult to attain. Additionally, individual socioeconomic status may mediate the relationship between neighborhood socioeconomic status and diabetes. Adjusting for individual socioeconomic status resulted in attenuation from age and sex adjusted hazard ratios of 37%. Future research should consider whether individual socioeconomic status is a mediator in the relationship between the neighborhood environment and health outcomes, instead of a confounder as assumed here, or whether individual socioeconomic status modifies the role of neighborhood socioeconomic status in shaping the risk of diabetes. If individual socioeconomic status is a mediator, those factors should not be included in models assessing the relationship between neighborhood socioeconomic status and diabetes. This would mean that our individual socioeconomic status-adjusted models underestimate the true relationship.

The relationship between neighborhood socioeconomic status and diabetes was observed for white participants in the ARIC study, but not for black participants. Because this could be an artifact of the way in which the neighborhood SES score variable was constructed as a summary of the entire ARIC population, this variable was recreated separately for black participants (i.e. the z-scores were re-centered using the mean and standard deviation for black participants only). In a post-hoc analysis with these race-specific nSES score variables, the main findings of this study remain unchanged. This suggests that the observed lack of an association among black participants is not due to the decision to center the neighborhood socioeconomic status summary score on the mean of the entire study sample. The decision to exclude the 55 black participants who lived in the primarily white study sites in Maryland and Minnesota may also be biasing our results towards the null in black participants, by excluding potentially higher socioeconomic status black participants from the sample. However, in sensitivity analyses including these patients, our results were robust to the decision to exclude them.

Alternatively, this finding could be due to a competing risk of death that is higher among black participants as compared to white participants. Median age at death is about two years lower among black participants, but a post-hoc analysis with death as a competing risk did not result in different inferences about the relationship between diabetes and neighborhood disadvantage among black participants. It is clear that black participants have higher prevalence of diabetes at baseline (P-value < 0.001), so perhaps the pool of black participants susceptible to diabetes was reduced prior to enrollment in the study. This would result in the null associations that we see in the prospective models, but does not explain why cross-sectional models also suggest a null relationship. Because the ARIC study enrolled older adults by design, survivor bias may be affecting our results. If there is more premature mortality among black Americans than white Americans, blacks susceptible to diabetes may have died prior to having the opportunity to enroll in ARIC, which could result in the null association that we see.

In the Black Women's Health Study, which followed Black women ages 21-69, neighborhood socioeconomic status was associated with risk of diabetes. They used a slightly different summary of neighborhood socioeconomic status than used in this study: [25] their summary score also included median household income, median housing value, percent of households receiving interest or rental income, percent of adults 25+ with college completed, percent of adults ages 16+ who have professional, managerial, executive jobs; but used percent of families with children not headed by single female instead of percent of adults over age 25 who completed high school. This suggests that there may be indicators of neighborhood socioeconomic status that are more salient in black populations than white populations. Or, there could be effect modification by sex that is not accounted for in the present analysis. Future research should consider the role that gender plays in this relationship.

The strengths of this study include the prospective nature of the data and the use of a factor analysis-based, multi-dimensional measure of neighborhood socioeconomic status. This study examines the association of neighborhood socioeconomic status with diabetes over time; previous studies have been limited by the cross-sectional nature of the design. Cross-sectional analysis of neighborhood effects is severely limited by the potential for selection bias—for example, illness leading to the inability to work, loss of income, and a subsequent move into a lower SES neighborhood.[13] In the ARIC study, neighborhood socioeconomic status was derived from the residential address reported by each participant at Visit 1 (1987-9), and then patients were followed until 2012, with outcome ascertainment at study visits and through annual telephone interviews. Interestingly, the baseline and prospective models demonstrate the same inferences about the relationship between neighborhood socioeconomic status and diabetes risk. However, the possibility of selection bias prior to enrollment in the ARIC study is still possible, and this analysis did not account for persons who moved during the study follow-up.

Key limitations of this analysis include that we defined neighborhoods as census tracts, which are administrative units that vary in geographic and population size and do not necessarily correspond to more sociologically realistic definitions of community or "neighborhood". Therefore, meaningful differences in socioeconomic status that exists within a single census tract may be obscured. This is called the "modifiable area unit problem (MAUP)" and is an issue in much of the published literature on neighborhood effects.[41]

Neighborhood socioeconomic status was measured at the first study visit, when the average age of participants was 54 years old. While we do know that address at visit 1 preceded the diagnosis of diabetes, we do not have a residential history for participants, and did not update the neighborhood exposure during the study period. Therefore, the duration of the neighborhood exposure is unknown. Additionally, adult neighborhood socioeconomic status may only be an approximation of the relevant exposure. It may be that neighborhood socioeconomic status during childhood and/or parental SES may be the more important exposure. Furthermore, we cannot rule out the possibility that neighborhood socioeconomic status is not causally related to diabetes risk. It is possible that people who are sick experience downward mobility that leads them to move to lower socioeconomic status neighborhoods. This would be an unmeasured source of confounding. Future studies should consider residential history and parental residential history when data are available.

Conclusion

We provide evidence that neighborhood socioeconomic status is associated with the risk of diabetes among whites in the ARIC study and that this association is independent of the individual socioeconomic status. With 1 in 3 American adults at increased risk for diabetes,[22] there is great need for population-wide strategies to reduce the incidence of diabetes. Future research should focus on community-level interventions that can reduce the risk of diabetes. Reducing the burden of chronic disease in the United States may depend on identifying and mitigating the effects of the residential environment.

Sociologist Patrick Sharkey describes the effect of neighborhoods as "…a multiplicative function of neighborhood characteristics, the timing and duration of individuals' exposure to the neighborhood, and individuals' vulnerability to the effects of the neighborhood".[42] This framework and the results of this study suggest that longitudinal data in cohort studies such as ARIC can and should be used to evaluate the association of neighborhood socioeconomic status and health outcomes over time. Increased attention to residential history, duration of residential exposure, and the role of race and gender as potential modifiers may improve future research.

Tables and Figures

Figure 1: Conceptual Framework

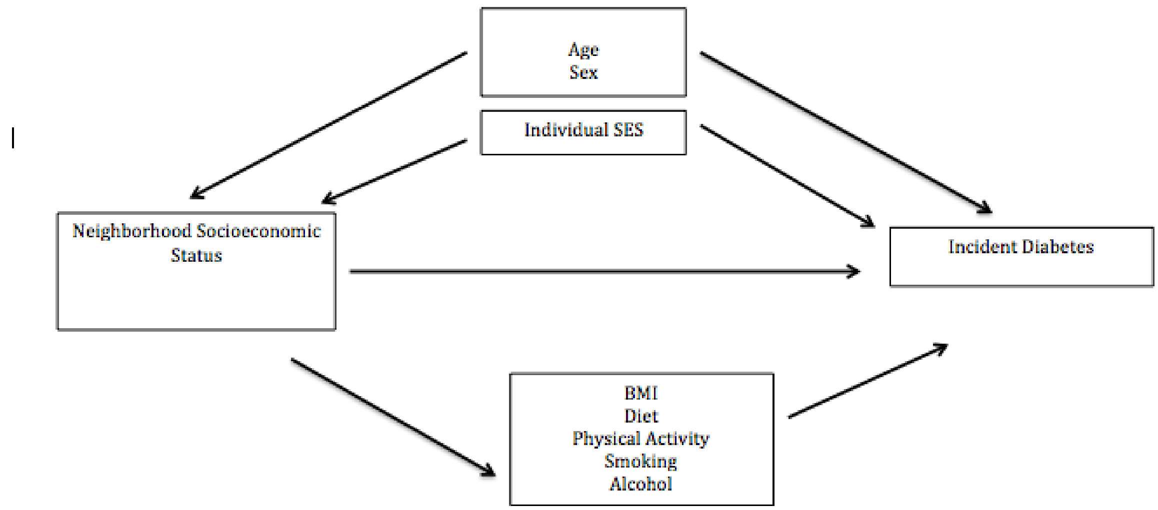

Figure 2: Frequency of Neighborhood Socioeconomic Status Summary Score by Race with Lowess Curve of the Log Odds of Diabetes by Neighborhood Socioeconomic Status Summary Score

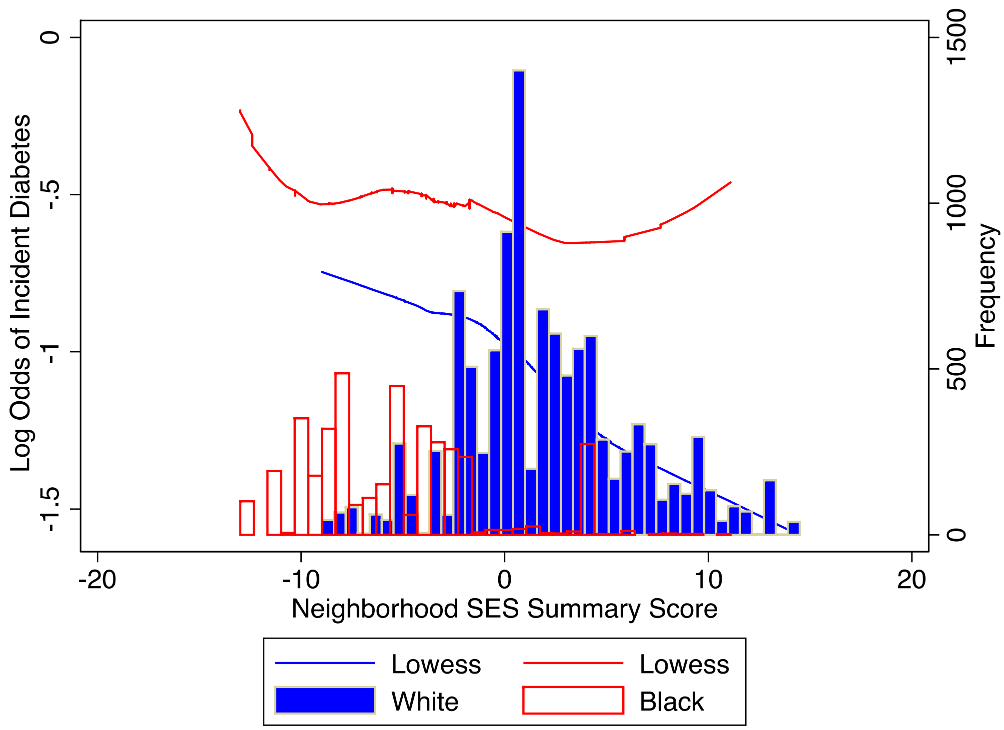

Figure 2: This histogram shows the frequency of neighborhood SES summary score value by race. A logistic lowess curve is included for both races to explore the linearity of the neighborhood summary score against the log-odds of incident diabetes. Among white participants, the neighborhood summary score is linearly related to the log odds of diabetes. Among Black participants, the relationship is approximately linear, although skewed by rare observations of neighborhood summary scores above 5.

Figure 3: Distribution of Neighborhood Socioeconomic Status Score by Study Site and Race

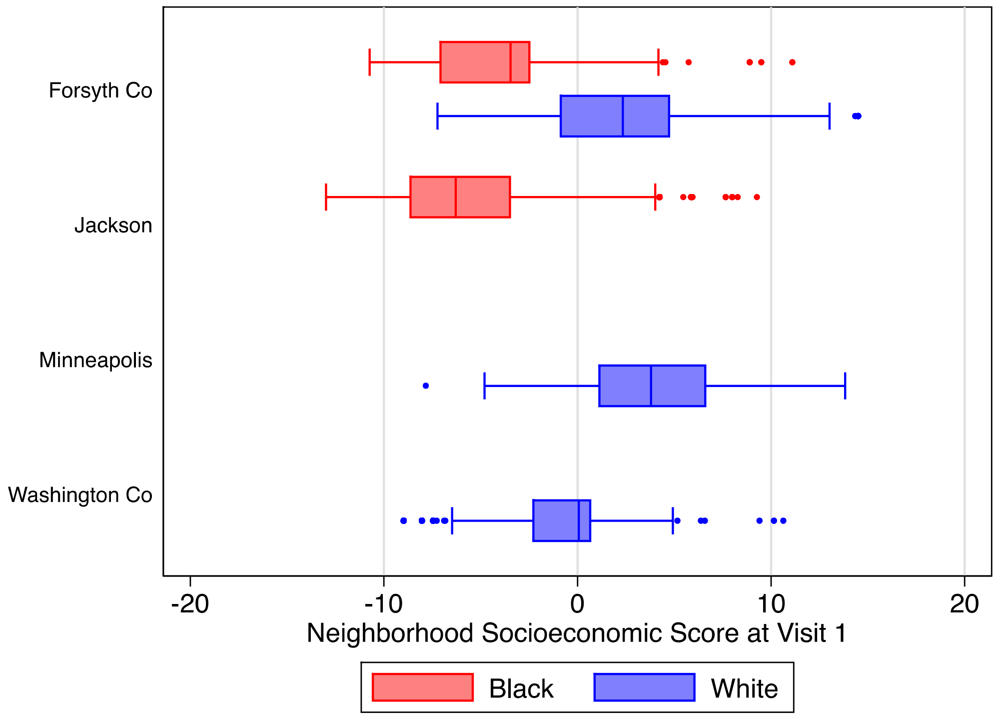

Figure 3: Overall, white participants have higher neighborhood SES scores than black participants. 22 Black participants were excluded from the Minneapolis site and 33 were excluded from the Washington County, MD site.

Table 1: Characteristics of ARIC Study Population

	White Participants			Black Participants			Overall
	Lowest Tertile of NB SES	**Middle Tertile of NB SES**	**Highest Tertile of NB SES**	**Lowest Tertile of NB SES**	**Middle Tertile of NB SES**	**Highest Tertile of NB SES**	
N	3,829	3,571	3,632	1,542	1,282	1,172	15,028
Mean Neighbor-hood SES Score (SD)	-2.12 (2.19)	1.60 (0.95)	6.79 (2.78)	-9.49 (1.40)	-5.51 (1.17)	-0.51(3.26)	-.0025 (5.40)
Mean Age (SD)	54.82 (5.71)	54.34 (5.71)	54.02 (5.74)	54.56 (5.79)	53.40 (5.83)	52.48 (5.67)	54.17 (5.76)
% Female	52.5%	53.5%	52.4%	63.4%	60.1%	55.0%	55.2%
BMI at Visit 1	27.58 (5.12)	26.99 (4.90)	26.40 (4.44)	30.01 (6.47)	29.78 (6.25)	29.03 (5.69)	27.71 (5.37)
Current or Former Smoker	61.2%	59.9%	60.3%	54.5%	49.0%	46.9%	58.4%
Alcohol Use (Mean Grams of Alcohol per Day (SD))**	5.07 (12.3)	6.80 (14.4)	8.13 (13.5)	5.00 (15.9)	3.80 (10.7)	4.34 (12. 8)	6.08 (13.8)
Physical Activity (Mean Sport Index Score)***	2.44 (0. .77)	2.50 (0.79)	2.66 (0.84)	2.07 (0.623)	2.17 (0.70)	2.22 (0.74)	2.43 (0.79)
% Income < 25,000*	48.2%	29.9%	14.7%	75.3%	59.8%	38.4%	38.9%
% Less than HS	45.5%	29.5%	16.8%	37.6%	19.8%	6.5%	23.9%
% College	16.6%	27.7%	55.2%	25.7%	45.2%	59.63%	35.4%

* Median Income in 1987 (in 1987 dollars, roughly equivalent to $51,000 in 2015)
** The CDC defines a "standard drink" as 14.0 grams of alcohol; i.e. one 12 oz beer, 5 oz of wine, or 1.5 oz of distilled spirits or liquor. [43]
*** Score values range from 1 (low activity) to 5 (high activity)

Table 2: Diabetes in the ARIC study

	White Participants			Black Participants			Overall
	Lowest Tertile of NB SES N (%)	Middle Tertile of NB SES N (%)	Highest Tertile of NB SES N (%)	Lowest Tertile of NB SES N (%)	Middle Tertile of NB SES N (%)	Highest Tertile of NB SES N (%)	N (%)
Number at Baseline	3,829	3,571	3,632	1,542	1,282	1,172	15,028
Diabetes at Baseline (%)	430 (11.2%)	316 (8.9%)	251 (6.9%)	332 (21.5%)	241 (18.8%)	199 (17.0%)	1863 (11.9%)
Incident Diabetes cases after Baseline	1,000	782	677	428	373	348	3782
Person-years of follow-up	55209.654	54745.93	59387.669	17171.116	15796.375	15241.667	226558
Diabetes Incidence Rates per 1,000 PY	18.11	14.28	11.40	24.93	23.61	22.83	16.69

Table 3: Odds Ratios (with 95% Confidence Intervals) from Logistic Models of Cross-Sectional Association of Neighborhood SES and Diabetes at Visit 1, Among White Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.943*** (0.929 - 0.958)	0.946*** (0.932 - 0.961)	0.978** (0.961 - 0.995)
Age, years		1.060*** (1.048 - 1.072)	1.045*** (1.033 - 1.059)
Female Sex		0.813*** (0.713 - 0.926)	0.740*** (0.644 - 0.849)
Income			
<$5000			REF
$5000-7999			1.048 (0.619 - 1.775)
$8000-11,999			0.697 (0.427 - 1.138)
$12,000-15,999			0.589** (0.367 - 0.944)
$16,000-24,999			0.450*** (0.289 - 0.700)
$25,000-34,999			0.413*** (0.266 - 0.641)
$35,000-49,999			0.395*** (0.253 - 0.616)
Over $50,000			0.317*** (0.201 - 0.500)
Education			
< High School			REF
High School			0.876 (0.734 - 1.047)
College			0.742*** (0.601 - 0.917)
Constant	0.110*** (0.103 - 0.117)	0.005*** (0.003 - 0.009)	0.029*** (0.012 - 0.067)
Observations	11,032	11,032	10,546
Number of Census Tracts	287	287	280

*** p<0.01, ** p<0.05

Table 4: Odds Ratios (with 95% Confidence Intervals) from Logistic Models of Cross-Sectional Association of Neighborhood SES and Diabetes at Visit 1, Among Black Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.969** (0.945 - 0.994)	0.982 (0.957 - 1.008)	0.999 (0.972 - 1.026)
Age		1.061*** (1.047 - 1.076)	1.046*** (1.030 - 1.062)
Female		1.188** (1.008 - 1.400)	1.167 (0.975 - 1.397)
<u>Income</u>			
< $5000			REF
$5000-7999			0.750 (0.557 - 1.009)
$8000-11,999			0.936 (0.709 - 1.235)
$12,000-15,999			0.870 (0.648 - 1.166)
$16,000-24,999			0.630*** (0.470 - 0.843)
$25,000-34,999			0.606*** (0.430 - 0.855)
$35,000-49,999			0.606** (0.409 - 0.898)
Over $50,000			0.409*** (0.243 - 0.686)
<u>Education</u>			
< High School			REF
High school			0.832 (0.677 - 1.022)
College			0.783 (0.611 - 1.004)
Constant	0.209*** (0.174 - 0.252)	0.008*** (0.004 - 0.018)	0.029*** (0.012 - 0.073)
Observations	3,996	3,996	3,597
Number of Census Tracts	109	109	108

*** p<0.01, ** p<0.05

Figure 4: Cumulative Hazard of Incident Diabetes by Race-Specific Tertile of Neighborhood Socioeconomic Status Summary Score

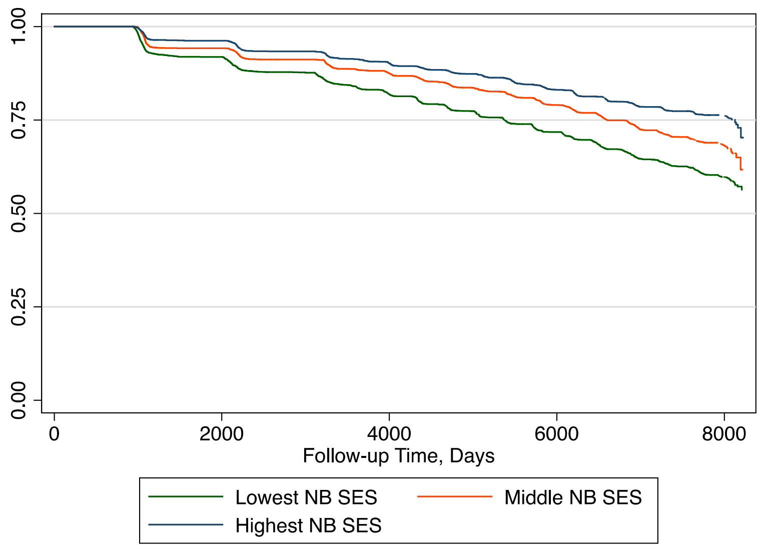

Figure 5: Kaplan Meier Log-Log Plot of Diabetes-Free Survival by Tertile of Neighborhood Socioeconomic Status Summary Score

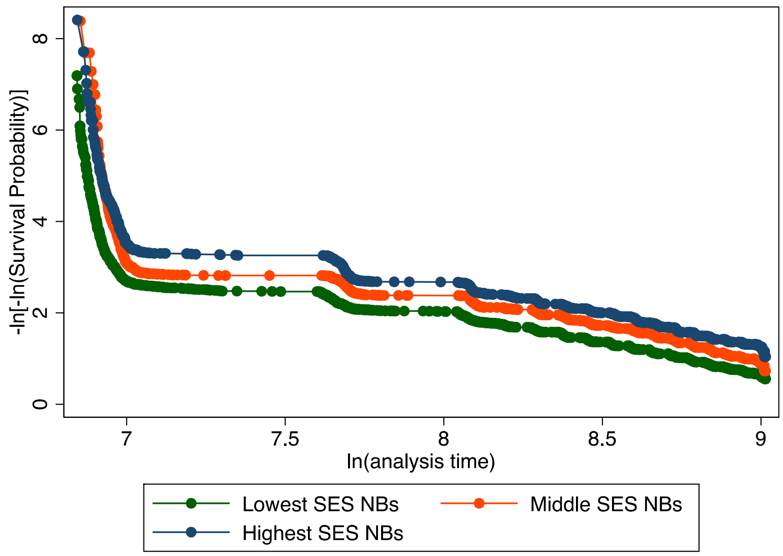

Figure 6: Cumulative Hazard of Incident Diabetes, by Race and Race-Specific Tertile of Neighborhood Socioeconomic Status Summary Score

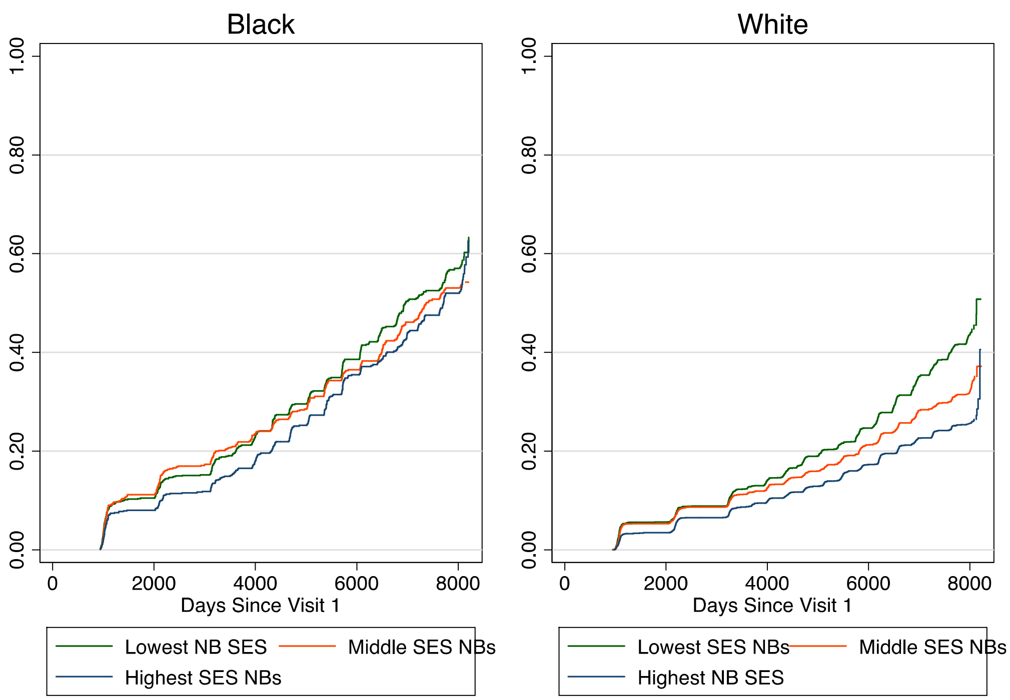

Table 5: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association between Neighborhood Socioeconomic Status and Incident Diabetes, Among White Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.955*** (0.947 - 0.964)	0.955*** (0.947 - 0.964)	0.972*** (0.962 - 0.982)
Age		1.004 (0.997 - 1.012)	1.000 (0.993 - 1.008)
Female		0.797*** (0.732 - 0.867)	0.775*** (0.708 - 0.847)
Income			
< $5000			REF
$5000-7999			0.682 (0.408 - 1.139)
$8000-11,999			0.858 (0.564 - 1.304)
$12-15,999			0.669** (0.454 - 0.984)
$16-24,999			0.685** (0.470 - 0.999)
$20 -34,999			0.707* (0.488 - 1.025)
$35-49,999			0.684** (0.471 - 0.994)
Over $50,000			0.644** (0.440 - 0.941)
Education			
< High School			REF
High school			0.830*** (0.743 - 0.926)
College			0.666*** (0.588 - 0.754)
Obs	9,779	9,779	9,351

*** p<0.01, ** p<0.05

Table 6: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association between Neighborhood Socioeconomic Status and Incident Diabetes, Among Black Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.986 (0.971 - 1.001)	0.985 (0.970 - 1.001)	0.999 (0.983 - 1.015)
Age		0.998 (0.991 - 1.006)	0.991** (0.983 - 0.999)
Female		0.995 (0.865 - 1.145)	0.954 (0.828 - 1.099)
<u>Income</u>			
< $5000			REF
$5000-7999			0.882 (0.694 - 1.121)
$8000-11,999			0.731*** (0.579 - 0.922)
$12-15,999			0.820 (0.660 - 1.019)
$16-24,999			0.728*** (0.581 - 0.912)
$20-34,999			0.804 (0.601 - 1.075)
$35-49,999			0.743** (0.563 - 0.981)
Over $50,000			0.691*** (0.534 - 0.895)
<u>Education</u>			
< High School			REF
High school			0.911 (0.784 - 1.059)
College			0.854** (0.757 - 0.963)
Obs	3,064	3,064	2,747

*** p<0.01, ** p<0.05

Table 7: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association Between Neighborhood Socioeconomic Status and Self-Reported Incident Diabetes, Among White Participants

	Unadjusted	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.954*** (0.945 - 0.964)	0.954*** (0.945 - 0.964)	0.970*** (0.959 - 0.981)
Age		1.001 (0.993 - 1.009)	0.997 (0.989 - 1.005)
Female		0.823*** (0.757 - 0.894)	0.799*** (0.733 - 0.872)
<u>Income</u>			
<$5000			REF
$5000-7999			0.614* (0.351 - 1.074)
$8000-11,999			0.733 (0.472 - 1.139)
$12,000-15,999			0.629** (0.421 - 0.940)
$16,000-24,999			0.616** (0.416 - 0.913)
$25,000-34,999			0.626** (0.425 - 0.924)
$35,000-49,999			0.607** (0.410 - 0.899)
Over $50,000			0.572*** (0.383 - 0.855)
<u>Education</u>			
Less than HS			REF
High School			0.817*** (0.727 - 0.917)
College			0.674*** (0.592 - 0.768)
Obs	9,779	9,779	9,351

*** p<0.01, ** p<0.05

Table 8: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association Between Neighborhood Socioeconomic Status and Self-Reported Incident Diabetes, Among Black Participants

	Unadjusted	Model 1	Model 2c
Neighborhood Socioeconomic Status Summary Score	0.990 (0.975 - 1.005)	0.990 (0.975 - 1.005)	1.006 (0.990 - 1.023)
Age		1.000 (0.991 - 1.008)	0.992 (0.981 - 1.002)
Female		0.999 (0.867 - 1.151)	0.940 (0.817 - 1.083)
<u>Income</u>			
<$5000			REF
$5000-7999			0.881 (0.678 - 1.144)
$8000-11,999			0.690*** (0.537 - 0.887)
$12,000-15,999			0.829 (0.647 - 1.061)
$16,000-24,999			0.683*** (0.547 - 0.852)
$25,000-34,999			0.802 (0.584 - 1.103)
$35,000-49,999			0.709** (0.538 - 0.935)
Over $50,000			0.622*** (0.466 - 0.829)
<u>Education</u>			
Less than HS			REF
High School			0.894 (0.762 - 1.049)
College			0.829*** (0.727 - 0.947)
Obs	3,064	3,064	2,747

*** p<0.01, ** p<0.05

Table 9: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association Between Black-Specific Neighborhood Socioeconomic Status and Incident Diabetes, Among Black Participants

	Model 1	Model 2
Black Neighborhood Socioeconomic Status Score*	0.985 (0.970 - 1.001)	0.999 (0.983 - 1.015)
Age	0.998 (0.991 - 1.006)	0.991** (0.983 - 0.999)
Female	0.995 (0.865 - 1.145)	0.954 (0.828 - 1.099)
Income		
<$5000		REF
$5000-7999		0.882 (0.694 - 1.121)
$8000-11,999		0.731*** (0.579 - 0.922)
$12,000-15,999		0.820 (0.660 - 1.019)
$16,000-24,999		0.728*** (0.581 - 0.912)
$25,000-34,999		0.804 (0.601 - 1.075)
$35,000-49,999		0.743** (0.563 - 0.981)
Over $50,000		0.691*** (0.534 - 0.895)
Education		
Less than HS		
High School		0.911 (0.784 - 1.059)
College and Above		0.854** (0.757 - 0.963)
Observations	3,064	2,747

* Calculated using data from black participants only.

Table 10: Results of Competing Risk Cox Models of Association between Neighborhood Socioeconomic Status and Incident Diabetes, Among Black Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.986 (0.971 - 1.001)	0.989 (0.975 - 1.004)	0.999 (0.983 - 1.015)
Age		0.995** (0.978 – 0.992)	0.979** (0.971 - 0.988)
Female		1.077 (0.935 - 1.241)	1.054 (0.910 - 1.220)
<u>Income</u>			
< $5000			REF
$5000-7999			0.940 (0.741 - 1.192)
$8000-11,999			0.804 (0.632 – 1.022)
$12-15,999			0.903 (0.738 - 1.106)
$16-24,999			0.837 (0.657 – 1.066)
$20-34,999			0.919 (0.692 - 1.226)
$35-49,999			0.876 (0.563 - 0.981)
Over $50,000			0.812 (0.624 – 1.057)
<u>Education</u>			
< High School			REF
High school			0.930 (0.802 - 1.077)
College			0.858** (0.753 - 0.978)
Obs	3,064	3,064	2,747

*** p<0.01, ** p<0.05

<u>**Table 10 is a competing risks analysis including the competing risk of death among black participants.**</u>

Table 11: Odds Ratios (with 95% Confidence Intervals) from Logistic Models of Cross-Sectional Association of Neighborhood Socioeconomic Status and Diabetes at Visit 1, Among All Black Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic	0.966***	0.978	0.996
Status Summary Score	(0.943 - 0.989)	(0.954 - 1.003)	(0.970 - 1.022)
Age		1.060***	1.045***
		(1.045 - 1.074)	(1.030 - 1.061)
Female		1.194**	1.166*
		(1.014 - 1.407)	(0.974 - 1.394)
Income			
< $5000			REF
$5000-7999			0.749
			(0.556 - 1.008)
$8000-11,999			0.939
			(0.712 - 1.239)
$12-15,999			0.855
			(0.638 - 1.146)
$16-24,999			0.619***
			(0.462 - 0.828)
$20-34,999			0.587***
			(0.416 - 0.828)
$35-49,999			0.620**
			(0.420 - 0.915)
Over $50,000			0.389***
			(0.231 - 0.654)
Education			
Less than HS			REF
High School			0.829
			(0.676 - 1.018)
College			0.792
			(0.618 - 1.015)
Constant	0.201***	0.008***	0.030***
	(0.169 - 0.240)	(0.004 - 0.018)	(0.012 - 0.074)
Observations	4,051	4,051	3,645
Number of census tracts	135	135	131

*** p<0.01, ** p<0.05

Table 11 includes 55 black participants who were excluded from the main models because they live in the primarily white sites in Maryland and Minnesota.

Table 12: Hazard Ratios (with 95% Confidence Intervals) from Cox Models of Association Between Race-Specific Neighborhood Socioeconomic Status and Incident Diabetes, Among All Black Participants

	Crude	Model 1	Model 2
Neighborhood Socioeconomic Status Summary Score	0.987* (0.972 - 1.002)	0.987* (0.972 - 1.002)	1.000 (0.984 - 1.016)
Age		0.999 (0.992 - 1.007)	0.992** (0.983 - 1.000)
Female		0.999 (0.869 - 1.149)	0.965 (0.837 - 1.113)
Income			
< $5000			REF
$5000-7999			0.884 (0.695 - 1.123)
$8000-11,999			0.734*** (0.582 - 0.926)
$12-15,999			0.828 (0.667 - 1.029)
$16-24,999			0.737*** (0.588 - 0.923)
$20-34,999			0.799 (0.598 - 1.068)
$35-49,999			0.757** (0.573 - 1.000)
Over $50,000			0.705*** (0.545 - 0.911)
Education			
Less than High School			REF
High School			0.898 (0.773 - 1.045)
College			0.852*** (0.755 - 0.961)
Observations	3,111	3,111	2,789

*** p<0.01, ** p<0.05

Table 12 includes 55 black participants who were excluded from the main models because they live in the primarily white sites in Maryland and Minnesota.

References

1. Diez-Roux AV, Kiefe CI, Jacobs DR, Jr., et al. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Annals of epidemiology.* Aug 2001;11(6):395-405.
2. Diez Roux AV, Mair C. Neighborhoods and health. *Annals of the New York Academy of Sciences.* Feb 2010;1186:125-145.
3. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. *Obesity reviews : an official journal of the International Association for the Study of Obesity.* Mar 2011;12(3):217-230.
4. Diez Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *The New England journal of medicine.* Jul 12 2001;345(2):99-106.
5. Borrell LN, Diez Roux AV, Rose K, Catellier D, Clark BL, Atherosclerosis Risk in Communities S. Neighbourhood characteristics and mortality in the Atherosclerosis Risk in Communities Study. *International journal of epidemiology.* Apr 2004;33(2):398-407.
6. Mujahid MS, Diez Roux AV, Borrell LN, Nieto FJ. Cross-sectional and longitudinal associations of BMI with socioeconomic characteristics. *Obesity research.* Aug 2005;13(8):1412-1421.
7. Chichlowska KL, Rose KM, Diez-Roux AV, Golden SH, McNeill AM, Heiss G. Individual and neighborhood socioeconomic status characteristics and prevalence of metabolic syndrome: the Atherosclerosis Risk in Communities (ARIC) Study. *Psychosomatic medicine.* Nov 2008;70(9):986-992.
8. Geiss LS, Wang J, Cheng YJ, et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. *Jama.* Sep 24 2014;312(12):1218-1226.
9. Centers for Disease Control and Prevention. *National Diabetes Statistics Report, 2014.* 2014.
10. Brown AF, Ettner SL, Piette J, et al. Socioeconomic position and health among persons with diabetes mellitus: a conceptual framework and review of the literature. *Epidemiologic reviews.* 2004;26:63-77.
11. Gary-Webb TL, Suglia SF, Tehranifar P. Social epidemiology of diabetes and associated conditions. *Current diabetes reports.* Dec 2013;13(6):850-859.
12. 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. *Archives of internal medicine.* Oct 12 2009;169(18):1698-1704.
13. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Social science & medicine.* May 2004;58(10):1929-1952.
14. Diez Roux AV. Neighborhoods and health: where are we and were do we go from here? *Revue d'epidemiologie et de sante publique.* Feb 2007;55(1):13-21.
15. Diez Roux AV. Investigating neighborhood and area effects on health. *American journal of public health.* Nov 2001;91(11):1783-1789.

16. Yen IH, Syme SL. The social environment and health: a discussion of the epidemiologic literature. *Annual review of public health.* 1999;20:287-308.
17. Diez Roux AV. Residential environments and cardiovascular risk. *Journal of urban health : bulletin of the New York Academy of Medicine.* Dec 2003;80(4):569-589.
18. Prevention CfDCa. Number (in Millions) of Civilian, Noninstitutionalized Persons with Diagnosed Diabetes, United States, 1980–2011. http://www.cdc.gov/diabetes/statistics/prev/national/figpersons.htm. Accessed 03/13/2015, 2015.
19. Prevention CfDCa. Prediabetes: Could it be You? .
20. Prevention CfDCa. Age-Adjusted Rate per 100 of Civilian, Noninstitutionalized Population with Diagnosed Diabetes, by Education, United States, 1980–2011. http://www.cdc.gov/diabetes/statistics/prev/national/figbyeducation.htm.
21. Prevention CfDCa. Age-Adjusted Rate per 100 of Civilian, Noninstitutionalized Population with Diagnosed Diabetes, by Race, United States, 1980–2011. http://www.cdc.gov/diabetes/statistics/prev/national/figbyrace.htm.
22. Centers for Disease Control and Prevention. Preventing Diabetes. http://www.cdc.gov/diabetes/basics/prevention.html. Accessed March 14, 2015.
23. American Diabetes A. Economic costs of diabetes in the U.S. in 2012. *Diabetes care.* Apr 2013;36(4):1033-1046.
24. Ludwig J, Sanbonmatsu L, Gennetian L, et al. Neighborhoods, obesity, and diabetes--a randomized social experiment. *The New England journal of medicine.* Oct 20 2011;365(16):1509-1519.
25. Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic status and incidence of type 2 diabetes: results from the Black Women's Health Study. *American journal of epidemiology.* Mar 1 2010;171(5):564-570.
26. Murray ET, Diez Roux AV, Carnethon M, Lutsey PL, Ni H, O'Meara ES. Trajectories of neighborhood poverty and associations with subclinical atherosclerosis and associated risk factors: the multi-ethnic study of atherosclerosis. *American journal of epidemiology.* May 15 2010;171(10):1099-1108.
27. Schootman M, Andresen EM, Wolinsky FD, et al. The effect of adverse housing and neighborhood conditions on the development of diabetes mellitus among middle-aged African Americans. *American journal of epidemiology.* Aug 15 2007;166(4):379-387.
28. Auchincloss AH, Diez Roux AV, Brown DG, O'Meara ES, Raghunathan TE. Association of Insulin Resistance with Distance to Wealthy Areas: The Multi-Ethnic Study of Atherosclerosis. *American journal of epidemiology.* 2006;165(4):389-397.
29. Cox M, Boyle PJ, Davey PG, Feng Z, Morris AD. Locality deprivation and Type 2 diabetes incidence: a local test of relative inequalities. *Social science & medicine.* Nov 2007;65(9):1953-1964.
30. Diez Roux A, Jacobs, DR & Kiefe, CI. Neighborhood characteristics and components of the insulin resistance syndrome in young adults: the coronary artery risk development in young adults (CARDIA) study. *Diabetes care.* 2002;25(11):1976-1982.
31. Liu L, Nunez AE. Multilevel and urban health modeling of risk factors for diabetes mellitus: a new insight into public health and preventive medicine. *Advances in preventive medicine.* 2014;2014:246049.

32. Gaskin DJ, Thorpe RJ, Jr., McGinty EE, et al. Disparities in diabetes: the nexus of race, poverty, and place. *American journal of public health.* Nov 2014;104(11):2147-2155.
33. Abeyta IM, Tuitt NR, Byers TE, Sauaia A. Effect of community affluence on the association between individual socioeconomic status and cardiovascular disease risk factors, Colorado, 2007-2008. *Preventing chronic disease.* 2012;9:E115.
34. Muller G, Kluttig A, Greiser KH, et al. Regional and neighborhood disparities in the odds of type 2 diabetes: results from 5 population-based studies in Germany (DIAB-CORE consortium). *American journal of epidemiology.* Jul 15 2013;178(2):221-230.
35. Sampson RJ, Sharkey P. Neighborhood selection and the social reproduction of concentrated racial inequality. *Demography.* Feb 2008;45(1):1-29.
36. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *American journal of epidemiology.* Apr 1989;129(4):687-702.
37. Life Course Socioeconomic Status (LCSES): Manual of Procedures and Overview of Data Sets Available http://www.lifecourseepi.info/lifecourse2/forms/m1704m_web.pdf.
38. United States Census Bureau. Geographic Terms and Concepts - Census Tract. http://www.census.gov/geo/reference/gtc/gtc_ct.html.
39. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association.* 1999;94(446):496-509.
40. Nau C, Schwartz BS, Bandeen-Roche K, et al. Community socioeconomic deprivation and obesity trajectories in children using electronic health records. *Obesity.* Jan 2015;23(1):207-212.
41. *Methods in Social Epidemiology.* Jossey-Bass; 2006.
42. Sharkey P, Faber JW. Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects. *Annual Review of Sociology.* 2014;40(1):559-579.
43. Centers for Disease Control and Prevention. Alcohol & Public Health: Frequently Asked Questions 2014; http://www.cdc.gov/alcohol/faqs.htm. Accessed April 15, 2015, 2015.

Curriculum Vitae

Emily A. Knapp

EDUCATION

Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD
May 2015

- Master of Health Science in Epidemiology; General Epidemiology and Methodology

Beloit College, Beloit, WI **December 2009**

- Bachelor of Arts in Health and Society; Summa Cum Laude, *Phi Beta Kappa*

AWARDS

Nancy Fink Award for Scholarship and Service, Johns Hopkins School of Public Health
May 2014
Claudette Cummings Community Health Award, Beloit College **May 2010**
Phi Beta Kappa **May 2009**

RESEARCH EXPERIENCE

Patient Registry Coordinator, Cystic Fibrosis Foundation, Bethesda, MD
September 2010- Present

Research Assistant, Welch Center for Prevention, Epidemiology, and Clinical Research
Johns Hopkins University School of Public Health, Baltimore, MD **June - September 2014**

TEACHING EXPERIENCE

Teaching Assistant, Epidemiology of Diabetes and Obesity course, Johns Hopkins School of Public Health **Spring 2015**

Tutor, Epidemiology and Biostatistics, Johns Hopkins School of Public Health
July 2014-present

Coordinator, Social Epidemiology Journal Club, Johns Hopkins School of Public Health
AY 2014-2015

Coordinator, General Epidemiology & Methodology Journal Club, Johns Hopkins School of Public Health **AY 2014-present**

PUBLICATIONS

Goss CH, MacNeill SJ, Quinton H, Marshall BC, Elbert A, **Knapp E**, Petren K, Sewall A, Gunn E, Osmond J, Bilton D. Children and Young Adults in the US Have Improved Lung Function Compared to the UK. Thorax. 2014 Sept 25: Epub ahead of print.

MacKenzie T, Gifford AH, Sabadosa KA, Quinton HB, **Knapp EA**, Goss CH, et al. Longevity of Patients With Cystic Fibrosis in 2000 to 2010 and Beyond: Survival Analysis of the Cystic Fibrosis Foundation Patient Registry. Ann Intern Med. 2014; 161:233-241.

Case Example 29. Challenges in transitions and changes in data collection. In Gliklich R, Dreyer N, Leavy M, eds. Registries for Evaluating Patient Outcomes: A User's Guide. Third edition. Two volumes. AHRQ Publication No. 13(14)-EHC111. Rockville, MD: Agency for Healthcare Research and Quality. April 2014. http://www.effectivehealthcare.ahrq.gov/registries-guide-3.cfm.

Maisonneuve P, Marshall BC, **Knapp EA**, Lowenfels AB. Cancer risk in cystic fibrosis: A 20-year nationwide study from the United States. *Journal of the National Cancer Institute*. 2013 Jan 16; 105(2): 122-9.

ABSTRACTS

Knapp EA, Dougherty GB, Bilal U, Burke BT, Glass TA. Obesogenic environments in childhood in PA: a network approach. Accepted for oral presentation at the 48th Annual Society for Epidemiologic Research meeting, June 2015.

Knapp E, Salsgiver E, Fink A, Sewall A, Marshall B, Saiman L. Trends in Respiratory Microbiology of People with Cystic Fibrosis in the United States, 2006-2012. Presented at the 28th Annual North American Cystic Fibrosis Conference, October 2014.

Dasenbrook E, Gemmen E, Choodnovskiy I, Sawicki GS, Hamblett NM, **Knapp EA**, Fink A, Goss CH. Association between the addition of inhaled hypertonic saline to RhDNase and clinical outcomes: propensity score matched analysis of the US CFF patient registry. Presented at the 28th Annual North American Cystic Fibrosis Conference, October 2014.

Fink AK, Marshall BC, **Knapp EA**, Engels EA. Risk of Cancer in Cystic Fibrosis Following Lung Transplantation. Presented at the 27th Annual North American Cystic Fibrosis Conference, October 2013.

Goss CH, MacNeill SJ, Quinton H, Marshall BC, Elbert A, **Knapp E**, Petren K, Sewall A, Gunn E, Osmond J, Bilton D. Children and Young Adults in the US Have Improved Lung Function Compared to the UK. Presented at the 26th Annual North American Cystic Fibrosis Conference, October 2012.