Associations Between Long-Term Ambient Air Pollution, Neighborhood Physical Disinvestment, and Incident Diabetes in the Hispanic Community Health Study/Study of Latinos

Cara Smith

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Joel Kaufman, Chair
Elena Austin
Amanda Fretts
Stephen Mooney

Program Authorized to Offer Degree:

Department of Environmental and Occupational Health Sciences

University of Washington

Abstract

Associations Between Long-Term Ambient Air Pollution, Neighborhood Physical Disinvestment, and Incident Diabetes in the Hispanic Community Health Study/Study of Latinos

Cara Smith

Chair of the Supervisory Committee:
Joel Kaufman
Department of Environmental and Occupational Health Sciences

Risk factors for type 2 diabetes (T2D) include genetics, age, adiposity, diet, physical activity (PA) levels and environmental conditions.[1–3] Putative environmental risk factors include exposure to air pollution-- specifically traffic-related air pollutants--and neighborhood conditions.[4–7] A meta-analysis by Yang et al.[5] of 11 studies reported an association between fine particulate matter ($PM_{2.5}$) and incident T2D (HR= 1.10, 95% CI: 1.04-1.17 per 10 $\mu g/m^3$ increment). Studies on NO_2 have been limited and associations more mixed for incident T2D.[5] There have previously been a limited number of studies on neighborhood environment, specifically neighborhood physical disinvestment (formerly referred to as disorder), and T2D. Disinvestment in the built environment is a result of systemic social and economic processes that deprive areas of resources needed to maintain and develop physical environments. Disinvestment is reflected in visual indicators of deterioration and neglect, and is a concern for health as areas of high disinvestment may reduce resident's ability and willingness to engage in PA.[8] Increasing neighborhood disorder has also been found to interact with genetic risk of T2D, as those with

higher polygenic score for T2D (a summary of genome-wide genetic risk) living in neighborhoods of higher levels of disorder had a greater risk of T2D.[9] Data on air pollution, neighborhood environment, and T2D risk is limited and even more so in Hispanic/Latino populations and subgroups at high risk of diabetes.

This dissertation addresses those data gaps in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), a longitudinal community-based cohort study of 16,415 self-identified Hispanic/Latinos recruited between 2008 and 2011 from the Bronx, Chicago, Miami, and San Diego. Yearly follow-up calls were conducted and a second visit occurred between 2014-2017. Laboratory analysis from both exams includes fasting plasma glucose (FPG) and Glycosylated hemoglobin (HbA1c). Diabetes was defined in two way: the primary definition for diabetes required a FPG ≥126 mg/dL; HbA1c level ≥6.5%; post-OGTT glucose ≥ 200 mg/dL; and/or self-reported use of antihyperglycemic drugs,[10] and the secondary definition included all those who fit the primary definition of T2D and additionally includes those who self-reported diabetes or high sugar in their blood. Additionally, different progression of diabetes were investigated and pre-diabetes was defined as having a FPG ≥100 mg/dL and <126 mg/dL; and HbA1c level ≥5.7% and <6.5%.[10] The different progression of diabetes were: 1) progression from pre-diabetes to diabetes, where the sample only included those with pre-diabetes at baseline, 2) Progression of those free of diabetes and pre-diabetes to diabetes and 3) progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes at visit 2.

The first aim of this dissertation focuses on estimating long-term exposure to ambient $PM_{2.5}$ and NO_2 in HCHS/SOL cohort members and the association of these pollutants with incident diabetes, and if these associations are modified by physical activity levels and/or adiposity. A spatiotemporal model was used to estimate air pollution exposure at residential

addresses. Individual survey-weighted Poisson regression models were used to assess the association between each air pollutant's average concentration over the follow-up period and incident diabetes at visit 2. Of the 8248 participants free of diabetes (primary definition) at visit 1, there were 950 incident diabetes cases at visit 2 with an age-centered adjusted incidence rate of 16.3 per 1000 person-years. For the secondary diabetes definition, 8129 participants were free of diabetes at visit 1 and there were 1354 incident cases by visit 2 with an age-centered adjusted incidence rate of 24.5 per 1000 person-years. When utilizing the primary definition of diabetes, a 2.48 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a 27% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.27, 95% CI: 1.09 – 1.49). An 8.69 ppb increase in NO_2 was associated with a 22% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.22, 95% CI: 1.08 – 1.37). Associations with both $PM_{2.5}$ and NO_2 were attenuated and no longer statistically significant when adjusting for years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ethnic heritage. There was no evidence of effect modification by physical activity or waist circumference. When adjusting for our main set of covariates at each study center, higher exposure to $PM_{2.5}$ and NO_2 were only associated with increased risk of incident diabetes at the Bronx study center. Future studies should continue to investigate the association between $PM_{2.5}$ and NO_2 and incident diabetes in a less spatially clustered cohort with a longer follow-up period. Future studies should also continue to investigate the potential modifying effects of physical activity and adiposity.

The second aim of this dissertation focused on generating neighborhood disinvestment scores for each participants' residential address in the HCHS/SOL and investigated the association between baseline neighborhood disinvestment and incident diabetes at visit 2. A

virtual street audit of Google Street View Imagery was conducted and an item response theory model was fit to indicators (litter, graffiti, under-maintained buildings, bars on windows, and abandoned buildings) to form a scale measuring a latent level of disinvestment. Ordinary kriging was used to estimate levels for each residential address within the HCHS/SOL census tracts via spatial interpolation. Covariate-adjusted and survey-weighted Poisson regression models were used to investigate the association between neighborhood disinvestment and incident diabetes at visit 2 among 9120 participants free of diabetes at visit 1. A sensitivity analysis among those who did not move during the follow-up period was conducted as well. A one-standard deviation increase in neighborhood disinvestment score was associated with a 13% (95% CI: 1%-23%) lower risk of incident diabetes when adjusting for age, sex, education, income, study center/heritage, years in the US, family history, and a neighborhood socioeconomic index. Our sensitivity analysis yielded qualitatively similar results with lower precision. Overall, our analysis does not support the hypothesis that neighborhood physical disinvestment is associated with incident type 2 diabetes in this Hispanic/Latino population. Future studies should continue to evaluate the role of food environment, greenspace, and physical activity when considering neighborhood scale infrastructure impacts on diabetes risk.

Table of Contents

List of Tables

List of Figures

List of Supplementary Tables

Acknowledgements

First, I would like to acknowledge my PhD Committee Chair Dr. Joel Kaufman. His guidance and feedback over the past five years has been immensely helpful in completing this dissertation. Dr. Kaufman has also successfully created a great research group to surround me and answer my questions whenever needed. Group members include Elizabeth Spalt, Amanda Gassett, Dustin Burham, Logan Piepmeier, Karen Jansen, Kevin Darole, and Xinmei Huang. This group not only helped in completing this work but also in providing fun, joy, and much needed random conversations during my PhD studies.

I would like to thank and acknowledge the rest of my PhD committee: Dr. Elena Austin, Dr. Mandy Fretts, Dr. Steve Mooney, and Dr. India Ornelas. All have been extremely helpful during this process and provided me with great guidance and suggestions.

The following undergraduate and graduate students at the University of Washington were a key part of completing the virtual street audits: Daanya Siddiqui, Dimple Vaidya, Emma Waddel, Greta Gunning, Hailey Hummel, Nede Oviebo, Ryan Brownell, Timilehin Toye Eyitayo

Many thanks to all the HCHS/SOL study participants and study staff, especially the field center staff members that helped up with the Ogawa sampling campaigns. Fellow HCHS/SOL collaborators have also been helpful over the past few years to provide feedback on various HCHS/SOL related work and papers: Dr. Matthew Allison, Dr. Jordan Carlson, Dr. Earle Chambers, Dr. Christina Cordero, Dr. Linda Gallo, Dr. Robert Kaplan, Dr. Qibin Qi.

Thank you to all the students, staff, and faculty in the Department of Environmental and Occupational Health Sciences at the University of Washington. Special shout out to our monthly bake-off members and our intramural soccer team.

Special shout-out to my non-UW academic community (i.e non-UW friends and family). You added more fun, enjoyment, distractions, support these past few years. I've enjoyed all the hangs, hikes, field hockey tournaments, various sport matches, and dessert consumption.

Shout out to all the dogs; you're the best and bring so much joy.

This work was supported by the National Institute of Environmental Health Sciences [R01ES030994: SOLAir: Environmental Factors and Diabetes Development in Latinos]; The National Institute of Health [5P30ES007033: Interdisciplinary Center for Exposures, Diseases, Genomics, and Environment (EDGE Center)]; New York Regional Center for Diabetes Translation Research [DK111022]; and the University of Washington Diabetes Research Center [P30 DK017047].

The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN2682013000001I / N01-HC-65233), University of Miami (HHSN268201300004I/N01-HC-65234), Albert Einstein College of Medicine

(HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago (HHSN268201300003I / N01- HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements

Lastly, to my favorite human in the entire world, my number one fan, my emotional support human, Oscar Agyei. Thank you for making me smile every day, for being there for me every day even when you are many miles away, for giving me more confidence, for always rooting and cheering for me, for listening, for helping and problem solving with me, for making me happier, and for loving me every day. Me dɔ wo paa.

Chapter 1: Introduction and Literature Review

In the United States, an estimated 38.4 million Americans were living with type 2 diabetes (T2D) and an estimated 97.6 million are living with prediabetes in 2021.[11] T2D has a high global prevalence as well, and reports estimate over 400 million people were living with the disease in 2021.[12] Global prevalence is trending upwards with an increase to 642 million people expected by 2040.[13] While type 2 diabetes mainly affects those over 45, prevalence among those younger is increasing.[14,15]

T2D is a chronic condition characterized by the body's inability to use insulin effectively.[15] Primary risk factors of T2D include diet, age, obesity, physical activity levels and genetic risk factors. Environmental risk factors include exposure to arsenic,[16,17] persistent organic pollutants such as organochlorine,[17] and air pollution, specifically traffic-related air pollutants.[4,5,18] Heidemann et al.[19] found those living by extremely busy roads had higher odds of developing diabetes than those living by very low or no traffic roads. A meta-analysis by Yang et al.[5] of 11 studies reported an association between $PM_{2.5}$ and incident T2D (HR= 1.10, 95% CI: 1.04-1.17 per 10 $\mu g/m^3$ increment).

One's neighborhood environment can affect their exposure to air pollution, their physical activity levels, and their risk of T2D. There have previously been a limited number of studies on neighborhood environment, specifically neighborhood physical disinvestment (formerly referred to as disorder), and T2D. Of the few published studies, heterogeneity among the study designs are high as researchers define and collect data on neighborhood environment in different manners. Disinvestment in the built environment is a result of systemic social and economic processes that deprive areas of resources needed to maintain and develop physical environments.

Disinvestment is reflected in visual indicators of deterioration and neglect, and is a concern for health as areas of high disinvestment may reduce resident's ability and desirability to engage in physical activity.[20–23] Increasing neighborhood disorder has also been found to heighten the genetic risk of T2D.[9]

As there is limited data on air pollution, neighborhood environment, and T2D, there is even less data on Hispanic/Latino populations and subgroups. The first aim of this dissertation was to estimate air pollution exposure at participants' residence and investigate the association between exposure and incident diabetes. The second aim was to measure neighborhood disinvestment and investigate the association between neighborhood disinvestment and incident diabetes. This dissertation addressed data gaps by investigating both air pollution and neighborhood-built environment as exposures in a diverse sample of Hispanics/Latinos across four US cities.[24]

Type 2 Diabetes

There are two main types of diabetes mellitus, type 1 and type 2. Type 1 is due to autoimmune destruction of pancreatic β-cells and is more common among children and young adults.[2] Type 2 diabetes is more common than type 1 with 90-95% of diabetes cases being type 2.[12,25] T2D is a chronic condition related to the body's ineffective use of insulin.[15] In the case of T2D, the cause of the pancreatic β-cell loss is non-immune and the insensitivity to insulin varies.[2] The pancreatic β-cell dysfunction and insulin resistance causes insulin deficiency in patients.[13]

Both genetic and environmental risk factors cause T2D with obesity appearing to be the predominant risk factor.[2,3] Schnurr et al.[3] found effects of having high genetic risk to be modest compared to the higher T2D risk of obesity, which has its own genetic risks. Obesity and

lifestyle choices related to smoking, alcohol consumption, limited physical activity, and diet were associated with higher T2D risk regardless of genetic risk.[3] A study comparing the HCHS/SOL cohort and the Framingham Heart Study found BMI, age, hypertension, and full-time employment to be independent predictors of diabetes risk.[26] A systematic review and meta-analysis noted the importance of accounting for obesity, physical activity, nutrition, smoking and second-hand smoke in diabetes risk studies investigating associations with air pollution exposure.[1]

Diet and exercise have long been linked to diabetes risk. Meat based diets are associated with a higher risk of diabetes than plant-based.[27] Drinking sugary beverages has also been linked to an increase in obesity and T2D risk.[27] In regards to the association between obesity and T2D risk, the location of excess adipose tissue may be important as visceral obesity and ectopic fat appear to increase risk more than high BMI.[27]

While it is difficult to study diet and physical activity in cohorts due to challenges in assessing and monitoring behavior, social desirability bias in questionnaire responses, reporting bias, and the temporal nature of these behaviors, a number of studies have found improving diet and physical activity can reduce risk of T2D.[3,27,28] A study of 3234 overweight or obese adults with impaired glucose tolerance found increasing physical activity and adopting a low-fat diet for weight reduction reduced T2D risk.[28] Even non-strenuous movements such as walking and completing daily tasks resulted in beneficial metabolic effects compared to sedentary behavior.[29] Those with a more sedentary lifestyle tend to have higher concentrations of circulating pro-inflammatory mediators while physical activity tends to attenuate low-grade inflammation and benefit β-cell function.[27,30]

High levels of physical activity are associated with lower risk of diabetes than low or no level of physical activity with one study estimating a 30% relative risk reduction.[13,31] In the HCHS/SOL cohort, a subset of participants wore an accelerometer for a week and found moderate-vigorous physical activity was inversely and nonlinearly associated with incident T2D risk.[32] However, benefits seemed to level off after 30 minutes of moderate-vigorous physical activity a day and the association was present among those over 50 but not those under 50 years old.[32]

Smoking has also been linked to T2D risk. A meta-analysis found the highest relative risk of T2D among heavy smokers compared to light or former smokers.[33] However, no association was found between smoking and incident T2D in a large multi-ethnic cohort.[34] A few studies have investigated a combination of these risk factors by calculating a "favorable lifestyle factor". A 5-point lifestyle score calculated from diet, physical activity, smoking status, alcohol consumption, and BMI was used in a Spanish cohort, and investigators found those with higher lifestyle scores had lower incident rates of T2D than those with the lowest scores.[35] However, the sample size was small and only 51 T2D cases occurred.[35] Schurr TM et al.[3] found those with unfavorable or intermediate lifestyle had higher hazard ratio of incidence T2D compared to those with favorable lifestyle. These modifiable lifestyle factors are generally easier to change than environmental risk factors such as air pollution and noise.

Air Pollution

Exposure to air pollution is associated with higher risk of cardiovascular and cardiometabolic diseases.[36–39] Associations and exposure levels vary based on the pollutant, which can come from a variety of sources and have various chemical makeups. Four commonly

studied air pollutants for health effects are particulate matter with an aerodynamic diameter of 2.5 microns ($PM_{2.5}$), nitrogen dioxide (NO_2), nitrogen oxides (NO_x), and ozone (O_3). As the criteria for $PM_{2.5}$ is the size, the particles can result from a number of sources and can contain various chemicals. Common sources of $PM_{2.5}$ are vehicle exhaust, construction, combustion of wood, heating oil, coal, wildfires, and unpaved roads. These small particles are able to penetrate deep into the lung tissue resulting in direct translocation into the bloodstream, oxidative stress, and systemic effects on other organs and systems such as the autonomic nervous system.[36,37] Short term health effects of $PM_{2.5}$ include nose, eye, throat, and lung irritation, increased blood pressure and decreased heart rate variability; long-term effects include cardiovascular and cardiometabolic diseases.[36,37] The chemical makeup and size of the particulate matter may affect the resulting health effects and the biological mechanism.[36] Components with a higher oxidative stress potential may increase the oxidative stress.[36] $PM_{2.5}$ from fuel combustion can increase black carbon and trace metals.[36] While different components of $PM_{2.5}$ may result in different biological reactions, identifying which ones are the most toxic is difficult and may not be the most productive aspect to focus on. Instead, focus should remain on the sources of these pollutants and limiting human exposure.[36]

Nitrogen oxides (NO_x) is a group of highly reactive gases containing NO_2 and nitrogen oxide (NO). NO_2 is of particular interest in urban air pollution epidemiology studies as the main source is emission from vehicles and is more specific to traffic-related pollution than $PM_{2.5}$.[40] Additional sources include construction equipment, power plants, and soil via denitrification. NO is a colorless, odorless insoluble gas while NO_2 is a brownish, water soluble gas with a chlorine-like odor. NO_2 is more toxic to humans than NO.[39] Nitrogen oxides are respiratory irritants associated with increased susceptibility to respiratory infections and asthma.[41] Short term health

effects range from eye and nose irritation to death.[39] The low solubility of NO_2 in water allows the gas to travel deep into the lower respiratory tract.[39] As with $PM_{2.5}$, focusing on the specific toxicity of the pollutant is not the most practical or efficient way to reduce adverse health outcomes on a population level.[36] These pollutants are best viewed as markers of traffic-related air pollution and the goal of investigating these associations with health outcomes is to better understand variations on small spatial scales. Adverse health outcomes will likely reduce when traffic reduces and when populations spend less time near traffic.

Ozone in the stratosphere occurs naturally and is beneficial to life on earth as it reduces the harmful effects of the sun's ultraviolet radiation. However, ozone at ground-level is hazardous to health. Ground-level O_3 is a secondary pollutant as there are no direct emission sources; O_3 needs to be formed. Ground-level O_3 forms when NO_x, volatile organic compounds (VOCs), and solar radiation are present. Due to the chemical reactions that occur and transportation by wind, NO_x tends to be higher in the mornings and closer to traffic while O_3 is higher in the afternoon and further from traffic.[39,42] As sunlight is required for the reaction, O_3 tends to be highest on hot sunny days.[39]

Ground-level ozone is a respiratory irritant that increases susceptibility to infection and aggravates asthma.[43] Similar to NO_2, acute health effects of O_3 range from eye and nose irritation to death.[39,42] Ozone's properties of being a high reactive, oxidative gas with low solubility in water relate to its health effects. As ozone has low solubility in water, it is not effectively removed by the upper respiratory tract, allowing it to travel to the lower respiratory tract where it can dissolve into the epithelial lining fluid.[42] The oxidative potential of the gas causes oxidative stress in the body which may affect cell signaling and cause cell damage.[42]

Air Pollution and Type 2 Diabetes

A recent meta-analysis by Yang et al.[5] on 11 studies showed an association between $PM_{2.5}$ and incident T2D (HR= 1.10, 95% CI: 1.04-1.17 per 10 $\mu g/m^3$ increment). For NO_2, seven studies investigating T2D incidence showed a lower pooled effect than $PM_{2.5}$ (HR = 1.01, 95% CI: 0.99-1.02 per 10 $\mu g/m^3$ increment).[5] Another recent meta-analysis by Liu et al.[18] found $PM_{2.5}$ to be associated with higher incidence of T2D (HR: 1.10, 95% CI: 1.04-1.16). NO_2 exposure showed a difference in incident T2D between men and women.[18] When looking at prevalence of T2D, higher levels of $PM_{2.5}$, PM_{10}, and NO_2 were associated.[18] Both meta-analyses highlighted the high level of heterogeneity among the studies.[5,18] Additionally, there were not enough studies on O_3 and NO_x to conduct a meaningful meta-analysis and discussion.[18] Previous studies on NO_2 and incident diabetes are limited, but association is generally positive.[1,44] However, a few studies on NO_2 and incident diabetes found no association.[45,46] NO_2 tends to have larger within-city variability then $PM_{2.5}$[40], making the resolution of the air pollution model even more important.

Heterogeneity is high among studies examining the association of air pollution and incident type 2 diabetes as research groups use different instruments and models to estimate exposure, estimate exposure on different spatial and temporal scales, and have different temporal definitions for 'long-term' exposure. Additionally, populations vary as does ascertainment of health data and relevant confounders, and the follow-up period. Table 1 summarizes relevant publications on air pollution and incident diabetes.

Very few studies on NO_x, O_3, and incident diabetes have been published. An analysis on a cohort of Danish nurses found a weak, positive association with NO_x in a fully adjusted model (HR: 1.01, 95% CI: 0.98-1.05).[44] A large cohort in Rome had a small positive association between incident diabetes and NO_x (HR: 1.009, 95% CI:1.002 – 1.015) and O_3 (HR: 1.012, 95%

CI:1.000 – 1.024).[46] Stronger association was found among those <50 years (HR: 1.051, 95% CI: 1.021, 1.083).[46] However, no data on BMI or physical activity was available for the cohort.[46] A longitudinal study on older Mexican Americans in Sacramento utilized a time-varying annual O_3 average over a 5-year period prior to the onset of an event for all in the risk set at said event time.[47] Ozone was positively associated with diabetes (HR: 1.13, 95% CI: 1.00, 1.28) and the relationship was stronger among those who reported being physically active outdoors (HR: 1.52, 95% CI: 1.21, 1.90).[47]

Table 1. Studies assessing the association between air pollution and incident diabetes.

Reference	Location and Population	Study Period	Exposure	Exposure Model	Statistical Model	Outcome	Outcome data source	Control Variables	Results	Pollution increment
Cervantes-Martínez et al., 2022[48]	Female teachers from Mexico City metro area	2008-	$PM_{2.5}$- monthly time-weighted exposure at home and work addresses	Spatiotemporal model with resolution of 1 km x 1 km	Cox PH model with time-varying exposure on monthly time scale	Incident T2D	Self-report and administrative databases	Time-dependent age, smoking, time-dependent physical activity, SES, wood smoking, time spent outdoors	HR: 1.72 (95% CI: 1.47-2.01)	10 µg/m^3
			NO_2- monthly time-weighted exposure at home and work addresses						HR: 1.52 (95% CI: 1.37-1.68)	10 ppb
Chen et al 2013[49]	Ontario, Canada; Aged 35+	1996-2010	6-year average concentrations of $PM_{2.5}$ at baseline resident postal code	Satellite-based estimates derived from aerosol optical depth data. Resolution 10 x 10 km	Stratified Cox PH model with strata defined as single-year age groups, cycle of survey, and region (north/south). Used time-weighted exposure	Incident Diabetes	Ontario Diabetes Database - a validated registry of diabetics	Sex, education, income, BMI, and stratifying on age, survey year, and region	HR: 1.11 (95% CI: 1.02, 1.21)	10 µg/m^3

Clark C et al. 2017[45]	Vancouver, Canada; Age 45-85 y old	1994-2002	$PM_{2.5}$ - 5-year exposure period at participant's postal code	LUR model - 30m resolution	Logistic regression with time-weighted exposure	Incident Diabetes	Administrative records - unable to distinguish between type 1 and 2	Gender, age, area-level household income	OR: 1.03 (95% CI: 1.01,1.05)	IQR of 1.6 $\mu g/m^3$
			NO_2						OR: 1.00 (95% CI: 0.98, 1.02)	IQR of 8.4 $\mu g/m^3$
Coogan PF et al 2016[50]	United States; Black Women ages 21-69 at baseline	1995-2011	Monthly estimates of $PM_{2.5}$ at residential addresses from 1998-2008. Accounted for spatial, but not temporal variation.	Hybrid LUR model with Maximum Entropy techniques.	Cox PH Model with time-weighted exposure	Incident T2D	Self-report of doctor-diagnosis. 96% accuracy	BMI, neighborhood SES, education, vigorous exercise, diet pattern	Fully adjusted HR: 0.99 (95% CI: 0.90-1.09); leaner women had an increased incidence of T2D (HR: 1.36, CI: 0.91-2.04) compared to heavier women (HR: 0.95 (95% CI: 0.79 – 1.13); Women under 40 had an increased incidence of T2D (HR: 1.19 95% CI: 0.94-1.51) compared to women aged 40-54 (HR: 0.98 95% CI: 0.86, 1.11)	IQR increase of 2.9 $\mu g/m^3$ $PM_{2.5}$

Hansen et al. 2016[44]	Denmark; Female Nurses; Age 44+ by 1999	1993-2013	$PM_{2.5}$ - 5-year running mean at residential address	Danish air pollution dispersion model - annual mean concentrations at residential addresses	Extended Cox PH model	Incident Diabetes	Danish National Diabetes Register	Age, BMI, smoking, physical activity, alcohol, diet, hypertension, myocardial infarction, employment status, marital status	Fully adjusted HR of 1.11 (95% CI: 1.02-1.22)	IQR of 3.1 $\mu g/m^3$
			NO_2 - 5-year running mean at residential address						HR of 1.05 (95% CI: 0.99,1.12)	IQR of 7.5 $\mu g/m^3$
			NO_X - 5-year running mean at residential address						HR of 1.01 (95% CI: 0.98, 1.05)	IQR of 10.2 $\mu g/m^3$
Lao Et al., 2019[51]	Taiwan - Adult 18+ of Chinese descent residing in Taiwan	2001-2014	$PM_{2.5}$ - 2-year average	Satellite-based spatiotemporal model with 1x1 km^2 resolution	Cox PH regression models with time-varying exposure and covariates	Incident T2D	Medical assessment or self-reported physician-diagnosis	Age, sex, BMI, education, season, year, smoking, alcohol, physical activity, diet, occupational exposure, hypertension	HR range: 1.16 – 1.28; BMI significant modifier	Reference 1st quartile <21.7 $\mu g/m^3$. IQR 21.7-28.0 $\mu g/m^3$
Li et al. 2022[52]	UK Biobank – Aged 40-69 at baseline	2007-2018	Annual average $PM_{2.5}$ concentration at participants residential address at baseline. Categorical as low (<9.5	Land use regression model - ESCAPE	Cox PH model with time-constant exposure and covariates	Incident T2D	Hospital records from administrative databases	Sex, age, race, education, income, smoking, alcohol, BMI, diet, family history, hypertension, CVD, depression, cancer, physical activity	HR: 1.47 (95% CI: 1.36-1.59)	5 $\mu g/m^3$

			µg/m^3), moderate (9.5 – 10.3 µg/m^3), and high (> 10.3 µg/m^3)							
			NO_2: Low (<24.6 µg/m^3), moderate (24.6- 31.5 µg/m^3), high (>31.5 µg/m^3)						HR: 1.08 (95% CI: 1.06-1.10)	10 µg/m^3
Park SK et al. 2015[53]	United States - Multiple cities; aged 45-84 at baseline	2000-2012	Annual average $PM_{2.5}$ concentration from 2000 at participants residential address	Spatiotemporal model	Cox PH model with time constant exposure	Incident diabetes	Measured fasting serum glucose levels >126 mg/dL after 12-hour overnight fast; or use of antidiabetic medication	Sex, age, race/ethnicity, family history, education, smoking, alcohol, physical activity, neighborhood SES, BMI, and study site	Overall HR: 1.05 (95% CI: 0.87, 1.26); Only Chicago: HR: 1.55 (95% CI: 1.03,2.32)	IQR of 2.43 µg/m^3
			NO_X						Overall HR: 1.04 (95% CI: 0.77,1.40)	IQR of 47.1 ppb
Renzi et al. 2018[46]	Rome, Italy; Resident in Rome at 2001 Census; aged 35+ by 2013	2008-2013	Annual $PM_{2.5}$ exposure at baseline at residential address	LUR model - developed using measurements from 2010	Cox PH model	Incident Diabetes	Administrative records - unable to distinguish between type 1 and 2	Sex, SES, marital status, education level, occupation, place of birth	HR: 1.002 (95% CI: 0.982, 1.022)	5 µg/m^3

			NO_2						HR: 1.003 (95% CI: 0.996, 1.015)	10 µg/m^3
			NO_x						HR: 1.009 (95% CI: 1.002, 1.015)	20 µg/m^3
			O_3	Three-dimensional Eulerian dispersion model on a 1 Km grid.					HR: 1.012 (95% CI: 1.000, 1.024)	10 µg/m^3
Yishak Sade et al., 2023[54]	United States; Medicare enrollees	2000-2016	$PM_{2.5}$ - Annual predictions at ZIP code level	Spatiotemporal model on a 1 km^2 grid	Cox-equivalent re-parameterized Poisson survival	Incident Diabetes	Medicare Chronic Conditions Warehouse	Age, race, sex, Medicaid insurance, annual ZIP code means of summer and winter temperature, and annual ZIP code level sociodemographic variables	HR: 1.07 (95% CI: 1.06, 1.09)	5 µg/m^3
			NO_2 - Annual predictions at ZIP code level						HR: 1.06 (95% CI: 1.05, 1.06)	5 ppb

			O_3 – Averaged predictions during May-October at ZIP code level						HR: 0.99 (95% CI: 0.99, 1.00)	5 ppb
Yu et al., 2021[47]	Sacramento, CA; Self-identified Latinos aged 60+	1998-2007	O_3 - average exposure modeled 5 years prior to event for everyone in risk set at event time	LUR model with data from saturation monitoring at 49 sites across study area	Cox PH model with time varying O_3 exposure.	Incident Diabetes	Self-report of doctor-diagnosis; antidiabetic medication us; or fasting glucose level > 126 mg/dL measured during visit	Age, sex, education, longest held occupation, Neighborhood SES, outdoor physical activity, smoking, and household income	Overall HR: 1.13 (95% CI: 1.00, 1.28); among those physical active HR: 1.52 (95% CI: 1.21, 1.90)	10 ppb

Biological Mechanism

As previously mentioned, different pollutants have different chemical make-ups and are of different sizes. This may result in different health effects and different biological mechanisms causing the effects. The composition of the pollutants, as well as susceptibility of the host may influence the biological mechanism and there is likely more than one pathway at play.[55] There is also uncertainty around the specific mechanism for how each pollutant may cause T2D. One of the main theories for $PM_{2.5}$ is the presence induces oxidative stress and pulmonary inflammation which then results in systemic inflammation and leads to insulin resistance.[1,51,56,57] Additionally, $PM_{2.5}$ can act as a hypothalamic stressor which induces peripheral inflammation and abnormal glucose metabolism.[1] There are also links between this mechanism and higher association among those living near major roads.[58,59] The oxidative potential of particles near major roads was shown to be significantly higher than particles at background locations further away from major roads.[58] High-sensitivity C-reactive protein (CRP), a blood marker for systemic inflammation, was found to be positively associated with traffic intensity of the closest street.[59] Those living near a busy road had increased CRP values 10.2% (95% CI: 2.4%, 18.8%) higher compared to those living near a quiet road.[59] Increased serum levels of CRP have also been associated with exposure to $PM_{2.5}$ and traffic.[59,60]

A number of animal studies have been conducted to look into the mechanistic relationship of air pollution and T2D. Experiments on rats found exposure to O_3 may induce insulin resistance via oxidative stress and systemic inflammation.[61,62] Inhaling O_3 was shown to induce autonomic system responses in the rats as the inhaled pro-oxidant molecules penetrate deep into the lungs and translocate from the alveolar fluid into circulation.[61,62] These animal models help explain the linked between exposure, systemic inflammation, and insulin resistance

seen in human. Additionally, personal physical characteristics may also play a role in the mechanistic process as those with larger tidal volumes have higher O_3 uptake efficiency as more O_3 molecules are able to penetrate deep into the lungs with each respiration cycle.[63]

Role of Physical Activity

While being physically active is generally positive for one's health, and known to reduce risk of T2D, there is concern about the potential adverse effects of being physically active in a polluted environment. Higher physical activity levels mean higher inhalation rates and higher inhaled dose of air pollutants. However, the long-term benefits of habitual physical activity may counteract any acute adverse effects of the increased dose during exercise.[31] A comparative analysis by Tainio M. et al.[64] found the benefits of active travel (cycling or walking) outweighed the adverse effects of air pollution in areas with $PM_{2.5}$ concentration less than 100 $\mu g/m^3$. If $PM_{2.5}$ levels exceed 100 $\mu g/m^3$, then it is estimated only an hour and a half of cycling per day would result in more harm than good.[64] However, there is still limited information on the association between individually assessed physical activity, $PM_{2.5}$, NO_2, NO_x, and O_3 with incident T2D and how much exposure to air pollution limits the benefits of physical activity.[31]

Researchers collected data on habitual physical activity via questionnaires and estimated $PM_{2.5}$ exposure based on the 2-year average of the year of and year before a medical examination on a Taiwanese cohort.[31] Higher levels of physical activity and lower levels of $PM_{2.5}$ were associated with lower risk of T2D.[31] Physical activity reduced the risk of T2D regardless of $PM_{2.5}$ levels while $PM_{2.5}$ increased the risk regardless of physical activity levels.[31] They were unable to distinguish between outdoor and indoor physical activity, which would affect exposure and inhaled dose during physical activity, but assumed most was outdoors based on results from

a previous survey.[31] Chen et al.[32] and Hansen et al.[44] also did not find evidence for effect modification by physical activity or BMI. However, a cohort of older Mexican Americans found higher levels of outdoor physical activity increased the positive association between ozone and diabetes.[47] A Swiss cohort found stronger effects between NO_2 and incident T2D among the physical active participants.[65]

Neighborhood Built Environment

Physical activity levels may also depend on the neighborhood built environment. Aspects of the built environment can improve health by facilitating physical activity and decreasing adverse exposures. For example, a bike path for both recreation and active transportation that is far from major roadways and contains protective barriers could help promote physical activity and limit exposure to air pollution. A systematic review by Rugel and Brauer[40], noted a lack of studies on the built environment and the relation to physical activity that incorporates co-varying exposures such as air pollution. There is specifically a need for more research to better understand if physical activity mediated the association of the built environment and diabetes risk factors. MESA residents with better physical environments had lower BMI when adjusting for age, race, education, and income (Mean Difference: -1.06 (kg/m^2), 95% CI: -1.99, -0.12).[66] However, associations were attenuated when adjusting for diet and physical activity (Mean Difference: -0.69 (kg/m^2), 95% CI: -1.67, 0.29), suggesting a mediating role of diet and/or physical activity.[66] While there have not been any studies linking neighborhood built environment, physical activity, and residential air pollution exposure with incident T2D, there have been a number of studies on built environment and physical activity and well as a few on built environment and incident T2D with physical activity as a mediator.

A recent HCHS/SOL study in San Diego (SOL CASAS) by Carlson et al.[67] investigated the association between the neighborhood built environment and physical activity. The investigators looked at different types of physical activity and found the cohort reported relatively high occupational MVPA compared to active transport and recreational MVPA. When looking cross-sectionally, those living in areas with higher retail density had 2 more min/day of overall MVPA compared to those in lower retail density areas. Those in areas with higher residential density, retail density, or recreation areas had higher reported active transportation than those in lower areas. Those in higher greenness and park count had higher reported recreational MVPA. Higher walkability and physical activity supportiveness' indices were associated with more active transportation among individuals in areas with higher socioeconomic deprivation, a measure of socioeconomic status focused on household income and ownership, while there was no association among those in areas with lower socioeconomic deprivation. Higher park count was associated with more recreational MVPA among individuals in areas with lower socioeconomic deprivation and not associated with those in higher socioeconomic deprivation. Authors noted those with higher deprivation may have less time to conduct recreational MVPA or are too tired from occupational MVPA. SOL CASAS was not able to measure changes in neighborhood environment over time, but was able to measure change in physical activity levels. They found overall MVPA only slightly decreased over a 6-year follow-up period, but no specific neighborhood feature was associated with the change in physical activity.

A mediation analysis of physical activity on local amenities and incident T2D found an indirect effect through physical activity of 1.01 (95% CI: 1.01 -1.02) with the proportion mediated through physical activity being 5.9%.[6] If all the participants reported not having local

amenities then the odds of T2D would be 1% higher for those with physical activity level less than 150 mins per week compared to those with more than 150 minutes per week.[6] Dendup et al.[6] results showed how the lack of access to local amenities can increase risk of T2D by increasing BMI and reducing physical activity. Gebreabe et al.[7] also found physical activity to be a potential mediator between neighborhood violence and problems and the prevalence of T2D.

In a cross-sectional survey in Jamaica, interviewers assessed the neighborhood environment and found increased levels of neighborhood disorder/disinvestment were associated with low/no physical activity (OR: 1.05, 95% CI: 1.02-1.07).[68] However, neither low/no physical activity nor overweight/obesity were found to act as mediators in relationship between neighborhood variables and diabetes.[68] Braver den et al.[69] also found no evidence of physical activity as a mediator between neighborhood walkability and HbA1C or fasting plasma glucose (FPG), but the initial association between neighborhood walkability and 1 year change in HbA1C or FPG was not meaningful (HbA1C: β: -0.16, 95% CI: -0.59, 0.27; FPG: β: -0.03, 95% CI: -0.13, 0.06).

In the San Diego site of the HCHS/SOL, greater neighborhood socioeconomic deprivation was found to be associated with higher BMI (Coeff: 1.08, 95% CI: 0.66, 1.49), HbA1c levels (Coeff: 0.08, 95% CI: 0.04, 0.12), and having worse diabetes status (pre-diabetes vs normoglycemia and diabetes vs pre-diabetes) at baseline (OR: 1.25, 95% CI: 1.06, 1.47).[70] At the second visit, about 6 years later, greater neighborhood socioeconomic deprivation was associated with increasing BMI (Coeff: 0.35, 95% CI: 0.11, 0.58) and worsening diabetes status (OR: 1.27, 95% CI: 1.10, 1.46).[70] Social disorder, a metric of crime rates, liquor store per capita, and vacant houses and land, was related to increasing BMI at visit 2 (Coeff: 0.31, 95% CI: 0.23, 1.52) but not worsening diabetes status (OR: 0.96, 95%CI: 0.67, 1.35).[70] There were no

associations between walkability, greenness, and residential stability and worsening diabetes status at visit 2.[70]

A longitudinal study on African-Americans in Jackson, Mississippi found a positive association between neighborhood violence (fights, violent arguments, sexual assault, rape, and robbery) and neighborhood problems (noise, heavy traffic, speeding cars, lack of access to adequate food or shopping, litter) and incident T2D (HR: 1.16, 95% CI: 0.95, 1.43 and 1.24, 95% CI: 1.00, 1.54, respectively).[7] The previously mentioned cross-sectional survey in Jamaica found no association between neighborhood disorder or availability of recreational facilities and diabetes (OR: 0.99, 95% CI: 0.95-1.03 and 1.01, 95% CI: 0.77, 1.32, respectively).[68]

As with air pollution epidemiology study, heterogeneity among built environment studies is high.[40,71] There are various built environment measures to investigate and various methods and approaches to quantify them.[40,71] A particularly interesting measure of the built environment is neighborhood physical disinvestment, previously referred to as neighborhood physical disorder. Neighborhood physical disinvestment in the built environment is a result of deterioration, neglect, and financial disinvestment due to social and economic processes redirecting resources needed to maintenance and development away from neighborhoods of lower-income populations. Disinvestment is reflected in visual indicators of deterioration and neglect such as deteriorated sidewalks, lack of street lighting, litter, and abandoned buildings. Living within disinvested neighborhoods may increase stress and reduce residents' ability and desire to engage in physical activity, which may in turn increase diabetes risk.[21–23,72–74] Additionally, among Hispanic residents in the Health and Retirement Study, those in more disordered neighborhoods were more likely to self-rate their health as poor (coef: -0.11, SE= 0.05).[74] Residents in more disordered areas may feel less safe and may find walking to be hazardous.[73,74]

To our knowledge, there has not been a study investigating the association between neighborhood physical disinvestment and type 2 diabetes incidence. There have been a couple studies on neighborhood disorder and prevalent diabetes[9,68] and a study on neighborhood disorder and diabetes distress conducted among those with T2D.[75] The Jamaican cross-sectional survey found no association between neighborhood disorder and diabetes prevalence (OR: 0.99, 95% CI: 0.95, 1.03).[68] However, perceived neighborhood disorder was associated with T2D prevalence in a non-Hispanic White subset of the Health and Retirement study cohort (OR: 1.11, 95% CI: 1.05, 1.17).[9] Robinette et al.[9] looked at the interaction between neighborhood disorder and genome-wide genetic risk and found neighborhood disorder heightens the genetic risk of T2D (OR: 1.10, 95% CI: 1.04, 1.16).[9] Among adults with T2D in Quebec, researchers found less neighborhood disorder was associated with lower diabetes distress (β: -0.03, 95% CI: -0.05, -0.01).[75] They estimated each additional aspect of neighborhood order was associated with a decreased diabetes distress score of 3.1%.[75]

While studies related to neighborhood disinvestment and T2D are limited, there have been a few studies on neighborhood disinvestment and physical activity. Limited physical activity due to the built environment is a modifiable risk factor for obesity, cardiometabolic disorders, and general health.[74,76,77] It is of particular concern for older residents as they may be more affected due to limited physical agility and mobility.[20,22,73] Additionally, older individuals spend more time in their residential neighborhood, increasing their exposure and sensitivity to the disorder.[20] A study on mobility limitation over a 2-year period found those living in neighborhoods with lower disorder had higher probability of recovering their mobility.[77] The report also estimated about 38% of the effect neighborhood disorder had on mobility recovery was mediated by physical activity.[77]

A study by Mooney et al.[20] using the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to assess neighborhood disorder and physical activity in older NYC adults found an inverse association between disorder and physical activity at baseline. Neighborhood disorder was also significantly associated with low or no physical activity in a Jamaican cross-sectional survey (OR: 1.05, 95% CI: 1.02-1.07).[68] A cross-sectional analysis on women in four Chicago neighborhoods did not find an association between disorder and total-moderate-to-vigorous physical activity (β: 52.7, 95% CI: -7.2, 112.7), but did find higher levels of disorder were associated with obesity (OR: 1.43, 95% CI: 1.01, 2.02).[76] A study in California using self-reported data on perceived neighborhood disorder and physical activity found no association between perceived neighborhood disorder and physical activity duration (PR: 1.01, 95% CI: 0.97, 1.05).[78] However, findings did show Latinas reporting higher perceived neighborhood disorder and attaining less than a high school diploma, higher income, and administered surveys in Spanish were more likely to report shorter duration of physical activity.[78]

Most of the previously mentioned studies on neighborhood-built environment and physical disorder were based on perceived, and not objective, exposure data reported by participants.[6,7,77–79,9,21,23,66,69,73–75] A main concern with utilizing exposure data assessed by study participants is 'same source bias' resulting in a spurious association. If participants are self-reporting both their exposures and outcomes, their experience with the outcome may influence how they report their exposure. The reverse could be true as well, with their exposure influencing how they report their outcome.

Other approaches to collecting built environment data that get around the issue of 'same source bias' include asking non-participants living in the same area their thoughts and opinions,[66]

collecting data from publicly available databases,[69,70,80] or in-person street audits[22] and assessment by researchers.[68] A less costly and time-consuming alternative to in-person street audits are virtual street audits.[20,76,81,82]

Virtual street audits have repeatedly shown to be reliable[83–88] and have the extra benefit of being able to assess previous street and neighborhood conditions.[8,20,83] Google started collecting street view data in 2007. Initial images were grainy and of lower quality, but improved in 2008. Using Google's 'Time Machine' feature, auditors can see the same location at multiple points in time allowing them to assess changes in the neighborhood environment over time. One concern with utilizing Google street view data is the images are not evenly available across time and place.[83] Google's street view cars do not drive each street the same amount of times. Some streets could be covered multiple times a year while others may have only been covered once or twice since 2007. Additionally, Google does not publish information on where, when, or why they drive their routes. It seems streets in urban areas and major city streets are prioritized over minor residential streets. It is possible Google is prioritizing streets with the higher potential for change over a shorter period of time.

Image availability was assessed in HCHS/SOL Bronx and San Diego recruitment area.[89] Images were more widely available across space and time in the Bronx compared to San Diego.[89] Bias due to systematic differences in image availability in the Bronx should not be a concern for researchers.[89] However, more caution should be taken when conducting and assessing longitudinal audits in San Diego as lower image availability appeared to be concentrated in neighborhoods with recent housing developments.[89]

Chapter 2: Motivation

T2D is of increasing concern as global prevalence and incidence is steadily increasing.[13] Studying Hispanic/Latinos is particularly important as they were 60% more likely to be diagnosed with diabetes by a physician than non-Hispanic white in the US in 2022.[105] Hispanics were also 1.5 times more likely to die from diabetes compared to non-Hispanic whites,[105] and more likely to develop T2D at a younger age.[106] While Hispanic/Latinos are more likely to have T2D than non-Hispanic whites, Hispanic/Latino is a large heterogeneous ethnic group and all subgroups are not at equal risk of T2D. Prevalence of T2D in the HCHS/SOL cohort at baseline ranged from 10.2% to 18.3% among six subgroups.[107] South Americans had the lowest prevalence (10.2%) followed by Cubans (13.4%) and Central Americans (17.7%).[107] Diabetes prevalence in Dominicans and Puerto Ricans was 18.0%, and prevalence was highest among Mexicans (18.3%).[107] Limited research on Hispanic/Latino subgroups limits our understanding of why these different rates occur. While genetic and cultural factors may be at play,[108] we seek to investigate if these differences among subgroups are a result of air pollution exposure and neighborhood built environment as one's Hispanic/Latino background may affect the neighborhood they live in due to familial, cultural, linguistic, and socioeconomic reasons.[100–102] HCHS/SOL's longitudinal cohort of self-identifying Hispanic/Latinos allow us to address an ethnic data gap as Hispanic/Latinos have been underrepresented in research. Importantly, it allows us to conduct analysis among Hispanic/Latino subgroups.

Air pollution, neighborhood disinvestment, and physical activity are all modifiable risk factors and important areas to focus on for T2D prevention. While physical activity is an individually modifiable risk factor, air pollution and neighborhood disinvestment are typically considered modifiable at the community level. Individuals have little control over their exposure

to ambient air pollution and neighborhood physical environment as these exposures are mainly a result of where participants live. While where one lives seems like a personal choice, one's SES, immigration status, the housing market and structural racism does not make this free choice a reality.[90,102] Historically, racist practices such as redlining exacerbated residential segregation and environmental injustices.[90] Neighborhoods were considered less desirable based on higher populations of people of color and environmental pollution.[93,94] These less desirable neighborhoods are often closer to industry and major roadways, increasing air pollution exposure levels.[93] Banks unwillingness to provide loans for residents in the less desirable neighborhoods limited the generation of wealth through home ownership in communities of color.[91] This is particularly true for undocumented Hispanic/Latino immigrants as their immigration status limits their ability to obtain a bank mortgage.[102] Limited wealth and political power of residents of color resulted in more hazardous industries in their neighborhood as they did not have representation or say in local government's land use decisions.[91]

Residents may feel as though they have less agency to pressure officials for changes due to generations of structural and interpersonal racism and their voiced concerns may be ignored due to racial discrimination. This is particularly true, and riskier, for undocumented immigrants. Additionally, residents of lower income neighborhoods may not have the disposable income available to make improvements on their private property. As redlining limited generational home-ownership in communities of color and immigration status impedes the ability to obtain a bank mortgage, the percentage of renters is higher, leaving the needed private investments to the landlords and out of resident's control.[91,102]

Reducing air pollution and investing in disinvested areas to increase physical activity can help reduce risk of T2D, however, the relationships between them need to be better understood.

This dissertation addressed shortcomings of previous research by utilizing a spatiotemporal model to estimate air pollution exposure at residential addresses and a virtual street audit to assess neighborhood physical disinvestment.

A number of previous studies on air pollution and T2D utilize models that do not factor in changes on the spatial and temporal scale[45,50,56,109] and may estimate exposure on a large spatial scale.[49,51,56,109] The spatiotemporal model allows for fine-scale estimates of $PM_{2.5}$, NO_x, NO_2, and O_3.[110] To account for pollutants variation over time and space, the spatiotemporal model builds upon universal kriging by assuming these variations are systematic and can be predicted based on geographic characteristics and spatial smoothing.[111] Data from regulatory monitors and snapshot monitoring campaigns focused in the participant's residential neighborhoods were fed into the model to estimate exposure levels at residential addresses on a 2-week timescale.[110] Previous studies on NO_x, NO_2, O_3 and T2D have been limited. Studies on NO_x and O_3 have been particularly sparse, preventing any previous meta-analysis or discussion.[18]

CANVAS allowed us to assess neighborhood disinvestment in a multiple city cohort using a virtual street audit. Previous studies on neighborhood disinvestment and T2D have been limited[9,68,75,79] and most have collected disorder data from study participants via questionnaires.[9,75,79] Using data from a virtual audit helps eliminate the issue of 'same source bias' present in a number of previous disinvestment studies.

Physical activity data collected by the HCHS/SOL cohort allowed us to investigate how physical activity modifies the relationship between air pollution and T2D. In general, if $PM_{2.5}$ levels are below 100 $\mu g/m^3$ the benefits of physical activity outweigh the negative effects of pollution exposure.[64] However, data is still limited on the modification of physical activity on the

association between $PM_{2.5}$, NO_2, NO_x, and O_3 with incident type 2 diabetes and how much exposure to air pollution limits the benefits of physical activity.[31]

Chapter 3: Long-Term Ambient Air pollution and Incident Diabetes between visits 1 and 2 of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

Introduction

Type 2 diabetes (T2D) is a chronic condition characterized by the body's inability to use insulin effectively,[15] and is a condition that is increasing globally.[13] In the United States, T2D is of particular concern among Hispanic/Latinos as they were 60% more likely to be diagnosed with diabetes by a physician than non-Hispanic white in the US in 2022,[105] and are more likely to develop T2D at younger ages.[106]

Well established risk factors of T2D include diet, age, obesity, physical activity levels and genetic risk factors. T2D has also been linked to environmental exposures such as arsenic,[16,17] persistent organic pollutants such as organochlorine,[17] and air pollution, specifically traffic-related air pollutants.[4,5,18,112] While there is some uncertainty around how different pollutants such as $PM_{2.5}$, NO_2, NO_x, and O_3 may increase risk of T2D, there are a few proposed pathways.[1,51,56,57] A widely accepted pathway for $PM_{2.5}$ is its presence induces oxidative stress and pulmonary inflammation which then results in systemic inflammation and leads to insulin resistance.[1,51,56,57] $PM_{2.5}$ can also act as a hypothalamic stressor which induces peripheral inflammation and abnormal glucose metabolism.[1] These biological mechanisms have been supported by animal models,[61,62] the significantly higher oxidative potential of particles near major roads compared to locations further from major roads [58] and evidence of higher levels of high-sensitivity C-reactive protein in those living near a busy road compared to a quiet road.[59,60]

Recent meta-analyses and reviews have highlighted the association between air pollutants and incident T2D in various countries and populations.[1,5,18,112] A meta-analysis by Yang et al.[5] of 11 studies reported an association between $PM_{2.5}$ and incident T2D (HR= 1.10, 95% CI: 1.04-

1.17 per 10 μg/m^3 increment), supporting another meta-analysis by Liu et al.[18] published a year earlier (HR: 1.10, 95% CI: 1.04-1.16 per 10 μg/m^3 increment). Neither meta-analysis found a significant association between NO_2 and incident T2D. A small meta-analysis by Eze et al.[1] on four NO_2 studies found a higher risk of T2D with increased exposure (pooled RR: 1.08, 95% CI: 1.00, 1.17 per 10 μg/m^3 increment), however, two of the four studies were longitudinal and two were cross-sectional. Overall, studies on NO_2 have been limited and associations more mixed for incident T2D than $PM_{2.5}$.

There is a high level of heterogeneity among previous study designs on air pollution and incident T2D, and these divergent findings may be related to divergent assessment stragies.[1,5,18] Studies can vary by how the pollutants are measured and exposure modeled, the spatial and temporal scales of the modeled exposures, and definition of 'long -term' exposure.[1,5] Study populations vary as does the ascertainment of diabetes, relevant confounders, and the follow-up period. As NO_2 concentrations vary spatially on a finer scale then $PM_{2.5}$ concentrations[40], NO_2 exposure estimates based on a spatial scale calibrated to $PM_{2.5}$ may be erroneous.

The association between air pollutants and T2D could be modified by physical activity levels and adiposity.[44,113–115] While being physically active generally lowers risk of T2D, being physically active in a polluted environment has the potential to decrease the benefits of physical activity.[113,114] A comparative analysis by Tainio M. et al.[64] found the benefits of active travel (cycling or walking) outweighed the adverse effects of air pollution in areas with $PM_{2.5}$ concentration less than 100 μg/m^3. However, consistent evidence of an interaction between long-term air pollution exposure and physical activity is lacking, especially on the association between individually assessed physical activity, $PM_{2.5}$, and NO_2 with incident T2D.[31,114] Adiposity could

be an effect modifier as being obese may put one at high risk of developing diabetes related to air pollution exposure compared to those who are not obese.[44,115]

In the presented analysis, we assessed the association between $PM_{2.5}$ and NO_2 and incident diabetes in the large longitudinal cohort of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) to address an ethnic data gap as Hispanic/Latino adults have been underrepresented in this research space. Additionally, we employ a fine-scale spatiotemporal air pollution model to predict long-term $PM_{2.5}$ and NO_2 exposure at residential addresses. This improves upon previous air pollution and T2D studies that utilized models that did not factor in changes on the spatial and temporal scale[45,50,56,109] and estimated exposure on a large spatial scale.[49,51,56,109] We further investigated how physical activity and adiposity modifies the relationship between air pollution and T2D.

Methods

Study Design and Population

Study participants are part of the HCHS/SOL, a longitudinal cohort made up of 16,415 self-identifying Hispanic/Latino adults at least 18 years in age at enrollment from the Bronx, Chicago, Miami, and San Diego.[24] More details on the HCHS/SOL sample and study design have been previously reported. [116,117] HCHS/SOL used a probability sampling strategy to recruit study participants with the goal of having a representative sample of US Hispanic/Latinos living in the communities around the four field centers. Participants were recruited from census tracts surrounding the four field centers.[116] Each field center targeted census tracts based on proximity to the field center, percentage of Hispanic/Latinos based on the 2000 census, and local information. [116] HCHS/SOL utilized a face-to-face sample recruitment strategy.[116] To remain

cost-effective, HCHS/SOL oversampled in geographic clusters and households most likely to be Hispanic/Latino.[116] They identified high geographic clusters of Hispanic/Latinos based on the 2000 census. Households were also clustered and addresses associated with a Hispanic/Latino surname were over sampled.[116] HCHS/SOL also strived for socio-economic representation and stratified geographic clustered based on the proportion of residents 25 years or old with minimum high school education.[116] Lastly, HCHS/SOL over-sampled households with adults aged 45-74 years old.[116] The goals of SOL are to provide information on health status and disease burden in US Hispanic/latinos and to enable investigation of relationships between various risk factors and disease incidence in a population historically underrepresented in research.

Cohort enrollment and in-person baseline visits spanned 2008 to 2011. Yearly follow-up phone calls were conducted and a second round of in-person visits (visit 2) occurred between 2014 and 2017. Participants filled out questionnaires and underwent clinical examinations during the in-person visits at baseline and visit 2. The questionnaires assessed demographics, physical activity levels, current health and medical history, and socioeconomic status.[117] Clinical examinations included collection of blood and urine for assays of cardiovascular risk factors.[117] The yearly follow-up phone calls contained questions on general health, potential updates from their doctors or health professional, and any potential hospitalization or emergency room visits that occurred since the last follow-up.[117]

Diabetes Ascertainment

Diabetes status was collected at baseline, during each follow-up phone call, and at the second in-person visit. At these times, participants were asked if a doctor or health care professional has told them they have diabetes or high sugar in their blood and if they are

receiving any treatment.[117] During the in-person visits, blood tests were administered to test fasting plasma glucose (FPG) and glycosylated hemoglobin (HbA1c).[117] A 2-hour oral glucose tolerance test (OGTT) was also conducted on participants who did not self-report a diabetes diagnosis and on participants with FPG < 150 mg/dL.[107] For this analysis, we defined diabetes in in two ways. Our primary definition was based the American Diabetes Association criteria: FPG ≥126 mg/dL; HbA1c level ≥6.5%; post-OGTT glucose ≥ 200 mg/dL; and/or self-reported use of antihyperglycemic drugs.[10] Our secondary definition extended this to further include those who self-reported diabetes or high sugar in their blood. Ascertainment of diabetes in the HCHS/SOL cohort was not able to distinguish between type 1 and type 2 diabetes.

Pre-diabetes classification was assessed based on laboratory assay conducted at baseline and visit 2. Based on the American Diabetes Association criteria, pre-diabetes was defined as having a FPG ≥100 mg/dL and <126 mg/dL; and HbA1c level ≥5.7% and <6.5%.[10]

Ogawa Campaigns

The campaigns involved setting up Ogawa samplers to measure NO_2, NO_x, and O_3 for two weeks on utility poles at gradient distances from roadways. Each city had 16-17 sites and each site had 6 Ogawa samplers. Ogawa samplers were set up in the University of Washington field laboratory about a week before deployment. Double chamber samplers contained NO_x and NO_2 pads while a single chamber sampler contained an O_3 pad. In addition to the main set of samplers, 10% laboratory blanks, 10% field blanks, and 10% duplicate samplers were set up. Laboratory blanks were prepared the same as the samples but remained in the field laboratory refrigerator. Ogawa samplers were shipped overnight to the HCHS/SOL field sites in coolers with ice. Once received by HCHS/SOL field site members, samplers were placed in the refrigerator until the SOLAir field work team arrived. The field work team set the samplers up

over a 2-day period. The field blanks remained with the samples until the samples were set up outside. The field blanks were then stored in a refrigerator at the field site during the 2-week sampling period. The sample and duplicate Ogawa samplers were clipped inside metal rain shelters to help protect them from wind and rain. Each shelter contained one single chamber sampler with an O_3 pad and one double chamber sampler with a NO_x and a NO_2 pad. The shelters were placed on utility poles along streets perpendicular to the target roadway and on a gradient of 0 – 50 m, 50 – 100 m, and 100 – 350 m on each side of the target streets. N-S and E-W streets were selected to allow for estimates that account for wind patterns. Sites also focused on various traffic volumes and population density. These 6 samplers transecting a target roadway are a "cluster". There were 16-17 clusters in each city (Chicago, Miami, and San Diego). After the 2-week sampling period, the Ogawa samplers were transported back in coolers to UW and analyzed.

Ogawa Analysis

Ogawa sampler pads were analyzed using two different methods depending on the pollutant. Ozone and NO_2 were analyzed using ion chromatography (IC) and NO_x was analyzed using Ultraviolet (UV) spectroscopy. The O_3 Ogawa samplers pads are coated with a sodium nitrite solution. The presence of O_3 oxidizes the nitrite to nitrate. After the two-week sampling period ends, the pads are submerged in milli-Q water to extract the nitrate ions. The Dionex ICS1000 with an AS40 autosampler and conductivity detector (Dionex, Sunnyvale, CA) is used to determine the nitrate ion concentration in the extract. IC standards for nitrate ions are used as well. The laboratory and field blanks are run through the IC and used for calibration. The nitrate concentration is then used to calculate ozone concentration over the two-week sampling period

by using a sampling rate of 21.8 ml of ozone per minute. NO_2 also uses IC analysis but the NO_2 pads are coated with triethanolamine (TEA). This forms a TEA/anion adduct and nitrite is the analyte of interest.

NO_x uses UV-Vis spectroscopy analysis. NO_x pads are coated with 2- Pyenyl-4,4,5,5-tetramethylimidazoline-3-oxide-1-oxyl (PTIO) which traps NO_2 and NO gas. Collected samples are extracted with milli-Q water and a color-producing reagent (CPR) is added. CPR is also added to the nitrite ion standards of various concentrations that are used for calibration. As with NO_2, nitrite is the analyte of interest. The Molecular Devices Spectromax 190 absorbance microplate reader (Molecular Devices Corp., Sunnyvale, CA) detects the nitrite ions colorimetrically at a wavelength of 545 nm. The chromatograms are analyzed using the Chromeleon software. Concentrations of NO_x over the sampling period were calculated from the mass using equations and tables provided by Ogawa & CO.[118] Quality control checks involved comparing laboratory blanks, field blanks, and duplicate samples as well as collocated AQS data.

Air Pollution Measures

Long-term exposure to $PM_{2.5}$ and NO_2 was estimated using a fine scale spatiotemporal model.[110] The spatiotemporal model has been described elsewhere.[119–121] Briefly, the model utilizes air pollution data from agency-run stationary monitors and two-week community monitoring campaigns. The model factors in variation in observations over space and time as well as geographic covariates to predict exposure for $PM_{2.5}$ and NO_2 at each participant's residential location every two weeks over the study period. The $PM_{2.5}$ predictions leverages a national spatiotemporal prediction model while NO_2 predictions used city specific models (i.e.

separate models for Bronx, Chicago, Miami, and San Diego). The hierarchical model structure and city specific models for NO_2 allows for unbalanced data.[119–121] The model can be written as

$$C(s,t) = \mu(s,t) + \nu(s,t)$$

where C(s,t) is the log-transformed 2-week average pollutant concentration at location s and time t. μ(s,t) is the spatiotemporal mean surface and can be broken down further to terms that contain the time trends, long-term means at location s, and spatially varying coefficients for the time trends.[119–121] ν(s,t) is the spatial correlation structure which is assumed independent for each time point.[119–121]

The spatiotemporal models utilized data from the EPA's Air Quality System (AQS) in all four cities as well as the additional New York City Community Air Survey (NYCCAS) monitoring in New York City. Additional data on NO_2 was collected during two-week community monitoring campaigns to capture within-city variability.[120] For the SOLAir project, campaigns were conducted in San Diego and Miami three times and Chicago twice during 2021 and 2022. Additional data for models in Chicago and the Bronx were previously collected in these cities as part of MESA Air[122] and SPIROMICS.[123] In addition, the Bronx had NYCCAS monitoring sites that use the same passive diffusion samples, Ogawa samplers (Ogawa and Company, USA, Inc., Pompano Beach, FL), for gaseous pollutants. Data from the EPA monitors, Ogawa samplers, and geographic covariates were incorporated into an established spatiotemporal model that was able to incorporate irregular monitoring information, reduce high dimensional covariate data using partial least squares regression, and leverages spatial smoothing to achieve satisfactory fine-scale predictions.[110]

<u>NO_2 Model Parameters</u>

We developed city-specific NO_2 spatiotemporal models in the Bronx, Chicago, Miami, and San Diego. All models were developed using data from long-term AQS monitors, defined as having data for over 2 years with over 80% coverage. The Bronx model was developed using additional long-term NYCCAS monitors, fixed monitors from MESA Air,[110] and Ogawa samples from MESA Air and SPIROMICS.[123] The Chicago model was developed using Ogawa samples from MESA Air and SOLAir, and Miami and San Diego only had additional monitoring from SOLAir Ogawa samples.

We ran various iterations of the city specific NO_2 models and assessed cross-validation performance to identify the best fitting model. For the Bronx, we assessed model performance using 10-fold with agency monitors, NYCCAS monitors, and fixed site monitor cross-validation to assess performance on the two-week time scale, on an annual time scale, and on the long-term average. Additionally, we used a leave one cluster out cross-validation for the MESA Air and SPIROMICS Ogawa samplers. We looked at performance *across time* by averaging all the predications at each Ogawa sampler location across the sampling campaigns and we looked at performance by campaign. The best performing model for the Bronx utilized 1 time trend, 4 PLS components, and an exponential kriging covariance matrix. R^2 for the annual, long-term average, and two-week time scale 10-fold cross-validation were all 0.8. R^2 for MESA Air leave one cluster out across time was 0.92, and was 0.88 for each sampling campaign. For SPIROMICS, model performance was also good with the across time R^2 of 0.67 and a R^2 of 0.88 for each sampling campaign.

For Chicago, we assessed model performance using leave one out with agency monitors and leave one out with agency monitors and fixed site monitor cross-validation to assess performance on the two-week time scale, on an annual time scale, and on the long-term average.

Additionally, we used a leave one cluster out cross-validation for the MESA Air and SOLAir Ogawa campaigns to assess performance across time for each cluster and for each sampling campaign date. The Chicago model used extra spatial covariates from the SOLAir snapshot locations and 5 additionally selected covariates (log meters to A1-A1 intersections, log of NO_x emissions within 30000 m, A2-A3 intersections within 150 m, and industrial land use within 1500 m and 5000 m). The best fitting model used 1 time trend, 4 PLS components, and an exponential kriging covariance matrix. For the leave one out for agency monitors, R^2 for the annual, long-term average, and two-week time scale 10-fold cross-validation were 0.81, 0.72, and 0.7, respectively. For the leave one out for agency monitors and fixed sites, R^2 for the annual, long-term average, and two-week time scale 10-fold cross-validation were 0.77, 0.60, and 0.68, respectively.

R^2 for MESA Air leave one cluster out across time was 0.73, and was 0.70 for each Ogawa campaign. For SOLAir, the Ogawa sampler predictions across time R^2 was 0.75. The performance by Ogawa sampling campaigns were weaker, R^2 of 0.33, because the Ogawa samplers between July and October are not correlated. Additionally, the gradient away from the roadway differed significantly between the two months, meaning they did not all follow the expected gradient patterns of higher predictions closer to the roadway. However, modeling results were consistent between previous modeling efforts and are we confident in our predictions based on the performance of the other cross validations.

For Miami, we assessed model performance using leave one out with agency monitors cross-validation to assess performance on the two-week time scale, on an annual time scale, and on the long-term average. We used a leave one cluster out cross-validation for the SOLAir Ogawa samplers and assessed performance across time for each cluster and for each sampling

campaign date. Additional spatial covariates from the SOLAir Ogawa sampler locations were added to the model. The best fitting model used 1 time trend, 3 PLS components, and an exponential kriging covariance matrix. R^2 for the annual, long-term average, and two-week time scale leave one out cross-validation were 0.63, 0.70, and 0.61, respectively. R^2 for SOLAir leave one cluster out across time was 0.60 and was 0.71 for each sampling campaign date.

San Diego used the same cross validation scheme as Miami and used the additional spatial covariates from the SOLAir Ogawa sampler locations. The best fitting model for San Diego used 1 time trend, 3 PLS components, and an exponential kriging covariance matrix. R^2 for the annual, long-term average, and two-week time scale leave one out cross-validation were 0.84, 0.83, and 0.84, respectively. R^2 for SOLAir leave one cluster out across time was 0.60 and was 0.86 over time for each sampling campaign date.

Combining Air Pollution Predictions and HCHS/SOL data

The spatiotemporal model provided air pollution predictions for each building located within census blocks where HCHS/SOL participants lived on a 2-week timescale. Predictions for the full follow-up period (2008-2017) were transferred into the HCHS/SOL secure research workspace (SRW). Predictions were then joined with study participant data using a spatial join of the closest building centroid and the residential geocodes of the study participants. Study participants needed to have air pollution predictions for 95% of their follow-up time in order to be included in analysis. Having less than 95% was a result of missing address history data or incomplete addresses that were not able to be geocoded. Participant exposure over the follow-up period was calculated as the average exposure between their baseline visit date and their visit 2

date. Average exposure during follow-up accounted for any residential moves during the follow-up period.

Physical Activity

Physical Activity data was collected by HCHS/SOL investigators at baseline. Investigators collected data using the Global Physical Activity Questionnaire (GPAQ) and accelerometers.[32,117] The questionnaire included questions on active transportation, recreational MVPA, and occupational MVPA. Overall physical activity was calculated by combining the recreational moderate-to-vigorous physical activity (MVPA), occupation MVPA, and active transport minutes per day. Overall physical activity was dichotomized based on the recommended amount of 150 minutes per week.[124]

HCHS/SOL participants wore the Actical accelerometer (model 198-0200-03, Minimiter Respironics, Bend, OR) for 7 days to collect data on overall MVPA.[32,117] Participants were instructed to wear the Accelerometer above the iliac crest and remove it only for swimming, showering, and sleeping.[32] Overall MVPA collected by the Actical was the minutes per day of activity equal to or greater than 1535 counts/min in participants with more than 3 days of wearing adherence. Adherence was defined as wearing an accelerometer for more than 10 hours. 77.7% (n = 12,750) met the adherence requirement for more than 3 days.[125] Overall MVPA by the Actical was also dichotomized based on the recommended amount of 150 minutes per week.[124]

Waist Circumference

Waist circumference was measured at baseline by measuring waist girth in centimeters. Waist circumference was dichotomized based on recommendations from the American Heart Association, National Heart, Lung and Blood Institute. Abdominal obesity was defined as greater than 88 cm for women and greater than 102 cm for men.[126]

Covariates

Age, sex, education level, income, Hispanic/Latino heritage, years living in the US, family history of diabetes, marital status, cigarette use, and alcohol use were all collected at baseline. Education was a 4-level categorical variable consisting of no high school diploma or GED (reference), at most a high school diploma or GED, high school (or GED) education, and University/college education. Income was a 5-level categorical variable (<$10,000 (ref), $10,001 - $20,000, $20,001 - $40,000, $0,001 - $75,000, and more than $75,000). As Hispanic/Latino heritage is highly correlated with study center, a 17-level combination variable of the two was created: Dominicans – Bronx (ref), Central American – Bronx, Central American – Chicago, Central American – Miami, Cubans – Miami, Mexicans – Bronx, Mexicans- Chicago, Mexicans – San Diego, Puerto Ricans – Chicago, South American – Bronx, South American – Chicago, South American – Miami, Other – Bronx, Other – Chicago, Other – Miami, and Other – San Diego . Years living in the US was a 3-level variable of less than 10 years, 10 years or more, or US born. The neighborhood SES index is a compositional measure of NSES at the census tract level that represents education, employment, housing, income/wealth, occupation, and residential stability.[127] Higher index scores indicate higher NSES disadvantage.[127] Marital status was a 3-level categorical variable of single (reference), married or living with a partner, or separated,

divorced, or widower. Both cigarette use and alcohol use were 3-level variables of never (reference), former, or current.

Inclusion Criteria

Starting with the 11,619 participants with ascertainment of diabetes status at visit 2, analysis was limited to those who were free of type 2 diabetes at baseline (N=9209 for primary diabetes outcome and N= 9,077 for secondary diabetes outcome). Analysis was further limited to participants with successfully geocoded residential addresses for the follow-up period and air pollution predictions that covered at least 95% of the follow-up period (N=8248 for primary diabetes outcome and N= 8129 for secondary diabetes outcome). Those without geocoded successfully geocoded residential addresses for 95% of follow-up period were younger, and more likely to be from the Bronx and single compared to the analytic sample. These characteristics indicate a population that is more likely to move residences multiple times over the follow-up period and may have an increased likelihood of underreporting address history.

Statistical Analysis

We computed demographic statistics on those with incident diabetes cases and those free of diabetes at visit 2 using both definitions of diabetes. We calculated age-centered adjusted incidence rates and descriptive statistics on $PM_{2.5}$ and NO_2 for each sample overall and stratified by study center.

For the primary analysis, we used individual Poisson regression models to assess the association between each air pollutant's average concentration over the follow-up period and incident diabetes at visit 2 using the primary diabetes definition (does not include self-reported

diagnosis). Our Poisson regression models used a log link and years between baseline and visit 2 as the offset for follow-up time to generate an incidence rate ratio (IRR) comparing groups with a one-IQR difference in either $PM_{2.5}$ or NO_2. Our base model controlled for age and gender. Our second model added relevant confounders in the air pollution – diabetes relationship: education, income, Hispanic/Latino heritage/Study Center, years in the US, and a neighborhood socioeconomic status index.[127] Next, we added family history of diabetes and waist circumference, two important factors in diabetes risk, for our main set of covariates. Interaction term of overall physical activity (mins/day) was added to the model to investigate physical activity as an effect modifier in the relationship between air pollution exposure and incident diabetes. Waist circumference was also added as an interaction term to investigate it as a potential effect modifier.

In addition to modeling incident diabetes, we also investigated the association between $PM_{2.5}$ and NO_2 and three different progressions of diabetes: 1) progression from pre-diabetes to diabetes, where the sample only included those with pre-diabetes at baseline, 2) progression of those free of diabetes and pre-diabetes to diabetes and 3) progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes at visit 2. Lastly, we modeled change in continuous HbA1C (%) levels over the follow-up period. The linear regression model for change in HbA1c also adjusted for baseline levels and antihyperglycemic medication usage at baseline and visit 2. We also stratified the models for change in continuous HbA1C by diabetes status at visit 1.

All regression models used complete cases and sampling weights to account for the HCHS/SOL complex sampling design and participant non-response at visit 2 by using the *survey* package in R.[116] The sampling weights were also calibrated to the age, sex, and Hispanic/Latino

heritage percentages of the 2010 US Census. Results from all models were exponentiated to report the IRR for each IQR increment in $PM_{2.5}$ or NO_2 exposure.

Secondary and Additional Analyses

For a secondary analysis, we repeated analyses using the T2D outcome definition that included self-reported diagnosis. As study participants' years of follow-up can vary and occur over a slightly different time-period, we investigated if having more or less than the median follow-up time resulted in different $PM_{2.5}$ and NO_2 exposure. Additionally, we investigated if there was an association between a 3-year average (July 01, 2011 – July 1, 2014) exposure to NO_2 at participant's residential addresses and incident diabetes at visit 2. We used the period between July 2011 and July 2014 as this was a 3-year period after all study participants completed visit 1 and before any study participant completed visit 2.

Missing Data

For the analytic sample, the only variable with greater than 5% missingness was income with 7.5% of participants missing data. Due to concerns about missing data and the computational intensity of multiple imputation by chained equations (MICE), we conducted MICE analysis as a sensitivity analysis on two model iterations. The MICE analysis consisted of 10 imputations and 50 iterations on the fully adjusted model for both the primary and secondary definitions of diabetes. Results from the MICE were very similar and did not change the overall interpretation of results. Since completing MICE analysis on all our models would be computationally intense and the MICE tested on our two iterations yielded very similar results, we used complete case analysis for our reported analyses.

All analyses were conducted in R Studio, with R version 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria). The 'mice' package (version 3.15.0) was used for the MICE analysis,[128] and the 'survey' package (version 4.2-1) was used for incorporating the survey design and weights into the regression models.[129]

Results

For our primary definition of diabetes (excluded self-reported diagnosis), the analytic sample comprised of 8248 participants who were free of diabetes at baseline. Of the 8248 participants, there were 950 incident diabetes cases at visit 2 with an age-centered adjusted incidence rate of 16.3 per 1000 person-years. For our secondary definition of diabetes (including self-reported diagnosis, there were 8129 participants free of diabetes at baseline and 1354 incident cases by visit 2 with an age-centered adjusted incidence rate of 24.5 per 1000 person-years. Weighted summary characteristics of participants by diabetes status, using the primary definition, at visit 2 are presented in Table 2.

Table 2. Weighted summary characteristics of participants by diabetes status at Visit 2 using primary definition of diabetes.

	Sample Weighted % or Mean (SD)	
Variables	Diabetes at Visit 2 (N= 950)	Free of Diabetes at Visit 2 (N = 7298)
Female	47.6	50.8
Age, years	47.3 (12.6)	38.5 (14.1)
Waist Circumference, cm	105 (12.2)	95.2 (13.4)
Hispanic/Latino Heritage – Study Center		
Central American - Bronx	0.8	1.4
Central American - Chicago	1.6	1.2
Central American - Miami	2.4	4.4
Cubans - Miami	15.9	17.2
Dominicans - Bronx	9.1	8.3

Mexicans- Miami	2.2	3.3
Mexicans- Chicago	12.5	11.8
Mexicans- San Diego	28.6	27.9
Puerto Rican - Bronx	13.0	8.9
Puerto Rican - Chicago	4.7	3.6
South American - Bronx	1.5	1.2
South American - Chicago	0.7	1.3
South American - Miami	0.6	2.5
Other - Bronx	1.9	1.5
Other – Chicago	0.5	0.9
Other - Miami	2.6	2.3
Other – San Diego	1.4	2.2
Years in the US		
Less than 10 years	23.6	29.0
10 years or more	58.6	48.9
US Born	17.9	22.0
Family History of Diabetes	49.9	36.4
Education		
No High School Diploma or GED	38.5	28.3
At most a High school diploma or GED	25.7	28.3
High school (or GED) education	12.5	13.4
University/college education	23.2	29.9
Income		
Less than $10,000	16.1	12.7
$10,001-$20,000	28.2	31.5
$20,001-$40,000	34.9	35.1
$40,001-$75,000	14.3	14.9
More than $75,000	6.4	5.7
Neighborhood SES Index	1.1 (0.8)	1.1 (0.8)
Marital Status		
Single	20.8	32.9
Married or living with a partner	60.5	53.1
Separated, divorced, or Widow(er)	18.6	14.0
Cigarette Use		
Never	58.2	64.5
Former	24.3	16.0
Current	17.5	19.5
Alcohol Use		
Never	17.3	17.5
Former	30.8	28.7
Current	52.0	53.8

Recommended GPAQ Physical Activity[a]		67.8	69.6
Recommended Actical Physical Activity[a]		34.4	43.6
Years between baseline and Visit 2		6.0 (0.8)	6.1 (0.8)
Study Center			
	Bronx	28.5	24.6
	Chicago	20.0	18.9
	Miami	21.5	26.5
	San Diego	30.0	30.0
$PM_{2.5}$ ($\mu g/m^3$)			
	Overall	9.2 (1.4)	9.1 (0.8)
	Bronx	10.5 (0.4)	10.5 (0.4)
	Chicago	10.0 (0.4)	10.0 (0.4)
	Miami	7.1 (0.3)	7.1 (0.3)
	San Diego	9.5 (0.8)	9.6 (0.7)
NO_2 (ppb)			
	Overall	13.6 (6.4)	13.0 (13.4)
	Bronx	22.2 (1.9)	22.0 (1.8)
	Chicago	15.8 (1.5)	15.9 (1.4)
	Miami	6.8 (1.4)	6.7 (1.5)
	San Diego	9.8 (1.8)	10.0 (1.8)

[a] Based on recommended 150 minutes per week

Maps of the average concentration of $PM_{2.5}$ between 2008 and 2011 for each building in a census block where HCHS/SOL participants in the Bronx, Chicago, Miami, and San Diego are presented in figures 1-4. Maps of the average concentration of NO_2 in each city are presented in figures 5-8. Among the HCHS/SOL participants included in this analysis, the mean $PM_{2.5}$ concentration at participants' residential addresses over the follow-up period ranged from 6.40 to 11.6 $\mu g/m^3$ across all 4 study centers with a median of 12.2 $\mu g/m^3$. $PM_{2.5}$ over the follow-up period was moderately correlated with population density (0.47) and the percentage of Hispanic in the 2010 census tract (r = -0.62). $PM_{2.5}$ was not correlated with neighborhood SES (r = 0.06). For NO_2, the average concentration at participants' residential addresses over the follow-up period ranged from 0.46-27.9 ppb across the 4 study centers. The median concentration was 12.2

ppb. NO_2 over the follow-up period was moderately correlated with population density (0.79) and the percentage of Hispanic in the 2010 census tract (r = -0.62). Similar to $PM_{2.5}$, NO_2 was not correlated with neighborhood SES (r = 0.35). Weighted mean and standard deviation of $PM_{2.5}$ and NO_2 for each of the 4 study centers is presented in Table 2.

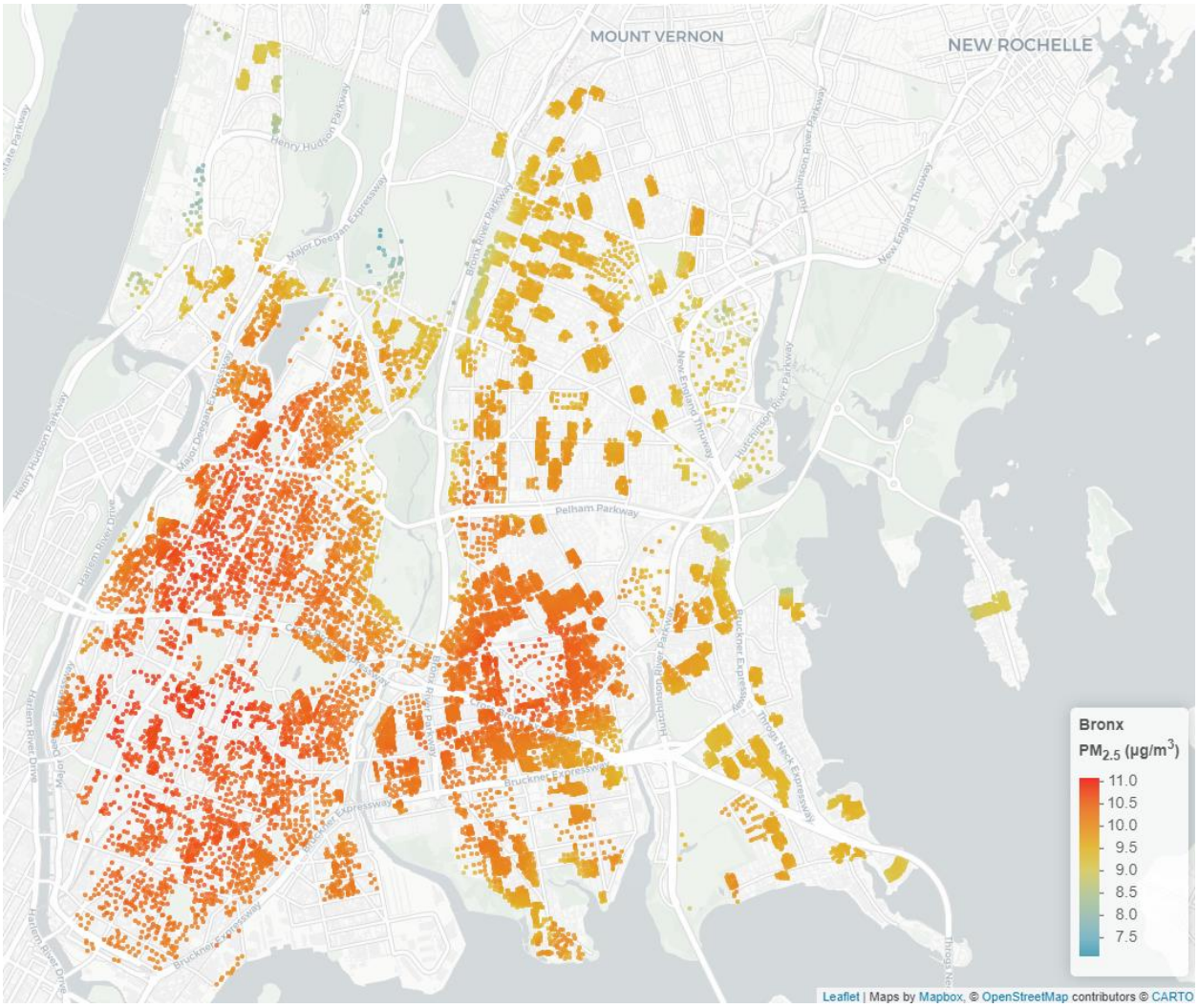


Figure 1. Average concentration of PM2.5 (µg/m3) between 2008 and 2017 at each census block in the Bronx site of the HCHS/SOL.

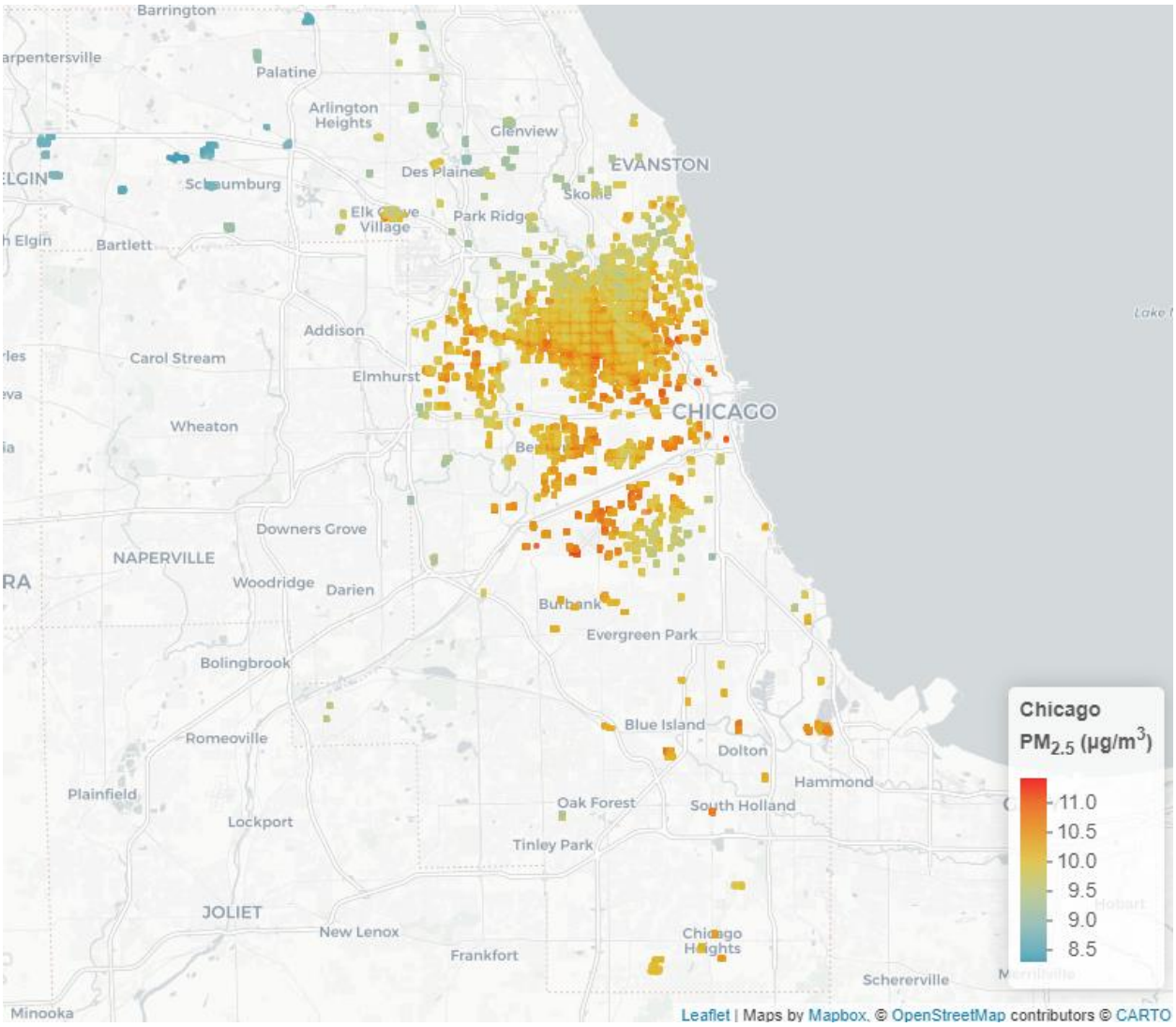


Figure 2. Average concentration of PM2.5 (µg/m3) between 2008 and 2017 at each census block in the Chicago site of the HCHS/SOL.

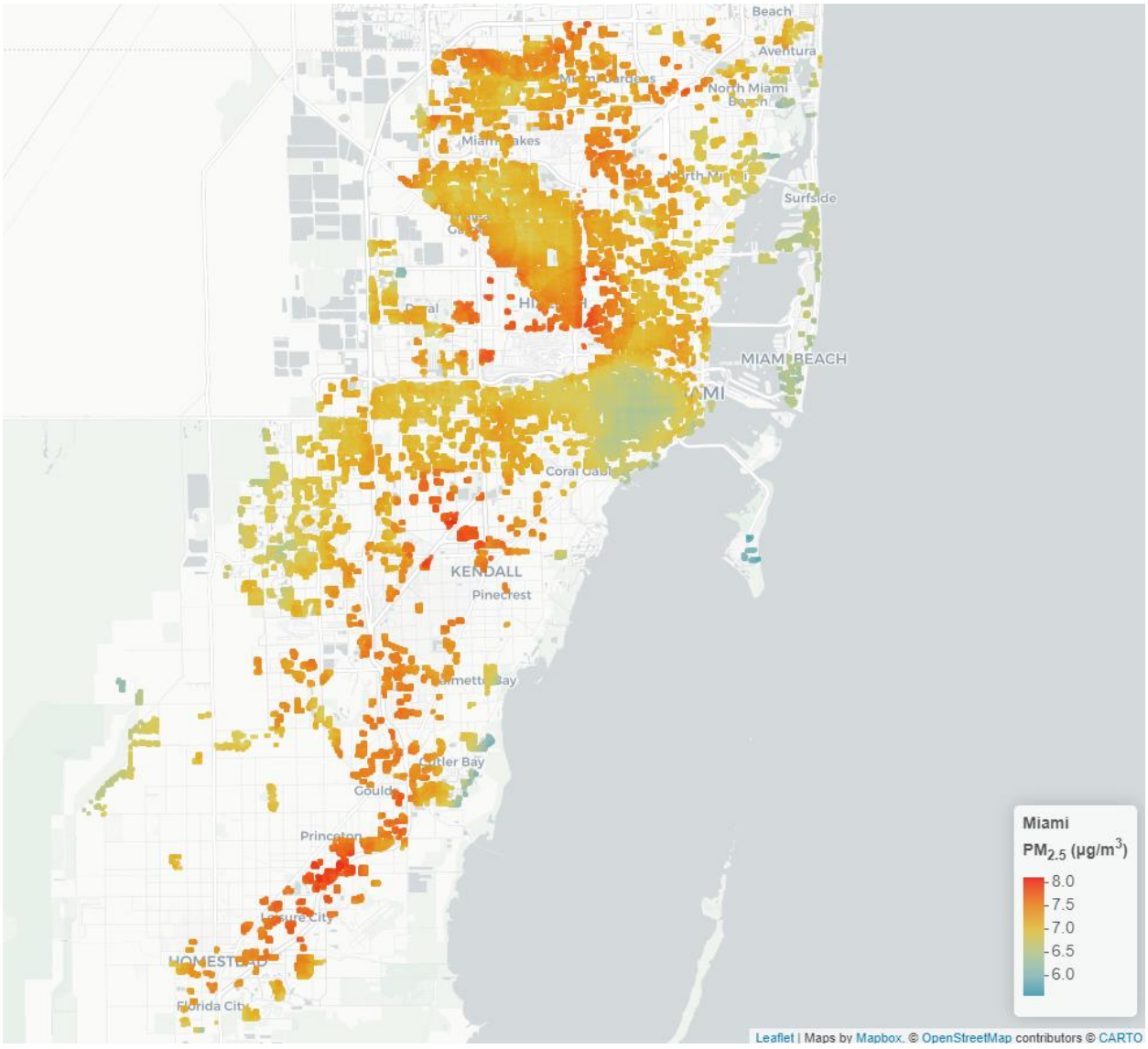


Figure 3. Average concentration of PM2.5 (µg/m3) between 2008 and 2017 at each census block in the Miami site of the HCHS/SOL.

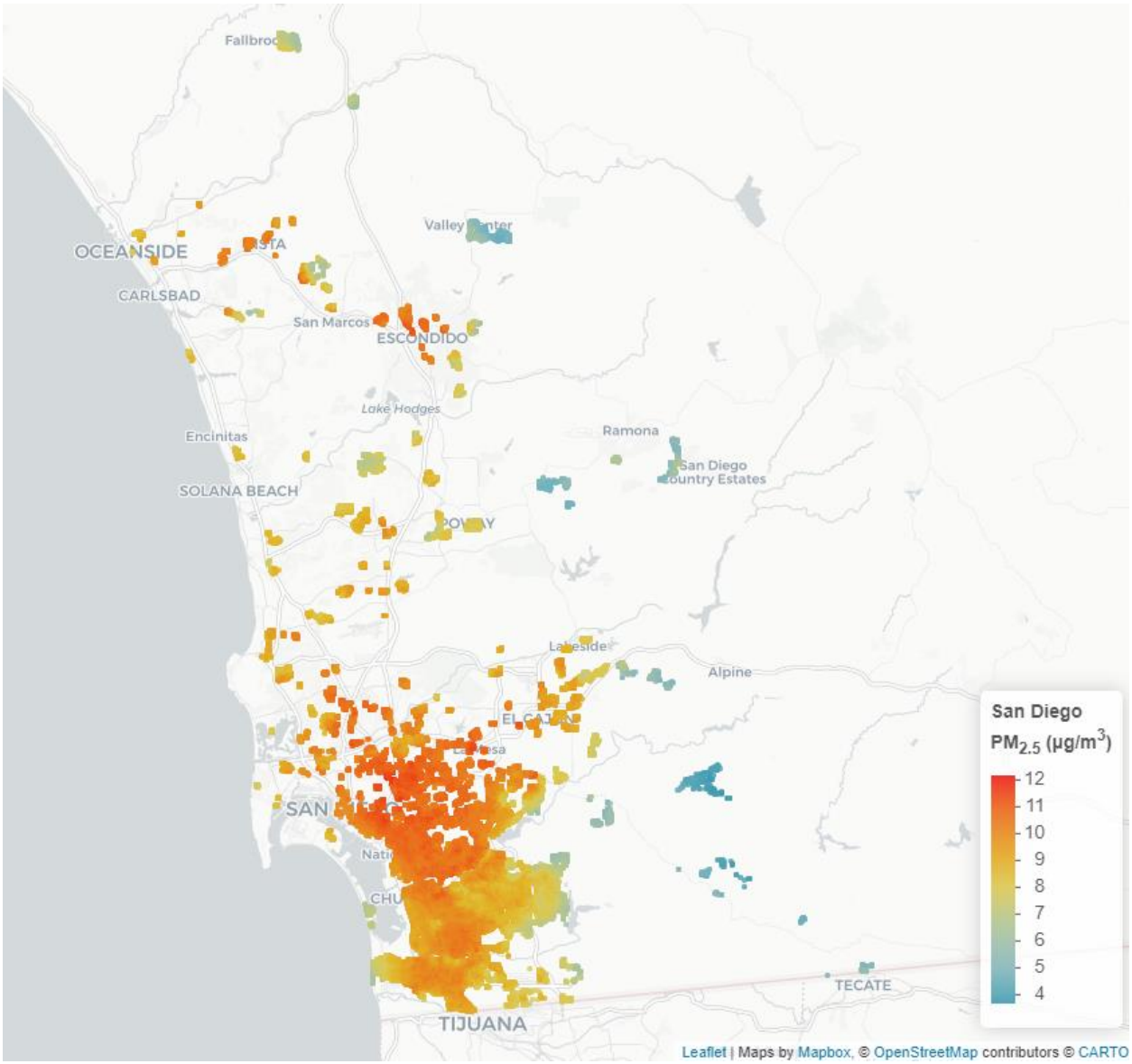


Figure 4. Average concentration of PM2.5 (µg/m3) between 2008 and 2017 at each census block in the San Diego site of the HCHS/SOL.

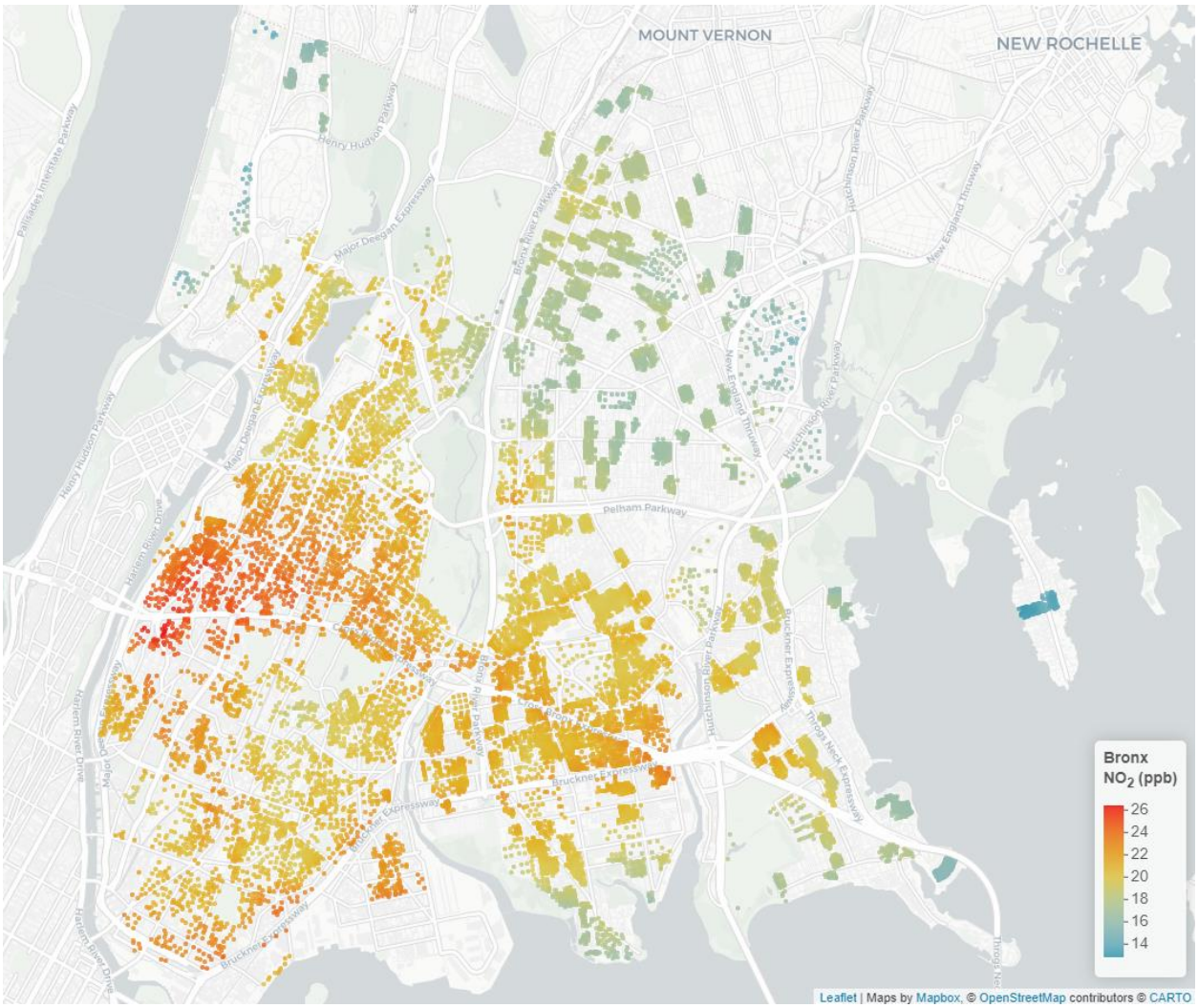


Figure 5. Average concentration of NO2 (ppb) between 2008 and 2017 at each census block in the Bronx site of the HCHS/SOL.

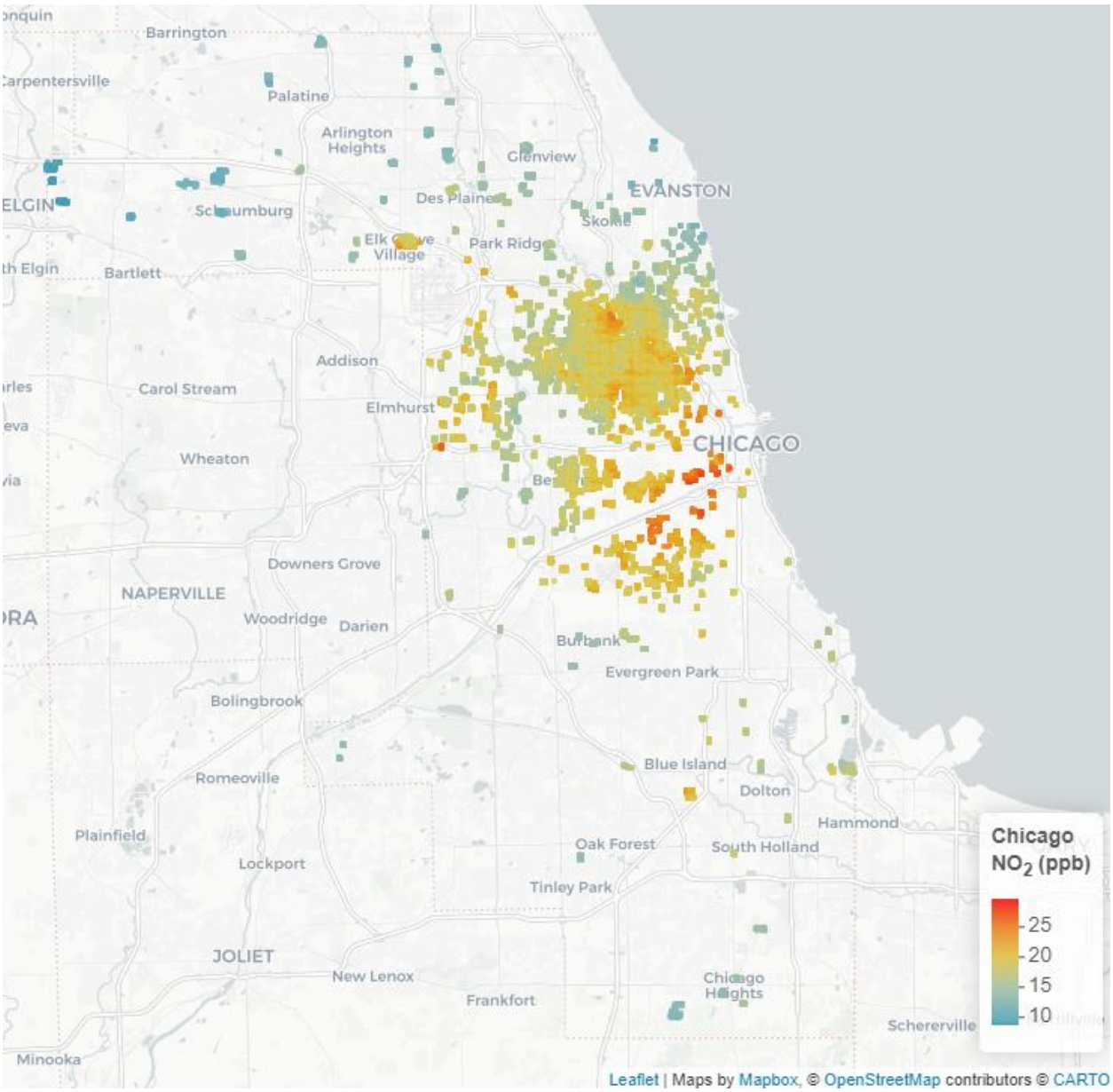


Figure 6. Average concentration of NO2 (ppb) between 2008 and 2017 at each census block in the Chicago site of the HCHS/SOL.

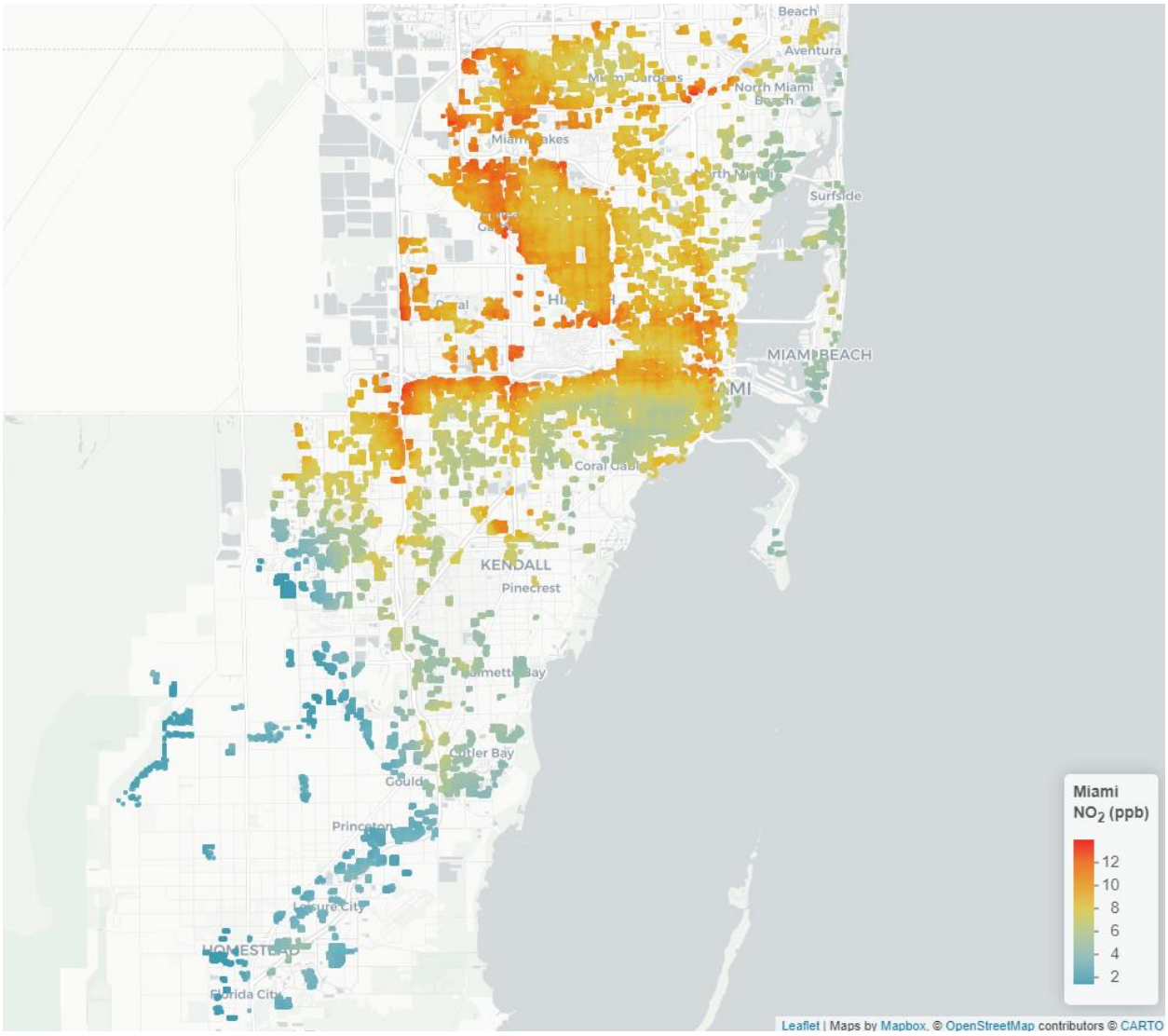


Figure 7. Average concentration of NO2 (ppb) between 2008 and 2017 at each census block in the Miami site of the HCHS/SOL.

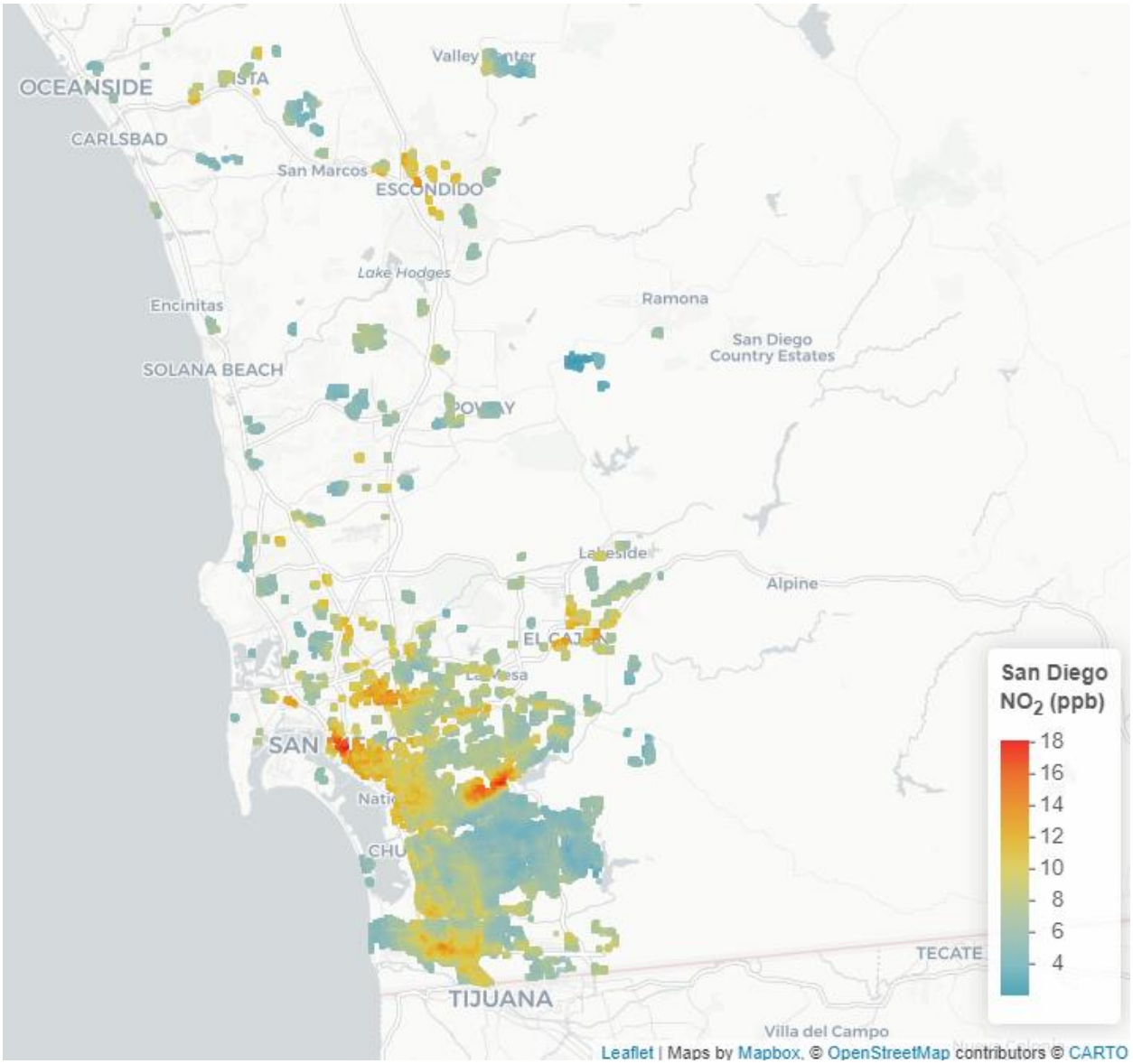


Figure 8. Average concentration of NO2 (ppb) between 2008 and 2017 at each census block in the San Diego site of the HCHS/SOL.

$PM_{2.5}$

When utilizing our primary definition of diabetes, a 2.48 μg/m^3 increase in $PM_{2.5}$ was associated with a 27% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.27, 95% CI: 1.09 – 1.49) (Table 3). This association was no longer significant when controlling for the additional covariates of years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ethnic heritage (Table 3). When utilizing our secondary definition of diabetes, a 2.48 μg/m^3 increase in $PM_{2.5}$ was associated with a 40% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.40, 95% CI: 1.23 – 1.59) (Table 3). This association was also attenuated when controlling for our main set of covariates (IRR = 1.19, 95% CI: 0.82 – 1.74). There was no evidence of effect modification in the association between $PM_{2.5}$ and incident diabetes by waist circumference, physical activity measured by the GPAQ, or physical activity measured by the Actical accelerometer (Supplemental Table 1). All associations between $PM_{2.5}$ and the different progressions of diabetes were not significant (Table 3).

For our analysis at each study center using the primary definition of diabetes and adjusting for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and ethnic heritage, a 1 μg/m^3 increase in $PM_{2.5}$ was associated with a 55% increased risk of incident diabetes in the Bronx (IRR = 1.55, 95% CI: 1.04 – 2.31) (Figure 9). Association within each of the 3 other cities were not significant (Figure 9).

Table 3. Minimally and fully adjusted incidence rate ratios (IRR) of diabetes (primary and secondary definitions) at visit 2 based on average concentration of PM2.5 (μg/m3) and NO2

(ppb) at participants' residential address over their follow-up period. For $PM_{2.5}$, IRR are reported per an IQR increase of 2.48 $\mu g/m^3$ for primary diabetes and 2.51 $\mu g/m^3$ for secondary diabetes. For NO_2, IRR are reported per an IQR increase of 8.69 ppb for primary diabetes and 8.73 ppb for secondary diabetes.

		Minimally Adjusted[a]			Adjusted[b]		
	Outcome	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
$PM_{2.5}$ ($\mu g/m^3$)							
	Primary diabetes (Lab results and/or self-reported medication)	8250	950	1.27 (1.09 – 1.49)*	7574	879	0.91 (0.55 – 1.51)
	Secondary diabetes (Primary + self-reported diagnosis)	8131	1355	1.40 (1.23 – 1.59)*	7464	1254	1.19 (0.82-1.74)
NO_2 (ppb)							
	Primary diabetes (Lab results and/or self-reported medication)	8248	950	1.22 (1.08 – 1.37)*	7572	879	1.09 (0.70-1.71)
	Secondary diabetes (Primary + self-reported diagnosis)	8129	1354	1.14 (1.04 – 1.26)*	7462	1253	1.19 (0.80 – 1.77)

Abbreviations: IRR, incident rate ratio
* $P < 0.05$
[a] Adjusted for age and sex
[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ ethnic heritage

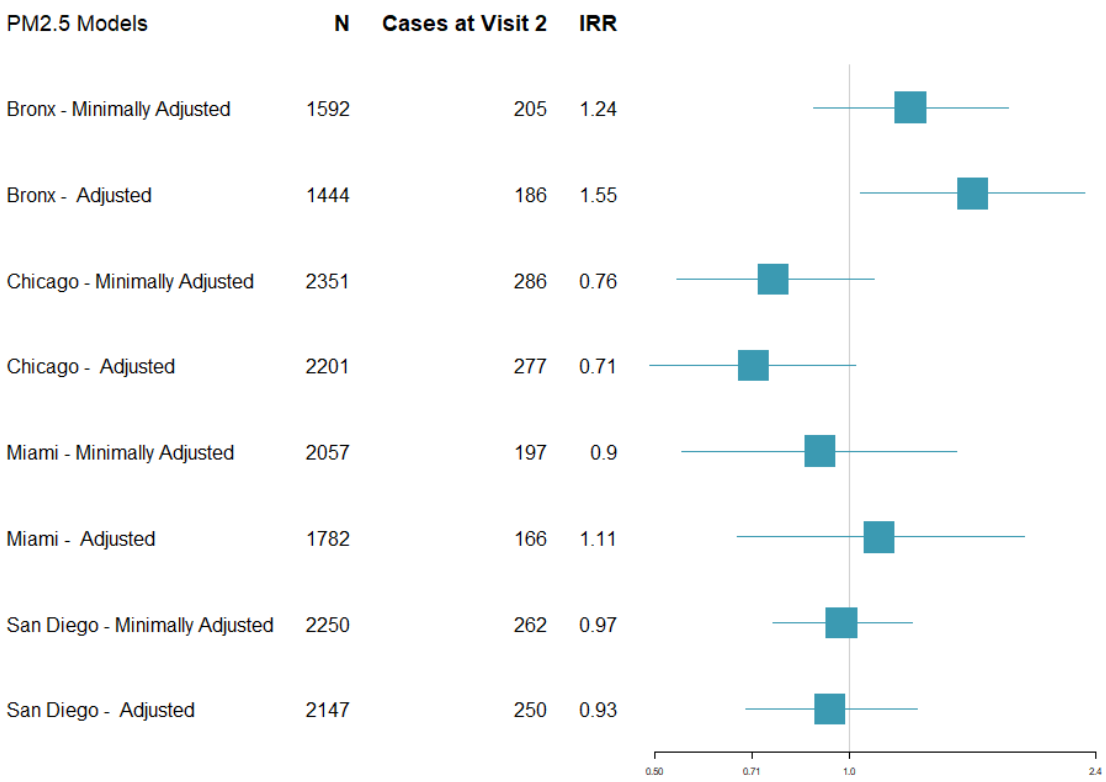


Figure 9. Association between average $PM_{2.5}$ concentration during follow-up and incident diabetes (primary diabetes definition) within each study center. Incident rate ratio (IRR) reported as 1 $\mu g/m^3$ increase in $PM_{2.5}$. Minimally adjusted model adjusts for age and sex. Adjusted model adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and ethnic heritage.

For the association between average concentration of $PM_{2.5}$ over the follow-up period and change in continuous HbA1C (%) levels, a 2.17 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with a mean change in HbA1C of 0.06% (95% CI: 0.02 – 0.09) when adjusting for age, sex, follow-up time, and baseline HbA1C levels (Table 4). When including adjustments for antihyperglycemic

medication usage at baseline and visit 2, a 2.17 μg/m^3 increase in $PM_{2.5}$ was associated with a mean change in HbA1C of 0.05% (95% CI: 0.01 – 0.08) (Table 4). The association was no longer significant when including additional adjustment for education, income, study center/ethnic heritage, years in the US, neighborhood SES, family history, waist circumference (Table 4). Among those free of diabetes at visit 1, the mean change in HbA1C levels was 0.01% for all model iterations and no associations were significant (Supplemental Table 2). Among those with diabetes at baseline, the mean change in HbA1A levels were minimal and not statistically significant (Supplemental Table 3).

Table 4. Association between average concentration of of $PM_{2.5}$ and change in continuous HbA1C (%) over the follow-up period. Mean change reported per 2.17 μg/m^3 increase in $PM_{2.5}$.

Model	**N**	**Mean Change$^+$ (%)**	**95% CI**	**P-value**
Model 1	10244	0.06	0.02 – 0.09	<0.001
Model 2	10010	0.05	0.01– 0.08	0.01
Model 3	9174	0.01	-0.10 – 0.13	0.80
Model 4	9133	0.01	-0.10 – 0.12	0.86

Model 1: adjusts for age, sex, time between visits, and visit 1 HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at visit 1 + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
$^+$ per 2.17 μg/m^3 increase in $PM_{2.5}$

<u>NO_2</u>

When utilizing our primary definition of diabetes, an 8.69 ppb increase in NO_2 was associated with a 22% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.22, 95% CI: 1.08 – 1.37) (Table 5). This association was less precise when controlling for the additional covariates of years living in the US, income, educational attainment, family history of

diabetes, neighborhood SES, waist circumference, and study center/ethnic heritage (Table 5). When utilizing our secondary definition of diabetes, an 8.73 ppb increase in NO_2 was associated with a 14% increased risk of incident diabetes when adjusting for age and sex (IRR = 1.14, 95% CI: 1.04 – 1.26) (Table 5). The IRR increased when controlling for our main set of covariates (IRR = 1.19, 95% CI: 0.80 – 1.77), however, the association was no longer significant. There was no evidence of effect modification in the association between NO_2 and incident diabetes by waist circumference, physical activity measured by the GPAQ, or physical activity measured by the Actical accelerometer (Supplemental Table 5). None of the associations between NO_2 and the different progressions of diabetes were significant (Table 5).

Those with the median follow-up time or less ($\leq$ 5.86 years) had significantly less average NO_2 (12.8 ppb) than those with longer follow-up time (13.6 ppb) (Supplemental Table 6). To standardize exposure time, we used the average NO_2 exposure between July 2011 and July 2014. Results from the associations between the 3-year average exposure and incidence diabetes in the minimally adjusted model were very similar to the results when using the average over each participant's full follow-up period (Supplemental Table 7). Effect sizes and confidence intervals were larger for the 3-year average exposure in the main adjusted model compared to the results for the exposure over the full follow-up period (Supplemental Table 7).

For our study center specific models using the primary definition of diabetes and adjusting for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and ethnic heritage, an increase in NO_2 of 1 ppb was associated with a 10% increased risk of incident diabetes in the Bronx (IRR = 1.10, 95% CI: 1.01 – 1.19) (Figure 10). Association within each of the 3 other cities were not significant (Figure 10).

Table 5. Minimally and fully adjusted incidence rate ratios (IRR) of difference progressions of diabetes (primary definition) at visit 2 based on average concentration of $PM_{2.5}$ ($\mu g/m^3$) and NO_2 (ppb) at participants' residential address over their follow-up period. For $PM_{2.5}$, IRR are reported per an IQR increase of 2.50 $\mu g/m^3$ for progression from pre-diabetes to diabetes, 2.47 $\mu g/m^3$ for progression of those free of diabetes and pre-diabetes to diabetes, and 2.47 $\mu g/m^3$ for progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes. For NO_2, IRR are reported per an IQR increase of 8.76 ppb for progression from pre-diabetes to diabetes, 8.60 ppb for progression of those free of diabetes and pre-diabetes to diabetes, and 8.60 ppb for progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes.

		Minimally Adjusted[a]			Adjusted[b]		
	Outcome	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
$PM_{2.5}$ ($\mu g/m^3$)							
	Progression from pre-diabetes to diabetes	4158	850	1.15 (0.97 – 1.36)	3831	788	0.75 (0.43 – 1.33)
	Progression of those free of diabetes and pre-diabetes to diabetes	4092	100	1.36 (0.84 – 2.20)	3743	91	0.56 (0.12 – 2.64)
	Progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes	4092	1713	1.01 (0.89 – 1.14)	3743	1587	0.91 (0.67 – 1.22)
NO_2 (ppb)							
	Progression from pre-diabetes to diabetes	4157	850	1.11 (0.98 – 1.27)	3830	788	0.88 (0.54 – 1.43)
	Progression of those free of diabetes and pre-diabetes to diabetes	4091	100	1.35 (0.91 – 2.01)	3742	91	0.46 (0.08 – 2.55)
	Progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes	4091	1712	1.02 (0.93 – 1.12)	3742	1586	1.14 (0.83 – 1.57)

Abbreviations: IRR, incident rate ratio
[a] Adjusted for age and sex
[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ ethnic heritage

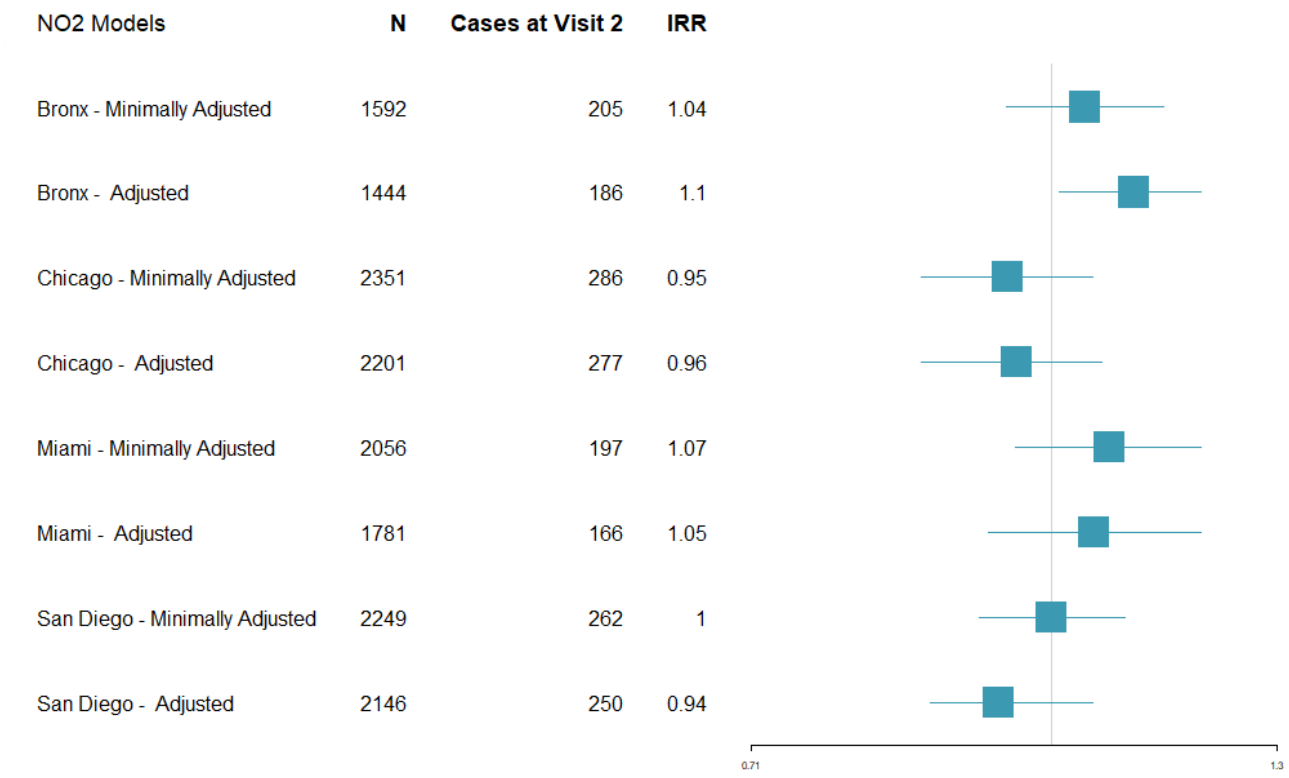


Figure 10. Association between average NO_2 concentration during follow-up and incident diabetes (primary diabetes definition) within each study center. Incident rate ratio (IRR) reported as 1 ppb increase in NO_2. Minimally adjusted model adjusts for age and sex. Adjusted model adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and ethnic heritage.

For the association between average concentration of NO_2 over the follow-up period and change in continuous HbA1C (%) levels, an 8.71 ppb increase in NO_2 was associated with a mean change in HbA1C of 0.05% (95% CI: 0.02 – 0.09) when adjusting for age, sex, follow-up time, and baseline HbA1C levels (Table 6). Results were similar when including adjustments for antihyperglycemic medication usage at baseline and visit 2 (Mean Change: 0.04, 95% CI: 0.01 – 0.08) (Table 6). Similar to the results for $PM_{2.5}$, the association between NO_2 and change in HbA1C was no longer significant when including additional adjustment for education, income, study center/ethnic heritage, years in the US, neighborhood SES, family history, waist circumference (Table 6). Among those free of diabetes at visit 1, the mean change in HbA1C levels were minimal (-0.02% – 0.01%) for all model iterations and no associations were significant (Supplemental Table 7). Among those with diabetes at baseline, the mean change in HbA1A levels higher (range: 0.12% - 0.29%) compared to those free of diabetes, and the change in continuous HbA1C (%) over the follow-up period was statistically significant when adjusting for age, sex, time between visits, and visit 1 HbA1c (Supplemental Table 8).

Table 6. Association between average concentration of NO_2 and change in continuous HbA1c (%) over the follow-up period. Mean change reported per 8.71 ppb increase in NO_2.

Model	**N**	**Mean Change$^+$ (%)**	**95% CI**	**P-value**
Model 1	10244	0.05	0.02 – 0.09	0.003
Model 2	10006	0.04	0.01– 0.08	0.01
Model 3	9171	0.02	-0.11 – 0.14	0.81
Model 4	9130	0.01	-0.11 – 0.14	0.86

Model 1: adjusts for age, sex, time between visits, and baseline HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at baseline + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
$^+$ per 8.71 ppb increase in NO_2

Discussion

In the presented analysis, higher $PM_{2.5}$ and NO_2 concentrations at residential addresses was associated with increased risk of incident diabetes after adjustment for age and sex. However, this association was attenuated, and no longer significant, when adjusting for additional sets of covariates. We found no evidence of effect modification by waist circumference or physical activity. In our study center level analysis, only higher $PM_{2.5}$ and NO_2 concentrations in the Bronx were associated with increased risk of incident diabetes when adjusting for our main set of covariates.

Our results were consistent with a prior analysis of air pollution and diabetes risk in the Multi-Ethnic Study on Atherosclerosis, a large multi-site cohort that utilized our fine-scale spatiotemporal modeling framework.[53] In that analysis, Park et al.[53] did not find an association between incident diabetes and long-term exposure to $PM_{2.5}$, NO_2, or NO_x after adjustment for study site. However, there was an association within the Chicago study site when investigating the association between the annual average concentration of $PM_{2.5}$ at residential addresses and incident diabetes (Park et al 2015). Renzi et al. (2018), did not find an association between annual exposure to $PM_{2.5}$ or NO_2, and incident diabetes in a Roman cohort. Our NO_2 results were also consistent with recent meta-analyses by Yang et al.[5] and Liu et al.[18], which did not find a significant association between NO_2 and incident T2D. Our point estimates were larger but less precise than a meta-analysis on two longitudinal and two cross-sectional studies by Eze et al.[1] which found a higher risk of T2D with increased exposure. Our results for $PM_{2.5}$ were also less precise than a recent meta-analysis suggesting a significant positive association between $PM_{2.5}$ and NO_2 and incident diabetes.[5,18]

Adjusting for the combination variable of study center and ethnic heritage strongly influenced the model results. In our main adjustment model for NO_2, when adjusting for all but study center/ethnic heritage the association remained positive and significant (IRR: 1.21, 95% CI: 1.00 – 1.47). The association was not significant for $PM_{2.5}$, however, not including study center/ethnic heritage in our main adjustment model limited the attenuation (IRR: 1.19, 95% CI: 0.91 – 1.54). The weakened association between air pollution and incident diabetes when adjusting for fixed-effects of study center/ethnic heritage was likely due to the low within-city variability of $PM_{2.5}$ and NO_2. The IQR for $PM_{2.5}$ was less than 1 $\mu g/m^3$ for each city. While the within city variability of NO_2 expectedly varied more, the IQR for each city was around 2 ppb. HCHS/SOL study participant's residences were not spread out across each city evenly, but rather spatially clustered. This was a byproduct of HCHS/SOL probability sampling strategy which targeted census tracts close to the study center clinics and tracts with a high percentage of Hispanics based on the 2000 US census.[116]

For $PM_{2.5}$, using the secondary definition of diabetes increased the effect estimates in both the minimally adjusted and main adjusted model when compared to the models with the primary diabetes definition. The main adjusted model for both definitions were not statistically significant, however, the secondary definition had a positive effect estimate while the primary definition had a negative effect estimate. For NO_2, using the secondary definition of diabetes, adjusting for only age and sex resulted in a lower effect estimate when compared to the primary definition, yet this relationship was still positive and significant. For the main adjusted model, using the secondary definition increased the effect estimate when compared to the primary definition, however the overall interpretation of no significant association remained. The

increased effect estimates when using the secondary diabetes definition may be due to the increased number of cases at visit 2 compared to primary diabetes cases.

When investigating the change in continuous HbA1C levels, higher $PM_{2.5}$ and NO_2 exposure slightly, yet significantly, increased the mean percent change in HbA1C over the follow-up period when adjusting for age, sex, follow-up time, baseline HbA1c levels, and use of antihyperglycemic medication at baseline and/or visit 2. However, the mean percent change was less than 0.1 %, which is below the clinically relevant change of 0.5%.[130] These associations were also attenuated and not significant when adjusting for study center/ethnic heritage.

We did not see evidence of effect modification by physical activity or waist circumference. It is possible we did not see evidence of effect modification due to the lack of association between $PM_{2.5}$ and NO_2 and incident diabetes in our main adjusted model. However, this does not explain the lack of evidence of effect modification as there was no evidence when adjusting for ethnic heritage instead of study center/ethnic heritage, or when only adjusting for the additional covariates of age and sex. There is a lack of consistent evidence on the interactions between long-term exposure to air pollution and physical activity and cardiometabolic outcomes.[44,114,131] More studies are needed to further investigate the relationship between air pollution exposure, physical activity, and diabetes risk. There is a particular need to investigate this relationship and account for the long-term air pollution exposure where the physical activity is occurring with a fine-scale spatial air pollution model.

Adjustment for study center is needed in analysis of multi-center studies to account for potential unmeasured confounding related to the different study centers. Characteristics of both the study center and the study center population can affect the association of interest. Variations in climate of each study center can affect health, health care access, diet, and physical activity

levels. Diets can also vary by region based on food availability and cultural preferences. There can be various noise and environmental stressors present in some study centers but not others. In relation to health care, health insurance coverage and quality of care can vary by study center. In HCHS/SOL, at visit 1 79.2% of study participants in the Bronx had health insurance coverage while only 29.3% of those in Miami had coverage. Health insurance coverage in Chicago and San Diego were 44.5% and 48.1%, respectively. Health insurance coverage increased substantially in each study center by visit 2, likely due to the enactment of the Affordable Care Act. Coverage increased to 90.9% in the Bronx, 63.0% in Chicago, 67.5% in Miami, and 77.9% in San Diego by visit 2. There can also be variations in the recruitment and administration of the study at the various study centers. While HCHS/SOL followed standard protocol, there were slight differences in recruitment strategies at each study center. It is also possible protocol for diabetes ascertainment varied slightly by study center (i.e. timing of blood draws). All the previously mentioned factors can affect the association between air pollution and diabetes incidence, necessitating the adjustment by study center in epidemiologic models.

In multicenter epidemiology studies, there has been discussion on how to best adjust for study center.[132] Researcher can adjust for study center using fixed effects, random effects, or analyzing each study center separately and then combine results into a meta-analysis.[132] The three approaches could provide similar or different conclusions depending on the study sample size, and the difference between the within- and between-center effects.[132] In the presented analysis, we adjusted for fixed effects and analyzed each study center separately. We adjusted for study center/ethnic heritage using fixed effects due to the large sample size within each study center. However, the low within-city variability may reduce the statistical power in the fixed effects model because the between-center variation is absorbed by the center-specific intercepts.

As between-center variability is larger than the within-center variability, using random effects adjustment in this analysis may be helpful as this approach utilizes the between-center exposure variations. A random effects model leverages both the between-center and within-center effects, whereas the fixed effects model only estimates the within-center effects.[132] However, random effect models can also lead to bias if unmeasured center-level confounders are correlated with the exposure and outcome.[132] Random effects do not completely control for confounding by study center as the fixed effect method does. Future analysis can investigate the use of random effects when controlling for study center/ethnic heritage.

While follow-up time was fairly standard among study participants (about 6 years), we investigated if having more or less than the median follow-up time affected exposure levels. Having longer than the median years of follow-up (> 5.86 years) resulted in significantly higher exposure to both $PM_{2.5}$ and NO_2. Differences in exposure based on length of follow-up could be related to changes in secular air pollution trends, however, this is the opposite of what we would expect as secular trends show a decrease in air pollution. The variation in exposure with follow-up time could be related to the study center as some centers were more likely to have the visit 2 clinic dates further from the visit 1 clinic dates than others. For example, in Chicago 61.4% of participants had longer than median follow-up time while in Miami only 28% of participants had a longer than median follow-up time. As a sensitivity analysis for secular trends in air pollution, we standardized the NO_2 exposure to the average between July 2011 and July 2014, a 3-year period after all visit 1 clinic dates and before any participant had their visit 2 clinic date. This sensitivity analysis did not change the overall interpretation of our main results, but did considerably increase the point estimate and confidence interval in our main adjusted model compared to our results using the average exposure over the participant's full follow-up period.

Our study has notable strengths including the use of fine-scale spatiotemporal air pollution models. These models, which incorporated agency monitoring and short-term community sampling campaigns, provide point estimates in the study regions on a 2-week time scale over the follow-up period. This, combined with yearly updates to residential addresses, allowed estimates of the average exposure at study participant's residences over their follow-up period. The large, multi-center cohort of the HCHS/SOL allowed us to investigate the relationship between $PM_{2.5}$ and NO_2 and incident diabetes in self-identified Hispanic/Latinos in four major US cities. HCHS/SOL collected data on diabetes status via self-reported diagnosis and antihyperglycemic medication usage, and via blood draws for FPG, HbA1C, and a 2-hour OGTT. The comprehensives of the HCHS/SOL dataset allowed us to investigate two definitions of diabetes and include relevant individual-level confounders in our analysis.

Nevertheless, our study was not without limitations. As noted previously, the spatial residential clustering of HCHS/SOL study participants limited the within-city variability of $PM_{2.5}$ and NO_2. We conducted within city analysis on the association between exposure and incident diabetes, but the reduced sample size at the city level may have limited our power to see significant effects. Future studies could investigate the associations with random effects for study center, increase within study center sample size, and recruit study participants from more spatially varying census tracts to limit spatial clustering. The average length of follow-up was 6 years, which may not be long enough to see effects as the biological development of T2D can take several years. We were unable to expand exposure estimates to capture exposure prior to visit 1 as we do not have residential address data prior to study enrollment. HCHS/SOL study participants underwent a third in-person visit between 2020 and 2024 and future analysis on this cohort can be expanded to include visit 3 once all study data becomes available. While we were

limited in our ability to distinguish between type 1 and type 2 diabetes in this cohort, the greater majority of diabetes cases are likely type 2 given the study participants' age. In addition to NO_2, SOLAir also collected NO_x and O_3 data during the Ogawa campaigns in Chicago, Miami, and San Diego. As of May 2025, center-specific spatiotemporal NO_x and O_3 models have not been developed. Future work can develop these models and investigate the associations between NO_x and O_3 and incident diabetes. Lastly, we estimated average ambient exposure at participant's residential addresses, a location where participants do not spend all of their time. Study participants can be exposed to NO_2 and $PM_{2.5}$ at their place of work, during transportation, and in their larger community environment. Exposures at non-residential locations were unaccounted for and may affect T2D risk.

In conclusion, our results did not support our hypothesis of an association between $PM_{2.5}$ and NO_2 and incident diabetes in the HCHS/SOL when adjusting for relevant confounders. Our results only supported our hypothesis in our main adjusted model in the Bronx, which had the highest levels of exposure and highest incident rates of diabetes. Future studies should continue to investigate the association between $PM_{2.5}$ and NO_2 and incident diabetes in a less spatially clustered within-city cohort with a longer follow-up period, and continue to investigate the potential modifying effects of physical activity and adiposity.

Supplemental Material

Supplemental Table 1. Effect modification in the association between $PM_{2.5}$ and incident diabetes by waist circumference, physical activity measured by the GPAQ, or physical activity measured by the Actical accelerometer. IRR are reported per an IQR increase of 2.48 $\mu g/m^3$ for primary diabetes and 2.51 $\mu g/m^3$ for secondary diabetes.

		Primary diabetes (Lab results and/or self-reported medication)		Secondary diabetes (Primary + self-reported diagnosis)	
	Parameter	**IRR (95% CI)**	**P-value**	**IRR (95% CI)**	**P-value**
Waist Circumference[a]					
	$PM_{2.5}$ ($\mu g/m^3$)	0.87 (0.51 – 1.50)	0.62	1.10 (0.71 – 1.69)	0.67
	Abdominal obesity (cm)	2.59 (0.66 – 10.2)	0.17	1.41 (0.41 – 4.92)	0.59
	$PM_{2.5}$ x Abdominal obesity	1.06 (0.74 – 1.51)	0.76	1.13 (0.82 – 1.57)	0.46
GPAQ Physical Activity[b]					
	$PM_{2.5}$ ($\mu g/m^3$)	0.88 (0.52 – 1.49)	0.64	1.16 (0.77 – 1.75)	0.49
	GPAQ Physical Activity ≥ 150 minutes per week	0.95 (0.21 – 4.25)	0.95	0.88 (0.25 – 3.14)	0.85
	$PM_{2.5}$ x GPAQ Physical Activity ≥ 150 minutes per week	1.01 (0.68 – 1.49)	0.96	1.03 (0.74 – 1.44)	0.87
Actical Physical Activity[b]					
	$PM_{2.5}$ ($\mu g/m^3$)	0.93 (0.54 – 1.62)	0.81	1.22 (0.81 – 1.86)	0.34
	Actical Physical Activity ≥ 150 minutes per week	0.20 (0.03 – 1.38)	0.10	0.36 (0.08 – 1.59)	0.18
	$PM_{2.5}$ X Actical Physical Activity ≥ 150 minutes per week	1.37 (0.84 – 2.25)	0.20	1.21 (0.83 – 1.78)	0.32

Note: Abdominal obesity was defined as greater than 88 cm for women and greater than 102 cm for men.[126]

Abbreviations: IRR, incident rate ratio

[a] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ ethnic heritage

[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and study center/ ethnic heritage

Supplemental Table 2. Association between average concentration of $PM_{2.5}$ and change in continuous HbA1C (%) over the follow-up period among those free of diabetes at visit 1. Mean change reported per 1.22 $\mu g/m^3$ increase in $PM_{2.5}$.

Model	N	Mean Change[+] (%)	95% CI	P-value
Model 1	8080	0.01	-0.01 – 0.02	0.43
Model 2	7907	0.01	-0.01 – 0.02	0.65
Model 3	7293	0.01	-0.03 – 0.05	0.68
Model 4	7264	0.01	-0.04 – 0.05	0.75

Model 1: adjusts for age, sex, time between visits, and baseline HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
[+] per 1.22 $\mu g/m^3$ increase in $PM_{2.5}$

Supplemental Table 3. Association between average concentration of $PM_{2.5}$ and change in continuous HbA1C (%) over the follow-up period among those with diabetes at visit 1. Mean change reported per 1.22 $\mu g/m^3$ increase in $PM_{2.5}$.

Model	N	Mean Change[+] (%)	95% CI	P-value
Model 1	2155	0.08	-0.01 – 0.17	0.08
Model 2	2103	0.06	-0.03 – 0.14	0.20
Model 3	1881	-0.01	-0.35 – 0.34	0.98
Model 4	1869	-0.02	-0.37 – 0.33	0.91

Model 1: adjusts for age, sex, time between visits, and baseline HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at baseline + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
[+] per 1.22 $\mu g/m^3$ increase in $PM_{2.5}$

Supplemental Table 4. Effect modification in the association between NO_2 and incident diabetes by waist circumference, physical activity measured by the GPAQ, or physical activity measured by the Actical accelerometer. IRR are reported per an IQR increase of 8.69 ppb for primary diabetes and 8.73 ppb for secondary diabetes.

		Primary diabetes (Lab results and/or self-reported medication)		Secondary diabetes (Primary + self-reported diagnosis)	
	Parameter	**IRR (95% CI)**	**P-value**	**IRR (95% CI)**	**P-value**
Waist Circumference[a]					
	NO_2 (ppb)	1.00 (0.63 – 1.59)	0.99	1.05 (0.69 – 1.58)	0.83
	Abdominal obesity (cm)	2.62 (1.53 – 4.50)	<0.01	1.65 (1.06 – 2.57)	0.03
	NO_2 x Abdominal obesity	1.14 (0.85 – 1.52)	0.39	1.22 (0.95 – 1.56)	0.11
GPAQ Physical Activity[b]					
	NO_2 (ppb)	1.12 (0.71 – 1.76)	0.63	1.19 (0.79 – 1.80)	0.40
	GPAQ Physical Activity ≥ 150 minutes per week	0.99 (0.59 – 1.66)	0.97	0.94 (0.62 – 1.43)	0.79
	NO_2 x GPAQ Physical Activity ≥ 150 minutes per week	1.00 (0.75 – 1.34)	0.98	1.03 (0.81 – 1.30)	0.82
Actical Physical Activity[b]					
	NO_2 (ppb)	1.32 (0.79 – 2.20)	0.29	1.44 (0.90 - 2.28)	0.12

	Actical Physical Activity ≥ 150 minutes per week	0.62 (0.35 – 1.08)	0.09	0.67 (0.43 – 1.03)	0.07
	NO_2 x Actical Physical Activity ≥ 150 minutes per week	1.07 (0.80 – 1.46)	0.62	1.08 (0.85 – 1.36)	0.53

Note: Abdominal obesity was defined as greater than 88 cm for women and greater than 102 cm for men.[126]

Abbreviations: IRR, incident rate ratio

[a] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ ethnic heritage

[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and study center/ ethnic heritage

Supplemental Table 5. Exposure characteristics among those with the median years of follow-up or less (≤ 5.86 years) and those with more than the median years of follow-up.

	<= 5.856 years of follow-up (n = 4095)	**> 5.856 years of follow-up (n= 4155)**	**Two sample t-test p-value**
Diabetes at Visit 2 (n)	476	474	
Free of Diabetes at Visit 2 (n)	3619	3681	
$PM_{2.5}$			
Range	6.4 – 11.1	6.4 – 11.6	
Median	9.5	9.9	
Mean (sd)	8.9 (1.4)	9.6 (1.2)	<0.01
IQR	2.8	1.1	
NO_2			
Range	0.6 – 26.5	0.5 – 27.9	
Median	10.9	13.6	
Mean (sd)	12.8 (6.0)	13.6 (5.3)	<0.01
IQR	9.1	7.7	

Supplemental Table 6. Minimally and fully adjusted incidence rate ratios (IRR) of diabetes (primary and secondary definitions) at visit 2 based on average concentration NO_2 (ppb) at participants' residential address between July 2011 and July 2014. IRR are reported per an IQR increase of 8.89 ppb for primary diabetes and 8.91 ppb for secondary diabetes.

		Minimally Adjusted[a]			Adjusted[b]		
	Outcome	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
NO_2 (ppb)							
	Primary diabetes (Lab results and/or self-reported medication)	8775	999	1.17 (1.05 – 1.32)*	8042	922	1.26 (0.79 – 2.03)
	Secondary diabetes (Primary + self-reported diagnosis)	8649	1424	1.12 (1.02 – 1.22)*	7927	1316	1.34 (0.88 – 2.03)

Abbreviations: IRR, incident rate ratio
* $P < 0.05$
[a] Adjusted for age and sex
[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, waist circumference, and study center/ ethnic heritage

Supplemental Table 7. Association between average concentration of $PM_{2.5}$ and change in continuous HbA1C (%) over the follow-up period among those free of diabetes at visit 1. Mean change reported per 8.70 ppb increase in NO_2.

Model	N	Mean Change$^+$ (%)	95% CI	P-value
Model 1	8087	0.01	-0.01 – 0.04	0.39
Model 2	7963	0.01	-0.02 – 0.04	0.49
Model 3	7341	-0.02	-0.10 – 0.07	0.68
Model 4	7262	-0.02	-0.11 – 0.06	0.59

Model 1: adjusts for age, sex, time between visits, and baseline HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
$^+$ per 8.70 ppb increase in NO_2

Supplemental Table 8. Association between average concentration of $PM_{2.5}$ and change in continuous HbA1C (%) over the follow-up period among those with diabetes at visit 1. Mean change reported per 8.79 ppb increase in NO_2.

Model	N	Mean Change$^+$ (%)	95% CI	P-value
Model 1	2153	0.15	0.01 – 0.29	0.04
Model 2	2101	0.12	-0.02 – 0.25	0.10
Model 3	1880	0.29	-0.30 – 0.88	0.34

Model 4	1869	0.26	-0.32 – 0.85	0.38

Model 1: adjusts for age, sex, time between visits, and baseline HbA1c values
Model 2: Model 1 + antihyperglycemic medication usage at baseline + antihyperglycemic medication usage at visit 2
Model 3: Model 2 + education + income + study center/ethnic heritage + years in the US + neighborhood SES
Model 4: Model 3 + family history + waist circumference
[+] per 8.79 ppb increase in NO_2

Chapter 4: Neighborhood Physical Disinvestment and Incident Diabetes between visits 1 and 2 of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL)

A version of this chapter is under review at SSM-Population Health as a manuscript for publication. The following are co-authors on the manuscript:

Cara M Smith, MPH[a*]; Elizabeth W Spalt, MS[a]; Linda C. Gallo, PhD[b]; Jordan A. Carlson, PhD[c]; Matthew Allison, MD, MPH[d]; Bharat Thyagarajan, PhD[e]; Earle C Chambers, PhD, MPH[f]; Amber Pirzada, MD[g]; Martha Daviglus, MD, PhD[g]; Christina Cordero, PhD[h]; Qibin Qi, PhD[i]; Elena Austin, ScD, MS[a]; Amanda M. Fretts, PhD, MPH[j]; India Ornelas, PhD, MPH[k]; Robert Kaplan, PhD[i]; Joel D Kaufman, MD, MPH[a]; Stephen J Mooney, PhD, MS[j]

[a] Department of Environmental and Occupational Health Sciences, School of Public Health, University of Washington, Seattle, Washington, USA
[b] Department of Psychology, San Diego State University, San Diego, California, USA
[c] Department of Pediatrics, Children's Mercy Hospital and University of Missouri, Kansas City, Missouri, USA
[d] Department of Family Medicine, University of California San Diego, La Jolla, California, USA
[e] Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA
[f] Department of Family and Social Medicine, Albert Einstein College of Medicine, Bronx, New York, USA
[g] Institute for Minority Health Research, University of Illinois at Chicago, Chicago, Illinois, USA
[h] Department of Psychology and Behavioral Medicine Research Center, University of Miami, Miami, Florida, USA
[i] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York, USA
[j] Department of Epidemiology, School of Public Health, University of Washington, Seattle Washington, USA
[k] Department of Health Systems and Population Health, School of Public Health, University of Washington, Seattle Washington, USA

Abstract

Neighborhood-scale environmental factors, including disinvestment in infrastructure, may impact cardiometabolic disease risk. To our knowledge, no prior studies have investigated the association between neighborhood disinvestment and incident diabetes. We used a virtual street audit of Google Street View Imagery to generate a neighborhood disinvestment score for participants' residential addresses in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). An item response theory model was fit to indicators (litter, graffiti, under-maintained buildings, bars on windows, and abandoned buildings) to form a scale measuring a latent level of disinvestment. Ordinary kriging was used to estimate levels for each residential address within the HCHS/SOL census tracts via spatial interpolation. HCHS/SOL is a longitudinal cohort study of self-identified Hispanic/Latino adults in the Bronx, Chicago, Miami, and San Diego. Using covariate-adjusted and survey-weighted Poisson regression models with data from 9120 participants free of diabetes at baseline (2008-2011), we investigated the association between neighborhood disinvestment and incident diabetes at visit 2 (2014-2017). A sensitivity analysis included only those who did not move during the follow-up period. A one-standard deviation increase in neighborhood disinvestment score was associated with a 13% (95% CI: 1%-23%) lower risk of incident diabetes when adjusting for age, sex, education, income, study center/heritage, years in the US, family history, and a neighborhood socioeconomic index. Our sensitivity analysis yielded qualitatively similar results with lower precision. Overall, our analysis does not support the hypothesis that neighborhood physical disinvestment is associated with incident type 2 diabetes in this Hispanic/Latino population.

Introduction

In the United States, around 1.1 million people aged 18 years and older were diagnosed with type 2 diabetes (T2D) in 2021.[25] Hispanic/Latinos are at higher risk compared to non-Hispanic Whites,[25] and in 2022 were found to be 60% more likely to be diagnosed with diabetes than non-Hispanic Whites.[105] Well-established T2D individual risk factors include genetics, age, adiposity, diet, and physical activity.[1–3] Modifiable environmental and neighborhood conditions may also contribute to T2D risk,[1–3] and addressing neighborhood level risk factors in diabetes prevention interventions may increase effectiveness.[133]

The neighborhood built environment is a research area where there are various measures to investigate and various methods and approaches to quantify said measures, however, many studies have focused on cross-sectional analysis of disease prevalence. Few studies have specifically focused on neighborhood built environment and T2D incidence.[40,71]

As diabetes prevalence is higher in low-income urban areas compared to high-income urban areas,[70,133,134] aspects of the built environment in those areas may contribute to increased risk for diabetes. Previous studies on the built environment have investigated how access to healthy food options[4,7,135] and green space[4,135,136] are associated with lower T2D risk. In addition to healthy food options and green space, neighborhood disinvestment, a social and economic process in which resources are redirected away from neighborhoods of lower-income populations, may also contribute to T2D risk.[133] Disinvestment is reflected in visual indicators of deterioration and neglect such as deteriorated sidewalks, lack of street lighting, litter, and abandoned buildings. Living within disinvested neighborhoods may increase stress and reduce

residents' ability and desire to engage in physical activity, which may in turn increase diabetes risk.[21–23,72–74]

However, understanding the spatial impacts of the sociopolitical process of being in a disinvested place can be challenging. While theory suggests sociopolitical disinvestment processes degrade diabetes-preventing mechanisms such as economic access to healthy foods and green space quality, these specific features can be challenging to measure directly. One alternate approach is to assess visual indicators of neighborhood disinvestment – features like litter and graffiti – as visible on Google Street View, and combine these indicators to create a disinvestment score.[8,81,137–140] Utilizing Google Street View increases the spatial availability of visual indicators and allows for comparison across space.

While prior work has not directly investigated the association between neighborhood physical disinvestment and T2D incidence, perceived neighborhood disorder – a related construct promoted by disinvestment and reflected in vandalism, trash, vacant buildings, and reports of feeling unsafe walking alone – was associated with T2D prevalence in a non-Hispanic White sample of the Health and Retirement Study cohort (OR: 1.11, 95% CI: 1.05, 1.17).[9] When investigating the interaction between neighborhood disorder and genome-wide genetic risk in the previously mentioned non-Hispanic White sample of the Health and Retirement Study cohort, neighborhood disorder was found to heighten the genetic risk of T2D (OR: 1.10, 95% CI: 1.04, 1.16). [9] However, a Jamaican cross-sectional survey found no association between neighborhood disorder and diabetes prevalence (OR: 0.99, 95% CI: 0.95, 1.03).[68]

These previous studies on neighborhood disorder and T2D have collected data from study participants via questionnaires[9,75] or in-person assessment by interviewers.[68] Collecting self-reported exposure and outcome data can lead to 'same source bias' as participant's experience

with the outcome may influence how they report their exposure. Those experiencing adverse outcomes may be more likely to notice and report potential exposures due to their beliefs and biases, resulting in spurious associations. Approaches to collecting neighborhood environment data to get around the issue of 'same source bias' include asking non-participants living in the same area their thoughts and opinions,[66] collecting data from publicly available databases,[69,70,80] or in-person street audits[22] and assessment by researchers.[68] A less costly and less time-consuming alternative to in-person street audits are virtual street audits.[8,76,81,82]

To investigate the relationship between an objective visual measure of the neighborhood built environment and T2D incidence, a virtual audit was conducted using Google Street View (Alphabet, Mountain View, CA) in the neighborhoods of HCHS/SOL study participants. The virtual audit was used to create a neighborhood physical disinvestment score at each participant's residential address. We used these neighborhood disinvestment scores to investigate the association between neighborhood disinvestment and incident diabetes between the first and second visits of the HCHS/SOL.

Methods

Study Design and Population

HCHS/SOL is a longitudinal cohort of 16,415 self-identifying Hispanic/Latino adults at least 18 years of age at enrollment from the Bronx, Chicago, Miami, and San Diego.[24] More details on the HCHS/SOL sample and study design have been previously reported.[116,117] Briefly, the first in-person visit and enrollment occurred between 2008 and 2011 in each of the four cities. Yearly follow-up phone calls were conducted, and a second round of in-person visits (visit 2) occurred between 2014 and 2017.

Questionnaires and clinical examinations were administered at both in-person visits. The questionnaires assessed demographics, current health and medical history, socioeconomic status, and residential information.[117] Clinical examinations included collection of blood and urine for various assays and analysis.[117] The yearly follow-up phone calls contained questions on general health, potential updates from their doctors or health professionals, and details on any potential hospitalization or emergency room visits that occurred since the last follow-up as well as any updates to their residential location.[117] The study population for this analysis included participants who completed the baseline and 2nd visit, were free of diabetes at the baseline visit, and had geocoded residential address data at both visits. All participating institutions received approval by their respective institutional review boards and written informed consent was received from all participants.

Diabetes Ascertainment

Ascertainment of diabetes occurred at baseline, during the yearly follow-up phone calls, and at visit 2. During calls and visits, participants were asked if a doctor or health care professional has told them they have diabetes or high sugar in their blood and if they were receiving any treatment.[117] However, only during the in-person visits was blood collected and assayed to measure fasting plasma glucose (FPG) and glycosylated hemoglobin (HbA1c) levels.[117] A 2-hour oral glucose tolerance test (OGTT) was also conducted on participants who did not self-report a diabetes diagnosis and on participants with FPG < 150 mg/dL.[107] Ascertainment of diabetes in the HCHS/SOL cohort was not able to distinguish between type 1 and type 2 diabetes.

For this analysis, we use two definitions of diabetes. Each definition is based on the American Diabetes Association criteria[10] and data collection practices. Our primary definition

for diabetes was having any of the following: a FPG ≥126 mg/dL; HbA1c level ≥6.5%; post-OGTT glucose ≥ 200 mg/dL; or use of antihyperglycemic drugs. Our secondary definition, while also limited to those with ascertainment of diabetes at visit 2, included all those who fit the primary definition of T2D and also includes those who self-reported diabetes or high sugar in their blood during the in-person visits or the annual follow-up phone calls.

Pre-diabetes was assessed at baseline and visit 2 as our criteria was dependent on laboratory assays. Based on the American Diabetes Association criteria, pre-diabetes was defined as having a FPG ≥100 mg/dL and <126 mg/dL; and HbA1c level ≥5.7% and <6.5%.[10]

Google Street View Audit

In order to generate an interpolated disinvestment measure for each residential address, we conducted a virtual street audit using Google Street View. For each of the four cities, we selected 1000 street locations to be virtually audited via Google Street View. Locations were selected within the 351 census tracts where participants lived at baseline using a 600-point grid. To increase precision in neighborhoods where most participants lived, we embedded an additional 600-point oversample in the census tracts with the top quintile of HCHS/SOL participants in each study center. All points were randomly jittered up to 250 meters. We used the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to conduct the virtual street audit.[87] We uploaded our set of coordinates to CANVAS which tried to match each coordinate to Google Street View imagery within 50 meters. If there were no matches within 50 meters, the location was not used. For each location, auditors answered a set of 53 questions related to neighborhood disinvestment based on the most recent image and oldest high-resolution image using the drop and spin method (Table S1). Images from 2007 and some from 2008 were of low-resolution, making it difficult to accurately answer the questions. To avoid this, we

instructed auditors to assess the oldest image of high-resolution after 2007. The drop and spin method refers to how CANVAS accessed Google Street View at each location and auditors were only able to spin around to answer the questions. They were not able to 'walk' up and down the street as this could change the date of the images.

Audit questions were taken from previously validated audits designed to assess neighborhood disinvestment and pedestrian safety.[81,141,142] Example of the questions auditors had to answer were "Is there garbage, litter, or broken glass in the street or on the sidewalk?", "Are there abandoned cars?", and "Do you see boarded up or abandoned buildings?". The full list of questions is available in supplementary material (Table S1). All responses were coded categorically and most were dichotomous. The virtual street audit was conducted by 12 undergraduate and graduate students at the University of Washington between July 2022 and May 2023. All auditors were trained on the same set of example locations. During training, all questions and examples were discussed to limit subjectivity between auditors. For example, multiple images of graffiti and murals were shown during training to highlight the differences between them. To assess agreement between the auditors, approximately 15% of the locations in each study center were randomly selected as a reliability subsample.

Neighborhood Disinvestment

To generate a neighborhood disinvestment measure for each participant's residential address, we used a validated approach involving an item response theory (IRT) model and ordinary kriging.[8,81] An IRT model was fit to indicators (litter, graffiti, under-maintained buildings, bars on windows, and abandoned buildings) to form a scale measuring a latent level of disinvestment. The IRT model's posterior probability of observing the indicators we actually observed was used to estimate a latent level of disinvestment at each audited location. We then used ordinary

kriging to estimate levels for each residential address within the HCHS/SOL census blocks via spatial interpolation. Maps showing the disinvestment levels at each address within a HCHS/SOL census block for the Bronx, Chicago, Miami, and San Diego, are shown in figures 11-14. Geocoded addresses with the disinvestment measure were then matched with participants' geocoded addresses in a secure research workspace. Our neighborhood disinvestment measure is unitless with higher values indicating more disinvestment. For this analysis, we have transformed the measure to its z-score, indicating a unit increase as a change in 1 standard deviation. More information on how the neighborhood disinvestment measure was developed is presented in the supplementary material.

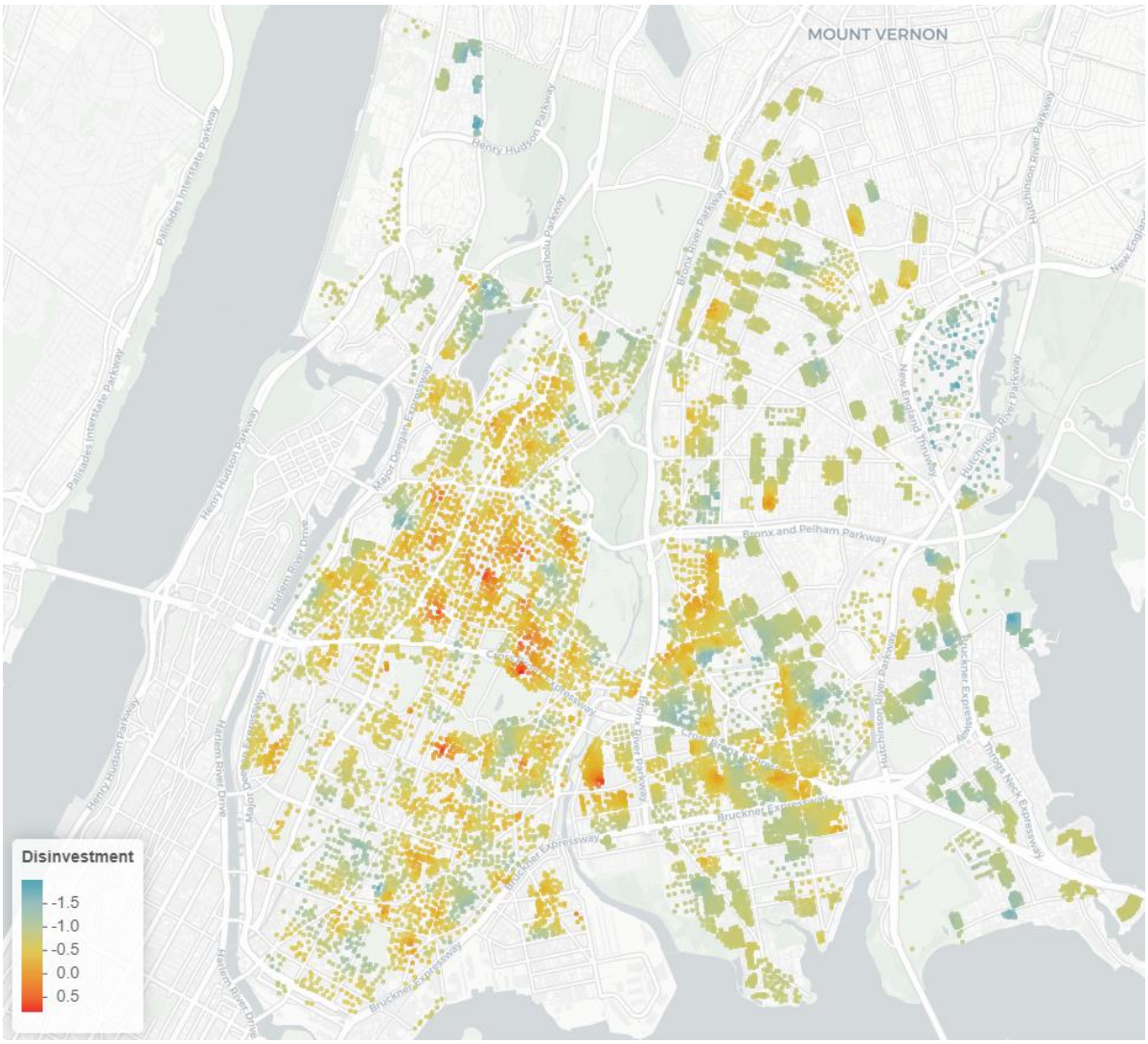


Figure 11. Neighborhood physical disinvestment levels at each building in a census block where HCHS/SOL reside in the Bronx.

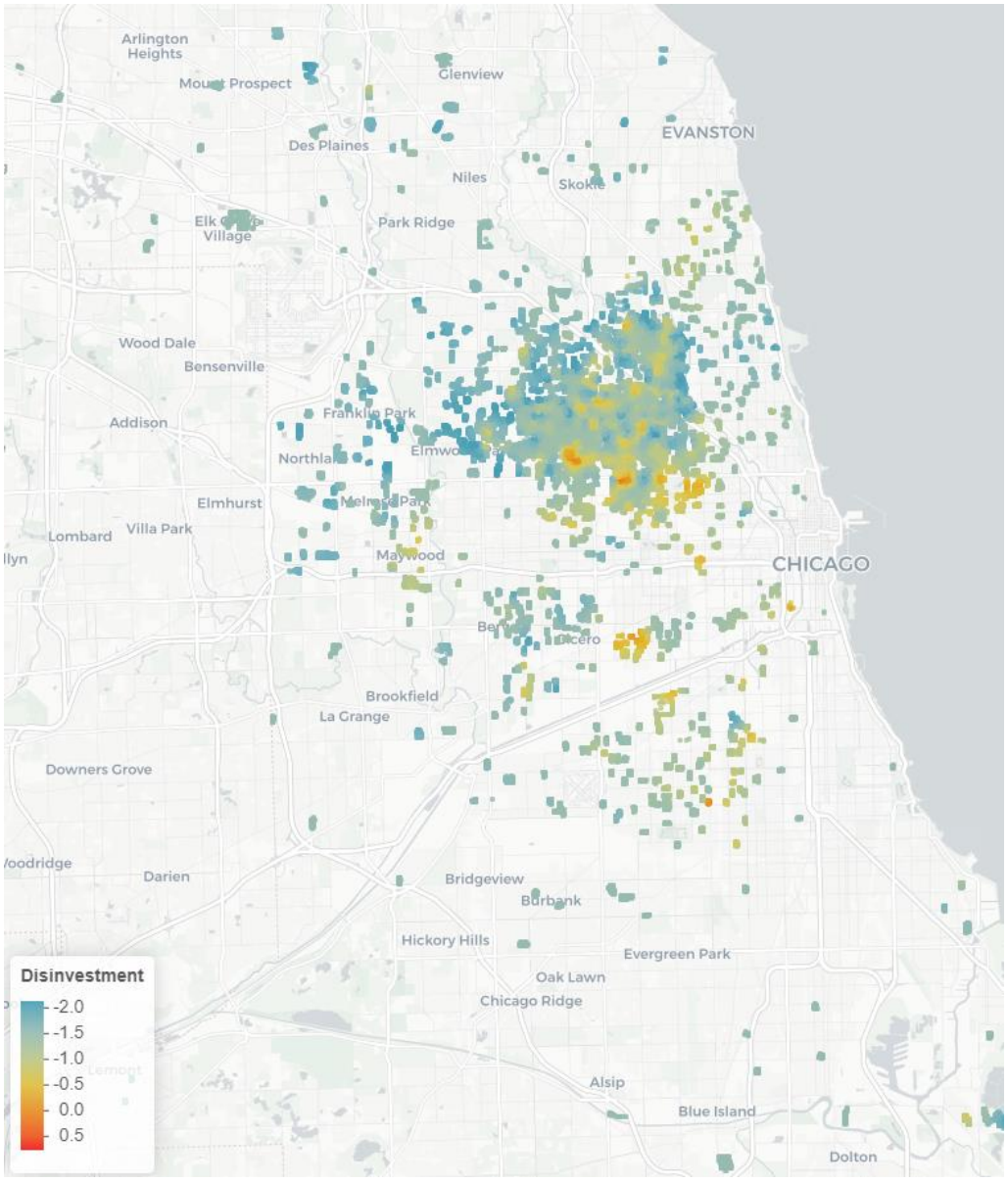


Figure 12. Neighborhood physical disinvestment levels at each building in a census block where HCHS/SOL reside in Chicago.

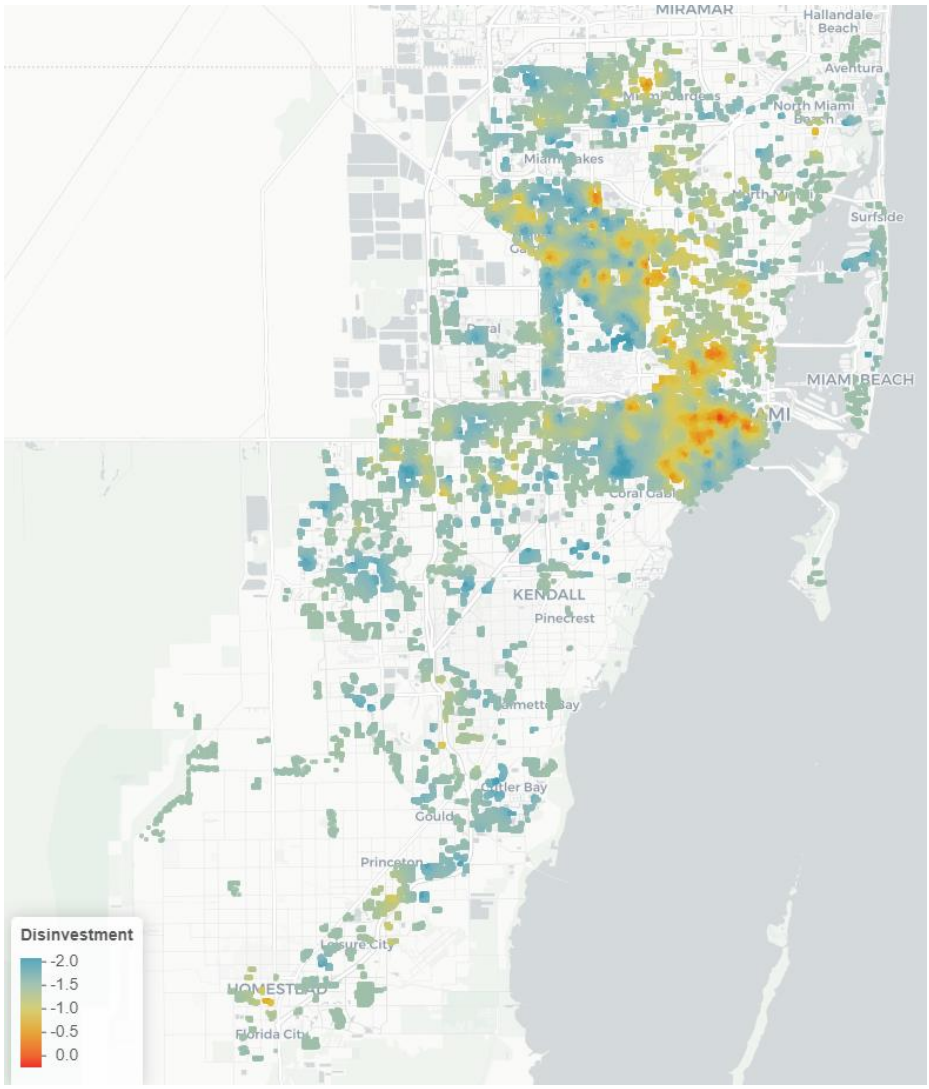


Figure 13. Neighborhood physical disinvestment levels at each building in a census block where HCHS/SOL reside in Miami.

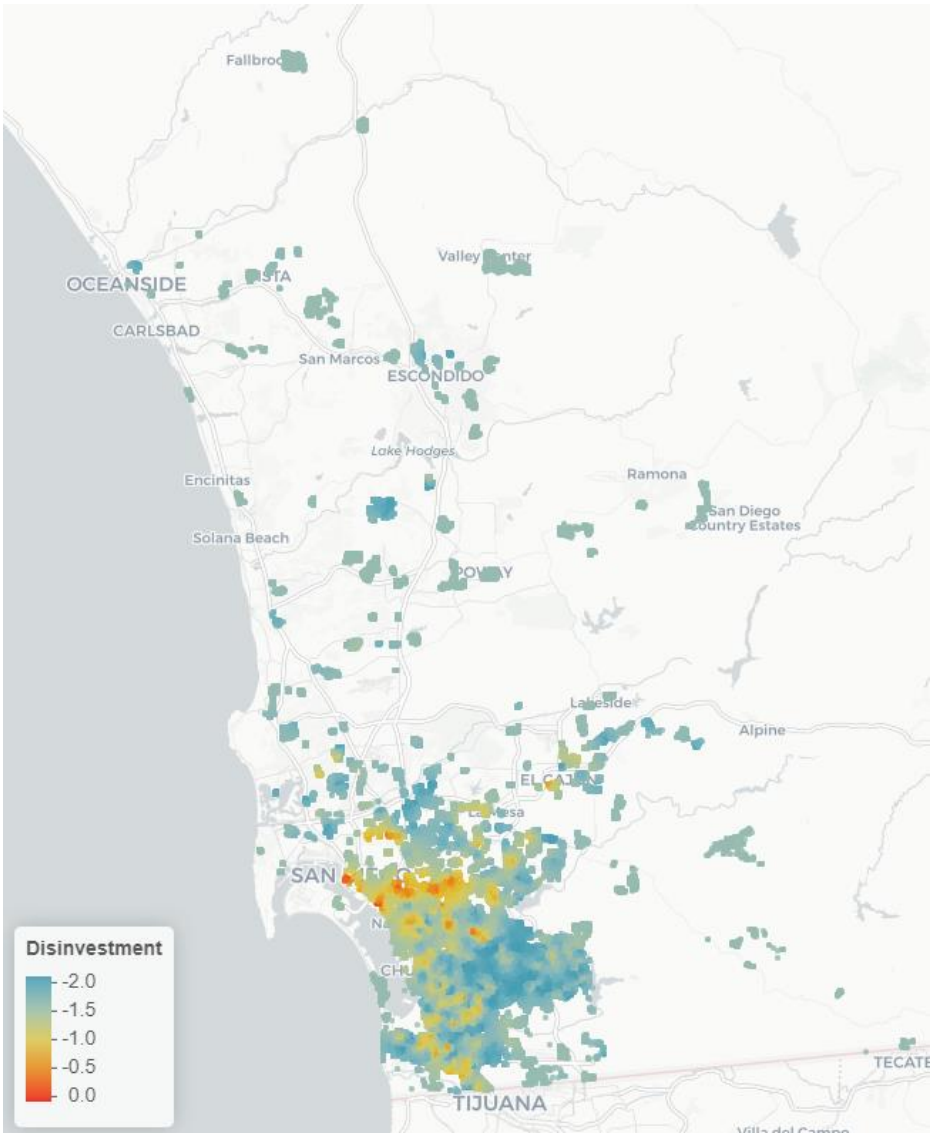


Figure 14. Neighborhood physical disinvestment levels at each building in a census block where HCHS/SOL reside in San Diego.

Inclusion Criteria

Of the 16,415 HCHC/SOL study participants, 11,619 had ascertainment of diabetes at visit 2. Of the 11,619 participants, we had a neighborhood disinvestment measure for 11,503 at their baseline residential address. For our primary analysis, participants needed to be free of diabetes at baseline based on the primary diabetes definition (N= 9120). For analysis with secondary diabetes definition, 8989 were free of diabetes at baseline.

Statistical Analysis

To compare the two definitions of diabetes, we computed demographic statistics on both samples and calculated an age-centered adjusted incidence rate. Range, median, mean, and standard deviation of neighborhood disinvestment were also calculated for each sample overall and stratified by study center. We ran Pearson's correlations on disinvestment at baseline and a neighborhood SES index,[127] population within a 1 km buffer, percent Hispanic at the census tract level. We ran Spearman's rank correlation on disinvestment and ordinal data on individual level education, income, and years living in the US. The neighborhood SES index is a compositional measure of NSES at the census tract level that represents education, employment, housing, income/wealth, occupation, and residential stability.[127] Higher index scores indicate higher NSES disadvantage.[127]

For our primary analysis, we investigated the association between neighborhood disinvestment at baseline residential addresses and incident diabetes at visit 2 using the primary diabetes definition (does not include self-reported diagnosis). We ran Poisson regression models using a log link and years between baseline and visit 2 as the offset for follow-up time to generate an incidence rate ratio (IRR) comparing groups with a one-standard deviation difference

in neighborhood disinvestment. All regression models used complete case analyses and sampling weights to account for the HCHS/SOL complex sampling design, those living in the same household, and participant non-response at visit 2 by using the *survey* package in R.[116] The sampling weights were also calibrated to the age, sex, and Hispanic/Latino heritage from the 2010 US Census.

Minimally adjusted models only adjusted for age and sex. Our fully adjusted model adjusted for age, sex, years living in the US (Born in the 50 US states, foreign born in the US for more than 10 year, or foreign born in the US for less than 10 years), a 5-level income variable (less than $10,000 as the reference), a 4-level educational attainment variable (no high school diploma or GED (reference), at most a high school diploma or GED, high school (or GED) education, and University/college education), family history of diabetes, a neighborhood SES index,[127] and a 17-level variable combining study center and Hispanic/Latino heritage group (i.e., Cuban, Dominican, Puerto Rican, Mexican, Central or South American, or other/more than one heritage; Dominicans from the Bronx as reference group). Study center-specific models were created using the same sets of covariates except replacing the 17-level combination variable with the 8-level Hispanic/Latino heritage variable (Cuban, Dominican, Puerto Rican, Mexican, Central or South American, more than one heritage, or other).

We repeated the regression models to investigate the association between neighborhood disinvestment and three different progressions of diabetes: 1) progression from pre-diabetes to diabetes, where the sample only included those with pre-diabetes at baseline, 2) Progression of those free of diabetes and pre-diabetes to diabetes and 3) progression of those free of diabetes and pre-diabetes at baseline to pre-diabetes at visit 2. Results from all models were

exponentiated to report the IRR for each standard deviation increment in neighborhood disinvestment.

Sensitivity Analyses

We repeated the Poisson regression models using the secondary definition of diabetes which included self-reported diagnosis. As 46.4% of participants moved residences during the follow-up period, we also conducted a sensitivity analysis restricted to those who did not move (N= 6164).

Missing Data

For the analytic sample, the only variable with greater than 5% missingness was income with 7.6% of participants missing data. Due to concerns about missing data and the computational intensity of multiple imputation by chained equations (MICE), we conducted MICE analysis as a sensitivity analysis on two model iterations. The MICE analysis consisted of 10 imputations and 50 iterations on the fully adjusted model for both the primary and secondary definitions of diabetes. Results from the MICE were very similar and did not change the overall interpretation of results. Since completing MICE analysis on all our models would be computationally intense and the MICE tested on our two iterations yielded very similar results, we used complete case analysis for our reported analyses.

All analyses were conducted in R Studio, with R version 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria). The 'ltm' package (version 1.2) was used for the IRT models. The 'mice' package (version 3.15.0) was used for the MICE analysis,[128] and the 'survey' package (version 4.2-1) was used for incorporating the survey design and weights into the regression models.[129]

Results

The analytic sample comprised 9,120 participants who were free of diabetes using our primary definition of diabetes (excluded self-reported diagnosis) at baseline. There were 1038 incident cases at visit 2 with an age-centered adjusted incidence rate of 15.9 per 1000 person-years. For our secondary definition of diabetes (including self-reported diagnosis), 8989 participants were free of diabetes at baseline and there were 1476 incident cases by visit 2 with an age-centered adjusted incidence rate of 23.5 per 1000 person-years. Weighted summary characteristics of participants by diabetes status at visit 2 are presented in Table 7.

Table 7. Weighted summary characteristics of participants by diabetes status at Visit 2 using primary definition of diabetes.

	Sample Weighted % or Mean (SD)	
Variables	Diabetes at Visit 2 (N= 1038)	Free of Diabetes at Visit 2 (N = 8082)
Female	47.0	50.9
Age, years	46.7 (12.8)	38.2 (14.1)
Hispanic/Latino Heritage		
Central American	4.9	7.5
Cuban	19.7	18.5
Dominican	9.4	9.6
Mexican	39.0	39.5
Puerto Rican	18.9	14.7
South American	3.3	5.5
More than one heritage	4.8	3.6
Other	0.2	1.0
Years in the US		
Less than 10 years	23.8	28.5
10 years or more	56.8	46.5
US Born	19.4	24.9
Family History of Diabetes	50.1	36.4

Education		
No High School Diploma or GED	36.2	27.9
At most a High school diploma or GED	28.2	28.8
High school (or GED) education	12.0	13.1
University/college education	24.7	30.2
Income		
Less than $10,000	15.3	13.8
$10,001-$20,000	31.4	31.6
$20,001-$40,000	34.3	33.9
$40,001-$75,000	13.1	15.1
More than $75,000	6.0	5.7
Neighborhood SES Index	1.1 (0.8)	1.1 (0.8)
Years between baseline and Visit 2	6.0 (0.8)	6.1 (0.9)
Study Center		
Bronx	30.1	28.1
Chicago	16.9	15.8
Miami	27.8	29.7
San Diego	25.3	26.4
Neighborhood Disinvestment		
Overall	-1.1 (0.5)	-1.1 (0.5)
Bronx	-0.6 (0.3)	-0.56 (0.3)
Chicago	-1.3 (0.4)	-1.2 (0.3)
Miami	-1.2 (0.4)	-1.2 (0.4)
San Diego	-1.5 (0.4)	-1.5 (0.4)

Baseline disinvestment distribution was slightly right-skewed and ranged from -2.07 to 0.81 with a median of -1.21 and a mean and standard deviation of -1.13 and 0.47, respectively. Baseline disinvestment z-score ranged from -2.00 to 4.12. Baseline disinvestment was moderately correlated with neighborhood SES (r = 0.56) and population within a 1 km buffer (r = 0. 64). Baseline disinvestment was not correlated with individual level education (ρ= -0.07), income (ρ = -0.18), years in the US (ρ = 0.04), or percentage of Hispanic/Latino population at the census tract level (r = -0.17).

For our primary analysis, utilizing the primary definition of diabetes, each one-standard deviation increase in neighborhood disinvestment was associated with a 13% decreased risk of incident diabetes (IRR= 0.87, 95% CI: 0.77 – 0.99) in the fully adjusted model (Table 8). IRR for all but two other model iterations, including stratification by study center, were not significant (Table S2 and S3). Among those free of pre-diabetes and diabetes at baseline, each one-standard deviation increase in neighborhood disinvestment was associated with a 34% decreased risk of incident diabetes (IRR= 0.66, 95% CI: 0.44 – 0.99) in the fully adjusted model (Table S2). Among those at the Miami study center free of diabetes at baseline, each one-standard deviation increase in neighborhood disinvestment was associated with a 20% decreased risk of incident diabetes (IRR= 0.80, 95% CI: 0.64 – 0.99) in the fully adjusted model (Table S3).

Table 8. Minimally and fully adjusted incidence rate ratios of diabetes (primary and secondary definitions) at visit 2 based on neighborhood disinvestment z-score at baseline.

	Minimally Adjusted[a]			Fully Adjusted[b]		
Outcome	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
Primary diabetes (Lab results and/or self-reported medication)	9120	1038	0.97 (0.89 – 1.06)	8346	956	0.87 (0.77 – 0.99)*
Secondary diabetes (Primary + self-reported diagnosis)	8989	1476	0.95 (0.88-1.02)	8230	1362	0.95 (0.86 – 1.06)
Primary diabetes (Subset of non-movers)	4727	584	1.05 (0.94 – 1.17)	4373	544	1.05 (0.90 – 1.23)

Abbreviations: IRR, incident rate ratio
* $P < 0.05$
[a] Adjusted for age and sex

[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and study center/ ethnic heritage

Results from our primary analysis were similar for our sensitivity analysis utilizing the secondary definition of diabetes (Table 8). The sensitivity analysis on a subset of participants who did not move during follow-up yielded similar non-significant results, with IRRs around 1, suggesting no association. For the fully adjusted model, utilizing the primary definition of diabetes, results were in the expected direction but not statistically significant (IRR = 1.05, 95% CI: 0.94 – 1.17) (Table 8). For the study center specific models among those who did not move, the fully adjusted model for Chicago was statistically significant with an IRR of 0.72 (95% CI: 0.56 – 0.96) (Table S4). All other study centers had a negative association between disinvestment and incidence diabetes except for the fully adjusted model for the Bronx (IRR: 1.05, 95% CI: 0.78 – 1.42) (Table S4).

Discussion

Using an image-based assessment of built environment conditions for a large urban Hispanic/Latino population, we found evidence of an association between neighborhood disinvestment and lower diabetes incidence in our primary analysis. These results were contrary to our hypothesized association and not robust to our additional analysis on diabetes progression nor our sensitivity analyses.

Ours is the first study relating an independently observed measure of visual indications of neighborhood physical disinvestment to incident diabetes. Our results are consistent with generally mixed findings of prior studies assessing the association between related aspects of the neighborhood and diabetes or HbA1c levels. For example, a previous study in Jamaica did not find an association between neighborhood disorder, assessed by an in-person audit, and diabetes

prevalence (OR: 0.99, 95% CI: 0.95, 1.03).[68] Another study in Jackson, Mississippi found a positive association between a survey-based measure of neighborhood problems (noise, heavy traffic, speeding cars, lack of access to adequate food or shopping, litter) and incident diabetes (HR: 1.24, 95% CI: 1.00, 1.54) among African Americans.[7] An analysis conducted at the San Diego site in the HCHS/SOL did not find an association between neighborhood indicators of social disorder, measured by per capita liquor stores, crime rates, vacant households, and vacant land, and change in HbA1c levels between baseline and visit 2.[70]

However, studies exploring neighborhood deprivation and disadvantage have more consistently found indications that disadvantage is associated with diabetes. The previously mentioned analysis at the San Diego site of the HCHS/SOL identified an association between higher neighborhood socioeconomic deprivation, a composite of educational attainment, income level, unemployment, crowded households, female-headed households with children, public health insurance, and households receiving public assistance, and worsening diabetes status at visit 2 (OR = 1.27, 95% CI = 1.10, 1.46).[70] More broadly, a harmonized analysis of three independent studies found increasing neighborhood disadvantage, measured by census tract data on income, education, unemployment, poverty, households on public assistance, households with no cars was generally associated with increased risk of T2D.[143]

There are several potential explanations for the overall lack of association between neighborhood disinvestment and incident diabetes in our analysis. For one, our neighborhood disinvestment measure had low variability overall (IQR = 0.65) and within each study center (each IQR < 0.60), which may have limited our ability to detect an association in an epidemiologic study. Future studies utilizing this method would benefit from greater variability in neighborhood conditions.

As we hypothesized neighborhood disinvestment would affect T2D risk via increasing stress and reducing residents' ability and desirability to engage in physical activity,[21–23,72–74] another potential explanation for our overall lack of association may be due to unmeasured confounding related to stressors and barriers to physical activity. Physical activity is a well-known protective factor for the development of T2D and antecedents of T2D including obesity. Individuals living in more disinvested neighborhoods may engage in less physical activity because they feel less safe and may find physical activity to be hazardous.[73,74] Individuals living in more disinvested neighborhoods may also have less access to resources that support physical activity. In the Multi-Ethnic Study of Atherosclerosis (MESA), long-term exposure to residential environments with greater resources to support physical activity was found to be associated with incident T2D.[144] Given this postulated mechanism, we did not adjust for physical activity as we believe physical activity mediates the disinvestment-T2D relationship. The relationship between disinvestment, physical activity, and T2D should be further investigated in future studies. Future studies should account for where physical activity occurs and how one's neighborhood environment may influence physical activity choices.

Disinvestment could also be related to other neighborhood factors such as the availability of healthy foods or support for active transportation; factors which we did not capture with Google Street View images. Given the positive correlation between population density and neighborhood disinvestment, dense neighborhoods may have more walkable destinations, and therefore residents may be more likely to use physically active forms of transportation.[145] Future studies should consider how the food environment, access to recreational spaces, and neighborhood walkability interact with neighborhood disinvestment to influence T2D risk.

Another potential explanation for our results may be related to the mixed effects of living in ethnic enclaves on health. Our measure could potentially reflect the beneficial effects of living in ethnically dense communities. Hispanic/Latino immigrants often settle in neighborhoods that have been established by other Hispanic/Latino residents due to their lower housing costs, high number of native Spanish speakers, and perceived safety from law enforcement arising from 'blending in' more.[99–102] Immigrants may find living in these enclaves to be protective and less stressful.[146] Living in neighborhoods with high immigrant populations may protect against obesity based on previous cross-sectional analyses, however, may be confounded by residential self-selection.[147] Li et al.,[147] found neighborhood co-ethnic density of Hispanic/Latinos to have an adverse effect on blood pressure and cholesterol level. Conversely, the researchers found the relationship between neighborhood immigrant concentration and cardiovascular health to be protective but strongly subjected to residential self-selection.[147] An analysis on residential segregation and incidence of metabolic syndrome in the HCHS/SOL cohort yielded mixed results on health effects of those living in areas with a high concentration of Hispanic/Latino residents.[148] Health effects of ethnic enclaves, which tend to have more residential disadvantages than neighborhoods with high non-Hispanic white populations,[102] may be dependent on nativity and how long one has been in the US. While we did adjust for years in the US, interrogating this relationship in HCHS/SOL is challenging due to the clustering of participants with similar backgrounds in small neighborhoods. Future studies need to investigate the structural relationship between neighborhood disinvestment, ethnic enclaves, and T2D in more detail.

Our study has notable strengths, including being the first study to investigate the relationship between visual indications of neighborhood disinvestment and incident type 2 diabetes. The HCHS/SOL allowed us to investigate this relationship in a large, diverse

population-based cohort of Hispanics/Latinos, an underrepresented group in research, in four major US cities. The comprehensiveness of the HCHS/SOL dataset allowed us to adjust for important individual-level confounders. Conducting a validated virtual audit system using Google Street View allowed us to generate a disinvestment measure in a quick and cost-effective manner compared to an in-person street audit.[8]

Our analysis was not without limitations. First, the baseline occurred between 2008 and 2011; however, the images used to derive the disinvestment variable were taken between July 2007 and April 2023 with most images from 2022, creating a temporal incompatibility. However, based on previous analysis there was little variation between the oldest and most recent images (citation redacted for blinding). Using images from this audit that do not temporally match our time of interest (baseline) is unlikely to bias our results. Second, the average length of follow-up was 6 years, which may be too short to see any effects as T2D can take several years to develop. Additionally, the mean age of participants free of diabetes at Visit 2, 38.2 years old, is younger than the mean age of diabetes diagnosis of Mexican-Americans, 44.9, from 2011-2018.[149] We are unable to expand analysis to include years before the baseline visit as we do not have residential data prior to their visit. Third, we are unable to distinguish between type 1 and type 2 diabetes in our ascertainment. However, given the age of the HCHS/SOL cohort, the greater majority of diabetes cases are likely type 2. Lastly, our neighborhood disinvestment measure was only focused on their residential neighborhood, an environment where participants do not spend all of their time. The neighborhood environment outside their residential area, such as their place of work and larger community environment, was unaccounted for and may affect T2D risk as well.

In conclusion, our results did not support our hypothesis of an association between visual indications of neighborhood disinvestment and incident diabetes in the HCHS/SOL cohort. Future studies should continue to evaluate the role of food environment, greenspace, and physical activity in considering neighborhood scale infrastructure impacts on diabetes risk.

Supplementary Material

Appendix on Disinvestment Measure

Overview

We developed our measure of visual indications of neighborhood disinvestment following the procedure developed by Raudenbush & Sampson,[142] adapted for Google Street View audits by Mooney et al.,[81] and further refined by Plascak et al.[140] Roughly, this process has 5 steps:

1) Conduct 'drop-and-spin' virtual audits assessing visual indicators of neighborhood disinvestment including graffiti, unrepaired buildings, etc., drawing from the Project for Human Development in Chicago Neighborhoods and Irvine-Minnesota systematic social observation inventories on a spatially dense sample of Google Street View images.
2) Assess inter-rater reliability of individual items, potentially dropping observations from auditors whose responses appear unreliable and/or audit items for which even modest reliability (Kappa > 0.2) could not be established
3) Fit an Item Response Theory model to remaining items, assessing model fit statistics, item discrimination, and range of assessed severity to select a final model. For each audited location, assign the modal posterior latent score as the visual indicators of disinvestment score for that location.

4) Then use the disinvestment scores in audited locations in an ordinary kriging model via spatial interpolation to estimate the visual indicators of disinvestment present at the centroid of each land parcel in each HCHS/SOL city.
5) In the HCHS/SOL secure research workspace, spatially link the disinvestment score to the land parcel containing the study participant's home address.

Step 1. Audit protocol and candidate indicators

The following items were used in virtual audits. Each item was assessed on the oldest and most recent image.

Question posed to auditors	Responses indicating disinvestment	Responses not indicating disinvestment
Is there vacant or undeveloped land?	Yes	No
Do you see any windows with bars on them?	Yes	No; Not applicable
Is there garbage, litter, or broken glass in the street or on the sidewalks?	Yes	No
Are there empty beer or liquor bottles visible in streets, yards, or alleys?	Yes	No
Are there abandoned cars?	Yes	No
Is there graffiti, or evidence of graffiti that has been painted over, on buildings, signs, or walls?	Yes	No
How many buildings, driveways, or other private infrastructure need non-trivial repairs?	One; Two or more	None; Not applicable
Do you see burned out buildings?	Yes	No
Do you see boarded up or abandoned buildings?	Yes	No

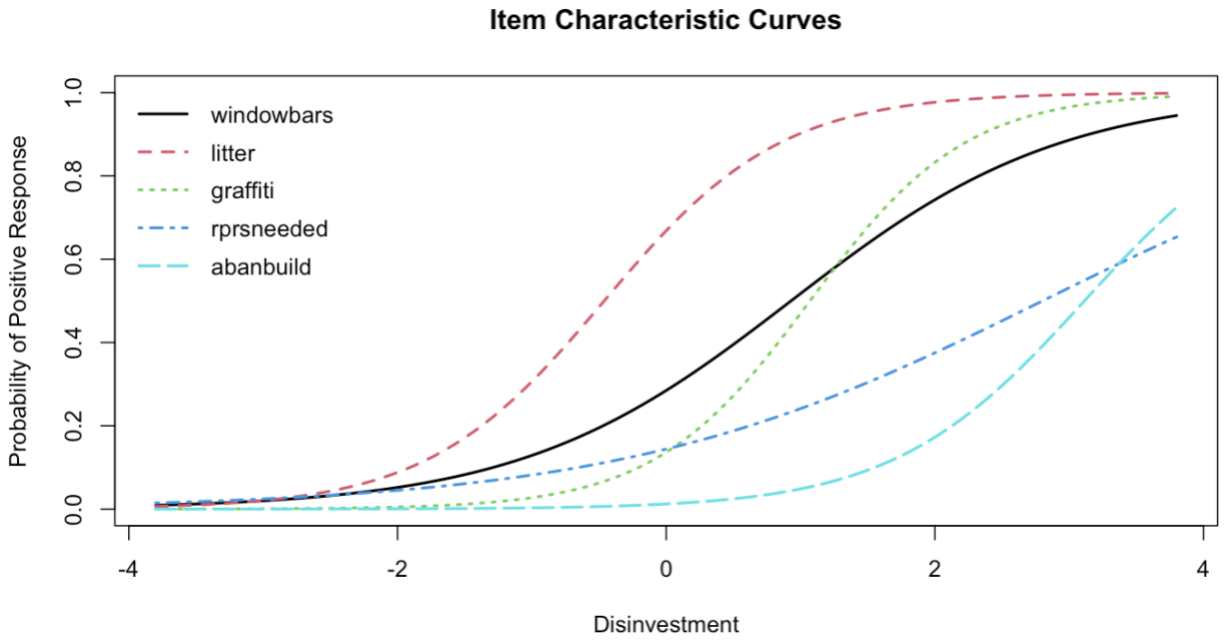
Step 2. Inter-rater reliability

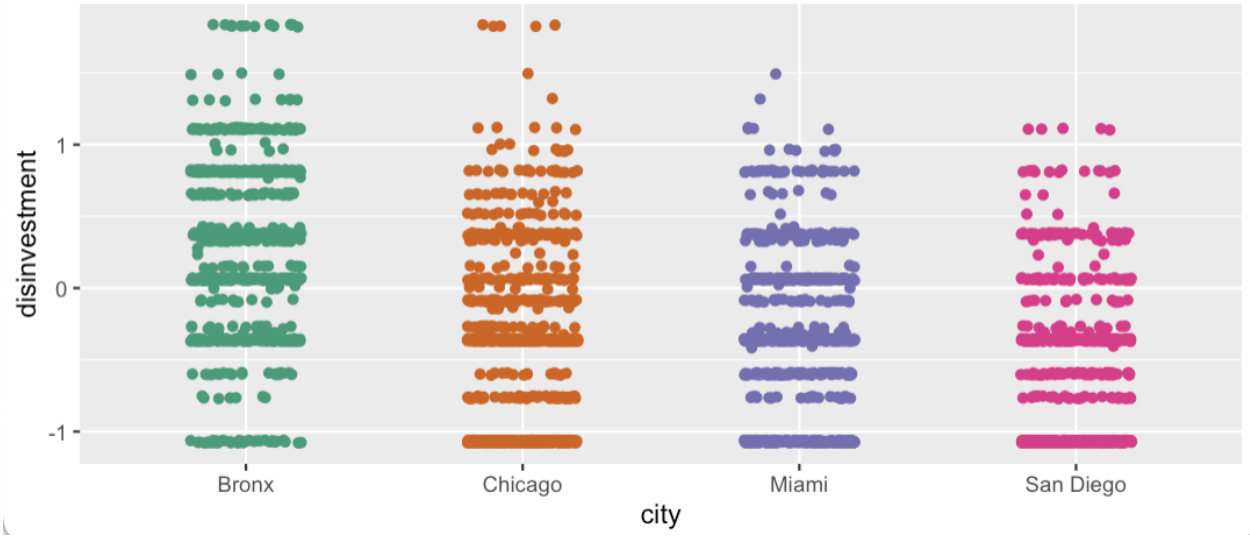
Four hundred and twenty-seven locations were randomly selected for an inter-rater reliability subsample. Answers where a rater indicated they could not tell or skipped over a question were treated as missing. No auditors' observations were dropped due to observable patterns of systematic errors

Audit Item	Kappa
Buildings Needing Repairs	0.22
Boarded Up Buildings	0.24
Vacant Lot	0.27
Empty Bottles	0.31
Litter	0.43
Burned Out Buildings	0.44
Abandoned Cars	0.55
Graffiti	0.64
Window Bars	0.71

Step 3. Item Response Theory scale development

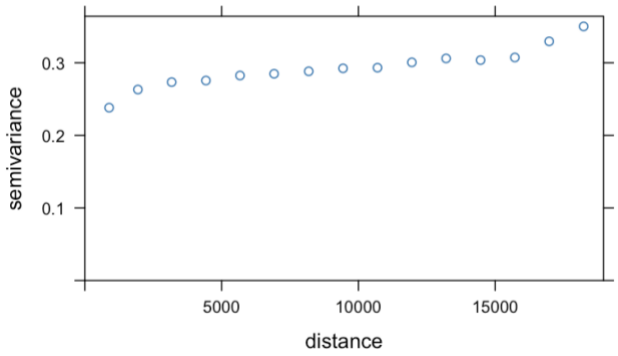
After exploring item frequency across the 4 HCHS/SOL cities and visual observation of item characteristic curves, we selected a five-item scale comprising litter, graffiti, bars on windows, abandoned buildings and presence of buildings needing repairs. Item characteristic curves and estimated disinvestment scores by city are shown below




disinvestment
1
0
-1
Bronx
Chicago
Miami
San Diego
city

Step 4. Spatial interpolation

Visual inspection of within-city variograms confirmed a pattern of spatial autocorrelation consistent with the use of ordinary kriging to spatially interpolate to the parcel level. For example, the upward curve of the variogram for the Bronx



Confirms that nearby observations predict disinvestment better than more remote observations. The corresponding map of physical disinvestment in the Bronx

Shows higher level of disinvestment near the West Farms/Crotona Park area of the central Bronx and lower levels in Co-op City in the northeast Bronx, consistent with face value expectations for spatial variability in disinvestment in the Bronx.

Step 5. Link to HCHS/SOL participants

Using parcel-level spatial linkage, we were able to link 16,177 (98.6%) participant baseline addresses to disinvestment measures. We were unable to match with 100% of participants as there were 175 participants without geocoded baseline addresses and there were 63 participants with baseline addresses more than 250 m from any land parcel with a disinvestment measure.

Supplementary Tables

Table S1. List of questions auditors answered for each location and the answer options.

Question	Answer Options
1) Is this view appropriate for auditing	Yes/No/Cannot Tell
2) Is there vacant or undeveloped land?*	Yes/No/Cannot Tell
3) Do you see any windows with bars on them?*	Yes/No/Cannot Tell
4) Is there garbage, litter, or broken glass in the street or on the sidewalks? *	Yes/No/Cannot Tell
5) Are there empty beer or liquor bottles visible in streets, yards, or alleys?*	Yes/No/Cannot Tell
6) Are there abandoned cars?*	Yes/No/Cannot Tell
7) Is there graffiti, or evidence of graffiti that has been painted over, on buildings, signs, or walls?*	Yes/No/Cannot Tell
8) How many buildings, driveways, or other private infrastructure need non-trivial repairs?*	None/One/Two or more/Not applicable/ Cannot Tell
9) Do you see burned out buildings?*	Yes/No/Cannot Tell
10) Do you see boarded up or abandoned buildings?*	Yes/No/Cannot Tell
11) Is there a sidewalk?	Yes/No/Cannot Tell
12) Is the sidewalk continuous?	Yes/No/Cannot Tell
13) Are there major bumps, cracks, holes, or weeds in the sidewalk?	Yes/No/Cannot Tell
14) Is the sidewalk unobstructed?	Yes/No/Cannot Tell
15) Should there be a sidewalk here?	Yes/No/Cannot Tell
16) Do you see any murals?	Yes/No/Cannot Tell
17) Do you see any speed bumps?	Yes/No/Cannot Tell
18) Can you see a bus stop and if so, what kind?	Bus stop with shelter/Bus stop with bench/ Bus top with signage only/ Multiple Types/ No bus stop/ Cannot Tell
19) Are there street lights present?	Yes/No/Cannot Tell

20) Do you see any bike facilities (lanes, bike racks, bike route signs, etc.)?	Yes/No/Cannot Tell
21) Do you see any businesses with Spanish language signs?	Yes/No/Cannot Tell
22) Do you see any businesses with English language signs?	Yes/No/Cannot Tell
23) Do you see any businesses with signs in a language other than Spanish or English?	Yes/No/Cannot Tell
24) Would you enjoy walking here?	Would greatly enjoy/ Would somewhat enjoy/ Would not enjoy/ Would prefer never to walk here/ I never enjoy walking/ Cannot Tell
25) Now consider the oldest image at this location: Is this image appropriate for auditing?	Yes/No/Cannot Tell
26) Consider the oldest image at this location: What month was this image recorded?	Jan/Feb/Mar/Apr/May/Jun/ Jul/Aug/Sep/Oct/Nov or Dec
27) Consider the oldest image at this location: What year was this image recorded?	2009 or earlier/2010/2011/2012/ 2013/2014/2015/2016/2017/ 2018/2019 or later
28) Consider the oldest image at this location: Is there vacant or undeveloped land?	Yes/No/Cannot Tell
29) Consider the oldest image at this location: Do you see any windows with bars on them?	Yes/No/Cannot Tell
30) Consider the oldest image at this location: Is there garbage, litter, or broken glass in the street or on the sidewalks?	Yes/No/Cannot Tell
31) Consider the oldest image at this location: Are there empty beer or liquor bottles visible in streets, yards, or alleys?	Yes/No/Cannot Tell
32) Consider the oldest image at this location: Is there graffiti, or evidence of graffiti that has been painted over, on buildings, signs, or walls?	Yes/No/Cannot Tell
33) Consider the oldest image at this location: Are there abandoned cars?	Yes/No/Cannot Tell
34) Consider the oldest image at this location: How many buildings, driveways, or other private infrastructure need non-trivial repairs?	None/One/Two or more/Not applicable/ Cannot Tell
35) Consider the oldest image at this location: Do you see burned out buildings?	Yes/No/Cannot Tell

36) Consider the oldest image at this location: Do you see boarded up or abandoned buildings?	Yes/No/Cannot Tell
37) Consider the oldest image at this location: Is there a sidewalk?	Yes/No/Cannot Tell
38) Consider the oldest image at this location: Is the sidewalk continuous?	Yes/No/Cannot Tell
39) Consider the oldest image at this location: Are there major bumps, cracks, holes, or weeds in the sidewalk?	Yes/No/Cannot Tell
40) Consider the oldest image at this location: Is the sidewalk unobstructed?	Yes/No/Cannot Tell
41) Consider the oldest image at this location: Should there be a sidewalk here?	Yes/No/Cannot Tell
42) Consider the oldest image at this location: Do you see any murals?	Yes/No/Cannot Tell
43) Consider the oldest image at this location: Do you see any speed bumps?	Yes/No/Cannot Tell
44) Consider the oldest image at this location: Can you see a bus stop and if so, what kind?	Bus stop with shelter/Bus stop with bench/ Bus top with signage only/ Multiple Types/ No bus stop/ Cannot Tell
45) Consider the oldest image at this location: Are there street lights present?	Yes/No/Cannot Tell
46) Consider the oldest image at this location: Do you see any bike facilities (lanes, bike racks, bike route signs, etc.)?	Yes/No/Cannot Tell
47) Consider the oldest image at this location: Do you see any businesses with Spanish language signs?	Yes/No/Cannot Tell
48) Consider the oldest image at this location: Do you see any businesses with English language signs?	Yes/No/Cannot Tell
49) Consider the oldest image at this location: Do you see any businesses with signs in a language other than Spanish or English?	Yes/No/Cannot Tell
50) Consider the oldest image at this location: Would you enjoy walking here?	Would greatly enjoy/ Would somewhat enjoy/ Would not enjoy/ Would prefer never to walk here/ I never enjoy walking/ Cannot Tell
51) Now consider both the oldest and newest images: are there signs of public investment such as improved sidewalks or repainted street markings between the oldest and newest?	Yes, a lot/ Yes, a little/ No/ Cannot Tell
52) Now consider both the oldest and newest images: are there signs of private	Yes, a lot/ Yes, a little/ No/ Cannot Tell

investment such as improved building maintenance between the oldest and newest?	
53) Is there anything else you'd like to say about this location?	(Open text box)

*Items used in Item Response Theory model/development of disinvestment score.

Table S2. Minimally and fully adjusted incidence rate ratios for different progression of diabetes (primary definition) from baseline to visit 2 based on neighborhood disinvestment z- score at baseline.

	Minimally Adjusted[a]			Fully Adjusted[b]		
Progression	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
Pre-diabetes to diabetes	4565	926	0.95 (0.87 – 1.04)	4192	854	0.91 (0.80 – 1.04)
Free of pre-diabetes and diabetes to diabetes	4555	112	1.02 (0.79 – 1.33)	4154	102	0.66 (0.44 – 0.99)*
Free of pre-diabetes and diabetes to either pre-diabetes or diabetes	4555	1888	0.99 (0.94 – 1.05)	4154	1744	0.95 (0.88 – 1.03)

Abbreviations: IRR, incident rate ratio

* $P < 0.05$

[a] Adjusted for age and sex

[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and study center/ ethnic heritage

Table S3. Minimally and fully adjusted incidence rate ratios of diabetes (primary definition) at visit 2 for each city based on neighborhood disinvestment z-score at baseline.

	Minimally Adjusted[a]			Fully Adjusted[b]		
City	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
The Bronx	1989	239	0.87 (0.70 – 1.10)	1805	216	0.93 (0.70 – 1.23)
Chicago	2414	295	0.89 (0.73 – 1.08)	2261	284	0.90 (0.74 – 1.09)
Miami	2312	225	0.87 (0.70 – 1.08)	1992	190	0.80 (0.64 – 0.99)*
San Diego	2405	279	0.91 (0.72 – 1.16)	2288	266	0.86 (0.66 – 1.13)

Abbreviations: IRR, incident rate ratio

* $P < 0.05$

[a] Adjusted for age and sex

[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and ethnic heritage

Table S4. Minimally and fully adjusted incidence rate ratios of diabetes (primary definition) at visit 2 for each city on a subset of participants that did not move during follow-up based on neighborhood disinvestment z-score at baseline.

	Minimally Adjusted[a]			Fully Adjusted[b]		
City	**N**	**Cases at Visit 2**	**IRR (95% CI)**	**N**	**Cases at Visit 2**	**IRR (95% CI)**
The Bronx	1293	174	0.91 (0.72 – 1.14)	1190	160	1.05 (0.78 – 1.42)
Chicago	1176	170	0.72 (0.52 – 1.00)*	1099	165	0.73 (0.56 – 0.96)*
Miami	948	86	0.86 (0.55 – 1.35)	833	73	0.83 (0.52 – 1.31)
San Diego	1310	154	0.99 (0.78 – 1.26)	1251	146	0.94 (0.70 – 1.26)

Abbreviations: IRR, incident rate ratio
* $P < 0.05$
[a] Adjusted for age and sex
[b] Adjusted for age, sex, years living in the US, income, educational attainment, family history of diabetes, neighborhood SES, and ethnic heritage

Chapter 5: Conclusion

This dissertation focused on how environmental exposures, specifically $PM_{2.5}$, NO_2, and neighborhood disinvestment affect type 2 diabetes risk in self-identified Hispanic/Latinos adults in the Bronx, Chicago, Miami, and San Diego. The first aim focused on how average exposure to $PM_{2.5}$ and NO_2 at study participants' residential addresses over the follow-up period affects risk of diabetes. Average exposure of $PM_{2.5}$ and NO_2 was estimated using fine-scale spatiotemporal air pollution models. These models provided point estimates on a two-week time scale over the full follow-up period. Study participants' exposure to $PM_{2.5}$ and NO_2 was calculated as the average exposure at their residential addresses over their follow-up period. There was evidence of a positive association between $PM_{2.5}$, NO_2, and incident diabetes when adjusting for age and sex when using either definitions of diabetes. However, there was no evidence of an association when adjusting for relevant covariates. There was a significant positive association between $PM_{2.5}$, NO_2, and incident diabetes in the Bronx study center in our full adjusted model. However, none of the other study centers exhibited a statistically significant association between $PM_{2.5}$, NO_2, and incident diabetes. There was also no evidence of effect modification by physical activity or waist circumference.

The second aim of this dissertation focused on the association between neighborhood disinvestment and incident diabetes. A virtual street audit was conducted using Google Street View imagery. Results from the virtual audit were used to fit an item response theory model to form a scale measuring latent levels of disinvestment. To estimate disinvestment levels at each residential address within the HCHS/SOL census tracts, ordinary kriging was used via spatial interpolation. A one-standard deviation increase in neighborhood disinvestment score was

associated with a 13% (95% CI: 1%-23%) lower risk of incident diabetes in the fully adjusted model. Results from this analysis were contrary to our hypothesis that neighborhood physical disinvestment is associated with incident diabetes, and were not robust to our additional analysis on disease progression nor our sensitivity analyses. This dissertation had notable strengths. For one, exposure to ambient $PM_{2.5}$ and NO_2 at residential addresses over the follow-up period was estimated using fine-scale spatiotemporal air pollution models. Second, a virtual Google Street View audit, an IRT model, and spatial interpolation were used to estimate neighborhood disinvestment at study participant's baseline residential address. The virtual street audit provided a quick and cost-effective way to generate a disinvestment measure compared to an in-person street audit.[8] The HCHS/SOL allowed for the investigation of the associations of these environmental exposures and incident diabetes in a large, multi-center cohort of self-identifying Hispanic/Latinos in four major US cities.

This dissertation had a few limitations of note. First, the spatial within-center residential clustering of study participants limited the within-center exposure variability. The HCHS/SOL used a probability sampling strategy which targeted census tracts close to the study center clinics and tracts with a high percentage of Hispanics based on the 2000 US census.[116] Therefore, participants within each study center were spatially clustered. Second, the average follow-up period for this analytic sample was 6 years, which may not have been long enough to see effects as T2D can take several years to develop. Third, we are unable to distinguish between type 1 and type 2 diabetes in our ascertainment. However, given the age of the HCHS/SOL cohort, the greater majority of diabetes cases are likely type 2. Lastly, our environmental exposures were measured at participant's residential addresses. Exposures experienced outside their residential

area, such as their place of work and larger community environment, were unaccounted for and may affect T2D risk as well.

There are a few directions future related work should take. For one, studies should increase within study center sample size and recruit study participants from more spatially varying census tracts to limit spatial residential clustering. Second, future studies should continue to investigate the potential modifying effects of physical activity and adiposity. Third, future studies should also incorporate information on air pollution levels where physical activity is occurring and how this relationship modifies the association with type 2 diabetes risk. Fourth, future studies on the HCHS/SOL can investigate the associations between NO_x and O_3 exposure and incident diabetes when data becomes available. Analysis can also be expanded to include visit 3 data once fully available. Lastly, future studies should continue to evaluate the impact of other neighborhood environment factors such as food environment and green space on diabetes risk.

Type 2 diabetes is an increasing global concern and research should continue to investigate the impacts of modifiable environmental risk factors such as air pollution and neighborhood disinvestment. Research should also continue to investigate the impacts of environmental risk factors on diabetes in Hispanic/Latino cohorts as they are an underrepresented group in research and have a higher risk of diabetes compared to non-Hispanic whites in the United States. [105]

References

1. Eze IC, Hemkens LG, Bucher HC, et al. Review Association between Ambient Air Pollution and Diabetes Mellitus in Europe. *Environ Health Perspect*. 2015;123(5):381-389. https://doi.org/10.1289/ehp.1307823.
2. Masharani U. CLASSIFICATION & PATHOGENESIS. *Curr Med Diagnosis Treat*. 2023;6(27):1-59.
3. Schnurr TM, Jakupović H, Carrasquilla GD, et al. Obesity, unfavourable lifestyle and genetic risk of type 2 diabetes: a case-cohort study. *Diabetologia*. 2020;63(7):1324-1332. doi:10.1007/s00125-020-05140-5
4. 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). doi:10.3390/ijerph15010078
5. Yang BY, Fan S, Thiering E, et al. Ambient air pollution and diabetes: A systematic review and meta-analysis. *Environ Res*. 2020;180(September 2019):108817. doi:10.1016/j.envres.2019.108817
6. Dendup T, Feng X, O'Shaughnessy P, Astell-Burt T. Perceived built environment and type 2 diabetes incidence: Exploring potential mediating pathways through physical and mental health, and behavioural factors in a longitudinal study. *Diabetes Res Clin Pract*. 2021;176:108841. doi:10.1016/j.diabres.2021.108841
7. Gebreab SY, Hickson DMA, Sims M, et al. Neighborhood social and physical environments and type 2 diabetes mellitus in African Americans: The Jackson Heart

Study. *Heal Place*. 2017;43(December 2016):128-137. doi:10.1016/j.healthplace.2016.12.001

8. Mooney SJ, Bader MDM, Lovasi GS, et al. Street Audits to Measure Neighborhood Disorder: Virtual or In-Person? *Am J Epidemiol*. 2017;186(3):265-273. doi:10.1093/aje/kwx004
9. Robinette JW, Boardman JD, Crimmins EM. Differential vulnerability to neighbourhood disorder: A gene×environment interaction study. *J Epidemiol Community Health*. 2019;73(5):388-392. doi:10.1136/jech-2018-211373
10. Association AD. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2013;36(SUPPL.1):67-74. doi:10.2337/dc13-S067
11. CDC. National Diabetes Statistics Report. Centers for Disease Control and Prevention.
12. Ong KL, Stafford LK, McLaughlin SA, et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet*. 2023;402(10397):203-234. doi:10.1016/S0140-6736(23)01301-6
13. Chatterjee S, Khunti K, Davies MJ. Type 2 diabetes. *Lancet*. 2017;389(10085):2239-2251. doi:10.1016/S0140-6736(17)30058-2
14. World Health Organisation (WHO). Diabetes. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/diabetes. Published 2022.
15. CDC. Type 2 DIabetes. Centers for Disease Control and Prevention. https://www.cdc.gov/diabetes/basics/type2.html. Published 2021.
16. Mohammed Abdul KS, Jayasinghe SS, Chandana EPS, Jayasumana C, De Silva PMCS. Arsenic and human health effects: A review. *Environ Toxicol Pharmacol*. 2015;40(3):828-

846. doi:10.1016/j.etap.2015.09.016

17. Thayer KA, Heindel JJ, Bucher JR, Gallo MA. Role of environmental chemicals in diabetes and obesity: A national toxicology program workshop review. *Environ Health Perspect*. 2012;120(6):779-789. doi:10.1289/ehp.1104597

18. Liu F, Chen G, Huo W, et al. Associations between long-term exposure to ambient air pollution and risk of type 2 diabetes mellitus: A systematic review and meta-analysis. *Environ Pollut*. 2019;252:1235-1245. doi:10.1016/j.envpol.2019.06.033

19. Heidemann C, Niemann H, Paprott R, Du Y, Rathmann W, Scheidt-Nave C. Residential traffic and incidence of type 2 diabetes: The german health interview and examination surveys. *Diabet Med*. 2014;31(10):1269-1276. doi:10.1111/dme.12480

20. Mooney SJ, Joshi S, Cerdá M, Kennedy GJ, Beard JR, Rundle AG. Neighborhood Disorder and Physical Activity among Older Adults: A Longitudinal Study. *J Urban Heal*. 2017;94(1):30-42. doi:10.1007/s11524-016-0125-y

21. Leon CFM De, Bienias JL, Barnes LL, Skarupski K a. Neighborhood Social Cohesion and Disorder in Relation to Walking in Community-Dwelling Older Adults. *J Aging Health*. 2009;21(1):155-171.

22. Qin W, Wang Y, Cho S. Neighborhood Social Cohesion, Physical Disorder, and Daily Activity Limitations Among Community-Dwelling Older Adults. *Arch Gerontol Geriatr*. 2021;93(October 2020):104295. doi:10.1016/j.archger.2020.104295

23. Wallace D, Chamberlain AW, Fahmy C. Changes in Neighborhood Social Control and Disorder and Their Relationship to Exercise Behavior. *Environ Behav*. 2019;51(6):717-748. doi:10.1177/0013916518761530

24. Daviglus ML, Talavera GA, Aviles-Santa L, et al. Prevalence of Major Cardiovascular

Risk Factors and Cardiovascular Diseases Among Hispanic/Latino Individuals of Diverse Backgrounds in the United States. *JAMA*. 2012;308(7):1775-1784. doi:10.1001/jama.2012.14517.Prevalence

25. CDC. Type 2 Diabetes. Centers for Disease Control and Prevention.
26. Kaplan RC, Song RJ, Lin J, et al. Predictors of Incident Diabetes in Two Populations: Framingham Heart Study and Hispanic Community Health Study/Study of Latinos. *BMC Public Health*. 2022.
27. Kolb H, Martin S. Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. *BMC Med*. 2017;15(1):1-11. doi:10.1186/s12916-017-0901-x
28. Knowler W., Barrett-Conner E, Folwer SE, et al. Reduction of the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002;346(6):393-403. doi:10.1097/01.ogx.0000055759.75837.e7
29. Levine JA. Sick of sitting. *Diabetologia*. 2015;58(8):1751-1758. doi:10.1007/s00125-015-3624-6
30. Li T, He S, Liu S, Kong Z, Wang J, Zhang Y. Effects of different exercise durations on Keap1-Nrf2-ARE pathway activation in mouse skeletal muscle. *Free Radic Res*. 2015;49(10):1269-1274. doi:10.3109/10715762.2015.1066784
31. Guo C, Yang HT, Chang L yun, et al. Habitual exercise is associated with reduced risk of diabetes regardless of air pollution: a longitudinal cohort study. *Diabetologia*. 2021;64(6):1298-1308. doi:10.1007/s00125-021-05408-4
32. Chen GC, Qi Q, Hua S, et al. Accelerometer-assessed physical activity and incident diabetes in a population covering the adult life span: The Hispanic Community Health Study/Study of Latinos. *Am J Clin Nutr*. 2020;112(5):1318-1327.

doi:10.1093/ajcn/nqaa232

33. Pan A, Wang Y, Talaei M, Hu FB, Wu T. Relation of active, passice, and quitting smoking with incident diabetes: a meta-analysis and systematic review. *Lanvet Diabetes Endocrinol*. 2015;3(12):958-967. doi:10.1016/S2213-8587(15)00316-2.Relation

34. Keith RJ, Al Rifai M, Carruba C, et al. Tobacco use, insulin resistance, and risk of type 2 diabetes: Results from the multi-ethnic study of atherosclerosis. *PLoS One*. 2016;11(6):1-15. doi:10.1371/journal.pone.0157592

35. Delgado-Velandia M, Gonzalez-Marrachelli V, Domingo-Relloso A, et al. Healthy lifestyle, metabolomics and incident type 2 diabetes in a population-based cohort from Spain. *Int J Behav Nutr Phys Act*. 2022;19(1):1-13. doi:10.1186/s12966-021-01219-3

36. Brook RD, Rajagopalan S, Iii CAP, et al. Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American. 2010. doi:10.1161/CIR.0b013e3181dbece1

37. Rajagopalan S, Al-Kindi SG, Brook RD. Air Pollution and Cardiovascular Disease: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2018;72(17):2054-2070. doi:10.1016/j.jacc.2018.07.099

38. Sun S, Cao W, Qiu H, et al. Benefits of physical activity not affected by air pollution: A prospective cohort study. *Int J Epidemiol*. 2020;49(1):142-152. doi:10.1093/ije/dyz184

39. Nieder R, Benbi DK. Reactive nitrogen compounds and their influence on human health: an overview. *Rev Environ Health*. 2022;37(2):229-246. doi:10.1515/reveh-2021-0021

40. Rugel EJ, Brauer M. Quiet, clean, green, and active: A Navigation Guide systematic review of the impacts of spatially correlated urban exposures on a range of physical health outcomes. *Environ Res*. 2020;185(February):109388. doi:10.1016/j.envres.2020.109388

41. EPA (U.S. Environmental Protection Agency). Nitrogen Dioxide (NO2) Pollution. EPA.

42. Nuvolone D, Petri D, Voller F. The effects of ozone on human health. *Environ Sci Pollut Res*. 2018;25(9):8074-8088. doi:10.1007/s11356-017-9239-3

43. EPA (U.S. Environmental Protection Agency). Health Effects of Ozone Pollution. *EPA*. 2022.

44. Hansen AB, Ravnskjær L, Loft S, et al. Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environ Int*. 2016;91:243-250. doi:10.1016/j.envint.2016.02.036

45. Clark C, Sbihi H, Tamburic L, Brauer M, Frank LD, Davies HW. Association of long-term exposure to transportation noise and traffic-related air pollution with the incidence of diabetes: A prospective cohort study. *Environ Health Perspect*. 2017;125(8):1-10. doi:10.1289/EHP1279

46. Renzi M, Cerza F, Gariazzo C, et al. Air pollution and occurrence of type 2 diabetes in a large cohort study. *Environ Int*. 2018;112:68-76. doi:10.1016/j.envint.2017.12.007

47. Yu Y, Jerrett M, Paul KC, et al. Ozone exposure, outdoor physical activity, and incident type 2 diabetes in the salsa cohort of older mexican americans. *Environ Health Perspect*. 2021;129(9):1-10. doi:10.1289/EHP8620

48. Cervantes-Martínez K, Stern D, Zamora-Muñoz JS, et al. Air pollution exposure and incidence of type 2 diabetes in women: A prospective analysis from the Mexican Teachers' Cohort. *Sci Total Environ*. 2022;818(655). doi:10.1016/j.scitotenv.2021.151833

49. Chen H, Burnett RT, Kwong JC, et al. Risk of incident diabetes in relation to long-term exposure to fine particulate matter in Ontario, Canada. *Environ Health Perspect*. 2013;121(7):804-810. doi:10.1289/ehp.1205958

50. Coogan PF, White LF, Burnett RT, Seto E. PM2.5 and Diabetes and Hypertension Incidence in the Black Women's Health Study. 2016;27(2):202-210. doi:10.1097/EDE.0000000000000418.PM

51. Lao XQ, Guo C, Chang LY, et al. Long-term exposure to ambient fine particulate matter (Pm2:5) and lung function in children, adolescents, and young adults: A longitudinal cohort study. *Environ Health Perspect*. 2019;127(12):759-769. doi:10.1289/EHP5220

52. Li ZH, Zhong WF, Zhang XR, et al. Association of physical activity and air pollution exposure with the risk of type 2 diabetes: a large population-based prospective cohort study. *Environ Heal A Glob Access Sci Source*. 2022;21(1):1-12. doi:10.1186/s12940-022-00922-3

53. Park SK, Adar SD, O'Neill MS, et al. Long-term exposure to air pollution and type 2 diabetes mellitus in a multiethnic cohort. *Am J Epidemiol*. 2015;181(5):327-336. doi:10.1093/aje/kwu280

54. Yitshak Sade M, Shi L, Colicino E, et al. Long-term air pollution exposure and diabetes risk in American older adults: A national secondary data-based cohort study. *Environ Pollut*. 2023;320(December 2022):121056. doi:10.1016/j.envpol.2023.121056

55. Rajagopalan S, Brauer M, Bhatnagar A, et al. Personal-Level Protective Actions against Particulate Matter Air Pollution Exposure: A Scientific Statement from the American Heart Association. *Circulation*. 2020:411-431. doi:10.1161/CIR.0000000000000931

56. Weinmayr G, Hennig F, Fuks K, et al. Long-term exposure to fine particulate matter and incidence of type 2 diabetes mellitus in a cohort study: Effects of total and traffic-specific air pollution. *Environ Heal A Glob Access Sci Source*. 2015;14(1):1-8. doi:10.1186/s12940-015-0031-x

57. Lim CC, Thurston GD. Air Pollution, Oxidative Stress, and Diabetes: a Life Course Epidemiologic Perspective. *Curr Diab Rep*. 2019;19(8). doi:10.1007/s11892-019-1181-y

58. Boogaard H, Janssen NAH, Fischer PH, et al. Contrasts in oxidative potential and other particulate matter characteristics collected near major streets and background locations. *Environ Health Perspect*. 2012;120(2):185-191. doi:10.1289/ehp.1103667

59. Lanki T, Hampel R, Tiittanen P, et al. Air pollution from road traffic and systemic inflammation in adults: A cross-sectional analysis in the European ESCAPE project. *Environ Health Perspect*. 2015;123(8):785-791. doi:10.1289/ehp.1408224

60. Chen J-C, Schwartz J. Metabolic Syndrome and Inflammatory Responses to Long-Term Particulate Air Pollutant. *Environ Health Perspect*. 2008;116(5):612-617. doi:10.1016/B978-0-12-386454-3.00341-9

61. Vella RE, Pillon NJ, Zarrouki B, et al. Ozone exposure triggers insulin resistance through muscle c-Jun N-terminal kinase activation. *Diabetes*. 2015;64(3):1011-1024. doi:10.2337/db13-1181

62. Bass V, Gordon CJ, Jarema KA, et al. Ozone induces glucose intolerance and systemic metabolic effects in young and aged brown Norway rats. *Toxicol Appl Pharmacol*. 2013;273(3):551-560. doi:10.1016/j.taap.2013.09.029

63. Ultman JS, Ben-Jebria A, Arnold SF. Uptake distribution of ozone in human lungs: intersubject variability in physiologic response. *Res Rep Health Eff Inst*. 2004;(125).

64. Tainio M, de Nazelle AJ, Götschi T, et al. Can air pollution negate the health benefits of cycling and walking? *Prev Med (Baltim)*. 2015. doi:10.1016/j.ypmed.2016.02.002

65. Eze IC, Schaffner E, Foraster M, et al. Long-term exposure to ambient air pollution and metabolic syndrome in adults. *PLoS One*. 2015;10(6):1-19.

doi:10.1371/journal.pone.0130337

66. Mujahid MS, Roux AVD, Shen M, et al. Relation between neighborhood environments and obesity in the multi-ethnic study of atherosclerosis. *Am J Epidemiol*. 2008;167(11):1349-1357. doi:10.1093/aje/kwn047

67. Carlson JA, Sallis JF, Jankowska MM, et al. Neighborhood built environments and Hispanic / Latino adults ' physical activity in the U . S .: The Hispanic community health study / study of Latinos community and surrounding areas study. *Prev Med (Baltim)*. 2022;160. doi:10.1016/j.ypmed.2022.107073

68. Cunningham-Myrie CA, Theall KP, Younger NO, et al. Associations between neighborhood effects and physical activity, obesity, and diabetes: The Jamaica Health and Lifestyle Survey 2008. *J Clin Epidemiol*. 2015;68(9):970-978. doi:10.1016/j.jclinepi.2014.08.004

69. Braver NR, Rutters F, Wagtendonk AJ, et al. Neighborhood walkability, physical activity and changes in glycemic markers in people with type 2 diabetes: The Hoorn Diabetes Care System cohort. *Heal Place*. 2021;69:102560. doi:10.1016/j.healthplace.2021.102560

70. Gallo LC, Savin KL, Jankowska MM, et al. Neighborhood Environment and Metabolic Risk in Hispanics/Latinos From the Hispanic Community Health Study/Study of Latinos. *Am J Prev Med*. 2022;000(000):1-9. doi:10.1016/j.amepre.2022.01.025

71. Ndjila S, Lovasi GS, Fry D, Friche AA. Measuring Neighborhood Order and Disorder: a Rapid Literature Review. *Curr Environ Heal reports*. 2019;6(4):316-326. doi:10.1007/s40572-019-00259-z

72. Mooney SJ, Joshi S, Cerda M, Kennedy GJ, Beard JR, Rundle AG. Contextual correlates of physical activity among older adults:Aneighborhood environment-wide association

study (ne-was). *Cancer Epidemiol Biomarkers Prev*. 2017;26(4):495-504. doi:10.1158/1055-9965.EPI-16-0827

73. Velasquez AJ, Douglas JA, Guo F, Robinette JW. What predicts how safe people feel in their neighborhoods and does it depend on functional status? *SSM - Popul Heal*. 2021;16:100927. doi:10.1016/j.ssmph.2021.100927

74. Velasquez AJ, Douglas JA, Guo F, Robinette JW. In the eyes of the beholder: Race, place and health. *Front Public Heal*. 2022;10(1). doi:10.3389/fpubh.2022.920637

75. Gariepy G, Smith KJ, Schmitz N. Diabetes distress and neighborhood characteristics in people with type 2 diabetes. *J Psychosom Res*. 2013;75(2):147-152. doi:10.1016/j.jpsychores.2013.05.009

76. Mayne SL, Jose A, Mo A, et al. Neighborhood disorder and obesity-related outcomes among women in Chicago. *Int J Environ Res Public Health*. 2018;15(7). doi:10.3390/ijerph15071395

77. Latham K, Williams MM. Does Neighborhood Disorder Predict Recovery from Mobility Limitation? Findings from the Health and Retirement Study. *J Aging Health*. 2015;27(8):1415-1442. doi:10.1177/0898264315584328

78. Plascak JJ, Hohl B, Barrington WE, Beresford SA. Perceived neighborhood disorder, racial-ethnic discrimination and leading risk factors for chronic disease among women: California Behavioral Risk Factor Surveillance System, 2013. *SSM - Popul Heal*. 2018;5(March):227-238. doi:10.1016/j.ssmph.2018.07.001

79. Robinette JW, Charles ST, Gruenewald TL. Neighborhood cohesion, neighborhood disorder, and cardiometabolic risk. *Soc Sci Med*. 2018;198(June 2017):70-76. doi:10.1016/j.socscimed.2017.12.025

80. Carroll SJ, Niyonsenga T, Coffee NT, Taylor AW, Daniel M. Does physical activity mediate the associations between local-area descriptive norms, built environment walkability, and glycosylated hemoglobin? *Int J Environ Res Public Health*. 2017;14(9):1-17. doi:10.3390/ijerph14090953

81. Mooney SJ, Bader MDM, Lovasi GS, Neckerman KM, Teitler JO, Rundle AG. Validity of an ecometric neighborhood physical disorder measure constructed by virtual street audit. *Am J Epidemiol*. 2014;180(6):626-635. doi:10.1093/aje/kwu180

82. Bader MDM, Mooney SJ, Bennett B, Rundle AG. The Promise, Practicalities, and Perils of Virtually Auditing Neighborhoods Using Google Street View. *Ann Am Acad Pol Soc Sci*. 2017;669(1):18-40. doi:10.1177/0002716216681488

83. Rzotkiewicz A, Pearson AL, Dougherty B V., Shortridge A, Wilson N. Systematic review of the use of Google Street View in health research: Major themes, strengths, weaknesses and possibilities for future research. *Heal Place*. 2018;52(July):240-246. doi:10.1016/j.healthplace.2018.07.001

84. Andrew G. Rundle, DrPH, Michael D.M. Bader, PhD, Catherine A. Richards, MPH, Kathryn M. Neckerman, PhD, and Julien O. Teitler P. Using Google Street View to Audit Neighborhood Environments. *Am J Prev Med*. 2011;40(1):94-100. doi:10.1016/j.amepre.2010.09.034.Using

85. Cheryl M. Kelly P. Using Google Street View to Audit the Built Environment: Inter- rater Reliability Results. *Ann Behav Med*. 2013;45(1):108-112. doi:10.1007/s12160-012-9419-9.Using

86. Li X, Zhang C, Li W, Ricard R, Meng Q, Zhang W. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban For Urban Green*.

2015;14(3):675-685. doi:10.1016/j.ufug.2015.06.006

87. Bader MDM, Mooney SJ, Lee YJ, et al. Development and deployment of the Computer Assisted Neighborhood Visual Assessment System (CANVAS) to measure health-related neighborhood conditions. *Heal Place*. 2015;31:163-172. doi:10.1016/j.healthplace.2014.10.012

88. Goel R, Garcia LMT, Goodman A, et al. Estimating city-level travel patterns using street imagery: A case study of using Google Street View in Britain. *PLoS One*. 2018;13(5):1-22. doi:10.1371/journal.pone.0196521

89. Smith CM, Kaufman JD, Mooney SJ. Google street view image availability in the Bronx and San Diego, 2007–2020: Understanding potential biases in virtual audits of urban built environments. *Heal Place*. 2021;72(June):102701. doi:10.1016/j.healthplace.2021.102701

90. Swope CB, Hernández D, Cushing LJ. The Relationship of Historical Redlining with Present-Day Neighborhood Environmental and Health Outcomes: A Scoping Review and Conceptual Model. *J Urban Heal*. 2022;(0123456789). doi:10.1007/s11524-022-00665-z

91. Rothstein R. *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York: Livright; 2017.

92. Greer J. The home owners' loan corporation and the development of the residential security maps. *J Urban Hist*. 2013;39(2):275-296. doi:10.1177/0096144212436724

93. Lane HM, Morello-Frosch R, Marshall JD, Apte JS. Historical Redlining Is Associated with Present-Day Air Pollution Disparities in U.S. Cities. *Environ Sci Technol Lett*. 2022;9(4):345-350. doi:10.1021/acs.estlett.1c01012

94. Nelson R. Mapping Inequality: Redlining in New Deal America.

95. Aaronson D, Hartley D, Mazumder B. The Effects of the 1930s HOLC "Redlining" Maps.

Am Econ J Econ Policy. 2021;13(4):355-392. http://www.chicagofed.org/webpages/publications/chicago_fed_letter/2006/april_225a.cfm.

96. Gonzalez DJX, Nardone A, Nguyen A V., Morello-Frosch R, Casey JA. Historic redlining and the siting of oil and gas wells in the United States. *J Expo Sci Environ Epidemiol*. 2022;(January):1-8. doi:10.1038/s41370-022-00434-9

97. Lee EK, Donley G, Ciesielski TH, et al. Health outcomes in redlined versus non-redlined neighborhoods: A systematic review and meta-analysis. *Soc Sci Med*. 2022;294(October 2021):114696. doi:10.1016/j.socscimed.2021.114696

98. Guidi J, Lucente M, Sonino N, Fava GA. Allostatic Load and Its Impact on Health: A Systematic Review. *Psychother Psychosom*. 2021;90(1):11-27. doi:10.1159/000510696

99. Asad AL, Rosen E. Hiding within racial hierarchies: how undocumented immigrants make residential decisions in an American city. *J Ethn Migr Stud*. 2019;45(11):1857-1882. doi:10.1080/1369183X.2018.1532787

100. Timberlake JM, Iceland J. Change in Racial and Ethnic Residential Inequality in American Cities, 1970–2000. *City Community*. 2007;6(4):335-365. doi:10.1111/j.1540-6040.2007.00231.x

101. Frey WH, Farley R. Latino, Asian, and black segregation in U.S. metropolitan areas: Are multiethnic metros different? *Demography*. 1996;33(1):35-50. doi:10.2307/2061712

102. Alba R, Deane G, Denton N, Disha I, McKenzie B, Napierala J. The Role of Immigrant Enclaves for Latino Residential Inequalities. *J Ethn Migr Stud*. 2014;40(1):1-20. doi:10.1080/1369183X.2013.831549

103. Iceland J, Weinberg DH, Steinmetz E. Racial and Ethnic Residential Segregation in the

United States: 1980-2000. *US Census 2000 Spec Reports*. 2002.

104. Lee MA, Ferraro KF. Neighborhood residential segregation and physical health among hispanic Americans: Good, bad, or benign? *J Health Soc Behav*. 2007;48(2):131-148. doi:10.1177/002214650704800203

105. Office of Minority Health. Diabetes and Hispanic Americans. U.S Department of Health and Human Services.

106. Center for Disease Control and Prevention. Hispanic or Latino People and Type 2 Diabetes. *CDC*. 2022.

107. Schneiderman N, Llabre M, Cowie CC, et al. Prevalence of diabetes among hispanics/latinos from diverse backgrounds: The Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *Diabetes Care*. 2014;37(8):2233-2239. doi:10.2337/dc13-2939

108. Aguayo-Mazzucato C, Diaque P, Hernandez S, Rosas S, Kostic A, Caballero AE. Understanding the growing epidemic of type 2 diabetes in the Hispanic population living in the United States. *Diabetes Metab Res Rev*. 2019;35(2):e3097. doi:10.1002/dmrr.3097.Understanding

109. O'Donovan G, Chudasama Y, Grocock S, et al. The association between air pollution and type 2 diabetes in a large cross-sectional study in Leicester: The CHAMPIONS Study. *Environ Int*. 2017;104(April):41-47. doi:10.1016/j.envint.2017.03.027

110. Kirwa K, Szpiro AA, Sheppard L, et al. Fine-Scale Air Pollution Models for Epidemiologic Research: Insights From Approaches Developed in the Multi-ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Curr Environ Heal Reports*. 2021;8(2):113-126. doi:10.1007/s40572-021-00310-y

111. Sampson PD, Szpiro AA, Sheppard L, Lindström J, Kaufman JD. Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmos Environ*. 2011;45(36):6593-6606. doi:10.1016/j.atmosenv.2011.04.073

112. Guo LH, Zeeshan M, Huang GF, et al. Influence of Air Pollution Exposures on Cardiometabolic Risk Factors: a Review. *Curr Environ Heal Reports*. 2023;10(4):501-507. doi:10.1007/s40572-023-00423-6

113. DeFlorio-Barker S, Lobdell DT, Stone SL, Boehmer T, Rappazzo KM. Acute effects of short-term exposure to air pollution while being physically active, the potential for modification: A review of the literature. *Prev Med (Baltim)*. 2020;139(May):106195. doi:10.1016/j.ypmed.2020.106195

114. Wu P, Guo Q, Zhao Y, et al. Emerging concern on air pollution and health: Trade-off between air pollution exposure and physical activity. *Eco-Environment Heal*. 2024;3(2):202-207. doi:10.1016/j.eehl.2024.01.012

115. Li X, Wang M, Song Y, et al. Obesity and the relation between joint exposure to ambient air pollutants and incident type 2 diabetes: A cohort study in UK Biobank. *PLoS Med*. 2021;18(8):1-16. doi:10.1371/journal.pmed.1003767

116. LaVange LM, Kalsbeek W, Sorlie PD, et al. Sample Design and Cohort Selection in the Hispanic Community Health Study/Study of Latinos. *Annu Epidemiol*. 2010;20(8):642-649. doi:10.1016/j.annepidem.2010.05.006.Sample

117. Sorlie PD, Aviles-Santa L, Wassertheil-Smoller S, et al. Design and Implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;20(8):629-641. doi:10.1016/j.annepidem.2010.03.015.Design

118. Ogawa & Co. Inc. NO, NO2, NOx and SO2 Sampling Protocol Using The Ogawa

Sampler. 2006;(2):1-28.

119. Szpiro AA, Sampson PD, Sheppard L, Lumley T, Adar SD, Kaufman JD. Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics*. 2010;21(6):606-631. doi:10.1002/env.1014

120. Keller JP, Olives C, Kim SY, et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environ Health Perspect*. 2015;123(4):301-309. doi:10.1289/ehp.1408145

121. Lindström J, Szpiro AA, Sampson PD, et al. A Flexible Spatio-Temporal Model for Air Pollution with Spatial and Spatio-Temporal Covariates. *Env Ecol Stat*. 2014;21(3):411-433. doi:10.1007/s10651-013-0261-4.A

122. Kaufman JD, Adar SD, Allen RW, et al. Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease. *Am J Epidemiol*. 2012;176(9):825-837. doi:10.1093/aje/kws169

123. Kirwa K, Gassett AJ, Sack C, et al. Estimating ambient air pollutant concentrations outside and inside homes in the Subpopulations and Intermediate outcomes in COPD air pollution (SPIROMICS air) cohort. *Environ Res*. 2024;259(June). doi:10.1016/j.envres.2024.119512

124. Center for Disease Control and Prevention. Adult Activity: An Overview. CDC. https://www.cdc.gov/physical-activity-basics/guidelines/adults.html#:~:text=According to the current Physical,and vigorous-intensity physical activity. Published 2023.

125. Evenson KR, Sotres-Alvarez D, Deng Y, et al. Accelerometer adherence and performance in a cohort study of US hispanic adults. *Med Sci Sports Exerc*. 2015;47(4):725-734.

doi:10.1249/MSS.0000000000000478

126. NHLBI. Assessing Your Weight and Health Risk. National Heart, Lung, and Blood Institute. https://www.nhlbi.nih.gov/health/educational/lose_wt/risk.htm. Published 2025.

127. Zewdie HY, Fahey CA, Harrington AL, et al. Racial residential segregation is associated with ambient air pollution exposure after adjustment for multilevel sociodemographic factors: Evidence from eight US-based cohorts. *Environ Epidemiol*. 2025;9(1):e367. doi:10.1097/EE9.0000000000000367

128. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45(3 SE-Articles):1-67. doi:10.18637/jss.v045.i03

129. Lumley T. Analysis of Complex Survey Samples. *J Stat Softw*. 2004;09. doi:10.18637/jss.v009.i08

130. Lenters-Westra E, Schindhelm RK, Bilo HJG, Groenier KH, Slingerland RJ. Differences in interpretation of haemoglobin A1c values among diabetes care professionals. *Neth J Med*. 2014;72(9):462-466.

131. Tainio M, Jovanovic Andersen Z, Nieuwenhuijsen MJ, et al. Air pollution, physical activity and health: A mapping review of the evidence. *Environ Int*. 2021;147(July 2020):105954. doi:10.1016/j.envint.2020.105954

132. Basagaña X, Pedersen M, Barrera-Gómez J, et al. Analysis of multicentre epidemiological studies: Contrasting fixed or random effects modelling and meta-analysis. *Int J Epidemiol*. 2018;47(4):1343-1354. doi:10.1093/ije/dyy117

133. Mujahid MS, Maddali SR, Gao X, Oo KH, Benjamin LA, Lewis TT. The impact of neighborhoods on diabetes risk and outcomes: Centering health equity. *Diabetes Care*. 2023;46(9):1609-1618. doi:10.2337/dci23-0003

134. Uddin J, Malla G, Long DL, et al. The association between neighborhood social and economic environment and prevalent diabetes in urban and rural communities: The Reasons for Geographic and Racial Differences in Stroke (REGARDS) study. *SSM - Popul Heal*. 2022;17(February):101050. doi:10.1016/j.ssmph.2022.101050

135. Feyissa TR, Wood SM, Vakil K, et al. The built environment and its association with type 2 diabetes mellitus incidence: A systematic review and meta-analysis of longitudinal studies. *Soc Sci Med*. 2024;361(April):117372. doi:10.1016/j.socscimed.2024.117372

136. Ccami-Bernal F, Soriano-Moreno DR, Fernandez-Guzman D, et al. Green space exposure and type 2 diabetes mellitus incidence: A systematic review. *Heal Place*. 2023;82(August 2022):103045. doi:10.1016/j.healthplace.2023.103045

137. Marco M, Gracia E, Martín-Fernández M, López-Quílez A. Validation of a Google Street View-Based Neighborhood Disorder Observational Scale. *J Urban Heal*. 2017;94(2):190-198. doi:10.1007/s11524-017-0134-5

138. Quinn JW, Mooney SJ, Sheehan DM, et al. Neighborhood physical disorder in New York City. *J Maps*. 2016;12(1):53-60. doi:10.1080/17445647.2014.978910

139. Mooney SJ, Bader MDM, Lovasi GS, Neckerman KM, Rundle AG, Teitler JO. Using Universal Kriging to Improve Neighborhood Physical Disorder Measurement. *Sociol Methods Res*. 2020;49(4):1163-1185. doi:10.1177/0049124118769103

140. Plascak JJ, Mooney SJ, Schootman M, et al. Validating a spatio-temporal model of observed neighborhood physical disorder. *Spat Spatiotemporal Epidemiol*. 2022;41(December 2021):100506. doi:10.1016/j.sste.2022.100506

141. Mooney SJ, Wheeler-Martin K, Fiedler LM, et al. Development and Validation of a Google Street View Pedestrian Safety Audit Tool. *Epidemiology*. 2020;31(2):301-309.

doi:10.1097/EDE.0000000000001124

142. Raudenbusch SW, Sampson RJ. Ecometrics, toward a science of assessing ecological settings. *Sociol Methodol*. 1999;29:1-41.

143. Uddin J, Zhu S, Adhikari S, et al. Age and sex differences in the association between neighborhood socioeconomic environment and incident diabetes: Results from the diabetes location, environmental attributes and disparities (LEAD) network. *SSM - Popul Heal*. 2023;24(November):101541. doi:10.1016/j.ssmph.2023.101541

144. Christine PJ, Auchincloss AH, Bertoni AG, et al. Longitudinal Associations Between Neighborhood Physical and Social Environments and Incident Type 2 Diabetes Mellitus: The Multi-Ethnic Study of Atherosclerosis (MESA). *JAMA Intern Med*. 2015;175(8):1311-1320. doi:10.1001/jamainternmed.2015.2691.Longitudinal

145. Mooney SJ, Hurvitz PM, Moudon AV, Zhou C, Dalmat R, Saelens BE. Residential neighborhood features associated with objectively measured walking near home: Revisiting walkability using the Automatic Context Measurement Tool (ACMT). *Heal Place*. 2020;63. doi:10.1016/j.healthplace.2020.102332.Residential

146. Pan SW, Carpiano RM. Immigrant density, sense of community belonging, and suicidal ideation among racial minority and white immigrants in Canada. *J Immigr Minor Heal*. 2013;15(1):34-42. doi:10.1007/s10903-012-9657-8

147. Li K, Wen M, Henry KA. Ethnic density, immigrant enclaves, and Latino health risks: A propensity score matching approach. *Soc Sci Med*. 2017;189:44-52. doi:10.1016/j.socscimed.2017.07.019

148. Pichardo CM, Pichardo MS, Gallo LC, et al. Association of neighborhood segregation with 6-year incidence of metabolic syndrome in the Hispanic community health

study/study of Latinos. *Ann Epidemiol*. 2023;78:1-8. doi:10.1016/j.annepidem.2022.11.003

149. Wang MC, Shah NS, Carnethon M, O'Brien MJ, Khan SS. Age at Diagnosis of Diabetes by Race and Ethnicity in the United Stated from 2011 to 2018. *JAMA Intern Med*. 2021;181(11). doi:10.1001/jamainternmed.2021.4945