The Creation of a Multidomain Neighborhood Environmental Vulnerability Index Across an Urban Center

Stephen P. Uong, MPH[1], Jiayi Zhou[1], Stephanie Lovinsky-Desir, MD, MS[2], Sandra S. Albrecht, PhD, MPH[1], Alexander Azan, MD[3], Earle C. Chambers, PhD, MPH[4], Perry E. Sheffield, MD, MPH[5], Azure Thompson, DrPH, MPH[6], Joseph Wilson, MHS[5], Jennifer Woo Baidal, MD, MPH[2], Jeanette A. Stingone, PhD, MPH[1]

[1]Columbia University, Mailman School of Public Health
[2]Columbia University, Vagelos College of Physicians and Surgeons
[3]NYU Langone Health, Department of Pediatrics at NYU Grossman School of Medicine
[4]Albert Einstein College of Medicine, Department of Family and Social Medicine
[5]Icahn School of Medicine at Mount Sinai, Department of Environmental Medicine and Public Health
[6]SUNY Downstate Health Sciences University, School of Public Health

ABSTRACT (243/250 words)

Compared to previous studies that have typically used a single summary score, we aimed to construct a multidomain neighborhood environmental vulnerability index (NEVI) to characterize the magnitude and variability of area-level factors with the potential to modify the health effects of environmental pollutants. Using the Toxicological Prioritization Index framework and data from the 2015-2019 U.S. Census American Community Survey and the 2020 CDC PLACES Project, we quantified census tract-level vulnerability overall and in 4 primary domains (demographic, economic, residential, and health status), 24 subdomains, and 54 distinct area-level features for New York City (NYC). Overall and domain-specific indices were calculated by summing standardized feature values within the subdomains and then aggregating and weighting subdomains within equally-weighted primary domains. In citywide comparisons, NEVI was correlated with both the Neighborhood Deprivation Index (r = 0.91) and the Social Vulnerability Index (r = 0.87) but provided additional information on features contributing to vulnerability. Vulnerability varied spatially across NYC, and hierarchical cluster analysis using subdomain scores revealed six patterns of vulnerability across domains: 1) low in all, 2) primarily low except residential, 3) medium in all, 4) high demographic, economic, and residential 5) high economic, residential, and health status, and 6) high demographic, economic and health status. Created using a tool that offers flexibility for theory-based construction, NEVI provided detailed metrics of vulnerability across domains that can inform targeted research and public health interventions aimed at reducing the health impacts from environmental exposures across an urban center.

Key Words: area-level vulnerability; data integration; environmental justice; social determinants of health; deprivation

INTRODUCTION

Previous research has linked environmental pollutants, such as air pollution, toxic metals, and plastics, to various health outcomes, including cancer, neurodevelopmental impacts, asthma, and heart disease across various urban settings.[1–4] Many prior studies conceptualized housing characteristics, poverty, health care access, and other neighborhood-level features individually as confounders or effect modifiers in evaluating the relationship between environmental pollutants and various health outcomes.[5–8] In these studies, "neighborhoods" are spatial units where individuals reside that aim to capture unique group-level properties.[9] These studies have demonstrated neighborhood-level factors can contribute to differences in environmental vulnerability, including greater likelihood of exposure to pollutants and greater susceptibility to health effects from environmental pollutants for individuals.

Composite neighborhood-level indices used to describe collective social and structural factors in previous work often measure a single domain of vulnerability or were not designed to focus on environmental vulnerability and flexible theory-based construction. For example, the commonly-used neighborhood deprivation index (NDI) that summarizes demographics, income, education, employment, and housing data have been used to measure neighborhood socioeconomic deprivation across different cities and health outcomes in environmental perinatal epidemiology studies.[10–14] Similarly, the Centers for Disease Control and Prevention's Social Vulnerability Index (SVI) summarizes several features to reflect neighborhood-level stressors relevant to community resilience to disasters or disease outbreaks.[15] While there are newer efforts to further characterize neighborhood-level vulnerability such as the Climate and Economic Justice Tool, these tools often emphasize a single summary metric and provide limited ways to compare the relative magnitude or importance of individual features or domains that

contribute to vulnerability between neighborhoods.[16] Information about specific features or domains provides a more nuanced understanding of vulnerability that would inform targeted environmental justice efforts and urban planning policies by describing both the location of vulnerable neighborhoods and type of neighborhood-level characteristics that may be contributing vulnerability in those neighborhoods. Furthermore, indices like the NDI and SVI do not include other features that may increase environmental vulnerability, including pre-existing health conditions and behaviors that may increase vulnerability to health impacts from air pollution, heat stress, and climate change.[10,17–20] Many indices have included neighborhood racial/ethnic composition as a component of vulnerability, although the measure does not contribute to environmental vulnerability but is rather a proxy for social and structural racist policies that led to segregation, disinvestment, and increased prevalence of other downstream outcomes that can increase vulnerability.[21,22] Composite metrics designed to understand the social and structural drivers of environmental vulnerability should include the downstream effects of racism rather than using the neighborhood racial composition as a proxy for these effects. Finally, some indices like the SVI and the EPA/CDC Environmental Justice Index disaggregate vulnerability across domains but do not provide a framework for users to adapt the indices for their specific purpose, limiting their flexibility for theory-based construction.[23]

Previous research has used an approach called the Toxicological Prioritization Index (ToxPi)[24] to combine multiple geospatial features into vulnerability indices across various spatial scales. In contrast to purely data-driven approaches commonly used in index construction,[10,14,25] ToxPi allows researchers to subjectively retain features based on the specific research question of interest. Furthermore, ToxPi allows characterization of contributions to vulnerability across distinct domains, providing an ability to investigate whether neighborhoods may be experiencing

different contributors to vulnerability and enabling tailored intervention approaches. For example, Bhandari et al applied ToxPi to the Houston-Galveston-Brazoria region to create a five-domain index designed to assist communities in developing plans to address health effects due to natural disasters and industrial activity.[26] A similar framework was used by the National Institute of Environmental Health Science when constructing their COVID Vulnerability Index, to visualize national vulnerability to COVID-19[27], again selecting domains based on their specific goals. Both prior approaches included environmental pollution as a component of their indices for overall vulnerability. While appropriate in some contexts, this limits the ability to investigate potential effect modification of the relationship between pollution and health outcomes by vulnerability.

The objective of this work was to apply the ToxPi approach to construct an adaptable multidomain neighborhood environmental vulnerability index (NEVI) in an urban center. We aimed to characterize the overall magnitude of and identify patterns in neighborhood environmental vulnerability within a large, demographically and socioeconomically diverse, and densely-populated urban area. This index was compared to previously-used indices to characterize the additional information gained from using a multidomain index and assess whether the index varies when constructed using downstream effects of racism rather than neighborhood racial/ethnic composition.

METHODS

Study Area

We focused on New York City (NYC), an urban center with intra-urban variability in a number of social and structural features that can potentially contribute to environmental vulnerability.[28] To capture variation within NYC, we conceptualized our neighborhoods as

census tracts, a relatively small spatial areal unit that could be aggregated into larger areas to correspond to other definitions of a neighborhood. In NYC, there are five distinct boroughs or counties (e.g. Bronx, Brooklyn, Manhattan, Queens, Staten Island), each with a unique development history that has shaped their current sociodemographic profile under their respective municipal jurisdictions. Thus, to maximize the policy and public health relevance of the NEVI, we describe its use in the context of these five boroughs.

Of the 2,167 census tracts in NYC, we excluded 51 tracts that had a population of less than twenty people and 30 tracts that had a population of at least twenty people but had missing data for at least one feature. The majority of these 30 census tracts were in non-residential areas, such as construction sites, parks, and areas with institutions like prisons or universities. As a result, we ultimately included 2,086 census tracts in the NEVI. This analysis is part of a research protocol reviewed and approved by the Columbia University Institutional Review Board.

Construction of Neighborhood Environmental Vulnerability Index

To construct the NEVI, we used data from the 2015-2019 U.S. Census American Community Survey (ACS) 5-year estimates and the 2020 data release from the Centers for Disease Control and Prevention (CDC) PLACES Project. ToxPi is a tool to integrate and visualize data across multiple domains.[29] Subject-matter knowledge is used throughout the process, including the selection of features, grouping the features into subdomains (and subdomains into domains), and weighting of all components to generate the overall index. To select our domains, subdomains, and distinct area-level features, we conducted a literature search of previous research on social and structural drivers of vulnerability to environmental pollution.[5,6,17,19,30–32] Furthermore, we reviewed previously published vulnerability indices to compile and adapt domains from previous applications of ToxPi for our purpose.[10,25–27,33] For

example, the HGBEnviroScreen tool included a social vulnerability domain that we disaggregated into demographic and economic domains and three other domains that measured environmental risks and exposures that we did not include to allow for later potential effect modification analyses by those environmental measures.[26] Guided by the literature and subject-matter knowledge of the research team, we selected 4 primary domains composed of 24 subdomains and 54 distinct area-level features that were hypothesized to contribute to potential environmental vulnerability. The four primary domains were 1) demographics, 2) economic indicators ("economic"), 3) residential characteristics and density ("residential"), and 4) health behavior, outcomes, preventative practices, and access ("health status") (Table 1). Features within the demographics, economic, and residential domains were from the U.S. Census American Community Survey, while features within the primary domain of health status were drawn from the CDC PLACES Project (model-based estimates).[34,35]

The demographic domain consisted of 7 subdomains: age distribution, female-led households, nativity, disability, single parent households, transportation behaviors, and single households. For our primary analyses, we did not include racial and ethnic composition because it is systemic racism manifested as racist policies rather than racial or ethnic identity that lead to greater social and economic stressors and environmental vulnerability.[36] Instead, we only included stressors in the index that reflect both individual-level aggregates and neighborhood-level characteristics. We conducted a sensitivity analysis to see if the NEVI would vary upon the inclusion of race and ethnicity composition. The economic domain included 6 subdomains: income and poverty status, occupation, income inequality, employment status, education, and vehicle availability. The residential domain included 7 subdomains: population density, group quarters, household density, age of housing structure, number of units in housing structure,

residential mobility, and housing vacancy. Finally, the health status domain included 4 subdomains: health behaviors, health conditions, prevention practices, and health insurance status.

We used the ToxPi Graphical User Interface to calculate the overall NEVI and domain-specific scores, conceptualizing each slice as a subdomain within our four domains.[37] We first standardized each feature with z-scores. Next, the software summed the values across the features within each subdomain before transforming the summed values by subtracting the minimum value and dividing by the range of the values.[38] The resulting subdomain scores were then multiplied by the weights (specified in Table 1) to calculate the overall NEVI.[38] While software calculates the overall NEVI and subdomain-specific scores, we manually calculated domain-specific scores by averaging the subdomain-specific scores within each of the four domains. Because ToxPi converts negative values to zero, we re-centered features so that the minimum value would equal zero, resulting in values that were zero or greater across all features. We coded features to ensure that greater (i.e., more positive) values would indicate greater vulnerability. The final NEVI overall and domain-specific scores ranged from 0 to 1.

Comparison to Existing Indices: NDI and SVI

For comparison, we constructed an adapted Neighborhood Deprivation Index (NDI) originally developed by Messer et al. from the 2015-2019 U.S. Census ACS and used the Social Vulnerability Index (SVI) prepared by CDC.[39] To construct a tract-level NDI that reflects heterogeneity across spatial strata of NYC, we adapted methods from both Messer et al. and Shmool et al (described in greater detail in the supplementary material).[14,39] Briefly, after conducting successive principal component analyses (PCA), 7 features were retained from a pool of candidate indicators of area-level deprivation: percent population with at least Bachelor's

level education, percent unemployed, adults in management or professional occupations, households in poverty (< 200% Federal Poverty Line), families with annual income less than $35,000, households with public assistance income, and proportion of residents who were persons of color (Supplemental Table S1). These seven features were included in a final PCA and the resulting first unrotated principal component was used as the tract-level NDI.

Descriptive and Statistical Analyses

We calculated and described the distributions of the NEVI (overall, by domain, and by borough), NDI, and SVI. Next, we mapped both the overall index and the domain-specific scores. To compare the NEVI to the NDI and SVI, we visualized distributions of quartiles of the indices in adjacent maps and calculated Spearman correlation coefficients between the NEVI and NDI/SVI across domain and borough. To identify common patterns in the NEVI subdomains across census-tracts, we conducted a hierarchical clustering analysis with complete linkage and selected the optimal number of clusters using the Gap-statistic.[40] The composition of resulting clusters was compared using heat maps of the median subdomain scores, standardized by feature to compare relative scores across clusters. As a sensitivity analysis, all analyses were repeated after including racial and ethnic composition features within the demographic domain of the NEVI. Specifically, we included 4 features describing the proportion of residents who identified as Hispanic/Latino of any race, Black non-Hispanic/Latino, Asian non-Hispanic/Latino, and other non-white Race non-Hispanic/Latino. Finally, we summarized overall and domain-specific vulnerability by high racial and ethnic composition in a given neighborhood, defined as having a racial and ethnic composition higher than the citywide median proportion. All data preprocessing and analyses were completed in R version 4.0.2. We used the *tidycensus* package to download data from the U.S. Census,[41] the *nycgeo* package to download shapefiles of NYC census tracts

included in our choropleth maps,[42] and the *psych* package to perform PCA.[43] All programming code and data are available on Github jstingone/nevi.

RESULTS

Overall NEVI and Domain-Specific Scores

Figure 1 displays the distribution of the overall NEVI along with the domain-specific scores in separate maps. Numeric estimates of the distribution of the overall index and domain-specific scores for all census tracts are provided in Supplemental Table S2.

As shown in Figure 1 A-D, the NEVI had variability in both overall magnitude and distribution of the domains when comparing the index across boroughs (Supplemental Table S2). The Bronx, an outer borough with mid-range population density and the lowest median income in NYC, had the highest median index (0.43, IQR = 0.35-0.48), while Staten Island, the borough with the lowest population-density and second highest median income, had the lowest median index (0.25, IQR = 0.23-0.30). Consistent with the overall index, median domain scores were the highest in the Bronx for the demographics (0.38, IQR = 0.32-0.41), economic (0.48, IQR = 0.37-0.57), and health status (0.49, IQR = 0.35-0.56) domains and the lowest in Staten Island for the demographics (0.21, IQR = 0.19-0.26), economic (0.25, IQR = 0.22-0.30), and residential (0.21, IQR = 0.18-0.25) domains (Supplemental Table S2, Supplemental Figure S1). However, the median residential domain score was the highest (0.41, IQR = 0.37-0.44) and median health status domain score was the lowest (0.24, IQR = 0.16-0.42) in Manhattan, the borough with the highest population-density but also the highest median income.

Comparisons Between NEVI, NDI, and SVI

Overall, there were strong, positive correlations between NEVI and NDI ($r = 0.91$, $p < 0.001$) and NEVI with SVI ($r = 0.87$, $p < 0.001$). However, there were some notable differences when visually comparing the distribution patterns of these three indices by borough, including less agreement between the indices in Queens, the most ethnically-diverse borough[44] (Figure 2). Correlations between the overall NEVI and NDI/SDI were lower in Queens than other areas, especially for the NDI (Supplemental Figure S2). The NDI did not explain as much variance in Queens as in other boroughs. Correlations between the domain-specific scores and the NDI and SVI were high, except between the NEVI residential score which consistently showed lower correlations with both the NDI and the SVI across boroughs.

Vulnerability Profiles of Census Tracts

As an example of how domain-specific features contribute differently to NEVI and support a targeted and adaptable public health approach, in Figure 3, we show vulnerability profiles for two sample census tracts in NYC with similar overall NEVI scores (0.239): Census Tract 671 and Census Tract 519. Despite having similar overall scores, the patterns across subdomain scores are different with greater economic vulnerability in Census Tract 519 (0.64 vs 0.23 economic sub-score) and slightly greater scores in residential and health-related features in Census Tract 671.

Vulnerability Profiles of NEVI Clusters

Our clustering analysis revealed 6 distinct patterns in NEVI domain-scores across census-tracts (Figure 3), specifically two low vulnerability clusters, one medium vulnerability cluster and three high vulnerability clusters: Cluster 1) low across all domains (median NEVI = 0.26, n = 435); Cluster 2) primarily low but greater residential vulnerability and very low health

vulnerability (median NEVI = 0.28, n = 331); Cluster 3) medium vulnerability across all domains (median NEVI = 0.35, n = 903); Cluster 4) high vulnerability across demographic, residential and economic domains (median NEVI = 0.44, n = 5); Cluster 5) high vulnerability across economic, residential and health status domains (median NEVI = 0.45, n = 241); and Cluster 6) high vulnerability across demographic, economic and health status domains (median NEVI = 0.47, n = 171). Figure 4 provides a heatmap to enable greater visualization of differences in specific subdomains across clusters. For example, the low vulnerability Cluster 2 had consistently lower health status vulnerability scores than Cluster 1 but tended to have higher scores in age of housing structure, units in housing structure, residential mobility, and vacancy within the residential domain. Additionally, the high vulnerability Cluster 6 has consistently lower residential vulnerability than the other high vulnerability clusters, except for the location of group quarters but higher vulnerability related to employment status and single-parent or female-led households.

Sensitivity Analyses by Race and Ethnicity

The NEVI did not change considerably when including racial and ethnic composition of census tracts as features. Supplemental Figure S3 illustrates that the mean values and distributions for the overall NEVI did not? change in NYC as a whole or by borough when including racial and ethnic composition. The Demographic domain scores changed slightly, particularly in Queens where inclusion of race/ethnicity features caused slightly higher mean domain scores. Cluster patterns changed slightly when including race/ethnicity features in the NEVI (Supplemental Figure S4). For example, some medium vulnerability census-tracts were regrouped with the low vulnerability clusters, leading to more low-vulnerability areas. Some previously high vulnerability census-tracts were reclassified as a medium vulnerability cluster.

This caused small changes in the overall vulnerability scores of clusters, but the majority of census-tracts remained within the same cluster.

In comparisons of vulnerability across race and ethnicity, overall vulnerability was lowest in White (neighborhood composition higher than the citywide median proportion) neighborhoods (median NEVI = 0.29), moderate in Asian, Multiple Race, and Native Hawaiian/Pacific Islander neighborhoods (median NEVI = 0.32), and highest in American Indian/Alaskan Native, Black, and Hispanic neighborhoods (median NEVI = 0.35, 0.37, and 0.38, respectively). When disaggregating by domain, residential vulnerability was similar across racial and ethnic groups, and economic and health status vulnerability was especially high among American Indian/Alaskan Native, Black, and Hispanic neighborhoods.

DISCUSSION

The NEVI characterized neighborhood environmental vulnerability across NYC and quantified contributions to vulnerability across four domains: 1) demographics, 2) economic indicators, 3) residential characteristics, and 4) health status. There was general agreement between the overall NEVI and previously developed indices for deprivation (NDI) and social vulnerability (SVI). However, the NEVI offered additional benefits, including its adaptable construction and additional information about the types of vulnerability across a geographic area. Together, these can inform efforts to enhance subsequent environmental justice research and interventions aimed at reducing vulnerability to environmental pollutant exposures in neighborhoods at the hyperlocal scale.

Our index characterized potential contributions towards vulnerability from different domains, in contrast to previous indices that only provide an overall score for vulnerability or deprivation.[10,45] Prior studies that used a single measurement for deprivation were limited in their ability to describe which aspect of the neighborhood environment were related to their results.[46,47] Knowing which domains contribute to higher neighborhood vulnerabilities would facilitate adaptable population health planning and research to allocate certain types of resources to targeted neighborhoods even if the domain or specific features are not themselves modifiable, such as tailoring interventions to specific vulnerable demographic groups in a given neighborhood. Our clustering analysis revealed six primary patterns of potential vulnerability, which differentiated boroughs in more detail than conveyed by the absolute magnitude. For example, clusters 1 and 2 both had low overall vulnerability, but cluster 2 had even lower vulnerability due to existing health conditions but greater vulnerability associated with residential characteristics. This could point to different routes of intervention to reduce and/or ameliorate related health impacts of environmental pollution. For example, when designing an intervention to reduce the health effects of air pollution, one may target better chronic disease management in Cluster 1 while focusing on improved ventilation in the older housing stock of Cluster 2.

Notably, we excluded measures of racial and/or ethnic composition in the construction of our final index. Communities of color disproportionately experience worse health outcomes resulting from exposure to environmental pollutants and greater amounts of social and economic stressors across the US. For example, racial segregation along with historic racist land use policies such as red-lining have led to concentrated poverty and structural disinvestment in racially minoritized and low income urban neighborhoods that could increase their vulnerability

to the health impacts of environmental pollutants.[31,32,48–51] Rather than include the neighborhood racial and ethnic composition as a feature in the NEVI, we included the downstream effects of racism, such as income inequality and degree of chronic disease burden. Our sensitivity analyses revealed that there was little difference in overall NEVI scores based on the inclusion of racial/ethnic composition, likely due to the inclusion of socioeconomic, residential and health-related outcomes that are highly correlated with race through the impacts of racism. A previous analysis of the Climate and Economic Justice Screening Tool, developed by the White House, also found similarity in community rankings regardless of whether racial and ethnic demographics were included.[52] However, other contexts seeking to use an index like the NEVI may want to include racial and/or ethnic composition as proxies for other measures of racism and structural inequities that remain unmeasured. This separation of race and ethnic composition from NEVI allowed us not only to identify higher overall vulnerability in communities of color but also that the variation in vulnerability was driven by differences in economic indicators and health status, reflective of racial and ethnic disparities seen in the US. We recommend that future studies further validate the NEVI across other exposures and in causal analyses with health outcomes.

The NEVI and NDI/SVI were generally similar across boroughs and domains, except the NEVI residential domain (i.e., lower correlations), likely because no or fewer housing-related features linked to environmental hazards was included in the NDI/SVI.[5,53,54] Overall, NEVI had a lower correlation with NDI/SVI in Queens and the NDI explained less total variation in Queens compared to other NYC boroughs, possibly because Queens is the most demographically and socioeconomically diverse borough in NYC.[44,55] A limitation of the NEVI is the need to exclude some Census tracts due to low population or missing data for at least one feature, limiting its

potential applicability in those areas. These areas are composed of institutions like prisons, in which people may have increased environmental vulnerability.[56,57] Additionally, we have currently only used the NEVI in one large, demographically and socioeconomically diverse, and densely populated urban center, potentially limiting the generalizability of our observed patterns in vulnerability to other urban areas. However, the adaptable NEVI creation process would facilitate the creation of the index in other urban areas to ascertain their distinct patterns of potential vulnerability.

The NEVI provides several advantages that increases its interpretability and utility. First, we were able to characterize the magnitude of potential environmental vulnerability, quantify the contributions from various domains, and identify patterns of vulnerability across an urban area. These provide additional interpretability that would better inform public health planning and additional domain-specific scores to evaluate in research studies incorporating vulnerability measures. Next, ToxPi enables the use of background knowledge to inform feature selection and weighting approaches, unlike indices that use data-driven methods to retain features and may drop features of theoretical importance but are highly correlated with other features. Being able to choose features and specify weights for contributions to vulnerability allows for a hybrid approach in which more hypothesis-driven information may be incorporated in the index. The process of customizing the NEVI in choosing features and specifying weights to fit specific hypotheses is transparent, facilitating discussion and critique. For example, we selected four domains that were hypothesized to reflect distinct area-level characteristics that could contribute to environmental vulnerability, distinct from the concentrations of environmental pollution that may exist in a neighborhood that might be included in other indices to identify areas most at-risk of specific health outcomes. In contrast to other indices like the SVI that includes four pre-set

themes, the use of ToxPi enables the addition of domains and/or features in a clear manner, that promotes flexibility for the investigator and transparency for communities and stakeholders seeking to interpret the vulnerability index.

Conclusion

We developed a neighborhood-level index to measure vulnerability to health impacts from environmental pollution for New York City and found that our NEVI was generally consistent with previously-developed deprivation and vulnerability scores. However, the NEVI was additionally able to characterize contributions to vulnerability across multiple domains, providing greater insight into intra-urban variation in vulnerability. Specifically, the customization option of this index-building approach allowed theory-based analysis of specific features/domain contribution to the index score (as we explored with racial composition). This metric can be used to inform targeted public and environmental health research and practice at the hyperlocal scale and improve our understanding of the impact of environmental exposures on communities with varying levels of vulnerability.

ACKNOWLEDGEMENTS

The authors would like to thank Denise Nguyen and Tamia E. Wellington for their help with preliminary analyses and members of Health Data for New York City (HD4NYC) for their feedback on the creation of the NEVI and this manuscript.

Funding

Research described in this article was conducted under contract to the Health Effects Institute (HEI), an organization jointly funded by the United States Environmental Protection Agency (EPA) (Assistance Award No. CR-83998101) and certain motor vehicle and engine manufacturers (#4985-RFA20-1B/21-8) The contents of this article do not necessarily reflect the views of HEI, or its sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and engine manufacturers. Additional funding was received from the National Institute of Environmental Health Sciences, (T32ES007322 (SPU), R00ES027022 (JAS, SPU), P30ES009089 (SLD, JAS), P30ES023515 (PES), and R01ES030717 (PES)), National Heart Lung and Blood Institute (K01HL140216 (SLD)), National Institute of Diabetes, Digestive and Kidney Disorders (K01DK107791 (SSA), P30DK111022 (EC)) and the Robert Wood Johnson Foundation-Amos Medical Faculty Development Award (SLD).) This study is a product of the HD4NYC project, a multi-institutional research platform to improve health equity in NYC, funded by Robert Wood Johnson Foundation to NYC DOHMH and New York Academy of Medicine (NYAM). Although the research described in this article was supported by an initiative funded in part by Robert Wood Johnson Foundation, it has not been subjected to review and, therefore, does not necessarily reflect the views of the foundation, and no official endorsement should be inferred.

REFERENCES

1. Brunekreef B, Holgate ST. Air pollution and health. *The Lancet*. 2002;360(9341):1233-1242. doi:10.1016/S0140-6736(02)11274-8

2. Castagna A, Mascheroni E, Fustinoni S, Montirosso R. Air pollution and neurodevelopmental skills in preschool- and school-aged children: A systematic review. *Neurosci Biobehav Rev*. 2022;136:104623. doi:10.1016/j.neubiorev.2022.104623

3. Halden RU. Plastics and Health Risks. *Annu Rev Public Health*. 2010;31(1):179-194. doi:10.1146/annurev.publhealth.012809.103714

4. Mudgal V, Madaan N, Mudgal A, Singh RB, Mishra S. Effect of toxic metals on human health. *Open Nutraceuticals J*. 2010;3(1):94-99. doi:10.2174/18763960010030100094

5. Rauh VA, Landrigan PJ, Claudio L. Housing and health: intersection of poverty and environmental exposures. *Ann N Acad Sci*. 2008;1136:276-288. doi:10.1196/annals.1425.032

6. Braubach M, Fairburn J. Social inequities in environmental risks associated with housing and residential location—a review of evidence. *Eur J Public Health*. 2010;20(1):36-42. doi:10.1093/eurpub/ckp221

7. Jerrett M, McConnell R, Wolch J, et al. Traffic-related air pollution and obesity formation in children: a longitudinal, multilevel analysis. *Environ Health*. 2014;13(1):49. doi:10.1186/1476-069X-13-49

8. Padula AM, Rivera-Núñez Z, Barrett ES. Combined Impacts of Prenatal Environmental Exposures and Psychosocial Stress on Offspring Health: Air Pollution and Metals. *Curr Environ Health Rep*. 2020;7(2):89-100. doi:10.1007/s40572-020-00273-6

9. Diez Roux AV, Mair C. Neighborhoods and health. *Ann N Y Acad Sci*. 2010;1186(1):125-145. doi:10.1111/j.1749-6632.2009.05333.x

10. Messer LC, Laraia BA, Kaufman JS, et al. The Development of a Standardized Neighborhood Deprivation Index. *J Urban Health*. 2006;83(6):1041-1062. doi:10.1007/s11524-006-9094-x

11. Lett E, Stingone JA, Claudio L. The Combined Influence of Air Pollution and Home Learning Environment on Early Cognitive Skills in Children. *Int J Environ Res Public Health*. 2017;14(11):1295. doi:10.3390/ijerph14111295

12. McGuinn LA, Windham GC, Messer LC, et al. Air pollution, neighborhood deprivation, and autism spectrum disorder in the Study to Explore Early Development. *Environ Epidemiol Phila Pa*. 2019;3(5):e067. doi:10.1097/ee9.0000000000000067

13. O'Lenick CR, Winquist A, Mulholland JA, et al. Assessment of neighbourhood-level socioeconomic status as a modifier of air pollution–asthma associations among children in Atlanta. *J Epidemiol Community Health*. 2017;71(2):129. doi:10.1136/jech-2015-206530

14. Shmool JLC, Bobb JF, Ito K, et al. Area-level socioeconomic deprivation, nitrogen dioxide exposure, and term birth weight in New York City. *Environ Res*. 2015;142:624-632. doi:10.1016/j.envres.2015.08.019

15. CDC SVI Documentation 2018 | Place and Health | ATSDR. Published June 22, 2021. Accessed June 30, 2021. https://www.atsdr.cdc.gov/placeandhealth/svi/documentation/SVI_documentation_2018.html

16. The White House Council on Environmental Quality. Climate and Economic Justice Screening Tool. Council on Environmental Quality. Published January 5, 2022. https://screeningtool.geoplatform.gov/en#3/33.47/-97.5

17. Balbus JM, Malina C. Identifying Vulnerable Subpopulations for Climate Change Health Effects in the United States. *J Occup Environ Med*. 2009;51(1). doi:10.1097/JOM.0b013e318193e12e

18. Lundgren K, Kuklane K, Gao C, Holmer I. Effects of Heat Stress on Working Populations when Facing Climate Change. *Ind Health*. 2013;51(1):3-15. doi:10.2486/indhealth.2012-0089

19. Makri A, Stilianakis NI. Vulnerability to air pollution health effects. *Int J Hyg Environ Health*. 2008;211(3):326-336. doi:10.1016/j.ijheh.2007.06.005

20. An R, Xiang X. Social Vulnerability and Obesity among U.S. Adults. *Int J Health Sci IJHS*. 2015;3(3). doi:10.15640/ijhs.v3n3a2

21. Gee GC, Ford CL. STRUCTURAL RACISM AND HEALTH INEQUITIES: Old Issues, New Directions. *Bois Rev Soc Sci Res Race*. 2011;8(1):115-132. doi:10.1017/S1742058X11000130

22. Williams DR, Priest N, Anderson NB. Understanding associations among race, socioeconomic status, and health: Patterns and prospects. *Health Psychol Off J Div Health Psychol Am Psychol Assoc*. 2016;35(4):407-411. doi:10.1037/hea0000242

23. Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry. Environmental Justice Index. Agency for Toxic Substances and Disease Registry. Published January 5, 2022. https://screeningtool.geoplatform.gov/en#3/33.47/-97.5

24. Reif David M., Martin Matthew T., Tan Shirlee W., et al. Endocrine Profiling and Prioritization of Environmental Chemicals Using ToxCast Data. *Environ Health Perspect*. 2010;118(12):1714-1720. doi:10.1289/ehp.1002180

25. Lalloué B, Monnez JM, Padilla C, et al. A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *Int J Equity Health*. 2013;12(1):21. doi:10.1186/1475-9276-12-21

26. Bhandari S, Lewis PGT, Craft E, Marvel SW, Reif DM, Chiu WA. HGBEnviroScreen: Enabling Community Action through Data Integration in the Houston-Galveston-Brazoria Region. *Int J Environ Res Public Health*. 2020;17(4):1130. doi:10.3390/ijerph17041130

27. Marvel SW, House JS, Wheeler M, et al. The COVID-19 Pandemic Vulnerability Index (PVI) Dashboard: Monitoring County-Level Vulnerability Using Visualization, Statistical Modeling, and Machine Learning. *Environ Health Perspect*. 2021;129(1):17701. doi:10.1289/EHP8690

28. New York City Department of Health and Mental Hygiene. *New York City Neighborhood Health Atlas*.; 2018. Accessed March 3, 2022. http://www1.nyc.gov/site/doh/health/neighborhood-health/nyc-neighborhood-health-atlas.page

29. Reif DM, Sypa M, Lock EF, et al. ToxPi GUI: an interactive visualization tool for transparent integration of data from diverse sources of evidence. *Bioinforma Oxf Engl*. 2013;29(3):402-403. doi:10.1093/bioinformatics/bts686

30. Diez Roux AV. Investigating Neighborhood and Area Effects on Health. *Am J Public Health*. 2001;91(11):1783-1789. doi:10.2105/ajph.91.11.1783

31. Gee GC, Payne-Sturges DC. Environmental Health Disparities: A Framework Integrating Psychosocial and Environmental Concepts. *Environ Health Perspect*. 2004;112(17):1645-1653. doi:10.1289/ehp.7074

32. Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding The Cumulative Impacts Of Inequalities In Environmental Health: Implications For Policy. *Health Aff (Millwood)*. 2011;30(5):879-887. doi:10.1377/hlthaff.2011.0153

33. Gay JL, Robb SW, Benson KM, White A. Can the Social Vulnerability Index Be Used for More Than Emergency Preparedness? An Examination Using Youth Physical Fitness Data. *J Phys Act Health*. 2016;13(2):121-130. doi:10.1123/jpah.2015-0042

34. Centers for Disease Control and Prevention, Division of Population Health. PLACES: Local Data for Better Health. CDC. Published January 5, 2022. https://www.cdc.gov/places/index.html

35. U.S.Census. American Community Survey Data. Published 2022. https://www.census.gov/programs-surveys/acs/data.html

36. Lett E, Asabor E, Beltrán S, Michelle Cannon A, Arah OA. Conceptualizing, Contextualizing, and Operationalizing Race in Quantitative Health Sciences Research. *Ann Fam Med*. 2022;20(2):157-163. doi:10.1370/afm.2792

37. Marvel SW, To K, Grimm FA, Wright FA, Rusyn I, Reif DM. ToxPi Graphical User Interface 2.0: Dynamic exploration, visualization, and sharing of integrated data models. *BMC Bioinformatics*. 2018;19(1):80. doi:10.1186/s12859-018-2089-2

38. Auerbach S, Filer D, Reif D, et al. Prioritizing Environmental Chemicals for Obesity and Diabetes Outcomes Research: A Screening Approach Using ToxCast™ High-Throughput Data. *Environ Health Perspect*. 2016;124(8):1141-1154. doi:10.1289/ehp.1510456

39. Messer LC, Laraia BA, Kaufman JS, et al. The Development of a Standardized Neighborhood Deprivation Index. *J Urban Health*. 2006;83(6):1041-1062. doi:10.1007/s11524-006-9094-x

40. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Ser B Stat Methodol*. 2001;63(2):411-423. doi:10.1111/1467-9868.00293

41. Walker K, Herman M, Eberwein K. tidycensus: Load US Census Boundary and Attribute Data as "tidyverse" and 'sf'-Ready Data Frames. Published online September 23, 2021. Accessed October 8, 2021. https://CRAN.R-project.org/package=tidycensus

42. nycgeo-package: nycgeo: Spatial Data Files for NYC in mfherman/nycgeo: Spatial Data Files for NYC. Published 2019. Accessed October 8, 2021. https://rdrr.io/github/mfherman/nycgeo/man/nycgeo-package.html

43. Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. Published online September 22, 2021. Accessed October 8, 2021. https://CRAN.R-project.org/package=psych

44. McGovern B, Frazier JW. Evolving ethnic settlements in Queens: Historical and current forces reshaping human geography. *Focus Geogr*. 2015;58(1):11.

45. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A Social Vulnerability Index for Disaster Management: *J Homel Secur Emerg Manag*. 2011;8(1). doi:10.2202/1547-7355.1792

46. Andrews MR, Tamura K, Claudel SE, et al. Geospatial Analysis of Neighborhood Deprivation Index (NDI) for the United States by County. *J Maps*. 2020;16(1):101-112. doi:10.1080/17445647.2020.1750066

47. Stimpson JP, Ju H, Raji MA, Eschbach K. Neighborhood deprivation and health risk behaviors in NHANES III. *Am J Health Behav*. 2007;31(2):215-222. doi:10.5555/ajhb.2007.31.2.215

48. Pastor M, Sadd J, Hipp J. Which Came First? Toxic Facilities, Minority Move-In, and Environmental Justice. *J Urban Aff*. 2001;23(1):1-21. doi:10.1111/0735-2166.00072

49. Phelan JC, Link BG. Is Racism a Fundamental Cause of Inequalities in Health? *Annu Rev Sociol*. 2015;41(1):311-330. doi:10.1146/annurev-soc-073014-112305

50. Williams DR, Lawrence JA, Davis BA. Racism and Health: Evidence and Needed Research. *Annu Rev Public Health*. 2019;40(1):105-125. doi:10.1146/annurev-publhealth-040218-043750

51. Williams DR, Collins C. Racial Residential Segregation: A Fundamental Cause of Racial Disparities in Health. *Public Health Rep*. 2001;116(5):404-416. doi:10.1093/phr/116.5.404

52. Naveena Sadasivam, Clayton Aldern. The White House excluded race from its environmental justice tool. We put it back in. *Grist*. Published online February 24, 2022. Accessed March 3, 2022. https://grist.org/equity/climate-and-economic-justice-screening-tool-race/

53. Jacobs DE. Environmental Health Disparities in Housing. *Am J Public Health*. 2011;101(S1):S115-S122. doi:10.2105/AJPH.2010.300058

54. Breysse P, Farr N, Galke W, Lanphear B, Morley R, Bergofsky L. The Relationship between Housing and Health: Children at Risk. *Environ Health Perspect*. 2004;112(15):1583-1588. doi:10.1289/ehp.7157

55. Office of the NY Comptroller. *An Economic Snapshot of Queens*.; 2018. Accessed July 19, 2022. https://www.osc.state.ny.us/files/reports/osdc/pdf/report-1-2019.pdf

56. Leon-Corwin M, McElroy JR, Estes ML, Lewis J, Long MA. Polluting our prisons? An examination of Oklahoma prison locations and toxic releases, 2011–2017. *Punishm Soc*. 2020;22(4):413-438. doi:10.1177/1462474519899949

57. Moran D, Jones PI, Jordaan JA, Porter AE. Does prison location matter for prisoner wellbeing? The effect of surrounding greenspace on self-harm and violence in prisons in England and Wales. *Wellbeing Space Soc*. 2022;3:100065. doi:10.1016/j.wss.2021.100065

TABLES

Table 1. Features Included in the Neighborhood Environmental Vulnerability Index

Domain	Subdomain Represented by Slice (Number of Features, Weight of Overall Index)	Features[a]
Demographics (7 subdomains, 12 features)	Age (2 features, 1/28)	• Age under 18 (%) • Age 65 or older (%)
	Female-led household (1 feature, 1/28)	• Female-led households (%)
	Immigration (4 features, 1/28)	• Speak English "not at all" or "not well" (%) • Foreign-born (%) • Entered U.S. 2010 or later (%) • Not a U.S. citizen (%)
	Disability (1 feature, 1/28)	• With a disability (%)
	Single parent households (1 feature, 1/28)	• Children (age younger than 18) living with one parent (%)
	Transportation (2 features, 1/28)	• Who take public transportation, taxicab, bicycle, walking, or other non-automobile means (%) • Aggregate travel time to work
	Single households (1 feature, 1/28)	• Live alone (%)
Economic Indicators (6 subdomains, 7 features)	Income and poverty (2 features, 1/24)	• Household Income (median) • Income in past 12 months below poverty level (%)
	Occupation (1 feature, 1/24)	• Work in service, natural resources, construction, maintenance, production, transportation, or material moving occupations (%)
	Income inequality (1 feature, 1/24)	• Gini Index of income inequality
	Employment status (1 feature, 1/24)	• Unemployed among those in labor force aged 20-64 (%)
	Education (1 feature, 1/24)	• With less than high school education (%)

	Vehicle availability (1 feature, 1/24)	• % with no vehicle available (%)
Residential Characteristics and Density (7 subdomains, 8 features)	Population density (1 feature, 1/28)	• Population per square mile (%)
	Group quarters (1 feature, 1/28)	• Live in group quarters (%); Group quarters include all locations not considered housing units. These include correctional facilities, nursing homes, mental hospitals, military barracks, group homes, missions, shelters, etc.
	Household density (1 feature, 1/28)	• More than one occupant per room (%)
	Age of housing structure (1 feature, 1/28)	• Age of housing in 2019 (years)
	Units in housing structure (2 features, 1/28)	• Living in structure with 1 attached housing unit, 2 or more units, or in a mobile home, boat, RV, or van (%) • Living in structure with 20 or more units (%)
	Residential mobility (1 feature, 1/28)	• Moved in the past year (%)
	Housing vacancy (1 feature, 1/28)	• Vacancy (%)
Health Behaviors, Conditions, Prevention Practices, and Insurance Status (4 subdomains, 27 features)	Health behaviors (4 features, 1/16)	• Currently smoke • Binge drink • Have no leisure-time physical activity • Sleep for less than 7 hours per day
	Health conditions (14 features, 1/16)	• High blood pressure, 2017 • High blood pressure medication use, among those with high blood pressure, 2017 • Obesity • Cancer, non-skin • Current asthma • Coronary heart disease • Stroke • COPD • Diabetes • High cholesterol, 2017 • Chronic Kidney Disease • Mental health not good for >=14 days • Physical health not good for >=14 days • All teeth lost

Prevention practices (8 features, 1/16)	• Routine checkup within past year • Adult men aged 65 or older who are on date with clinical preventive services (flu shot past year, PPV shot ever, colorectal cancer screening) • Adult women aged 65 or older who are on date with clinical preventive services (flu shot past year, PPV shot ever, colorectal cancer screening, mammogram within past 2 years) • Visit to dentist or dental clinic • Cervical cancer screening among women aged 21-65 • Cholesterol screening, 2017 • Colonoscopy screening among adults aged 50-75 years • Mammography use among women aged 50-74 years
Health insurance status (1 feature, 1/16)	• Lack of health insurance

[a]**Data sources**: U.S. Census American Community Survey (ACS) 2015-2019 5-year estimates, Centers for Disease Control and Prevention (CDC) 2020 data release from the PLACES Project

TABLES

Table 1. Features Included in the Neighborhood Environmental Vulnerability Index[1]

Domain	Subdomain Represented by Slice (Number of Features, Weight of Overall Index)	Features
Demographics (7 subdomains, 12 features)	Age (2 features, 1/28)	• Age under 18 (%) • Age 65 or older (%)
	Female-led household (1 feature, 1/28)	• Female-led households (%)
	Immigration (4 features, 1/28)	• Speak English "not at all" or "not well" (%) • Foreign-born (%) • Entered U.S. 2010 or later (%) • Not a U.S. citizen (%)
	Disability (1 feature, 1/28)	• With a disability (%)
	Single parent households (1 feature, 1/28)	• Children (age younger than 18) living with one parent (%)
	Transportation (2 features, 1/28)	• Who take public transportation, taxicab, bicycle, walking, or other non-automobile means (%) • Aggregate travel time to work
	Single households (1 feature, 1/28)	• Live alone (%)
Economic Indicators (6 subdomains, 7 features)	Income and poverty (2 features, 1/24)	• Household Income (median) • Income in past 12 months below poverty level (%)
	Occupation (1 feature, 1/24)	• Work in service, natural resources, construction, maintenance, production, transportation, or material moving occupations (%)
	Income inequality (1 feature, 1/24)	• Gini Index of income inequality
	Employment status (1 feature, 1/24)	• Unemployed among those in labor force aged 20-64 (%)
	Education (1 feature, 1/24)	• With less than high school education (%)

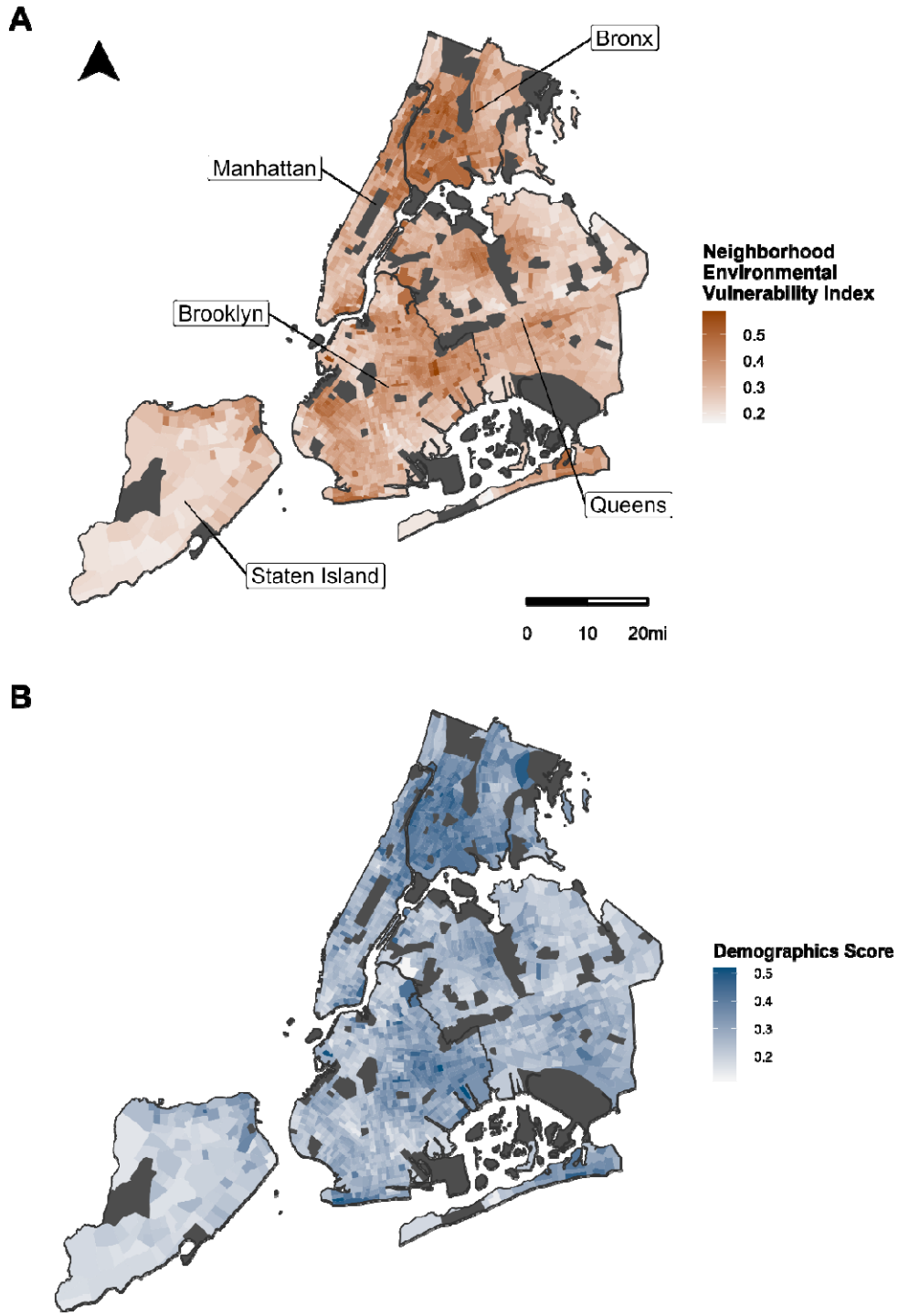
	Vehicle availability (1 feature, 1/24)	• % with no vehicle available (%)
Residential Characteristics and Density (7 subdomains, 8 features)	Population density (1 feature, 1/28)	• Population per square mile (%)
	Group quarters (1 feature, 1/28)	• Live in group quarters (%); Group quarters include all locations not considered housing units. These include correctional facilities, nursing homes, mental hospitals, military barracks, group homes, missions, shelters, etc.
	Household density (1 feature, 1/28)	• More than one occupant per room (%)
	Age of housing structure (1 feature, 1/28)	• Age of housing in 2019 (years)
	Units in housing structure (2 features, 1/28)	• Living in structure with 1 attached housing unit, 2 or more units, or in a mobile home, boat, RV, or van (%) • Living in structure with 20 or more units (%)
	Residential mobility (1 feature, 1/28)	• Moved in the past year (%)
	Housing vacancy (1 feature, 1/28)	• Vacancy (%)
Health Behaviors, Conditions, Prevention Practices, and Insurance Status (4 subdomains, 27 features)	Health behaviors (4 features, 1/16)	• Currently smoke • Binge drink • Have no leisure-time physical activity • Sleep for less than 7 hours per day
	Health conditions (14 features, 1/16)	• High blood pressure, 2017 • High blood pressure medication use, among those with high blood pressure, 2017 • Obesity • Cancer, non-skin • Current asthma • Coronary heart disease • Stroke • COPD • Diabetes • High cholesterol, 2017 • Chronic Kidney Disease • Mental health not good for >=14 days • Physical health not good for >=14 days • All teeth lost

Prevention practices (8 features, 1/16)	• Routine checkup within past year • Adult men aged 65 or older who are on date with clinical preventive services (flu shot past year, PPV shot ever, colorectal cancer screening) • Adult women aged 65 or older who are on date with clinical preventive services (flu shot past year, PPV shot ever, colorectal cancer screening, mammogram within past 2 years) • Visit to dentist or dental clinic • Cervical cancer screening among women aged 21-65 • Cholesterol screening, 2017 • Colonoscopy screening among adults aged 50-75 years • Mammography use among women aged 50-74 years
Health insurance status (1 feature, 1/16)	• Lack of health insurance*

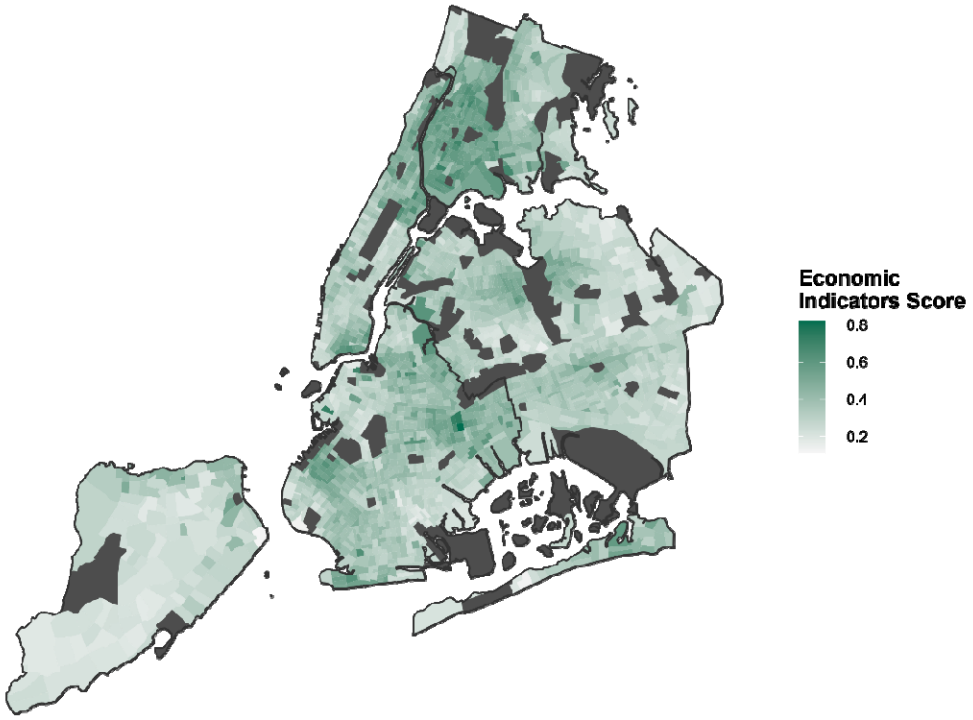
[1]**Data sources**: U.S. Census American Community Survey (ACS) 2015-2019 5-year estimates, Centers for Disease Control and Prevention (CDC) 2020 data release from the PLACES Project

FIGURE CAPTIONS

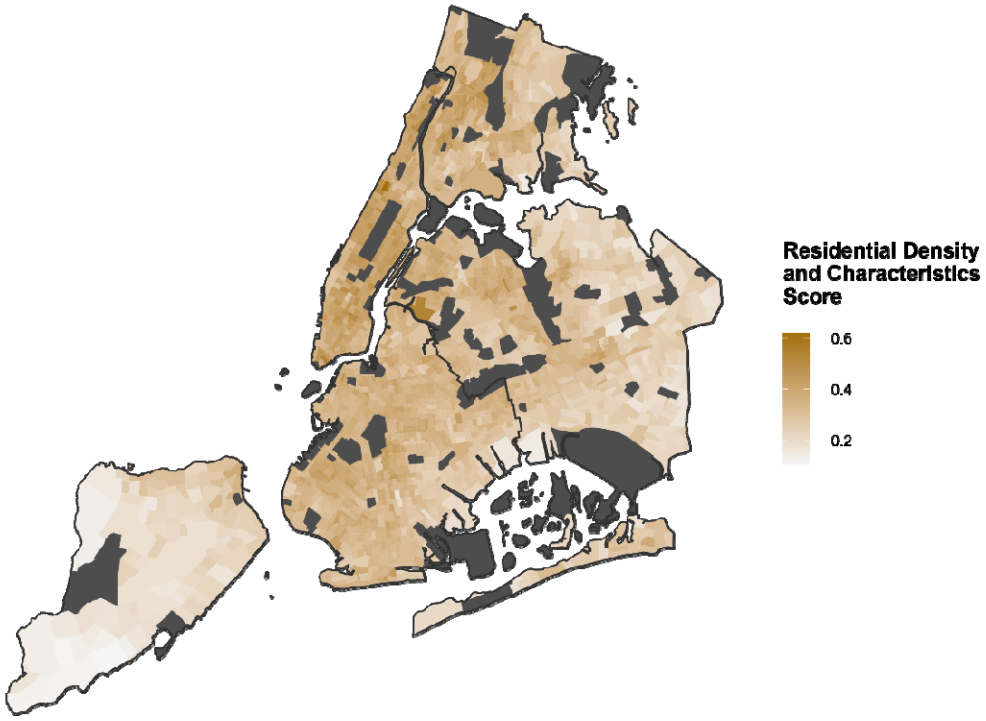
Figure 1. Maps of Neighborhood Environmental Vulnerability Index: Overall Index and Domain-Specific Scores Across New York City by Census Tract, 2015-2019

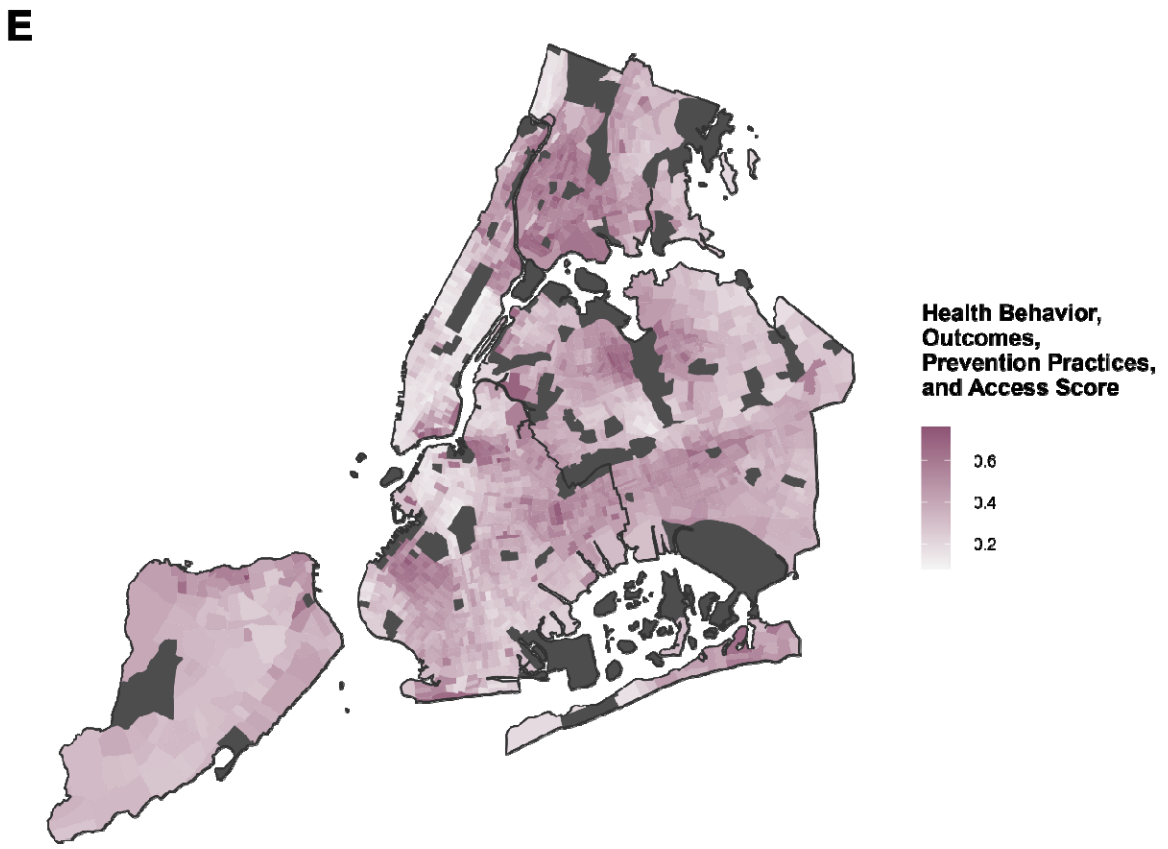


C


Economic
Indicators Score
0.8
0.6
0.4
0.2

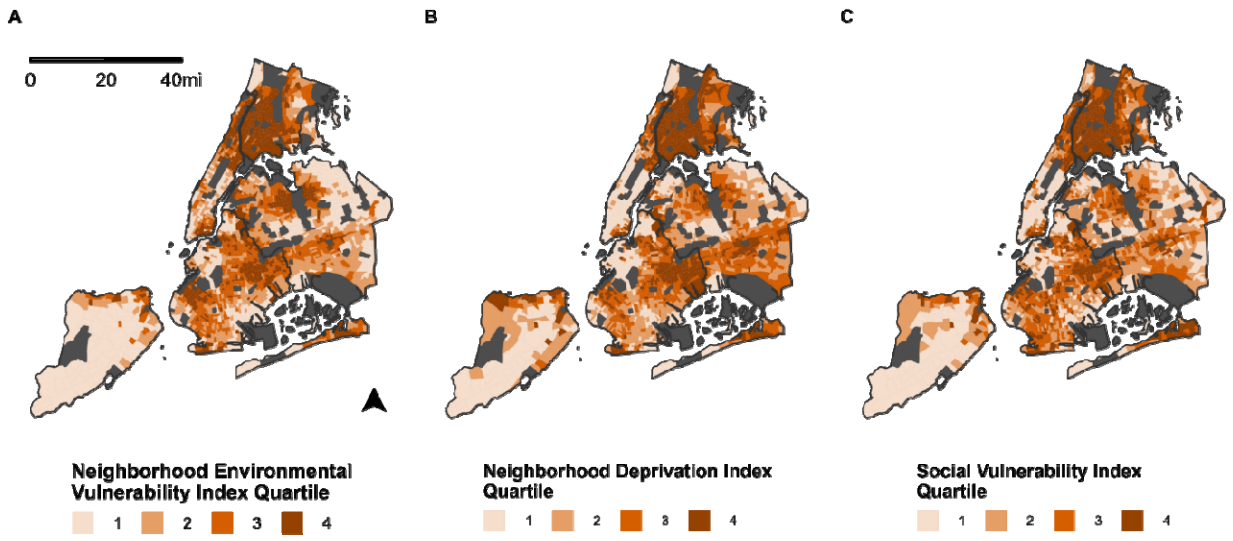
D


Residential Density
and Characteristics
Score
0.6
0.4
0.2



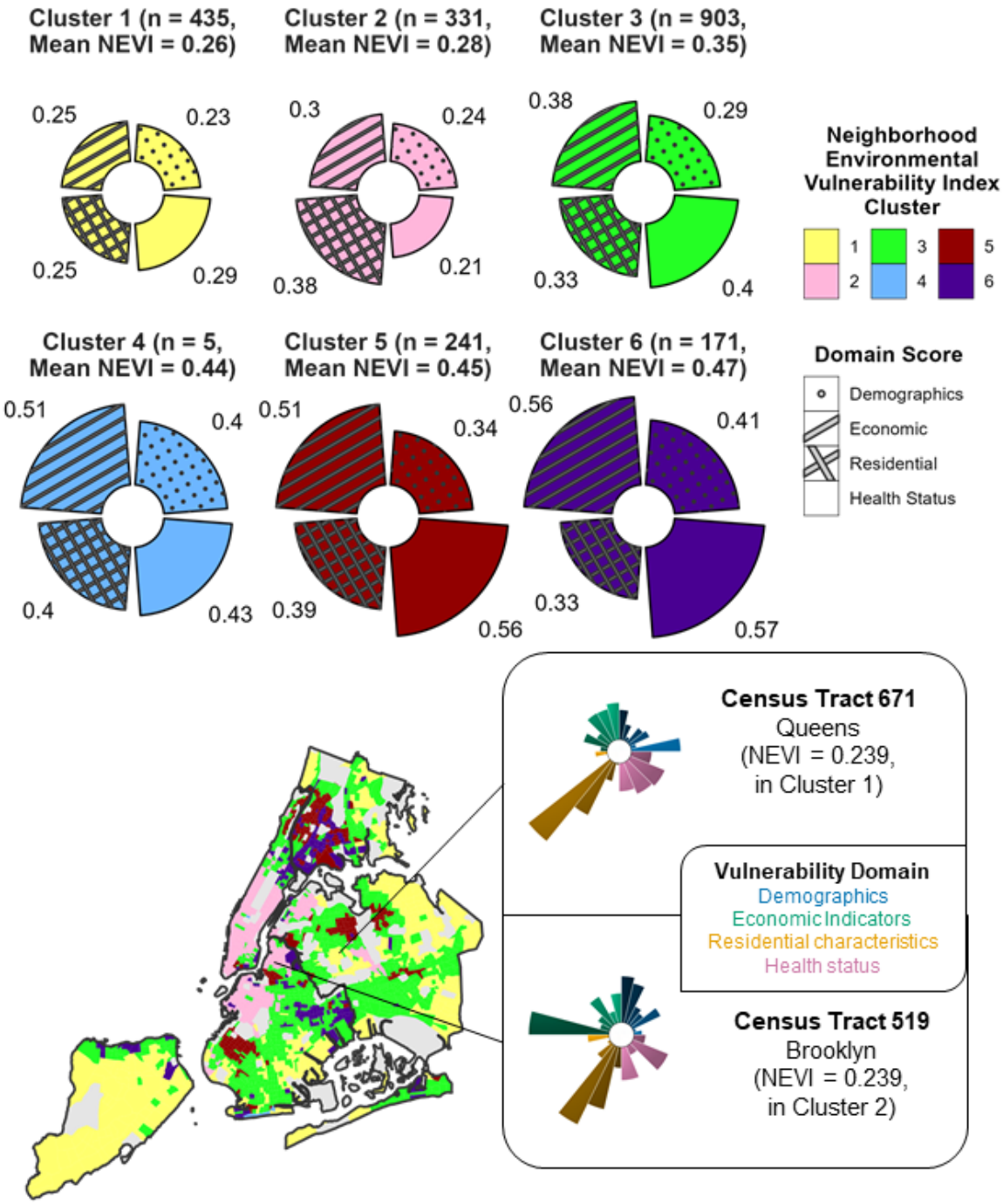
The maps display the distribution of the (A) overall Neighborhood Environmental Vulnerability Index across New York City along its (B-E) domain-specific scores. Areas that were excluded due to low population counts or missing features are shown in dark gray. Data Sources: U.S. Census American Community Survey 2015-2019 5-Year Estimates and Centers for Disease Control and Prevention PLACES Project 2020 Release

Figure 2. Maps of Neighborhood Environmental Vulnerability Index, Neighborhood Deprivation Index, and Social Vulnerability Index Across New York City by Census Tract, 2015-2019



The maps display the distribution of the (A) overall Neighborhood Environmental Vulnerability Index, (B) the Neighborhood Deprivation Index, and (C) the Social Vulnerability Index across New York City. Areas that were excluded due to low population counts or missing features are shown in dark gray. Data Sources: U.S. Census American Community Survey 2015-2019 5-Year Estimates, Centers for Disease Control and Prevention PLACES Project 2020 Release, and Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry 2018 Social Vulnerability index

Figure 3. Map of Clusters with Example Tract Vulnerability Profiles for the Neighborhood Environmental Vulnerability Index across New York City, 2015-2019



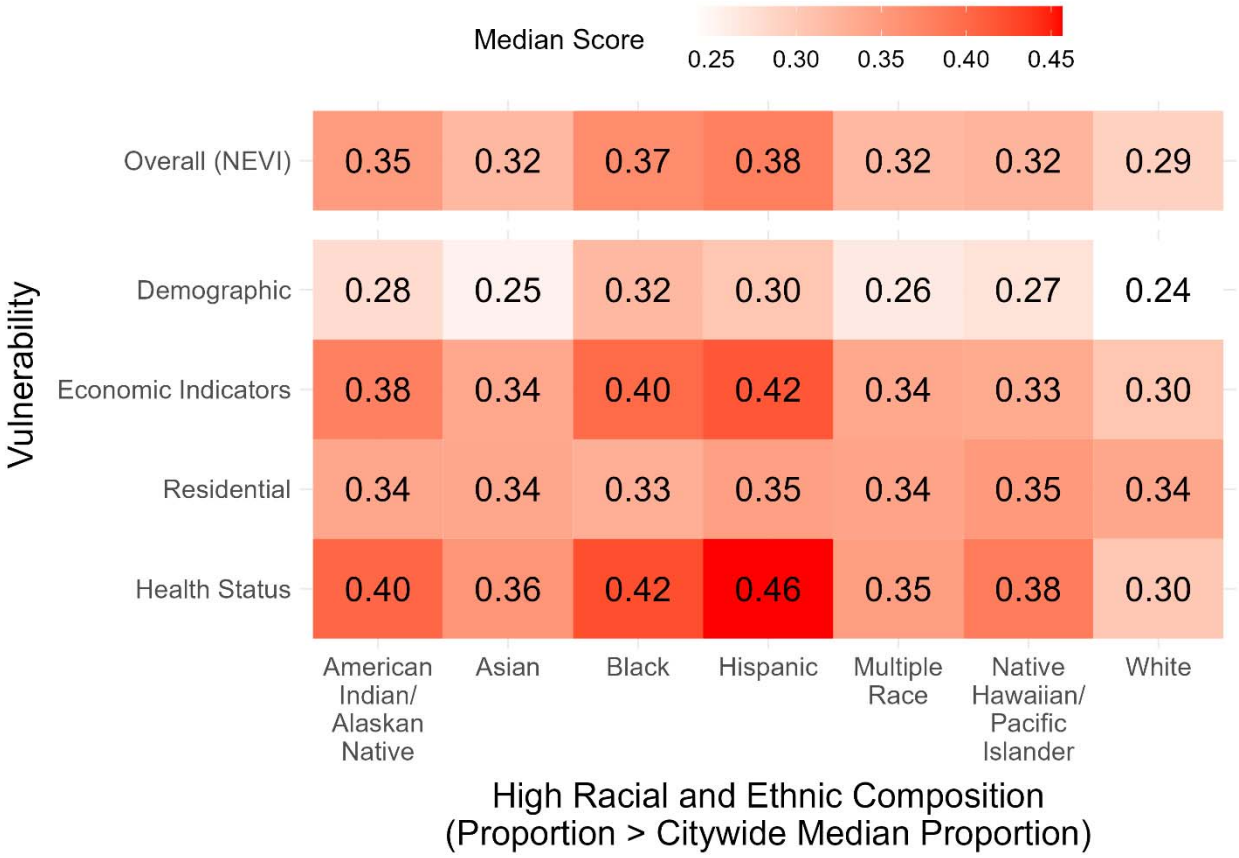
The vulnerability profiles for the six clusters (on top) and two census tracts are shown: Census Tract 671 in Queens and Census Tract 519 in Brooklyn (on bottom right). The different colors represent different clusters in the cluster vulnerability profiles, and the different colors in the Census tract vulnerability profiles represent different domains. Within each domain of the

Census tract vulnerability profiles, the different shades of each slice represent various subdomains, with larger slices representing greater vulnerability. The map (bottom left) shows the distribution of the NEVI clusters across NYC, with areas that were excluded due to low population counts or missing features shown in dark gray. Data Sources: U.S. Census American Community Survey 2015-2019 5-Year Estimates and Centers for Disease Control and Prevention PLACES Project 2020 Release.

Figure 4. Heatmap of Median Subdomain Scores by Neighborhood Environmental Vulnerability Index Cluster

The median subdomain scores were standardized by feature to emphasize the relative magnitude of the subdomain scores across clusters

Figure 5. Median Overall and Domain Scores by High Racial and Ethnic Composition



The median NEVI and domain-specific scores are shown by high racial and ethnic composition across neighborhoods, defined as having a racial and ethnic composition higher than the citywide median proportion. A given neighborhood may have a high composition across multiple racial and ethnic groups (e.g., a neighborhood with both 1) % Hispanic higher than the citywide median % Hispanic and 2) % Native Hawaiian/Pacific Islander higher than the citywide median % Native Hawaiian/Pacific Islander).