BMJ Open

Temporal Trends in the Association of Social Vulnerability and Race/Ethnicity with County-Level COVID-19 Incidence and Outcomes in the United States

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-048086
Article Type:	Original research
Date Submitted by the Author:	18-Dec-2020
Complete List of Authors:	Islam, Shabatun; Emory University, Nayak, Aditi; Emory University School of Medicine Atlanta, Department of Medicine Hu, Yingtian; Emory University Atlanta, Department of Biostatistics and Bioinformatics Mehta, Anurag; Emory University School of Medicine Atlanta, Department of Medicine Dieppa, Katherine Almuwaqqat, Zakaria; Emory University School of Medicine Atlanta, Department of Medicine Ko, Yi-An; Emory University Atlanta, Department of Biostatistics and Bioinformatics Patel, Shivani; Emory University Goyal, Abhinav; Emory University Sullivan, Samaah; Emory University School of Public Health, Epidemiology Lewis, Tene; Rollins School of Public Health, Emory University, Atlanta, GA, Department of Epidemiology Vaccarino, Viola; Emory University School of Public Health, Epidemiology Morris, Alanna; Emory University Quyyumi, Arshed; Emory University School of Medicine Atlanta, Department of Medicine
Keywords:	COVID-19, PUBLIC HEALTH, EPIDEMIOLOGY, SOCIAL MEDICINE

Temporal Trends in the Association of Social Vulnerability and Race/Ethnicity with County-Level COVID-19 Incidence and Outcomes in the United States

Running Title: Social Vulnerability, Race and the COVID-19 Pandemic in U.S.

Shabatun J. Islam, MD[1],* Aditi Nayak, MD[1],* Yingtian Hu, BS[3],* Anurag Mehta, MD[1],* Katherine Dieppa, M.Eng.[2], Zakaria Almuwaqqat, MD, MPH[1], Yi-An Ko, PhD[3], Shivani A. Patel, MPH, PhD[4], Abhinav Goyal, MD, MHS[1], Samaah Sullivan, PhD[4], Tené T. Lewis, PhD[4], Viola Vaccarino, MD, PhD[4], Alanna A. Morris, MD[1], MSc, Arshed A. Quyyumi, MD[1]

*Dr. Islam, Dr. Nayak, Yingtian Hu and Dr. Mehta contributed equally to this work.

[1] Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, GA
[2] City Operations, Getaround Incorportated, San Francisco, CA
[3] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA
[4] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA

Date: December 16, 2020
Word Count: 2983
Tables/Figures: 3/3

Funding Source: Abraham J. & Phyllis Katz Foundation, Byron Williams Jr, MD Fellowship Fund and the National Institutes of Health T32 HL130025 & T32 HL007745-26A1

Address for correspondence:

Arshed A. Quyyumi, MD
Professor of Medicine, Division of Cardiology
Emory University School of Medicine
Director, Emory Clinical Cardiovascular Research Institute
1462 Clifton Road Northeast, Suite 507
Atlanta, Georgia 30322
Tel: 404-727-3655
Fax: 404 712-8785
E mail: aquyyum@emory.edu

ABSTRACT

Background

The COVID-19 pandemic has disparately affected socially vulnerable and minority communities in the U.S but there is a paucity of data examining these temporal associations throughout the duration of the pandemic.

Objective

We examined the temporal association of county-level Social Vulnerability Index (SVI), a percentile-based measure of county-level social vulnerability to disasters, and its subcomponents and race/ethnic composition, with COVID-19 incidence and death per capita in the U.S in the 6 months between March 22nd, 2020 and September 26th, 2020.

Methods

Counties (n=2764) with $\geq$ 50 confirmed COVID-19 cases as of September 26th, 2020 were included in the study. The overall associations between SVI (and its subcomponents) and county level racial composition with the cumulative outcome variables including incidence and death per capita were assessed by fitting a negative-binomial mixed-effects model by treating SVI or racial composition as fixed and county as random effects. In addition, the same model was used to examine potential time varying associations between weekly number of cases/deaths and SVI or racial composition after accounting for SVI or racial composition, time (in weeks), and their interaction as fixed effects. Further adjustments were made for percentage of population aged ≥65 years, state level testing rate, and comorbidities using the average Hierarchical Condition Category (HCC) score.

Results

Higher SVI, indicative of greater social vulnerability, was independently associated with higher COVID-19 incidence (adjusted incidence rate ratio [IRR] per-10 percentile increase: 1.06, (95% CI 1.06, 1.07, p<0.001), and death per capita (adjusted incidence rate ratio [IRR] per-10 percentile increase: 1.05, (95% CI 1.04, 1.07, p<0.001). SVI became an independent predictor

of incidence starting on April 19th, 2020, peaking in mid-July, and then declining such that it was not a predictor during the last week of the study period. Of the SVI sub-components, minority status/language was an independent predictor of outcomes since the start of the analysis, with counties with higher proportion of Black and Hispanic/Latino residents having worse outcomes.

Conclusion

Counties with greater social vulnerability, especially high minority composition, continue to have worse COVID-19-related outcomes.

Article Summary/Strengths & Limitations

- Assessed a multitude of factors that can cause communities to be socially vulnerable including socioeconomic status, proportion of minorities, household composition, and crowding/access to transportation and COVID-19 outcomes
- Assessed nationwide temporal trends of such associations from the start of the pandemic till the end of September for a duration of 6 months
- Further delineated nationwide association and temporal trends of worse COVID-19 outcomes in predominantly Black and Hispanic communities
- While this ecological study allows us to explore nationwide county-level associations, we are not able to account for individual characteristics that may drive COVID-19 outcomes in socially vulnerable communities

Introduction

Community-level social disadvantage and vulnerability to disasters, as well as race/ethnic composition can influence the incidence of COVID-19 and its adverse outcomes in several ways. For example, lower socioeconomic status (SES) is associated with uncertain healthcare access, poor health status and higher risk factor burden that together contribute to a greater risk of adverse outcomes.[1] Labor inequalities and household overcrowding may decrease the ability to adhere to social-distancing guidelines.[2] Blacks and Hispanics are more likely to work in front-line jobs with lack of workplace protections that may additionally increase exposure risk.[3] Additionally, race/ethnic minorities and immigrants are less likely to have access to appropriate and timely healthcare.[3-5] Evidence suggests that these inequalities also contributed to disease spread and adverse outcomes during the H1N1 influenza pandemic.[6,7]

The Social Vulnerability Index (SVI), created and maintained by the Geospatial Research, Analysis, and Services Program (GRASP) at the Centers for Disease Control and Prevention (CDC) and Agency for Toxic Substances and Disease Registry, is a percentile-based index of county-level vulnerability to disasters and was designed for resource allocation to vulnerable communities during times of duress such as the COVID-19 pandemic.[8,9] The SVI includes measures of county-level socioeconomic status, housing composition and disability, minority status and language, and housing type and transport, and thus allows for a dynamic understanding of challenges encountered by communities. Emerging data during the COVID-19 pandemic has demonstrated that socially vulnerable neighborhoods have had worse outcomes during the early stages of the pandemic, even given the fact that the SVI had been designed to mitigate such adverse outcomes for vulnerable communities.[10-14] There has been a paucity of data, however, examining the temporal association between the SVI and COVID-19 outcomes during the 6 month period between March and September 2020 in the U.S. In addition, neighborhoods are socially vulnerable for various reasons (i.e. socioeconomic status, high proportion of minorities, etc.), and it is important to identify the exact determinants of poor outcomes to guide equitable public health policy and healthcare resource allocation. Similarly, it is now well known that Black[15,16] and Hispanic[17] individuals are especially susceptible to worse COVID-19 outcomes, but the temporal trend of these associations throughout the course of the pandemic remains unknown. Herein, we first report on the temporal trends of the association of county-level SVI and its subcomponents with COVID-19 incidence and death per capita in the U.S. Secondly, since the SVI subcomponent of minority status and language does not delineate specific racial ethnic composition, we also examine the temporal trends of the association of county-level proportion of Black and Hispanic residents and COVID-19 outcomes.

Methods

Study population & time frame: All U.S. counties (n=2146) with at least 50 confirmed COVID-19 cases and greater than 4-week of follow-up data were included in the analysis. Data was analyzed for a period of 28 weeks starting from March 22, 2020 to September 26th, 2020.

Patient and Public Involvement Statement: The patient and the public were not involved in the design of this study.

Outcomes: Primary outcomes of interest were county-level weekly COVID-19 incidence and death per capita of a county. Data were obtained from the Johns Hopkins Center for Systems Science and Engineering database.[1]

Exposures: Exposures studied were (a) 2018 county-level SVI and its subcomponents obtained from the CDC GRASP database,[2,3] and (b) racial composition data (reported as proportion of Black and Hispanic residents in a county (with data collected between 2015-2019) from the U.S. Census Bureau ACS database.[4] The SVI was developed by the CDC as a measure of community resilience to stresses on human health such as disease outbreaks and natural or human-caused disasters, to help public health officials and emergency response planners identify communities that are likely to need support before, during, and after a disaster. [2,3] The index combines statistical data from the U.S. Census on 15 variables, grouped together into four related themes: socioeconomic status (SES), housing composition and disability, minority status and language, and housing type and transport (**Table 1**). Each of these variables are ranked from lowest to highest vulnerability across census tracts in the U.S. and a county-level percentile rank is calculated for each variable, theme, and the overall SVI, with higher percentiles indicating higher social vulnerability.

Confounders: Covariates included in all models were proportion of county population aged >65 years, state-level COVID-19 testing rate (tests administered per 1000 population) obtained from the COVID Tracking Project database,[5] and 2018 Hierarchical Condition Category (HCC) risk score acquired from the Centers for Medicare and Medicaid Services (CMS) database as a proxy for county-level medical comorbidity.[6,7] The HCC risk score, based on medical risk profiles and demographics of county Medicare beneficiaries, was developed by CMS to risk-adjust Medicare spending for beneficiary health status.[6,7] The score compares favorably to other comorbidity indices in prediction of outcomes, [6] and aggregate county-level scores are publicly available.[7] All data sources used in this analysis are publicly available, and are listed in **Table S1**.

Statistical analysis:

The overall associations between SVI (and its subcomponents) with the cumulative outcome variables including incidence and death per capita were assessed by fitting a negative-binomial mixed-effects model accounting for SVI as fixed effects with county specific random intercepts. The time-varying associations between SVI (and its subcomponents) of a county with the weekly outcome variables were assessed by fitting a negative-binomial mixed-effects model with weekly total confirmed case numbers or weekly total death numbers as the outcome and county-specific random intercepts to account for overdispersion, correlation in the outcome within counties, and heterogeneity across counties. The fixed effects included SVI, time (in weeks), and the interaction between time and SVI. Time was expressed using natural cubic splines with 3 degrees of freedom to allow for nonlinear relationships. Similarly, overall associations and time-varying associations between country-specific Black and Hispanic race/ethnic composition and weekly outcome variables were evaluated using the same model by replacing SVI with the respective race/ethnic composition variable. Total population in each county was used as the offset in all models. We further adjusted for covariates including percentage aged ≥65 years, state level testing rate, and the HCC risk score as described above. All analyses were performed using R, version 3.6.1 (R Foundation for Statistical Computing). All *P* values were 2-sided, with a significance level of 0.05.

RESULTS

Among the 3,141 counties in the U.S., 2764 (88%) had ≥50 confirmed COVID-19 cases as of September 26th, 2020; accounting for a total of 6,951,774 cases from 100,575,907 administered tests, and 201,321 deaths. The median state-level COVID-19 testing rate was 27.8 per 1000 people [Range 15.5-79.1].The median SVI for counties included in this analysis was 0.53 [Range: 0.001-1.00], the median county-level COVID-19 incidence was 17.1 per 1000 people [Range: 0.83 – 173.1] and death per capita was 0.29 [Range 0-4.92] per 1000 people. The median proportions of Blacks and Hispanics per county were 2.9% [Range: 0.00-87.4%] and 4.2% [Range: 0.00-99.1%], respectively.

Overall and Temporal Associations between SVI and COVID-19 incidence

The incidence of COVID-19 infections was significantly higher in counties with greater SVI or higher social vulnerability, (adjusted incidence rate ratio [IRR] per-10 percentile increase: 1.06 [95% CI 1.06, 1.07], p<0.001) after adjusting for aforementioned confounders as of September 26thth, 2020. Thus, the most socially vulnerable counties (SVI ≥90th percentile) had an adjusted 1.6-fold higher COVID-19 incidence compared to the least vulnerable (SVI ≤10th percentile) counties. Among the SVI sub-components, indices of SES (adjusted IRR per 10-percentile increase: 1.04, [95% CI 1.03, 1.05], p<0.001), minority status and language (adjusted IRR per 10-percentile increase: 1.10 [95% CI 1.09, 1.11], p<0.001), and housing type and transportation (adjusted IRR per 10-percentile increase: 1.04 [95% CI 1.03, 1.04], p<0.001) were independently associated with COVID-19 incidence (**Table 2**). Counties with highest indices for SES (i.e. low SES), minority status and language, and housing type and transportation ($\geq$90th percentile for each) had an adjusted 1.3 to 2.2 higher COVID-19 incidence as compared to counties with the lowest ($\leq$10th percentile) corresponding indices.

Figure 1 demonstrates the temporal trends in the incidence of infections in relation to the overall SVI and its components. As shown, overall county-level SVI was an independent predictor of COVID-19 incidence starting in mid-April, 2020 (Week 5), with the association becoming stronger over time. However, the association weakened after mid-July, 2020 (Week 18) and by late September, 2020 (Week 27), there was no association between overall SVI and COVID-19 incidence.

SVI Subcomponents (**Figure 1**): The SVI subcomponent of minority status and language was an independent predictor of incidence from the beginning (March 22nd, 2020). The association was attenuated after adjustment for comorbidities using the HCC, but remained significantly associated with COVID-19 incidence until mid-September, 2020 (Week 26) when it was no longer associated. The SVI subcomponent of socioeconomic status was an independent predictor of incidence after accounting for co-morbidities starting in early May, 2020 (Week 8), with a strengthening association until mid-July, 2020 (Week 18), after which the association was weaker and became insignificant by late September, 2020 (Week 28). The indices of county-level household composition and disability and housing type and transportation become independent predictors of incidence of COVID-19 in late June, 2020 (Week 14) and late April, 2020, (Week 6), respectively. The associations weakened over time and were no longer significant predictors by late September, 2020.

Overall and Temporal Associations between SVI and COVID-19 death per capita

The average death per capita from COVID-19 over the 28-week duration of the study was significantly higher in counties with greater SVI or higher social vulnerability (adjusted IRR per-10 percentile increase: 1.05, (95% CI 1.04, 1.07, p<0.001) after adjusting for aforementioned confounders (**Table 2**). Thus, the most socially vulnerable counties (SVI $\geq$90th percentile) had an adjusted 1.6-fold higher COVID-19 death per capita compared to the least vulnerable (SVI ≤10th percentile) counties. All the SVI sub-components including indices of SES (adjusted IRR per 10-percentile increase: 1.03, 95% CI 1.01, 1.04, p<0.001), minority status and language (adjusted IRR per 10-percentile increase: 1.12, 95% CI 1.10, 1.14, p<0.001), housing type and transportation (adjusted IRR per 10-percentile increase: 1.02, 95% CI 1.03, 1.04, p<0.001), and household composition and disability (adjusted IRR per 10-percentile increase: 1.02, 95% CI 1.01, 1.03, p<0.001) were independently associated with COVID-19 death per capita (**Table 2**). Counties with highest indices for SES (i.e. low SES), minority status and language, housing type and transportation, and household composition and disability (≥90th percentile for each) had an adjusted 1.2 to 2.5 higher COVID-19 incidence as compared to counties with the lowest (≤10th percentile) corresponding indices.

Figure 2 demonstrates the temporal trends in death per capita in relation to the overall SVI and its components. As shown, overall county-level SVI was an independent predictor of COVID-19 death per capita starting in early June, 2020 (Week 12), with the association becoming stronger over time. However, the association weakened after mid-August, 2020 (Week 22) but remained significantly associated at the study end.

SVI Subcomponents (**Figure 2)**: The SVI subcomponent of minority status and language was significantly associated with COVID-19 death per capita from the beginning. While the association was attenuated slightly after adjustment for comorbidities, it remained significant throughout the entire 28-week duration of the analysis albeit the strength of the association decreased starting in late-July, 2020 (Week 19). The SVI subcomponent of socioeconomic status was independently associated with death per capita after accounting for co-morbidities starting in late June, 2020 (Week 14). The indices of county-level household composition and disability and housing type and transportation become associated with death per capita in mid- or early July, 2020. For all four subcomponents of the SVI, the associations weakened starting mid to late July, 2020 but remined significantly associated till the end of the study period.

Overall and Temporal Associations between race/ethnicity and COVID-19 Incidence & Death per Capita

In order to further investigate the association of minority status with worse COVID-19 outcomes, we compared the rate of infections and death per capita according to the proportion of whites, Hispanics and blacks within each county, based on county level data from 2015-2019 from the U.S. Census Bureau ACS database.[4] County-level increase in proportion of Black residents (adjusted IRR per 10% increase: 1.08, 95% CI 1.09, 1.14, p<0.001) and Hispanic residents (adjusted IRR per 10% increase 1.17, 95% CI 1.15, 1.19, p<0.001) were both independently associated with higher COVID-19 incidence (**Table 3**). Thus, counties with a high proportion (≥90%) of Blacks and Hispanics had a nearly 1.8-fold and 3.6-fold higher adjusted COVID-19 incidence, respectively, as compared to counties with a low proportion (≤10%). Similarly, county-level increase in proportion of Black residents (adjusted IRR per 10% increase: 1.11, 95% CI 1.08, 1.15, p<0.001) and Hispanic residents (adjusted RR per 10% increase: 1.15, 95% CI 1.11, 1.18, p<0.001) were both independently associated with higher COVID-19 death per capita (**Table 3**). Thus, counties with a high proportion (≥90%) of Black and Hispanic residents had a nearly 2.4-fold and 3-fold higher COVID-19 death per capita, respectively, as compared to counties with a low proportion (≤10%). Counties with a greater proportion of White residents had the lowest incidence of infection and lowest death per capita throughout the 28-week duration of the analysis.

Temporal Trends: Counties with a greater proportion of Black residents had both the highest incidence of infection and death per capita throughout the 28-week study period (**Figure 3**). Counties with a higher proportion of Hispanic/Latino residents also had higher rates of infection and death per capita throughout the study period when compared with counties with higher proportion of White residents. However, there was a steep rise in the incidence of infections in counties with a higher proportion of Hispanic/Latino residents around mid-June, 2020 (Week 13), such that these counties approached incidence levels observed in predominantly Black counties. The death per capita rates in counties with greater proportion of Hispanic/Latino residents were also higher compared to White residents throughout the study period, but have declined starting in late August (Week 21). Counties with predominantly black residents had the highest death per capita throughout the entire duration of the analysis.

Discussion

Herein, we demonstrate that U.S. counties with higher social vulnerability had an overall higher incidence of infection and death per capita from COVID-19. For every 10-percentile increase in SVI, indicative of increased social vulnerability, the incidence and death per capita of COVID-19 is 6% and 5% higher, respectively. However, there is a great deal of temporal variation in the association between SVI and COVID-19 outcomes. SVI became an independent predictor of incidence of infections in mid-April, 2020, peaking in mid-July, and was trending towards not being a significant predictor 6 months into the pandemic. The association between SVI and death rates shows similar trends but lags approximately 4-5 weeks behind that of incidence of COVID-19 infection. The inverted U-shaped association between SVI and COVID-19 outcomes is of great interest because it demonstrates that while the infection started in more affluent/privileged counties, socially vulnerable ones ended up bearing a disproportionate share of the burden during the earlier phases of the pandemic and thus, should have had more resource allocation to mitigate the disastrous effects of the pandemic. However, eventually, the disease spread to all communities, as demonstrated by the decreasing association between SVI and COVID-19 outcomes.

The temporal associations between the subcomponents of the SVI are of special interest because communities are vulnerable due to a multitude of factors. Our analysis demonstrates that especially during the early phases of the pandemic, communities with a greater share of minority populations rather than socioeconomic disadvantage or crowding, were disproportionately bearing the disastrous effects of the pandemic. Our more in-depth analysis of racial composition data, beyond the SVI subcomponent of minority status and language,

demonstrated nationwide trends that counties with a large share of Black residents had especially worse outcomes throughout the pandemic. Hispanic/Latino communities started becoming disproportionately
affected around mid-June, which coincides with the broadscale re-openings across the country.[18] Majority white communities have had better outcomes throughout the pandemic. Our analysis suggests that in the future, resource allocation should be prioritized for communities that are socially vulnerable due to high proportion of minority populations. These results are congruent with a more in-depth analysis completed in Cuyahoga County, OH and Wayne County, MI – both counties are socially vulnerable, but Wayne County has a higher proportion of Black residents and thus, had worse COVID-19 outcomes including death and hospital utilization.[10] We are demonstrating such trends on a nationwide basis but further analysis needs to be completed using more granular, patient-centric data.

Shortly following the re-openings across the country, the socioeconomic component of the SVI became an independent predictor of worse COVID-19 outcomes. The mechanisms are difficult to decipher but there is emerging evidence using cell phone data demonstrating that low income communities have been less able to socially distance during the COVID-19 pandemic likely due to a multitude of factors including less capacity to work from home, take paid or unpaid time off from work, and limited savings.[19,20] While another argument is that low socioeconomic status communities haver higher burden of preexisting health conditions[19], our findings are independent of comorbidities. Finally, lower income communities are also more likely to live in multi-family crowded environments,[21] yet we are seeing associations of death with housing type and transport and household composition much later into the pandemic. While there have been select studies completed in restricted geographic locales such as New York City showing the importance of these vulnerability markers[22], our analysis of nationwide data suggests that other components of the social vulnerability index such as minority and socioeconomic status are much more predictive of outcomes during pandemics.

Some of the limitations of our study include the use of county-level data, which does not allow us to account for individual characteristics that may drive COVID-19 outcomes in socially vulnerable communities. However, we are uniquely demonstrating strong nationwide temporal trends in the association between SVI, it's subcomponents, and county-level proportion of Black and Hispanic/Latino residents and COVID-19 outcomes continuing into the later phases of the pandemic in the U.S. which warrant further examination using individual-level data to understand mediating factors that can be intervened upon to improve outcomes in vulnerable communities.

Reference:

1. Link BG. Epidemiological sociology and the social shaping of population health. *Journal of health and social behavior.* 2008;49(4):367-384.
2. Drago R, Miller K. Sick at work: infected employees in the workplace during the H1N1 pandemic. *Institute for Women's Policy Research.* 2010(B264).
3. Hutchins SS, Fiscella K, Levine RS, Ompad DC, McDonald M. Protection of racial/ethnic minority populations during an influenza pandemic. *American journal of public health.* 2009;99(S2):S261-S270.
4. Leclere FB, Jensen L, Biddlecom AE. Health care utilization, family context, and adaptation among immigrants to the United States. *Journal of health and social behavior.* 1994:370-384.
5. Buchmueller TC, Lo Sasso AT, Lurie I, Dolfin S. Immigrants and employer-sponsored health insurance. *Health services research.* 2007;42(1p1):286-310.
6. Quinn SC, Kumar S, Freimuth VS, Musa D, Casteneda-Angarita N, Kidwell K. Racial disparities in exposure, susceptibility, and access to health care in the US H1N1 influenza pandemic. *American journal of public health.* 2011;101(2):285-293.
7. Soyemi K, Medina-Marino A, Sinkowitz-Cochran R, et al. Disparities among 2009 pandemic Influenza A (H1N1) hospital admissions: A mixed methods analysis–Illinois, April–December 2009. *PloS one.* 2014;9(4).
8. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. *Journal of homeland security and emergency management.* 2011;8(1).
9. Flanagan B, Hallisey E, Adams E, Lavery A. Measuring community vulnerability to natural and anthropogenic hazards: The Centers for Disease Control and Prevention's social vulnerability index. *Journal of Environmental Health.* 2018;80:34-36.
10. Gaynor TS, Wilson ME. Social Vulnerability and Equity: The Disproportionate Impact of COVID-19. *Public administration review.* 2020;80(5):832-838.
11. Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. *American journal of preventive medicine.* 2020;59(3):317-325.
12. Khazanchi R, Beiter ER, Gondi S, Beckman AL, Bilinski A, Ganguli I. County-level association of social vulnerability with COVID-19 cases and deaths in the USA. *Journal of General Internal Medicine.* 2020;35(9):2784-2787.
13. Kim SJ, Bostwick W. Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. *Health education & behavior.* 2020;47(4).
14. Nayak A, Islam SJ, Mehta A, et al. Impact of Social Vulnerability on COVID-19 Incidence and Outcomes in the United States. *medRxiv.* 2020.
15. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Network Open.* 2020;3(9):e2021892-e2021892.
16. Yancy CW. COVID-19 and African Americans. *JAMA.* 2020;323(19):1891-1892.
17. Podewils LJ, Burket TL, Mettenbrink C, et al. Disproportionate Incidence of COVID-19 Infection, Hospitalizations, and Deaths Among Persons Identifying as Hispanic or Latino - Denver, Colorado March-October 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(48):1812-1816.
18. Kaufman BG, Whitaker R, Lederer NM, Lewis VA, McClellan MB. Comparing Associations of State Reopening Strategies with COVID-19 Burden. *Journal of General Internal Medicine.* 2020.

19. Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences.* 2020;117(33):19658.
20. Orhun AY, Palazzolo M. Frugality Is Hard to Afford. *Journal of Marketing Research.* 2019;56(1):1-17.
21. Melki IS, Beydoun HA, Khogali M, Tamim H, Yunis KA. Household crowding index: a correlate of socioeconomic status and inter-pregnancy spacing in an urban setting. *Journal of Epidemiology and Community Health.* 2004;58(6):476.
22. Emeruwa UN, Ona S, Shaman JL, et al. Associations Between Built Environment, Neighborhood Socioeconomic Status, and SARS-CoV-2 Infection Among Pregnant Women in New York City. *JAMA.* 2020;324(4):390-392.

FIGURE LEGENDS

Figure 1: Temporal association between COVID-19 incidence and the a) county-level Social Vulnerability Index (SVI) and its subcomponents: b) Socioeconomic Status, c) Household Composition and Disability, d) Minority Status and Language, and e) Housing Type and Transportation between March 22nd, 2020 and September 26th, 2020. The base model (red lines) is adjusted for proportion of population age $\geq$65 years and state-level COVID-19 testing. The blue lines are additionally adjusted for CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities). Of note, proportion age $\geq$65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable

Figure 2: Temporal association between COVID-19 death per capita and the a) county-level Social Vulnerability Index (SVI) and its subcomponents: b) Socioeconomic Status, c) Household Composition and Disability, d) Minority Status and Language, and e) Housing Type and Transportation between March 22nd, 2020 and September 26th, 2020. The base model (red lines) is adjusted for proportion of population age $\geq$65 years and state-level COVID-19 testing. The blue lines are additionally adjusted for CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities). Of note, proportion age $\geq$65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable

Figure 3: Temporal association of county-level racial composition (Black, Hispanic/Latino, White) and COVID-19 a) incidence and b) death per capita between March 22nd, 2020 and September 26th, 2020 after adjusting for proportion of population age $\geq$65 years, state-level COVID-19 testing, and CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities)

Table 1. Components of the social vulnerability index

Socioeconomic Status	Below Poverty Unemployed Income No High School Diploma
Household composition and Disability	Age 65 years or older Age 17 years or younger Older than Age 5 years with a Disability
Minority Status and Language	Minority Speak English "Less than Well"
Housing Type and Transportation	Multi-Unit Structures Mobile Homes Crowding No Vehicle Group Quarters

Table 2. Overall association of county-level social vulnerability index (Incidence Risk Ratio [IRR] per-10 percentile increase) with incidence and death per capita of COVID-19 as of September 26th, 2020

	Model 1[a]		Model 2[b]	
	IRR (95% CI)	**P-value**	**IRR (95% CI)**	**P-value**
Incidence				
Overall Social Vulnerability Index*	1.09(1.08, 1.10)	<0.001	1.06(1.04, 0.07)	<0.001
Socioeconomic Status	1.06(1.05, 1.07)	<0.001	1.04(1.03, 1.05)	<0.001
Minority Status & Language	1.11(1.10, 1.12)	<0.001	1.10(1.09, 1.11)	<0.001
Housing Type & Transport	1.05(1.04, 1.05)	<0.001	1.04(1.03, 1.04)	<0.001
Household Composition & Disability*	1.00(0.99, 1.01)	0.464	0.98(0.97, 0.99)	<0.001
Death Per Capita				
Overall Social Vulnerability Index*	1.12(1.11, 1.14)	<0.001	1.05(1.04, 1.07)	<0.001
Socioeconomic Status	1.10(1.08, 1.11)	<0.001	1.03(1.01, 1.04)	<0.001
Minority Status & Language	1.15(1.13, 1.17)	<0.001	1.12(1.10, 1.14)	<0.001
Housing Type & Transport	1.06(1.05, 1.08)	<0.001	1.02(1.00, 1.03)	0.003
Household Composition & Disability*	1.08(1.06, 1.11)	<0.001	1.02(1.01, 1.03)	<0.001

[a] Base Model - adjusted for proportion of population age >65 years and state-level COVID-19 testing
[b] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities)
* Proportion age >65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable.

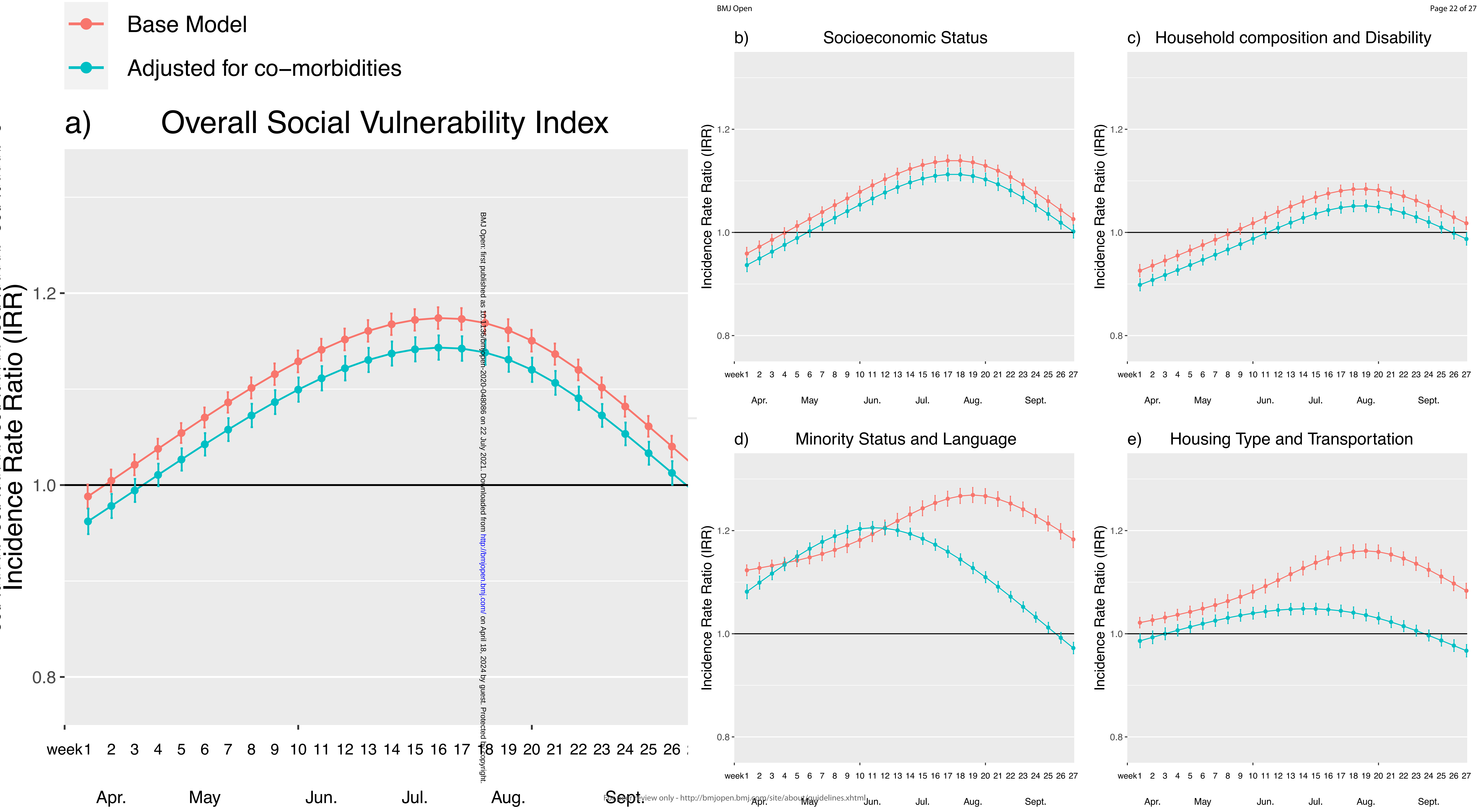
Table 3. Overall association of county-level race/ethnic composition (Incidence Risk Ratio [IRR] per 10% increase) with incidence and case fatality rate of COVID-19 as of September 26th, 2020

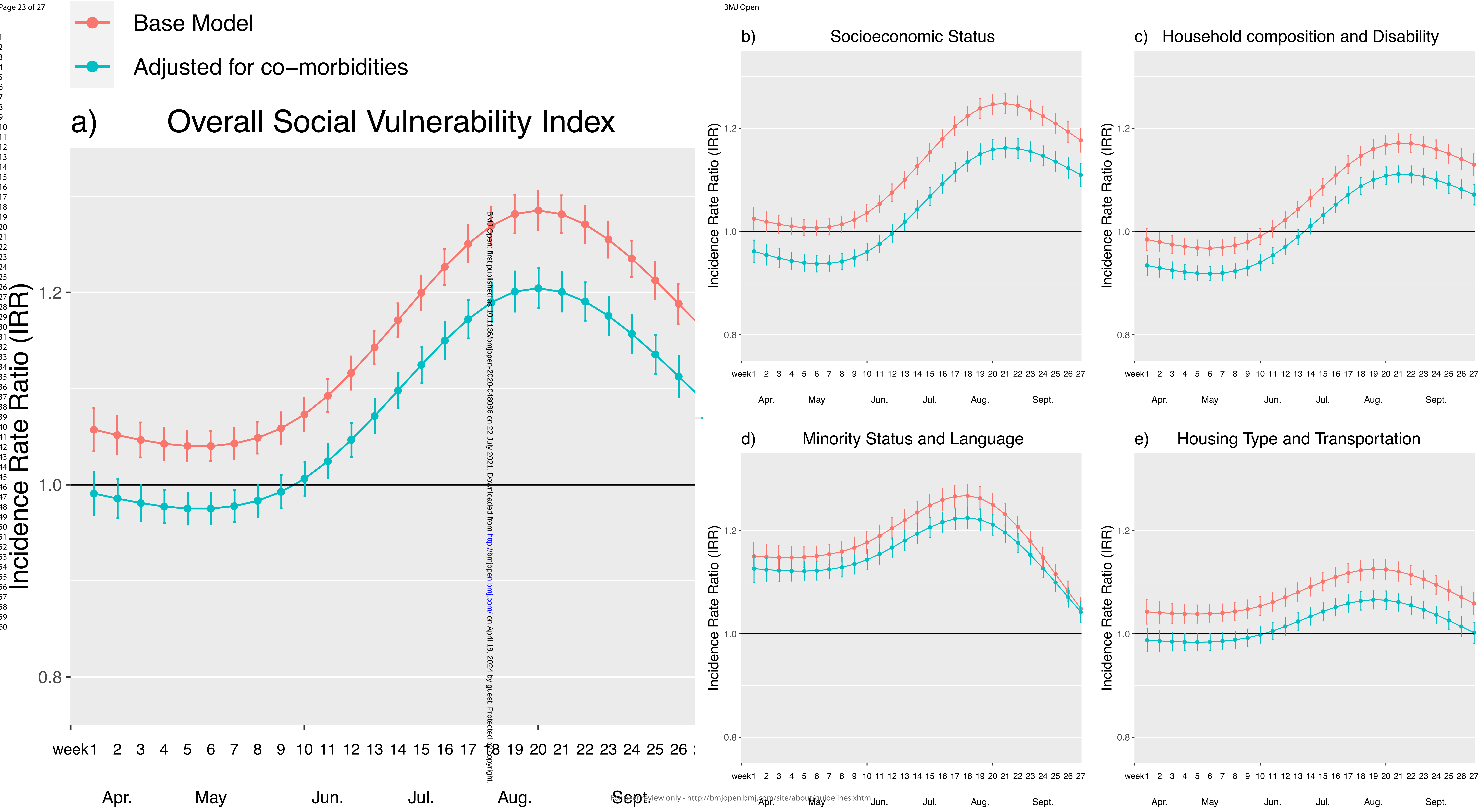
	Model 1[a]		**Model 2[b]**	
	IRR (95% CI)	**P-value**	**IRR (95% CI)**	**P-value**
Incidence				
Black	1.11(1.09, 1.13)	<0.001	1.08(1.09, 1.14)	<0.001
Hispanic	1.18(1.16, 1.21)	<0.001	1.17(1.15, 1.19)	<0.001
Death Per Capita				
Black	1.21(1.18, 1.25)	<0.001	1.11(1.08, 1.15)	<0.001
Hispanic	1.20(1.16, 1.24)	<0.001	1.15(1.11, 1.18)	<0.001

[a] Base Model - adjusted for proportion of population age $\geq$65 years and state-level COVID-19 testing
[b] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities)

Author contributions:

Shabatun J. Islam, MD - Concept, drafting and editing of manuscript, statistical analysis
Aditi Nayak, MD - Concept, drafting and editing of manuscript
Yingtian Hu, BS – Statistical analysis, review of manuscript
Anurag Mehta, MD- Concept, drafting and editing of manuscript
Katherine Dieppa, M.Eng. – Statistical analysis, review of manuscript
Zakaria Almuwaqqat, MD, MPH - Statistical analysis, review of manuscript
Yi-An Ko, PhD - Statistical analysis, review of manuscript
Shivani A. Patel, MPH, PhD - Concept, review of manuscript
Abhinav Goyal, MD, MHS - Concept, review of manuscript
Samaah Sullivan, PhD - Concept, review of manuscript
Tené T. Lewis, PhD - Concept, drafting and editing of manuscript
Viola Vaccarino, MD, PhD - Concept, drafting and editing of manuscript
Alanna A. Morris, MD, MSc - Concept, drafting and editing of manuscript
Arshed A. Quyyumi, MD - Concept, drafting and editing of manuscript, review of manuscript


Base Model
Adjusted for co−morbidities
a) Overall Social Vulnerability Index
b) Socioeconomic Status
c) Household composition and Disability
d) Minority Status and Language
e) Housing Type and Transportation
Incidence Rate Ratio (IRR)
1.2
1.0
0.8
week1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Apr.
May
Jun.
Jul.
Aug.
Sept.


Base Model
Adjusted for co−morbidities
a) Overall Social Vulnerability Index
b) Socioeconomic Status
c) Household composition and Disability
d) Minority Status and Language
e) Housing Type and Transportation
Incidence Rate Ratio (IRR)
1.2
1.0
0.8
week1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Apr.
May
Jun.
Jul.
Aug.
Sept.


a)
Incidence
Incidence Rate Ratio (IRR)
1.50
1.25
1.00
0.75
0.50
week1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Apr.
May
Jun.
Jul.
Aug.
Sept.
b)
Death per capita
Incidence Rate Ratio (IRR)
1.50
1.25
1.00
0.75
0.50
week1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Apr.
May
Jun.
Jul.
Aug.
Sept.
race
Black
White
Hispanic/Latino

SUPPLEMENTARY MATERIAL

Table S1. Data sources used in the analysis (publicly available)

Data	Source
Outcomes	
Case Fatality Rate	Johns Hopkins Center for Systems Science and Engineering database, accessed October 01, 2020: https://coronavirus.jhu.edu/us-map
Incidence	Johns Hopkins Center for Systems Science and Engineering database, accessed October 01, 2020: https://coronavirus.jhu.edu/us-map
Exposures	
Social Vulnerability Index	Centers for Disease Control and Prevention (CDC) Geospatial Research, Analysis, and Services Program (GRASP) database, accessed October 01, 2020: https://svi.cdc.gov/
Racial Composition	US Census Bureau, accessed October 01, 2020: https://www.census.gov/data.html
Confounders	
Proportion age ≥65 years	Centers for Disease Control and Prevention (CDC) Geospatial Research, Analysis, and Services Program (GRASP) database, accessed October 01, 2020: https://svi.cdc.gov/
Hierarchical Condition Category score	Centers for Medicare and Medicaid Services (CMS), accessed October 01, 2020 https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Geographic-Variation
Tests administered per state	The COVID Tracking Project, accessed October 01, 2020: https://covidtracking.com/

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
Page 1-3	√	(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale Page 4	2 √	Explain the scientific background and rationale for the investigation being reported
Objectives Page 5	3 √	State specific objectives, including any prespecified hypotheses
Methods		
Study design Page 5	4 √	Present key elements of study design early in the paper
Setting Page 5	5 √	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants Page 5	6 √	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables Page 5	7 √	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement Page 5-6	8* √	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias Page 6	9 √	Describe any efforts to address potential sources of bias
Study size Page 5	10 √	Explain how the study size was arrived at
Quantitative variables Page 6	11 √	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
Page 6	√	(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants Page 7	13* √	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data Page 7	14* √	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
Outcome data Page 7-10	15* √	Report numbers of outcome events or summary measures
Main results Page 7-10	16 √	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were

		adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses Page 10-11	17 √	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results Page 11	18 √	Summarise key results with reference to study objectives
Limitations Page 13	19 √	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation Page 11-12	20 √	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability Page 13	21 √	Discuss the generalisability (external validity) of the study results
Other information		
Funding Page 1	22 √	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Ecological Analysis of the Temporal Trends in the Association of Social Vulnerability and Race/Ethnicity with County-Level COVID-19 Incidence and Outcomes in the United States

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-048086.R1
Article Type:	Original research
Date Submitted by the Author:	07-Jun-2021

Complete List of Authors:	Islam, Shabatun; Emory University, Nayak, Aditi; Emory University School of Medicine Atlanta, Department of Medicine Hu, Yingtian; Emory University Atlanta, Department of Biostatistics and Bioinformatics Mehta, Anurag; Emory University School of Medicine Atlanta, Department of Medicine Dieppa, Katherine Almuwaqqat, Zakaria; Emory University School of Medicine Atlanta, Department of Medicine Ko, Yi-An; Emory University Atlanta, Department of Biostatistics and Bioinformatics Patel, Shivani; Emory University Goyal, Abhinav; Emory University Sullivan, Samaah; Emory University School of Public Health, Epidemiology Lewis, Tene; Rollins School of Public Health, Emory University, Atlanta, GA, Department of Epidemiology Vaccarino, Viola; Emory University School of Public Health, Epidemiology Morris, Alanna; Emory University Quyyumi, Arshed; Emory University School of Medicine Atlanta, Department of Medicine

Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Infectious diseases
Keywords:	COVID-19, PUBLIC HEALTH, EPIDEMIOLOGY, SOCIAL MEDICINE

SCHOLARONE™ Manuscripts

Ecological Analysis of the Temporal Trends in the Association of Social Vulnerability and Race/Ethnicity with County-Level COVID-19 Incidence and Outcomes in the United States

Running Title: Social Vulnerability, Race and the COVID-19 Pandemic in U.S.

Shabatun J. Islam, MD[1],* Aditi Nayak, MD[1],* Yingtian Hu, BS[3],* Anurag Mehta, MD[1],* Katherine Dieppa, M.Eng.[2], Zakaria Almuwaqqat, MD, MPH[1], Yi-An Ko, PhD[3], Shivani A. Patel, MPH, PhD[4], Abhinav Goyal, MD, MHS[1], Samaah Sullivan, PhD[4], Tené T. Lewis, PhD[4], Viola Vaccarino, MD, PhD[4], Alanna A. Morris, MD[1], MSc, Arshed A. Quyyumi, MD[1]

*Dr. Islam, Dr. Nayak, Yingtian Hu and Dr. Mehta contributed equally to this work.

[1] Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, GA
[2] City Operations, Getaround Incorportated, San Francisco, CA
[3] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA
[4] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA

Date: June 4, 2021
Word Count: 3837
Tables/Figures: 3/4

Funding Source: Abraham J. & Phyllis Katz Foundation, Byron Williams Jr, MD Fellowship Fund and the National Institutes of Health T32 HL130025 & T32 HL007745-26A1

Competing Interest Statement: Authors do not have any competing interests to disclose

Data Availability Statement: Data are available in a public, open access repository. A table listing data sources is provided in the Supplement.

Ethics/IRB statement: Since data is publicly available and on the population level, this study was exempt from IRB approval at Emory University

Address for correspondence:

Arshed A. Quyyumi, MD
Professor of Medicine, Division of Cardiology
Emory University School of Medicine
Director, Emory Clinical Cardiovascular Research Institute
1462 Clifton Road Northeast, Suite 507
Atlanta, Georgia 30322
Tel: 404-727-3655
Fax: 404 712-8785
E mail: aquyyum@emory.edu

ABSTRACT

Background

The COVID-19 pandemic adversely affected the socially vulnerable and minority communities in the U.S. initially, but the temporal trends during the year-long pandemic remain unknown.

Objective

We examined the temporal association between the county-level Social Vulnerability Index (SVI), a percentile-based measure of social vulnerability to disasters, its subcomponents and race/ethnic composition with COVID-19 incidence and mortality in the U.S. in the year starting in March 2020.

Methods

Counties (n=3091) with $\geq$ 50 COVID-19 cases by March 6th, 2021 were included in the study. Associations between SVI (and its subcomponents) and county level racial composition with the incidence and death per capita were assessed by fitting a negative-binomial mixed-effects model. This model was also used to examine potential time varying associations between weekly number of cases/deaths and SVI or racial composition. Data was adjusted for percentage of population aged ≥65 years, state level testing rate, comorbidities using the average Hierarchical Condition Category (HCC) score, and environmental factors including average fine particulate matter ($PM_{2.5}$), temperature and precipitation.

Results

Higher SVI, indicative of greater social vulnerability, was independently associated with higher COVID-19 incidence (adjusted incidence rate ratio [IRR] per-10 percentile increase:1.02, (95% CI 1.02, 1.03, p<0.001), and death per capita (1.04, (95% CI 1.04, 1.05, p<0.001). SVI became an independent predictor of incidence starting from March 2020, but this association became weak or insignificant by the winter, a period that coincided with a sharp increase in infection rates and mortality, and when counties with higher proportion of White residents were

disproportionately represented ("third wave"). By Spring of 2021, SVI was again a predictor of COVID-19 outcomes. Counties with greater proportion of Black residents also observed similar temporal trends COVID-19-related adverse outcomes. Counties with greater proportion of Hispanic residents had worse outcomes throughout the duration of the analysis.

Conclusion

Except for the winter "third wave" when majority White communities had the highest incidence of cases, counties with greater social vulnerability and proportionately higher minority populations, experienced worse COVID-19 outcomes.

Article Summary/Strengths & Limitations

- Examined full 12 months of county-level data in the US delineating the temporal trends in the association between social vulnerability index and COVID-19 outcomes
- Investigated COVID-19 outcomes in predominantly Black and Hispanic communities in comparison to White communities in the US
- Analysis is ecological, descriptive, and on the county-level rather than on an individual level
- Analysis adjusted for confounders including county level age $\geq$ 65, comorbidities, and environmental factors
- Analysis limited to the US

Introduction

Community-level social disadvantage and vulnerability to disasters, as well as race/ethnic composition can influence the incidence of COVID-19 and its adverse outcomes in several ways. For example, lower socioeconomic status (SES) is associated with uncertain healthcare access, poor health status and higher risk factor burden that together contribute to a greater risk of adverse outcomes.[1] Labor inequalities and household overcrowding may decrease the ability to adhere to social-distancing guidelines.[2] Black and Hispanic individuals are more likely to work in front-line jobs with lack of workplace protections that may additionally increase exposure risk.[3] Additionally, race/ethnic minorities and immigrants are less likely to have access to appropriate and timely healthcare.[3-5] Evidence suggests that these inequalities also contributed to disease spread and adverse outcomes during the H1N1 influenza pandemic.[6,7]

The Social Vulnerability Index (SVI), created and maintained by the Geospatial Research, Analysis, and Services Program (GRASP) at the Centers for Disease Control and Prevention (CDC) and Agency for Toxic Substances and Disease Registry, is a percentile-based index of county-level vulnerability to disasters and was designed for resource allocation to vulnerable communities during times of duress such as the COVID-19 pandemic.[8,9] The SVI includes measures of county-level socioeconomic status, housing composition and disability, minority status and language, and housing type and transport, and thus allows for a dynamic understanding of challenges encountered by communities. Emerging data during the COVID-19 pandemic has demonstrated that socially vulnerable neighborhoods have had worse outcomes during the early stages of the pandemic,[10-14] even given the fact that the SVI had been designed to mitigate such adverse outcomes for vulnerable communities. Data has shown that the association between SVI and COVID-19 outcomes temporally varied, with the trend reversing by October 2020,[15] but whether this continued to the latter durations of the pandemic in unknown. Similarly, it is now well known that Black[16,17] and Hispanic[18] individuals, who represent the largest minority groups within the U.S., are especially susceptible to worse COVID-19 outcomes, but the temporal trend of these associations throughout the course of the pandemic remains unknown. Herein, we first report on the temporal trends in the association of county-level SVI and its subcomponents with COVID-19 incidence and death per capita in the U.S. for the entire year from March 2020 to March 2021. Secondly, since the SVI subcomponent of minority status and language does not delineate specific racial ethnic composition, we also examine the temporal trends of the association of county-level proportion of Black and Hispanic residents and COVID-19 outcomes.

Methods

Study population & time frame: All U.S. counties (n=3091) with at least 50 confirmed COVID-19 cases and greater than 4-week of follow-up data were included in the analysis. Data was analyzed for a period of 50 weeks starting from March 22, 2020 to March 6th, 2021.

Patient and Public Involvement Statement: The patient and the public were not involved in the design of this study.

Outcomes: Primary outcomes of interest were county-level weekly COVID-19 incidence and death per capita of a county. Data were obtained from the Johns Hopkins Center for Systems Science and Engineering database.[1]

Exposures: Exposures studied were (a) 2018 county-level SVI and its subcomponents obtained from the CDC GRASP database,[2,3] and (b) racial composition data (reported as proportion of Black and Hispanic residents in a county) from the U.S. Census Bureau ACS database.[4] The SVI was developed by the CDC as a measure of community resilience to stresses on human health such as disease outbreaks and natural or human-caused disasters, to help public health officials and emergency response planners identify communities that are likely to need support before, during, and after a disaster. [2,3] The index combines statistical data from the U.S. Census on 15 variables, grouped together into four related themes: socioeconomic status (SES), housing composition and disability, minority status and language, and housing type and transport (**Table 1**). Each of these variables are ranked from lowest to highest vulnerability across census tracts in the U.S. and a county-level percentile rank is calculated for each variable, theme, and the overall SVI, with higher percentiles indicating higher social vulnerability. In terms of the racial composition data, we focused on the proportion of Black and Hispanic residents in a county since they represent the largest minority groups across a broader geographic region in the U.S.. The data was collected by the U.S. Census Bureau as self-reported race/ethnicity between 2015-2019. [4]

Confounders: Covariates included in all models were proportion of county population aged $\geq$65 years[4], state-level COVID-19 testing rate obtained from the COVID Tracking Project database,[5] 2018 Hierarchical Condition Category (HCC) risk score acquired from the Centers for Medicare and Medicaid Services (CMS) database as a proxy for county-level medical comorbidity, and environmental factors. State-level COVID-19 testing rate is calculated as all tests completed (whether symptomatic and asymptomatic, voluntary or contact tracing) divided by the state-level population. The HCC risk score, based on medical risk profiles and demographics of county Medicare beneficiaries, was developed by CMS to risk-adjust Medicare spending for beneficiary health status.[19,20] While the score was designed to reflect healthcare access and hospital admissions in a geographic area, it does compare favorably to other comorbidity indices in prediction of outcomes,[19] and aggregate county-level scores are publicly available.[20] As such, we are using the HCC risk score as a proxy for county-level comorbidities. For environmental factors, we included average daily temperature (degrees Fahrenheit),[21] average daily precipitation,[21] and average particulate matter of diameter $\geq$ 2.5 micrometers ($PM_{2.5}$).[22] All data sources used in this analysis are publicly available and are listed in **Table S1**. Given that all data is publicly available and no patient-level data is used, this study was exempt from Institutional Board Review (IRB) by Emory University.

Statistical analysis: The overall associations between SVI (and its subcomponents) with the cumulative outcome variables including incidence and death per capita were assessed by fitting a negative-binomial mixed-effects model accounting for SVI as fixed effects with county specific random intercepts. The time-varying associations between SVI (and its subcomponents) of a county with the weekly outcome variables were assessed by fitting a negative-binomial mixed-effects model with weekly total confirmed case numbers or weekly total death numbers as the outcome and county-specific random intercepts to account for overdispersion, correlation in the outcome within counties, and heterogeneity across counties. The fixed effects included SVI, time (in weeks), and the interaction between time and SVI. Time was expressed using natural cubic splines with 3 degrees of freedom to allow for nonlinear relationships. Similarly, overall associations and time-varying associations between country-specific White, Black and Hispanic race/ethnic composition and weekly outcome variables were evaluated using the same model by replacing SVI with the respective race/ethnic composition variable. Total population in each county was used as the offset in all models. We further adjusted for covariates including percentage aged ≥65 years, state-level testing rate, HCC risk score, average daily temperature (degrees Fahrenheit),[21] average daily precipitation,[21] and average particulate matter of diameter ≥ 2.5 micrometers ($PM_{2.5}$)[22] as described above. All analyses were performed using R, version 3.6.1 (R Foundation for Statistical Computing). All *P* values were 2-sided, with a significance level of 0.05.

RESULTS

Among the 3142 counties in the U.S., 3091 (98.38%) had >50 confirmed COVID-19 cases as of March 6th, 2021 and at least 4-week follow up data; accounting for a total of 28,547,384 cases from 362,058,535 administered tests, and 517,733 deaths. The median SVI for counties included in this analysis was 0.44 [Range: 0.17-0.85]. The median county-level COVID-19 incidence was 90.7 per 1000 people [Range: 2.61 – 368.2] and death per capita was

1.64 per 1000 people [Range 0.00-7.89]. The median proportions of White, Black and Hispanic residents per county were 89.4% [Range: 3.9-99.9%], 2.4% [Range: 0.0-87.4%] and 4.1% [Range: 0.0-99.1%], respectively. Overall SVI correlated most strongly with the subcomponent of socioeconomic status and least with minority status and language (Supplemental Figure 1). Proportion of Black residents correlated modestly (r=0.4) and proportion of Hispanic residents correlated slightly (r=0.2) with overall SVI (Supplemental Figure 1).

Overall and Temporal Associations between SVI and COVID-19 incidence

The incidence of COVID-19 infections was significantly higher in counties with greater SVI or higher social vulnerability, (adjusted incidence rate ratio [IRR] per-10 percentile increase: 1.02 [95% CI 1.02, 1.03], p<0.001) after adjusting for aforementioned confounders as of March 6th, 2021. Among the SVI sub-components, indices of SES (adjusted IRR per 10-percentile increase: 1.02, [95% CI 1.01, 1.03], p<0.001), minority status and language (adjusted IRR per 10-percentile increase: 1.02 [95% CI 1.01, 1.02], p<0.001), and household composition and disability (adjusted IRR per 10-percentile increase: 1.01 [95% CI 1.01, 1.02], p<0.001) were independently associated with COVID-19 incidence **(Table 2).**

Temporal Trends: **Figure 1** demonstrates the temporal trends in the incidence of infections in relation to the overall SVI and its components. As shown, overall county-level SVI was positively associated with COVID-19 incidence starting from the beginning of our analysis on March 22nd, 2020 (Week 1), with the association becoming stronger over time. However, the association weakened after mid-July, 2020 (Week 17) and there was no significant association between overall SVI and COVID-19 incidence between late October, 2020 (Week 32) and early December, 2020 (Week 37). This coincided with a large increase in cases within the U.S. (“third wave”). However, once the overall case load started to decrease from the peak by January, 2021 (Week 40) to early March 2021, overall SVI again demonstrated strong associations with COVID-19 incidence.

<u>SVI Subcomponents (**Figure 1**):</u> The SVI subcomponent of minority status and language was an independent predictor of incidence from the beginning (March 22nd, 2020). While the association attenuated after adjustment for comorbidities using the HCC, it strengthened after additionally adjusting for environmental factors and remained positively associated with COVID-19 incidence until mid-October, 2020 (Week 30) when it started to be negatively associated with COVID-19 incidence with a rise in cases in the U.S ("third wave"). Similar to the overall SVI, it again became positively associated once the cases decreased in the U.S. around January-February 2021 (Week 41-46).

The SVI subcomponent of socioeconomic status was an independent predictor of incidence after accounting for co-morbidities and environmental factors starting in early May, 2020 (Week 6), with a strengthening association until mid-July, 2020 (Week 17), after which the association weakened and became insignificant by early November, 2020 (Week 33) but again became significant by mid December 2021 (Week 39).

The indices of county-level household composition and disability and housing type and transportation become independent predictors of incidence of COVID-19 in early June, 2020 (Week 11) and late March, 2020, (Week 1), respectively. The strength of the positive association varied for county-level household composition and disability but remained significant throughout the duration of our analysis. County-level housing type and transport remained a positive predictor of incidence until mid September 2020 (Week 26), became negative afterwards and became positive again in February 2021 (Week 48). Of note, in analysis additionally adjusting for percentage of residents under the federal poverty limit, for SVI subthemes of minority status and language, household composition and disability, and housing type and transport, similar trends are noted (Supplemental Figure 2).

Overall and Temporal Associations between SVI and COVID-19 death per capita

The average death per capita from COVID-19 over the 50-week duration of the study was significantly higher in counties with greater SVI or higher social vulnerability (adjusted IRR per-10 percentile increase: 1.04, (95% CI [1.04, 1.05], p<0.001) after adjusting for aforementioned confounders (**Table 2**). All the SVI sub-components including indices of SES (adjusted IRR per 10-percentile increase: 1.05, 95% CI [1.04, 1.05], p<0.001), minority status and language (adjusted IRR per 10-percentile increase: 1.01, 95% CI [1.00, 1.02], p=0.004), housing type and transportation (adjusted IRR per 10-percentile increase: 1.05, 95% CI [1.04, 1.05], p<0.001), and household composition and disability (adjusted IRR per 10-percentile increase: 1.02, 95% CI [1.01, 1.02], p<0.001) were independently associated with COVID-19 death per capita (**Table 2**).

Temporal Trends: **Figure 2** demonstrates the temporal trends in death per capita in relation to the overall SVI and its components. As shown, overall county-level SVI firstly was an independent predictor of COVID-19 death per capita starting in early May, 2020 (Week 6) and the association became stronger over time. However, the association weakened after July, 2020 (Week 17) and became insignificant between November 2020 – January 2021 (Week 33 – 40) when cumulative deaths were at their highest. The association became significant starting in early January 2021 (Week 41) once the deaths started decreasing.

SVI Subcomponents (**Figure 2)**: The SVI subcomponent of minority status and language was significantly and positively associated with COVID-19 death per capita from the beginning of the analysis. The strength of the association decreased starting in late-July, 2020 (Week 17), and became negatively-associated starting in mid October, 2020 (Week 30) when the third peak in deaths were observed. It became positively associated again starting in early January 2021 (Week 41), with a decrease in deaths.

The SVI subcomponent of socioeconomic status was independently and positively associated with death per capita after accounting for co-morbidities and environmental factors starting in late May, 2020 (Week 9). While the strength of the association fluctuated throughout the duration of the pandemic, it remained a positive predictor and became more strongly associated with the death per capita starting early January 2021 (Week 41).

The index of county-level household composition and disability became positively associated with death per capita in early June 2020 (Week 11) and remained associated throughout the duration of the pandemic. Housing type and transportation became positively associated starting mid-May 2020 (Week 7), with the association weakening around early August (Week 19) and diminishing by early October 2020 (Week 28), but it became positively associated with death rate around February 2021 (Week 46). Of note, in analysis additionally adjusting for percentage of residents under the federal poverty limit, for SVI subthemes of minority status and language, household composition and disability, and housing type and transport, similar trends are noted (Supplemental Figure 2).

Overall and Temporal Associations between race/ethnicity and COVID-19 Incidence & Death per Capita

In order to further investigate the association of minority status with worse COVID-19 outcomes, we compared the rate of infections and death per capita according to the proportion of whites, Hispanics and blacks within each county, based on county level data from 2015-2019 from the U.S. Census Bureau ACS database.[4] Proportionately more blacks reside in the southern U.S. and Hispanics in southwestern states. Cumulatively for the full year analysis, county-level increase in proportion of Black residents (adjusted IRR per 10% increase: 0.99, 95% CI [0.98, 1.00], p=0.01) was associated with lower, while increase in proportion of Hispanic residents (adjusted IRR per 10% increase 1.07, 95% CI [1.05, 1.08], p<0.001) was associated with higher COVID-19 incidence after adjusting for comorbidities using the HCC score and environmental factors (**Table 3**). While county-level increase in proportion of Black residents (adjusted IRR per 10% increase: 1.00, 95% CI [0.99,1.02], p=0.85) was not associated with COVID-19 death per capita, county-level increase in Hispanic residents (adjusted RR per 10% increase: 1.07, 95% CI [1.05, 1.08], p<0.001) was independently associated with higher COVID-19 death per capita (**Table 3**).

Temporal Trends: Counties with a greater proportion of Black residents had the highest incidence of infection and death per capita from the start of the study period till about mid to late August 2020 (Week 22) (**Figure 3**). For a period of 10 weeks between early November 2020 and early January (for incidence) and 14 weeks between early November 2020 and mid-January 2021 (for death per capita), counties with greater proportion of Black residents had lower incidence and death per capita. By January 2021 (Week 42), this trend reversed such that counties with higher proportion of Black residents again had worse outcomes. Out of the 50 weeks included in our analysis, counties with higher proportion of Black residents had higher incidence during 40 weeks (80% of the analysis time frame) and higher death per capita during 36 weeks (72% of the analysis time frame). Counties with higher proportion of White residents showed opposite trends. During a period of 8 weeks between November 2020 and early January 2020, counties with higher proportion of White residents had higher incidence, which coincided with higher overall cases in the U.S. (“third wave”). Similarly, for a period of 11 weeks between November 2020 and January 2020, these communities had higher death per capita. Counties with higher proportion of Hispanic residents had both higher incidence and death per capita throughout the entire duration of the analysis.

When the geographical changes in the incidence and mortality rates from COVID-19 are examined over the year, it is apparent that whereas the early part of the pandemic affected the northeastern U.S. and areas in the south-east and southwestern U.S., areas that are enriched for minority populations, by the end of 2020 when the pandemic was at its peak (“third wave’), the mid-western states, with predominantly White populations, had the highest prevalence and mortality rates. By the Spring of 2021, the geographic distribution of cases again changed back to the areas affected initially during the pandemic **(Figure 4).**

Discussion

Herein, we demonstrate that U.S. counties with higher social vulnerability had an overall higher incidence of infection and death per capita from COVID-19. For every 10-percentile increase in SVI, indicative of increased social vulnerability, the incidence and death per capita of COVID-19 is 2% and 5% higher, respectively. However, there is a great deal of temporal variation in the association between SVI and COVID-19 outcomes, throughout the duration of the pandemic. SVI became an independent predictor of incidence of infections starting in April 2020, becoming an increasingly important predictor until August 2020. By late October, when the pandemic was at its third peak, SVI was no longer a predictor of the incidence of infections. The absence of SVI being a predictor of COVID-19 outcomes coincided with a sharp increase in cases and deaths within the U.S. between early November 2020 and early January 2021, when counties with higher proportion of White residents were disproportionately represented in COVID-19 cases and deaths. Once the winter peak in cases reversed, higher SVI communities again began to experience worse COVID-19 outcomes.

While overall we demonstrate that socially vulnerable communities bear a disproportionate share of the burden of worse outcomes, during the time with the highest COVID-19 incidence and deaths (i.e the "third wave") in the U.S., it is of great interest that SVI became a non-significant predictor of incidence and death. There are potentially several explanations for these trends and our analyses with the temporal associations between the subcomponents of the SVI are of special interest. Our analysis demonstrates that especially during the early phases of the pandemic, communities with a greater share of minority populations rather than socioeconomic disadvantage or crowding, were disproportionately bearing the disastrous effects of the pandemic. Our more in-depth analysis of racial composition data, beyond the SVI subcomponent of minority status and language, demonstrated nationwide trends that counties with a large share of Black and Hispanic residents had especially worse outcomes during the pandemic prior to the third wave between November 2020 to January 2021. During this period, majority white communities demonstrated the highest incidence and death rates. Our nationwide results are congruent with a more in-depth analysis completed in Cuyahoga County, OH and Wayne County, MI – both counties that are socially vulnerable, but

Wayne County has a higher proportion of Black residents and thus, had worse COVID-19 outcomes including death and hospital utilization.[10] In addition, these temporal changes may at least partly also be due to the geographical spread of COVID-19 infections in the U.S.. Whereas the pandemic affected the northeastern, southeastern and southwestern states during the early and late phases of the year of study, it was predominantly affecting the mid-western and central states, that are proportionately less diverse, during the third peak observed in the Winter months of 2020. While we controlled for community-level comorbidity burden, communities with a higher proportion of minority populations are vulnerable to worse health outcomes due to other factors above and beyond what is measured in the SVI, including structural racism, marginalization, and poor healthcare access.[23] These factors need to be further studied.

Shortly following re-openings across the country in late May 2020, the socioeconomic component of the SVI became an independent predictor of worse COVID-19 outcomes and follows a similar trend to the overall SVI throughout the duration of our analysis. There is emerging evidence using cell phone data demonstrating that low income communities have been less able to socially distance during the COVID-19 pandemic, likely due to a multitude of factors including less capacity to work from home, or to take paid or unpaid time off from work, and limited savings.[24,25] During the COVID-19 pandemic, data has suggested that Hispanic communities in the U.S. are particularly vulnerable to financial insecurities compared to other racial/ethnic groups due to their disproportionate representation in industries that have been most affected by the pandemic and having jobs that cannot be performed from home.[26] We notice that temporal trends in incidence and death per capita for communities with greater proportion of Hispanic residents closely mirror that of the socioeconomic component of the SVI and thus, low SES may partially explain why Hispanic communities have had the worst overall COVID-19 outcomes for the duration of our analysis.

Another potential mechanism may be that low socioeconomic status communities have a higher burden of preexisting health conditions[24], but our findings are independent of community-level comorbidities. Finally, lower income communities are also more likely to live in multi-family crowded environments[27] and studies completed in restricted geographic locales such as New York City show the importance of these vulnerability markers[28] in COVID-19 outcomes. In accordance with these studies, we are seeing similar associations with both household composition and disability and housing type and transportation with COVID-19 outcomes throughout the duration of the analysis time period. [24,25] However, future studies using individual patient level information across the U.S. need to be conducted to further clarify these associations.

Limitations:

One of the major limitations of our study is that is mainly descriptive, ecological, and uses only county-level data, which does not allow us to account for individual characteristics that may drive COVID-19 outcomes in socially vulnerable communities. In addition, we used data collected from different data sources, each of which was gathered at slightly varying time points and as such, may not completely represent the features of the community at the time of our analysis. We attempted to account for as many confounders as possible but recognize that we may not have been able to adjust for all confounders (including vaccinations) driving the associations seen in this analysis. In terms controlling for county-level comorbidity, we use the HCC risk score, which was designed to reflect healthcare access and hospital admissions in a geographic area, as a proxy which may impact the associations seen. Therefore, studies incorporating individual patient level data which includes more confounders are needed to further delineate associations seen in this ecological study. Finally, we focused our analysis on county-level proportion of Black and Hispanic residents within the U.S. and do not extend it to include Asian or Native American residents. Future studies that encompass other minority groups and examine trends presented in our study on a worldwide basis are needed.

Reference:

1. Link BG. Epidemiological sociology and the social shaping of population health. *Journal of health and social behavior.* 2008;49(4):367-384.
2. Drago R, Miller K. Sick at work: infected employees in the workplace during the H1N1 pandemic. *Institute for Women's Policy Research.* 2010(B264).
3. Hutchins SS, Fiscella K, Levine RS, Ompad DC, McDonald M. Protection of racial/ethnic minority populations during an influenza pandemic. *American journal of public health.* 2009;99(S2):S261-S270.
4. Leclere FB, Jensen L, Biddlecom AE. Health care utilization, family context, and adaptation among immigrants to the United States. *Journal of health and social behavior.* 1994:370-384.
5. Buchmueller TC, Lo Sasso AT, Lurie I, Dolfin S. Immigrants and employer-sponsored health insurance. *Health services research.* 2007;42(1p1):286-310.
6. Quinn SC, Kumar S, Freimuth VS, Musa D, Casteneda-Angarita N, Kidwell K. Racial disparities in exposure, susceptibility, and access to health care in the USU.S. H1N1 influenza pandemic. *American journal of public health.* 2011;101(2):285-293.
7. Soyemi K, Medina-Marino A, Sinkowitz-Cochran R, et al. Disparities among 2009 pandemic Influenza A (H1N1) hospital admissions: A mixed methods analysis–Illinois, April–December 2009. *PloS one.* 2014;9(4).
8. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A social vulnerability index for disaster management. *Journal of homeland security and emergency management.* 2011;8(1).
9. Flanagan B, Hallisey E, Adams E, Lavery A. Measuring community vulnerability to natural and anthropogenic hazards: The Centers for Disease Control and Prevention's social vulnerability index. *Journal of Environmental Health.* 2018;80:34-36.
10. Gaynor TS, Wilson ME. Social Vulnerability and Equity: The Disproportionate Impact of COVID-19. *Public administration review.* 2020;80(5):832-838.
11. Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the U.S.: an analysis of spatially varying relationships. *American journal of preventive medicine.* 2020;59(3):317-325.
12. Khazanchi R, Beiter ER, Gondi S, Beckman AL, Bilinski A, Ganguli I. County-level association of social vulnerability with COVID-19 cases and deaths in the U.S.A. *Journal of General Internal Medicine.* 2020;35(9):2784-2787.
13. Kim SJ, Bostwick W. Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. *Health education & behavior.* 2020;47(4).
14. Nayak A, Islam SJ, Mehta A, et al. Impact of Social Vulnerability on COVID-19 Incidence and Outcomes in the United States. *medRxiv.* 2020.
15. Neelon B, Mutiso F, Mueller NT, Pearce JL, Benjamin-Neelon SE. Spatial and temporal trends in social vulnerability and COVID-19 incidence and death rates in the United States. *Plos one.* 2021;16(3):e0248702.
16. Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial Disparities in Incidence and Outcomes Among Patients With COVID-19. *JAMA Network Open.* 2020;3(9):e2021892-e2021892.
17. Yancy CW. COVID-19 and African Americans. *JAMA.* 2020;323(19):1891-1892.
18. Podewils LJ, Burket TL, Mettenbrink C, et al. Disproportionate Incidence of COVID-19 Infection, Hospitalizations, and Deaths Among Persons Identifying as Hispanic or Latino - Denver, Colorado March-October 2020. *MMWR Morb Mortal Wkly Rep.* 2020;69(48):1812-1816.

19. Li P, Kim MM, Doshi JA. Comparison of the performance of the CMS Hierarchical Condition Category (CMS-HCC) risk adjuster with the Charlson and Elixhauser comorbidity measures in predicting mortality. *BMC Health Serv Res.* 2010;10:245.
20. CMS. *Medicare data for the geographic variation public use file: a methodological overview.* Centers for Medicare Medicaid Services; May 27, 2021 2021.
21. National Centers for Environmental Information. https://www.ncei.noaa.gov/. Accessed.
22. Robert Wood Johnson Foundation. County Health Rankings & Roadmaps: Rankings Data & Documentation
. https://www.countyhealthrankings.org/explore-health-rankings/rankingsdata-documentation. Published 2020. Accessed.
23. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the U.S.A: evidence and interventions. *Lancet.* 2017;389(10077):1453-1463.
24. Weill JA, Stigler M, Deschenes O, Springborn MR. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proceedings of the National Academy of Sciences.* 2020;117(33):19658.
25. Orhun AY, Palazzolo M. Frugality Is Hard to Afford. *Journal of Marketing Research.* 2019;56(1):1-17.
26. Gonzalez D, Karpman M, Kenney GM, Zuckerman S. Hispanic adults in families with noncitizens disproportionately feel the economic fallout from COVID-19. *Washington, DC: Urban Institute.* 2020.
27. Melki IS, Beydoun HA, Khogali M, Tamim H, Yunis KA. Household crowding index: a correlate of socioeconomic status and inter-pregnancy spacing in an urban setting. *Journal of Epidemiology and Community Health.* 2004;58(6):476.
28. Emeruwa UN, Ona S, Shaman JL, et al. Associations Between Built Environment, Neighborhood Socioeconomic Status, and SARS-CoV-2 Infection Among Pregnant Women in New York City. *JAMA.* 2020;324(4):390-392.

FIGURE LEGENDS

Figure 1: Temporal association between COVID-19 incidence and the a) county-level Social Vulnerability Index (SVI) and its subcomponents: b) Socioeconomic Status, c) Household Composition and Disability, d) Minority Status and Language, and e) Housing Type and Transportation between March 22nd, 2020 and March 6th, 2021. The base model (red lines) is adjusted for proportion of population age >65 years and state-level COVID-19 testing. The green lines are additionally adjusted for CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities). The blue lines are additionally adjusted for environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter > 2.5 micrometers ($PM_{2.5}$).Of note, proportion age >65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable

Figure 2: Temporal association between COVID-19 death per capita and the a) county-level Social Vulnerability Index (SVI) and its subcomponents: b) Socioeconomic Status, c) Household Composition and Disability, d) Minority Status and Language, and e) Housing Type and Transportation between March 22nd, 2020 and March 6th, 2021. The base model (red lines) is adjusted for proportion of population age >65 years and state-level COVID-19 testing. The green lines are additionally adjusted for CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities). The blue lines are additionally adjusted for environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter > 2.5 micrometers ($PM_{2.5}$). Of note, proportion age >65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable

Figure 3: Temporal association of county-level racial composition (Black, Hispanic/Latino, White) and COVID-19 a) incidence and b) death per capita between March 22nd, 2020 and September 26th, 2020 after adjusting for proportion of population age >65 years, state-level COVID-19 testing, CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities) and environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter > 2.5 micrometers ($PM_{2.5}$).

Figure 4: County-level map of the U.S. showing a) incidence and b) death per capita for COVID-19 across three timepoints – July 2020, December 2020, and March 2021. County-level proportion of Black, Hispanic/Latino, and White residents is shown in (c). As shown, Black and Hispanic residents are disproportionately represented in the southeast and southwestern US, where outcomes were worst in July 2020 and again in March 2021. Midwestern states where there are less diverse communities (higher proportion of White residents) showed worst outcomes in December 2020.

Table 1. Components of the social vulnerability index

Socioeconomic Status	Below Poverty Unemployed Income No High School Diploma
Household composition and Disability	Age 65 years or older Age 17 years or younger Older than Age 5 years with a Disability
Minority Status and Language	Minority Speak English "Less than Well"
Housing Type and Transportation	Multi-Unit Structures Mobile Homes Crowding No Vehicle Group Quarters

Table 2. Overall association of county-level social vulnerability index (Incidence Risk Ratio [IRR] per-10 percentile increase) with incidence and death per capita of COVID-19 as of March 6th, 2021

	Model 1[a]		Model 2[b]		Model 3[c]	
	IRR (95% CI)	P-value	IRR (95% CI)	P-value	IRR (95% CI)	P-value
Incidence						
Overall Social Vulnerability Index*	1.03 (1.03, 1.04)	<0.001	1.02 (1.02, 1.03)	<0.001	1.02 (1.02, 1.03)	<0.001
Socioeconomic Status	1.02 (1.02, 1.03)	<0.001	1.02 (1.01, 1.02)	<0.001	1.02 (1.01, 1.02)	<0.001
Minority Status & Language	1.02 (1.01, 1.02)	<0.001	1.02 (1.01, 1.02)	<0.001	1.02 (1.01, 1.02)	<0.001
Housing Type & Transport	1.01 (1.00, 1.01)	0.003	0.99 (0.99, 1.00)	0.07	0.99 (0.99, 1.00)	0.29
Household Composition & Disability*	1.02 (1.01, 1.02)	<0.001	1.01(1.01, 1.02)	<0.001	1.01 (1.01, 1.02)	<0.001
Death Per Capita						
Overall Social Vulnerability Index*	1.07 (1.06, 1.08)	<0.001	1.05 (1.04, 1.06)	<0.001	1.04 (1.04, 1.05)	<0.001
Socioeconomic Status	1.07 (1.07, 1.08)	<0.001	1.05 (1.04, 1.06)	<0.001	1.05 (1.04, 1.05)	<0.001
Minority Status & Language	1.03 (1.02, 1.03)	<0.001	1.01 (1.00, 1.02)	0.003	1.01 (1.00, 1.02)	0.004
Housing Type & Transport	1.06 (1.05, 1.06)	<0.001	1.05 (1.04, 1.05)	<0.001	1.05 (1.04, 1.05)	<0.001
Household Composition & Disability*	1.04 (1.03, 1.05)	<0.001	1.02 (1.01, 1.02)	<0.001	1.02 (1.01, 1.02)	<0.001

[a] Base Model - adjusted for proportion of population age $\geq$65 years and state-level COVID-19 testing
[b] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities)
[c] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities) + Environmental Factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter $\geq$ 2.5 micrometers ($PM_{2.5}$)

* Proportion age $\geq$65 years not included as a covariate for models for overall social vulnerability index and household composition/disability because these indices contain the age variable.

Table 3. Overall association of county-level race/ethnic composition (Incidence Risk Ratio [IRR] per 10% increase) with incidence and case fatality rate of COVID-19 as of March 06, 2021

	Model 1[a]		Model 2[b]		Model 3[c]	
	IRR (95% CI)	P-value	IRR (95% CI)	P-value	IRR (95% CI)	P-value
Incidence						
White	0.99 (0.98, 0.99)	<0.001	1.00 (1.00, 1.01)	0.45	1.01 (1.00, 1.01)	0.20
Black	1.01 (0.99, 1.01)	0.24	0.99 (0.98, 1.00)	0.01	0.99 (0.98, 1.00)	0.01
Hispanic	1.06 (1.05, 1.07)	<0.001	1.06 (1.05, 1.07)	<0.001	1.07 (1.05, 1.08)	<0.001
Death Per Capita						
White	0.94 (0.93, 1.00)	<0.001	0.99 (0.97,1.00)	0.05	0.99 (0.97, 1.00)	0.05
Black	1.06 (1.04, 1.08)	<0.001	1.00 (0.99, 1.02)	0.91	1.00 (0.99, 1.02)	0.85
Hispanic	1.10 (1.08, 1.11)	<0.001	1.067(1.049, 1.085)	<0.001	1.07 (1.05, 1.08)	<0.001

[a] Base Model - adjusted for proportion of population age ≥65 years and state-level COVID-19 testing
[b] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities)
[c] Base Model + CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities) + Environmental Factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter ≥ 2.5 micrometers ($PM_{2.5}$)

Author contributions:

Shabatun J. Islam, MD - Concept, drafting and editing of manuscript, statistical analysis
Aditi Nayak, MD - Concept, drafting and editing of manuscript
Yingtian Hu, BS – Statistical analysis, review of manuscript
Anurag Mehta, MD- Concept, drafting and editing of manuscript
Katherine Dieppa, M.Eng. – Statistical analysis, review of manuscript
Zakaria Almuwaqqat, MD, MPH - Statistical analysis, review of manuscript
Yi-An Ko, PhD - Statistical analysis, review of manuscript
Shivani A. Patel, MPH, PhD - Concept, review of manuscript
Abhinav Goyal, MD, MHS - Concept, review of manuscript
Samaah Sullivan, PhD - Concept, review of manuscript
Tené T. Lewis, PhD - Concept, drafting and editing of manuscript
Viola Vaccarino, MD, PhD - Concept, drafting and editing of manuscript
Alanna A. Morris, MD, MSc - Concept, drafting and editing of manuscript
Arshed A. Quyyumi, MD - Concept, drafting and editing of manuscript, review of manuscript

Figure 1

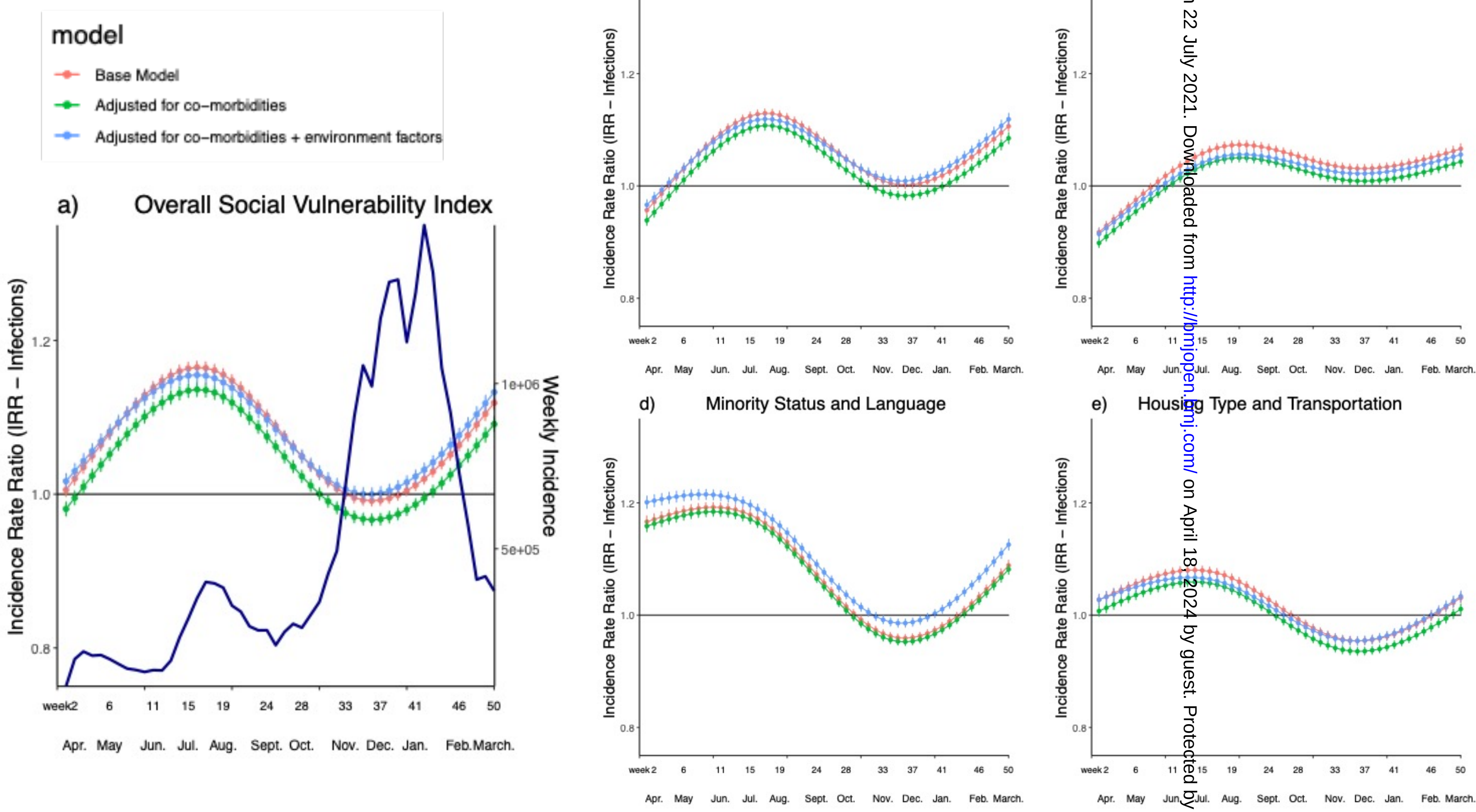

model
Base Model
Adjusted for co–morbidities
Adjusted for co–morbidities + environment factors
a) Overall Social Vulnerability Index
b) Socioeconomic Status
c) Household composition and Disability
d) Minority Status and Language
e) Housing Type and Transportation
Incidence Rate Ratio (IRR – Infections)
Weekly Incidence
1e+06
5e+05
0.8
1.0
1.2
week2 6 11 15 19 24 28 33 37 41 46 50
Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec. Jan. Feb. March.

Figure 2

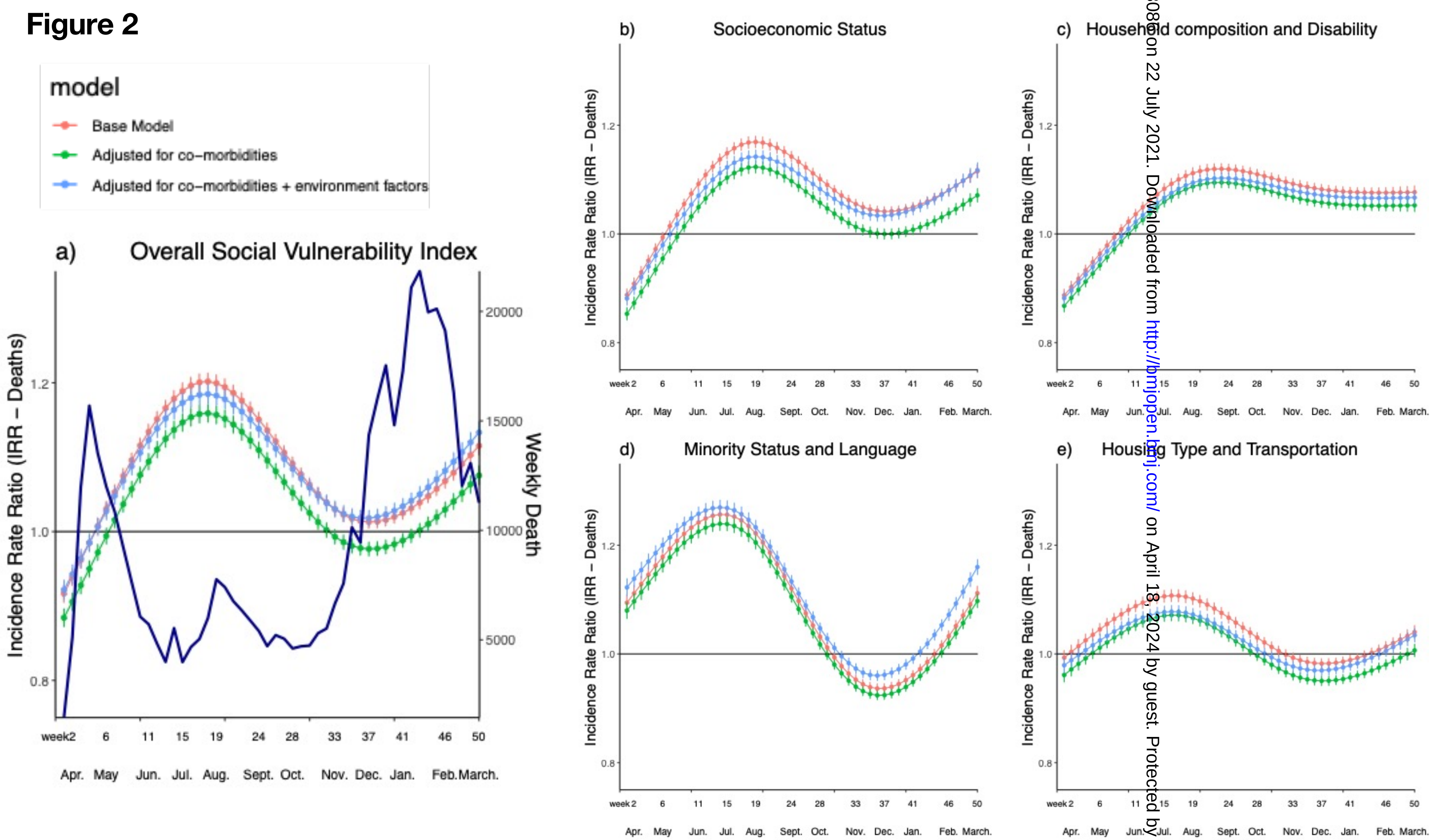

model
Base Model
Adjusted for co−morbidities
Adjusted for co−morbidities + environment factors
a) Overall Social Vulnerability Index
Incidence Rate Ratio (IRR – Deaths)
Weekly Death
week2 6 11 15 19 24 28 33 37 41 46 50
Apr. May Jun. Jul. Aug. Sept. Oct. Nov. Dec. Jan. Feb. March.
b) Socioeconomic Status
c) Household composition and Disability
d) Minority Status and Language
e) Housing Type and Transportation

Figure 3

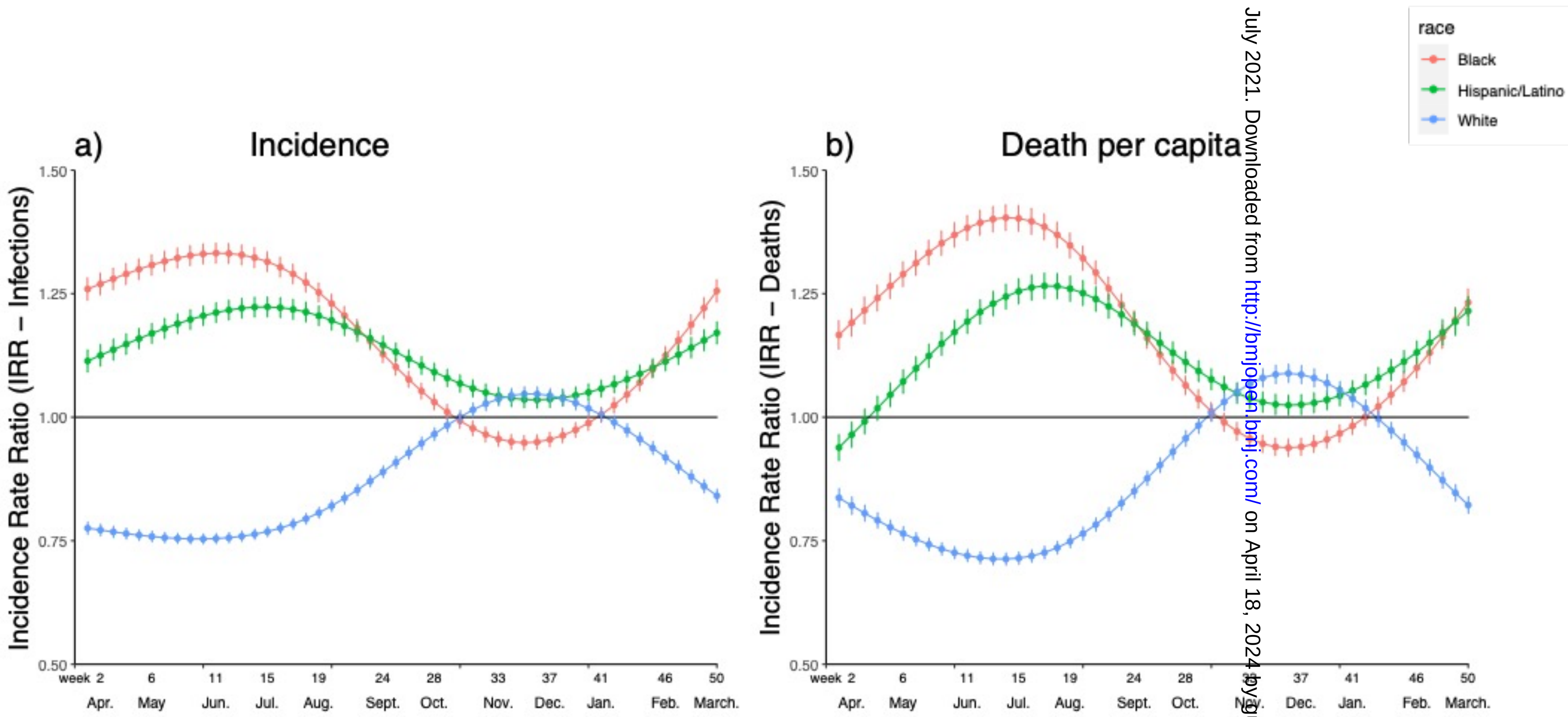
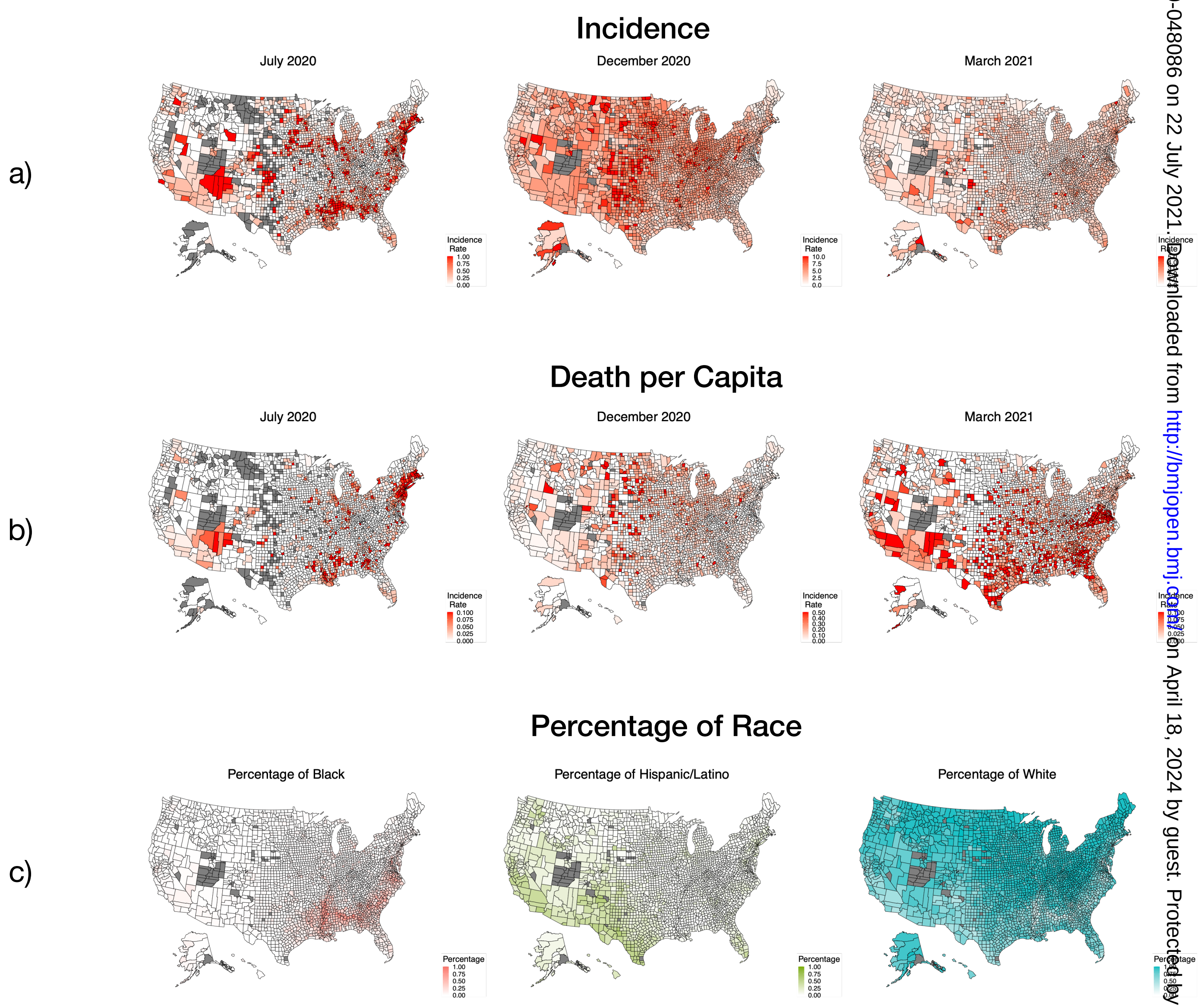

race
Black
Hispanic/Latino
White
a) Incidence
Incidence Rate Ratio (IRR – Infections)
b) Death per capita
Incidence Rate Ratio (IRR – Deaths)
1.50
1.25
1.00
0.75
0.50
week 2
6
11
15
19
24
28
33
37
41
46
50
Apr.
May
Jun.
Jul.
Aug.
Sept.
Oct.
Nov.
Dec.
Jan.
Feb.
March.

Figure 4


Incidence
a)
July 2020
December 2020
March 2021
Incidence Rate
1.00
0.75
0.50
0.25
0.00
10.0
7.5
5.0
2.5
0.0
Death per Capita
b)
0.100
0.075
0.050
0.025
0.000
0.50
0.40
0.30
0.20
0.10
0.00
Percentage of Race
c)
Percentage of Black
Percentage of Hispanic/Latino
Percentage of White
Percentage

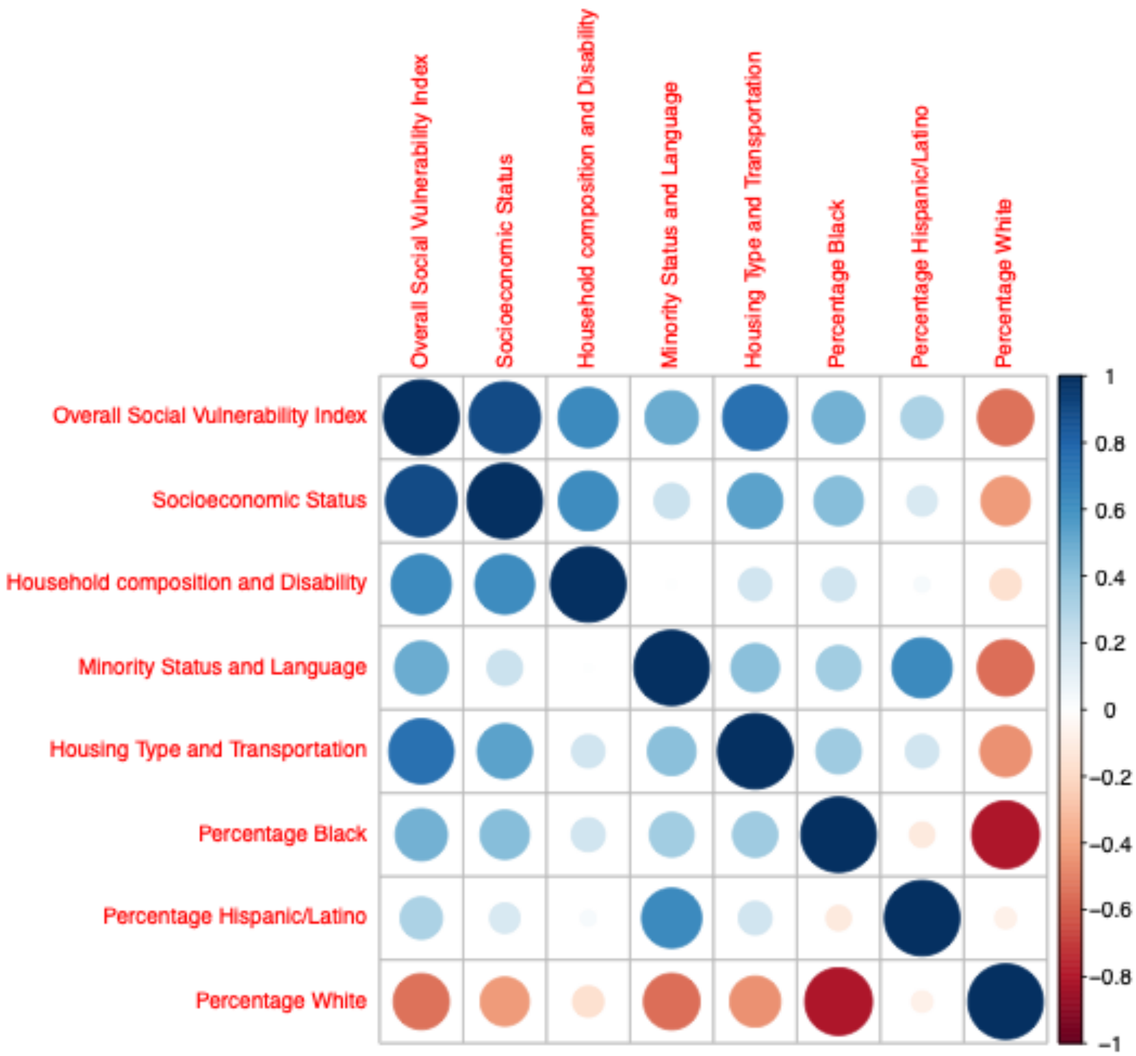
SUPPLEMENTARY MATERIAL

Table S1. Data sources used in the analysis (publicly available)

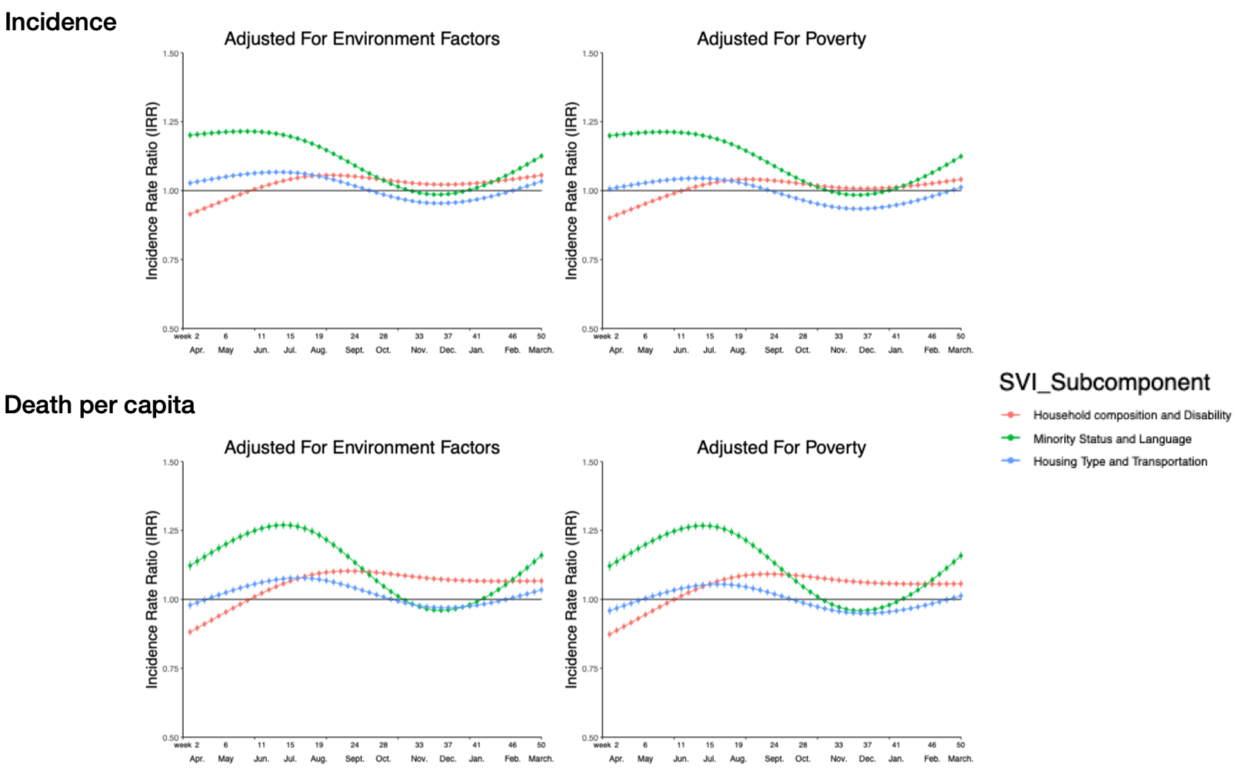
Data	Source
Outcomes	
Case Fatality Rate	Johns Hopkins Center for Systems Science and Engineering database, accessed May 28, 2021: https://coronavirus.jhu.edu/us-map
Incidence	Johns Hopkins Center for Systems Science and Engineering database, accessed May 28, 2021: https://coronavirus.jhu.edu/us-map
Exposures	
Social Vulnerability Index	Centers for Disease Control and Prevention (CDC) Geospatial Research, Analysis, and Services Program (GRASP) database, accessed May 28, 2021: https://svi.cdc.gov/
Racial Composition	US Census Bureau, accessed May 28, 2021: https://www.census.gov/data.html
Confounders	
Proportion age ≥65 years	Centers for Disease Control and Prevention (CDC) Geospatial Research, Analysis, and Services Program (GRASP) database, accessed May 28, 2021: https://svi.cdc.gov/
Hierarchical Condition Category score	Centers for Medicare and Medicaid Services (CMS), accessed May 28, 2021: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Geographic-Variation
Tests administered per state	The COVID Tracking Project, accessed May 28, 2021: https://covidtracking.com/

Environmental Factors	Temperature	National Centers for Environmental Information, accessed May 28, 2021: https://www.ncei.noaa.gov/
	Precipitation	National Centers for Environmental Information, accessed May 28, 2021: https://www.ncei.noaa.gov/
	$PM_{2.5}$	Robert Wood Johnson Foundation, accessed May 28, 2021: https://www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation

Supplemental Figure 1: Correlation plot of county-level overall SVI, its subcomponents, and county-level proportion of White, Black and Hispanic/Latino residents



Supplemental Figure 2: Temporal association between COVID-19 incidence (top) and death (bottom) and SVI subcomponents of a) Household Composition and Disability, b) Minority Status and Language, and c) Housing Type and Transportation between March 22nd, 2020 and March 66th, 2020. The models labelled *adjusted for environmental factor* were adjusted for proportion of population age >65 years, state-level COVID-19 testing, CMS average Hierarchical Condition Category (HCC) score (proxy for comorbidities), and environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation, and average particulate matter of diameter > 2.5 micrometers ($PM_{2.5}$). Of note, proportion age >65 years not included as a covariate for household composition/disability because this index contains the age variable.



STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
Page 1-3	√	(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale Page 4	2 √	Explain the scientific background and rationale for the investigation being reported
Objectives Page 5	3 √	State specific objectives, including any prespecified hypotheses
Methods		
Study design Page 5	4 √	Present key elements of study design early in the paper
Setting Page 5	5 √	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants Page 5	6 √	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables Page 5	7 √	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement Page 5-6	8* √	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias Page 6	9 √	Describe any efforts to address potential sources of bias
Study size Page 5	10 √	Explain how the study size was arrived at
Quantitative variables Page 6	11 √	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
Page 6	√	(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
Page 7	√	(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
Page 7	√	(b) Indicate number of participants with missing data for each variable of interest
Outcome data Page 7-10	15* √	Report numbers of outcome events or summary measures
Main results Page 7-10	16 √	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were

		adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses Page 10-11	17 √	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results Page 11	18 √	Summarise key results with reference to study objectives
Limitations Page 13	19 √	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation Page 11-12	20 √	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability Page 13	21 √	Discuss the generalisability (external validity) of the study results
Other information		
Funding Page 1	22 √	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.