

# BMJ Open Temporal trends in the association of social vulnerability and race/ethnicity with county-level COVID-19 incidence and outcomes in the USA: an ecological analysis

Shabatun J Islam <sup>1</sup>, Aditi Nayak,<sup>1</sup> Yingtian Hu,<sup>2</sup> Anurag Mehta,<sup>1</sup> Katherine Dieppa,<sup>3</sup> Zakaria Almuwaqqat,<sup>1</sup> Yi-An Ko,<sup>2</sup> Shivani A Patel,<sup>4</sup> Abhinav Goyal,<sup>1</sup> Samaah Sullivan,<sup>4</sup> Tené T Lewis,<sup>4</sup> Viola Vaccarino,<sup>4</sup> Alanna A Morris,<sup>1</sup> Arshed A Quyyumi<sup>1</sup>

**To cite:** Islam SJ, Nayak A, Hu Y, *et al*. Temporal trends in the association of social vulnerability and race/ethnicity with county-level COVID-19 incidence and outcomes in the USA: an ecological analysis. *BMJ Open* 2021;**11**:e048086. doi:10.1136/bmjopen-2020-048086

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-048086>).

SJI, AN, YH and AM contributed equally.

Received 18 December 2020  
Accepted 24 June 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Arshed A Quyyumi;  
[aquyyum@emory.edu](mailto:aquyyum@emory.edu)

## ABSTRACT

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

**Objective** We examined the temporal association of 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 USA in the year starting in March 2020.

**Methods** Counties (n=3091) with ≥50 COVID-19 cases by 6 March 2021 were included in the study. Associations between SVI (and its subcomponents) and county-level racial composition with 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 were adjusted for percentage of population aged ≥65 years, state-level testing rate, comorbidities using the average Hierarchical Condition Category score, and environmental factors including average fine particulate matter of diameter ≥2.5 µm, temperature and precipitation.

**Results** Higher SVI, indicative of greater social vulnerability, was independently associated with higher COVID-19 incidence (adjusted incidence rate ratio per 10 percentile increase: 1.02, 95% CI 1.02 to 1.03, p<0.001) and death per capita (1.04, 95% CI 1.04 to 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 in COVID-19-related adverse outcomes. Counties with greater proportion of Hispanic residents had worse outcomes throughout the duration of the analysis.

## Strengths and limitations of this study

- The study examined the full 12 months of county-level data in the USA, delineating the temporal trends in the association between Social Vulnerability Index and COVID-19 outcomes.
- The study investigated COVID-19 outcomes in predominantly black and Hispanic communities in comparison with white communities in the USA.
- The analysis is ecological, descriptive and at a county level rather than at an individual level.
- The analysis adjusted for confounders including county level age ≥65, comorbidities and environmental factors.
- The analysis was limited to the USA.

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

## 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.<sup>1</sup> Labour inequalities and household overcrowding may decrease the ability to adhere to social distancing guidelines.<sup>2</sup> Black and Hispanic individuals are more likely to work in front-line jobs with lack of workplace protections, which may

additionally increase exposure risk.<sup>3</sup> Additionally, race/ethnic minorities and immigrants are less likely to have access to appropriate and timely healthcare.<sup>3-5</sup> Evidence suggests that these inequalities also contributed to disease spread and adverse outcomes during the H1N1 influenza pandemic.<sup>6,7</sup>

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 the 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.<sup>8,9</sup> The SVI includes measures of county-level SES, 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 have demonstrated that socially vulnerable neighbourhoods have had worse outcomes during the early stages of the pandemic,<sup>10-14</sup> even given the fact that the SVI had been designed to mitigate such adverse outcomes for vulnerable communities. Data have shown that the association between SVI and COVID-19 outcomes temporally varied, with the trend reversing by October 2020,<sup>15</sup> but whether this continued to the latter durations of the pandemic is unknown. Similarly, it is now well known that black<sup>16,17</sup> and Hispanic<sup>18</sup> individuals, who represent the largest minority groups within the USA, 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 USA for the entire year from March 2020 to March 2021. Second, 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 between county-level proportion of black and Hispanic residents and COVID-19 outcomes.

## METHODS

### Study population and time frame

All US counties (n=3091) with at least 50 confirmed COVID-19 cases and greater than 4 weeks of follow-up data were included in the analysis. Data were analysed for a period of 50 weeks starting from 22 March 2020 to 6 March 2021.

### Patient and public involvement statement

Patients and the public were not involved in the design of this study.

### Outcomes

The primary outcomes of interest were county-level weekly COVID-19 incidence and death per capita of a county.

**Table 1** Components of the Social Vulnerability Index

Socioeconomic status	<ul style="list-style-type: none"> <li>▶ Below poverty.</li> <li>▶ Unemployed.</li> <li>▶ Income.</li> <li>▶ No high school diploma.</li> </ul>
Household composition and disability	<ul style="list-style-type: none"> <li>▶ Age 65 years or older.</li> <li>▶ Age 17 years or younger.</li> <li>▶ Older than age 5 years with a disability.</li> </ul>
Minority status and language	<ul style="list-style-type: none"> <li>▶ Minority.</li> <li>▶ Speak English 'less than well'.</li> </ul>
Housing type and transportation	<ul style="list-style-type: none"> <li>▶ Multiunit structures.</li> <li>▶ Mobile homes.</li> <li>▶ Crowding.</li> <li>▶ No vehicle.</li> <li>▶ Group quarters.</li> </ul>

Data were obtained from the Johns Hopkins Center for Systems Science and Engineering database.<sup>1</sup>

## Exposures

Exposures studied were (1) 2018 county-level SVI and its subcomponents obtained from the CDC GRASP database,<sup>2,3</sup> and (2) racial composition data (reported as proportion of black and Hispanic residents in a county) from the US Census Bureau American Community Survey (ACS) database.<sup>4</sup> 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.<sup>2,3</sup> The index combines statistical data from the US Census on 15 variables, grouped together into four related themes: SES, housing composition and disability, minority status and language, and housing type and transport (table 1). Each of these variables is ranked from lowest to highest vulnerability across census tracts in the USA 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 geographical region in the USA. The data were collected by the US Census Bureau as self-reported race/ethnicity between 2015 and 2019.<sup>4</sup>

## Confounders

Covariates included in all models were proportion of county population aged  $\geq 65$  years,<sup>4</sup> state-level COVID-19 testing rate obtained from the COVID-19 Tracking Project database,<sup>5</sup> 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.<sup>19,20</sup> While the score was designed to reflect healthcare access and hospital admissions in a geographical area, it does compare favourably with other comorbidity indices in prediction of outcomes,<sup>19</sup> and aggregate county-level scores are publicly available.<sup>20</sup> 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),<sup>21</sup> average daily precipitation<sup>21</sup> and average particulate matter of diameter  $\geq 2.5 \mu\text{m}$  ( $\text{PM}_{2.5}$ ).<sup>22</sup> All data sources used in this analysis are publicly available and are listed in online supplemental 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 df to allow for non-linear relationships. Similarly, overall associations and time-varying associations between county-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  $\geq 65$  years, state-level testing rate, HCC risk score, average daily temperature (degrees Fahrenheit),<sup>21</sup> average daily precipitation<sup>21</sup> and average  $\text{PM}_{2.5}$ <sup>22</sup> as described above. All analyses were performed using R V.3.6.1 (R Foundation for Statistical Computing). All p values were two-sided, with a significance level of 0.05.

## RESULTS

Among the 3142 counties in the USA, 3091 (98.38%) had  $>50$  confirmed COVID-19 cases as of 6 March 2021 and at least 4 weeks of 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 SES and least with minority status and language (online supplemental figure 1). The proportion of black residents correlated modestly ( $r=0.4$ ) and the proportion of Hispanic residents correlated slightly ( $r=0.2$ ) with overall SVI (online 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 to 1.03,  $p<0.001$ ) after adjusting for aforementioned confounders as of 6 March 2021. Among the SVI subcomponents, indices of SES (adjusted IRR per 10 percentile increase: 1.02, 95% CI 1.01 to 1.03,  $p<0.001$ ), minority status and language (adjusted IRR per 10 percentile increase: 1.02, 95% CI 1.01 to 1.02,  $p<0.001$ ), and household composition and disability (adjusted IRR per 10 percentile increase: 1.01, 95% CI 1.01 to 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 22 March 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 USA ('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.

### SVI subcomponents

The SVI subcomponent of minority status and language was an independent predictor of incidence from the beginning (22 March 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 USA ('third wave'). Similar to the overall SVI, it again became positively associated once the cases decreased in the USA around January–February 2021 (weeks 41–46) (figure 1).

**Table 2** Overall association of county-level Social Vulnerability Index (IRR per 10 percentile increase) with incidence and death per capita of COVID-19 as of 6 March 2021

	Model 1*		Model 2†		Model 3‡	
	IRR (95% CI)	P value	IRR (95% CI)	P value	IRR (95% CI)	P value
<b>Incidence</b>						
Overall Social Vulnerability Index§	1.03 (1.03 to 1.04)	<0.001	1.02 (1.02 to 1.03)	<0.001	1.02 (1.02 to 1.03)	<0.001
Socioeconomic status	1.02 (1.02 to 1.03)	<0.001	1.02 (1.01 to 1.02)	<0.001	1.02 (1.01 to 1.02)	<0.001
Minority status and language	1.02 (1.01 to 1.02)	<0.001	1.02 (1.01 to 1.02)	<0.001	1.02 (1.01 to 1.02)	<0.001
Housing type and transport	1.01 (1.00 to 1.01)	0.003	0.99 (0.99 to 1.00)	0.07	0.99 (0.99 to 1.00)	0.29
Household composition and disability§	1.02 (1.01 to 1.02)	<0.001	1.01 (1.01 to 1.02)	<0.001	1.01 (1.01 to 1.02)	<0.001
<b>Death per capita</b>						
Overall Social Vulnerability Index§	1.07 (1.06 to 1.08)	<0.001	1.05 (1.04 to 1.06)	<0.001	1.04 (1.04 to 1.05)	<0.001
Socioeconomic status	1.07 (1.07 to 1.08)	<0.001	1.05 (1.04 to 1.06)	<0.001	1.05 (1.04 to 1.05)	<0.001
Minority status and language	1.03 (1.02 to 1.03)	<0.001	1.01 (1.00 to 1.02)	0.003	1.01 (1.00 to 1.02)	0.004
Housing type and transport	1.06 (1.05 to 1.06)	<0.001	1.05 (1.04 to 1.05)	<0.001	1.05 (1.04 to 1.05)	<0.001
Household composition and disability§	1.04 (1.03 to 1.05)	<0.001	1.02 (1.01 to 1.02)	<0.001	1.02 (1.01 to 1.02)	<0.001

\*Base model: adjusted for proportion of population aged ≥65 years and state-level COVID-19 testing.

†Base model+CMS average HCC score (proxy for comorbidities).

‡Base model+CMS average HCC score (proxy for comorbidities)+environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation and average particulate matter of diameter ≥2.5 µm.

§Proportion aged ≥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.

CMS, Centers for Medicare and Medicaid Services; HCC, Hierarchical Condition Category; IRR, incidence rate ratio.

The SVI subcomponent of SES was an independent predictor of incidence after accounting for comorbidities 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).

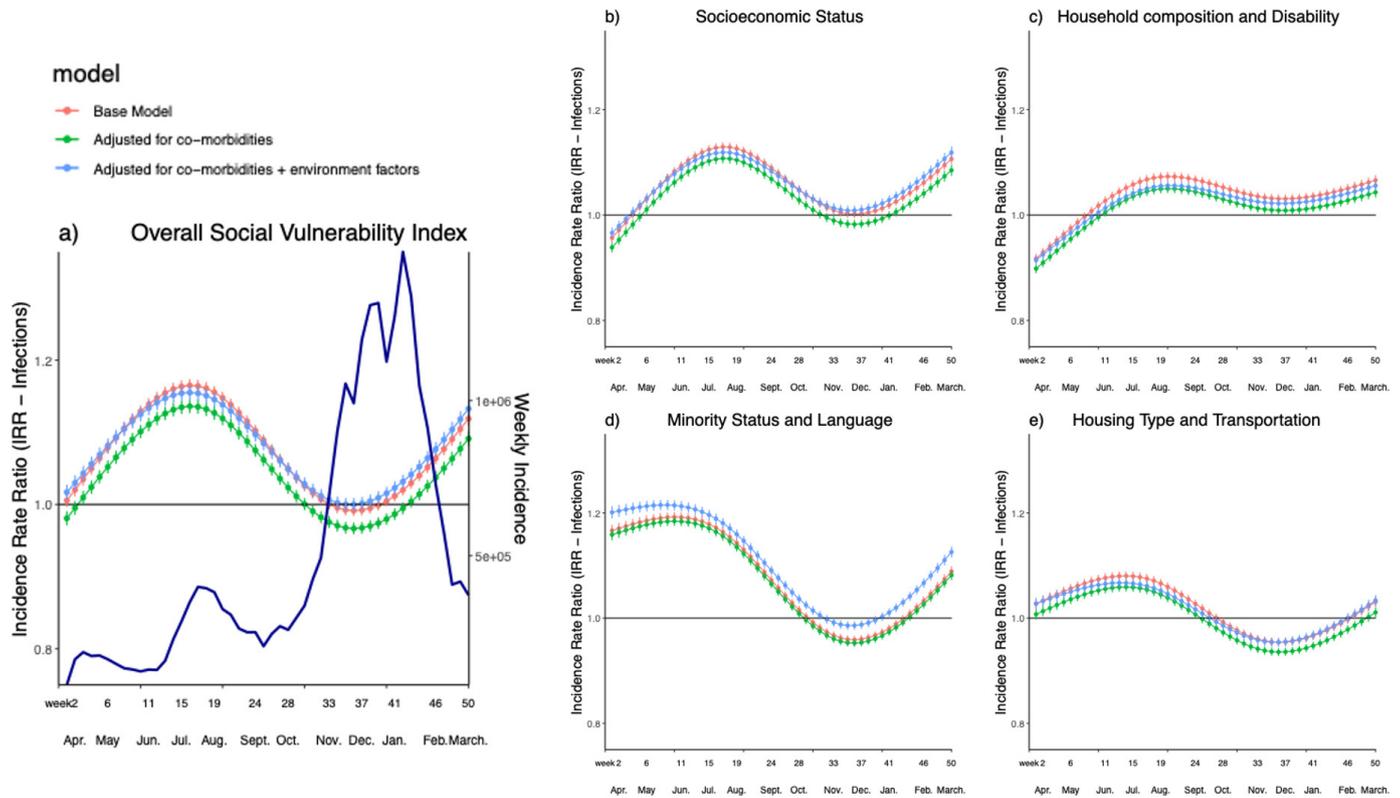
The indices of county-level household composition and disability and housing type and transportation become independent predictors of incidence of COVID-19 in early June 2020 (week 11) and late March 2020 (week 1), respectively. The strength of the positive association varied for county-level household 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 the 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 (online 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 to 1.05,  $p<0.001$ ) after adjusting for aforementioned confounders (table 2). All the SVI subcomponents, including indices of SES (adjusted IRR per 10 percentile increase: 1.05, 95% CI 1.04 to 1.05,  $p<0.001$ ), minority status and language (adjusted IRR per 10 percentile increase: 1.01, 95% CI 1.00 to 1.02,  $p=0.004$ ), housing type and transportation (adjusted IRR per 10 percentile increase: 1.05, 95% CI 1.04 to 1.05,  $p<0.001$ ), and household composition and disability (adjusted IRR per 10 percentile increase: 1.02, 95% CI 1.01 to 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 first 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



**Figure 1** Temporal association between COVID-19 incidence and (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 22 March 2020 and 6 March 2021. The base model (red lines) is adjusted for proportion of population aged  $\geq 65$  years and state-level COVID-19 testing. The green lines are additionally adjusted for CMS average Hierarchical Condition Category 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  $\geq 2.5 \mu\text{m}$ . Of note, the proportion aged  $\geq 65$  years was not included as a covariate for models for overall Social Vulnerability Index and household composition/disability because these indices contain the age variable. CMS, Centers for Medicare and Medicaid Services.

between November 2020 and January 2021 (weeks 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

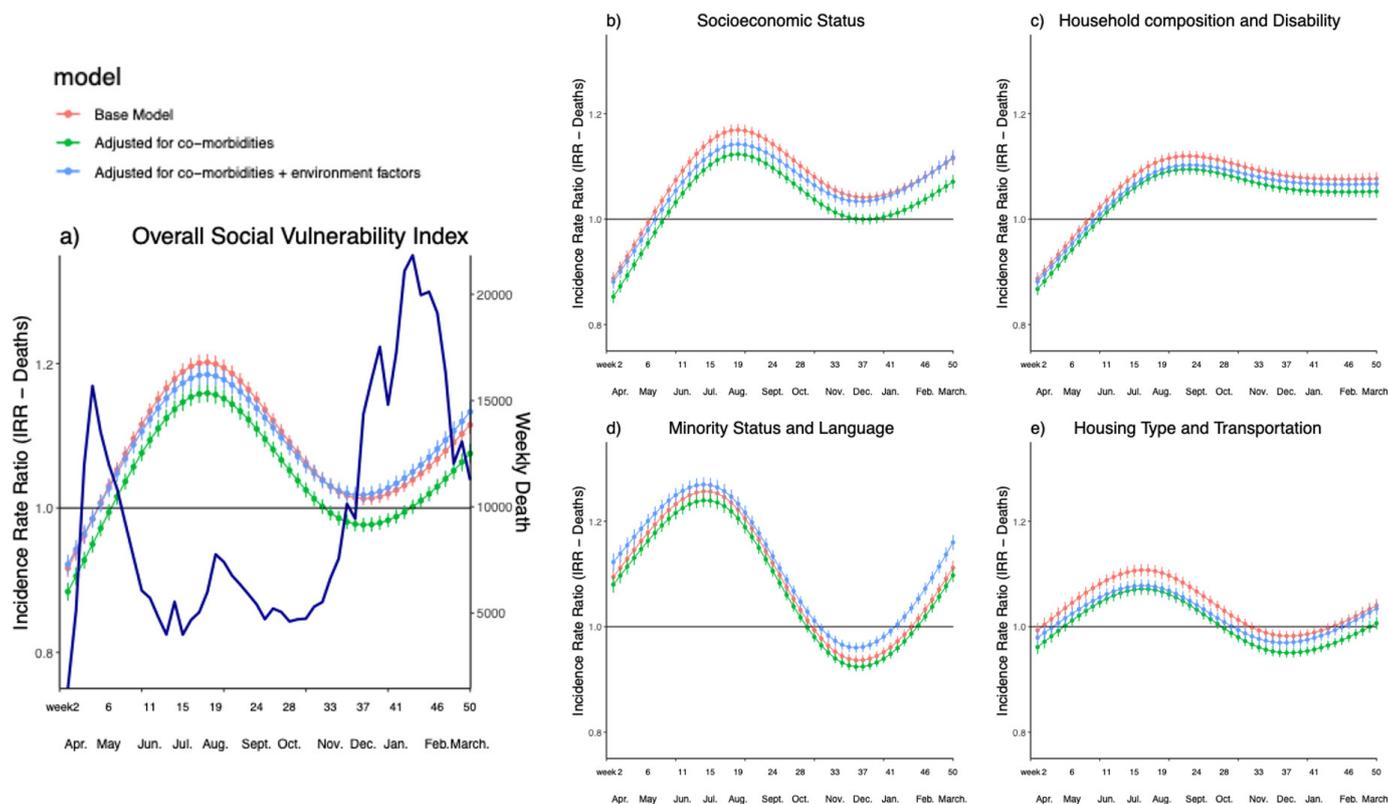
The SVI subcomponent of minority status and language was significantly and positively associated with COVID-19 death per capita from the beginning of the analysis. The strength of the association decreased starting in late July 2020 (week 17) and became negatively associated starting in mid-October 2020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

The SVI subcomponent of SES was independently and positively associated with death per capita after accounting for comorbidities 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 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 the 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 (online supplemental figure 2).

### Overall and temporal associations between race/ethnicity and COVID-19 incidence and 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 white, Hispanic and black residents within each county, based on county-level data from 2015 to



**Figure 2** Temporal association between COVID-19 death per capita and (A) county-level Social Vulnerability Index and its subcomponents (B) socioeconomic status, (C) household composition and disability, (D) minority status and language, and (E) housing type and transportation between 22 March 2020 and 6 March 2021. The base model (red lines) is adjusted for proportion of population aged  $\geq 65$  years and state-level COVID-19 testing. The green lines are additionally adjusted for CMS average Hierarchical Condition Category 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  $\geq 2.5 \mu\text{m}$ . Of note, the proportion aged  $\geq 65$  years was not included as a covariate for models for overall Social Vulnerability Index and household composition/disability because these indices contain the age variable. CMS, Centers for Medicare and Medicaid Services.

2019 from the US Census Bureau ACS database.<sup>4</sup> Proportionately more black residents reside in southern USA 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 to 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 to 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 to 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 to 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 until 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 2021, counties with higher proportion of white residents had higher incidence, which coincided with higher overall cases in the USA ('third wave'). Similarly, for a period of 11 weeks between November 2020 and January 2021, these communities had higher death per capita. Counties with higher proportion of Hispanic residents had

**Table 3** Overall association of county-level race/ethnic composition (IRR per 10% increase) with incidence and case fatality rate of COVID-19 as of 6 March 2021

	Model 1*		Model 2†		Model 3‡	
	IRR (95% CI)	P value	IRR (95% CI)	P value	IRR (95% CI)	P value
<b>Incidence</b>						
White	0.99 (0.98 to 0.99)	<0.001	1.00 (1.00 to 1.01)	0.45	1.01 (1.00 to 1.01)	0.20
Black	1.01 (0.99 to 1.01)	0.24	0.99 (0.98 to 1.00)	0.01	0.99 (0.98 to 1.00)	0.01
Hispanic	1.06 (1.05 to 1.07)	<0.001	1.06 (1.05 to 1.07)	<0.001	1.07 (1.05 to 1.08)	<0.001
<b>Death per capita</b>						
White	0.94 (0.93 to 1.00)	<0.001	0.99 (0.97 to 1.00)	0.05	0.99 (0.97 to 1.00)	0.05
Black	1.06 (1.04 to 1.08)	<0.001	1.00 (0.99 to 1.02)	0.91	1.00 (0.99 to 1.02)	0.85
Hispanic	1.10 (1.08 to 1.11)	<0.001	1.067 (1.049 to 1.085)	<0.001	1.07 (1.05 to 1.08)	<0.001

\*Base model: adjusted for proportion of population aged ≥65 years and state-level COVID-19 testing.

†Base model+CMS average HCC score (proxy for comorbidities).

‡Base model+CMS average HCC score (proxy for comorbidities)+environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation and average particulate matter of diameter ≥2.5 µm.

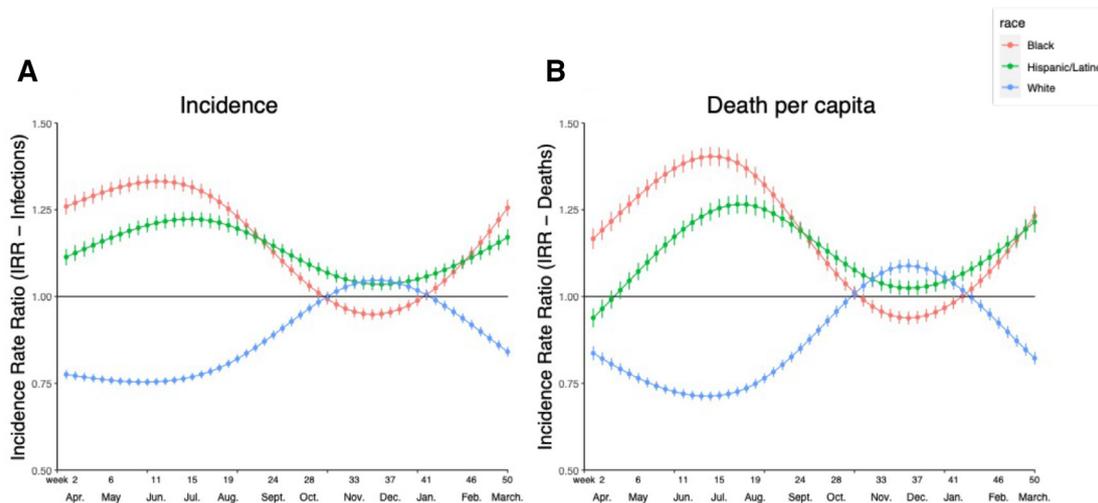
CMS, Centers for Medicare and Medicaid Services; HCC, Hierarchical Condition Category; IRR, incidence rate ratio.

both higher incidence and death per capita throughout the entire duration of the analysis.

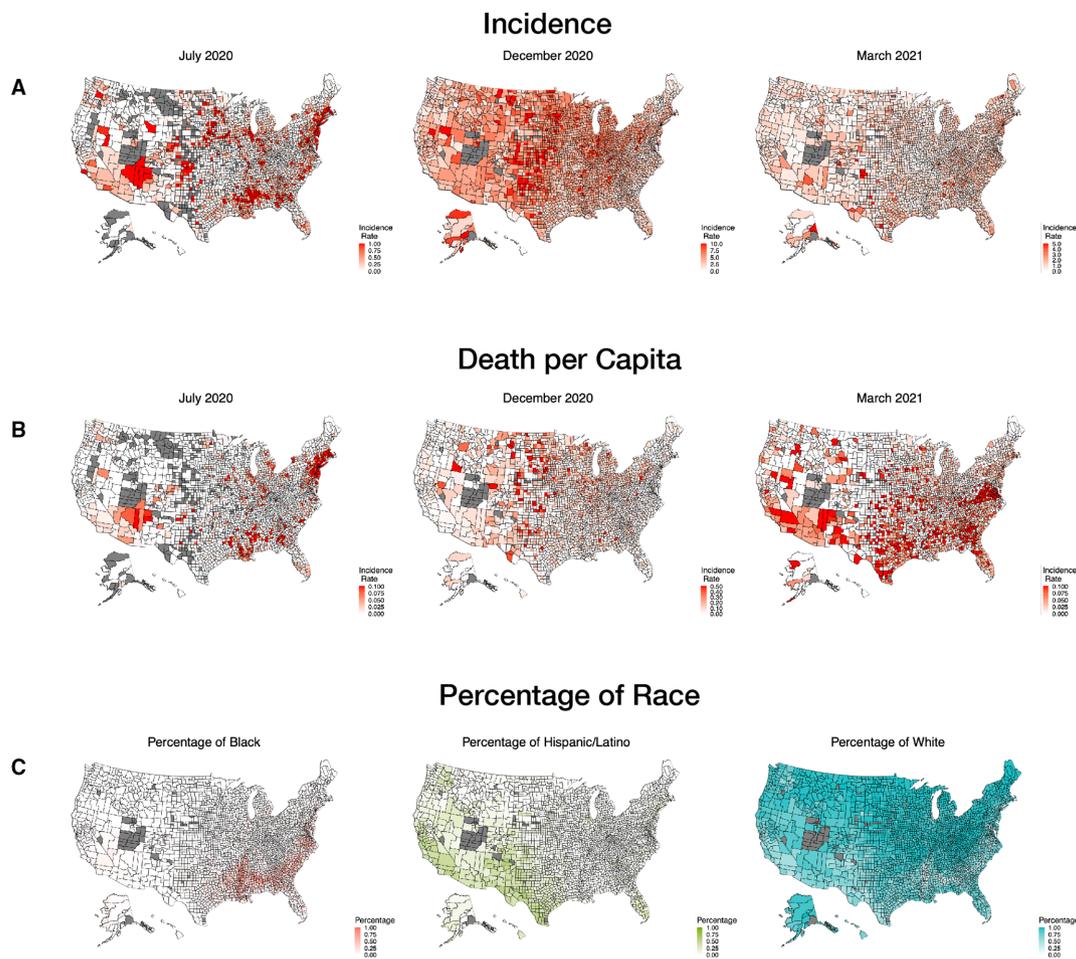
When the geographical changes in the incidence and mortality rates from COVID-19 are examined over the year, it is apparent that whereas the early part of the pandemic affected the northeastern USA and areas in the southeast and southwestern USA, areas that are enriched for minority populations, by the end of 2020 when the pandemic was at its peak ('third wave'), the midwestern states, with predominantly white populations, had the highest prevalence and mortality rates. By the spring of 2021, the geographical distribution of cases again changed back to the areas affected initially during the pandemic (figure 4).

### DISCUSSION

Herein, we demonstrate that US 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 are 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



**Figure 3** Temporal association of county-level racial composition (black, Hispanic/Latino, white) and COVID-19 (A) incidence and (B) death per capita between 22 March 2020 and 26 September 2020 after adjusting for proportion of population aged ≥65 years, state-level COVID-19 testing, CMS average Hierarchical Condition Category score (proxy for comorbidities), and environmental factors including average daily temperature (degrees Fahrenheit), average daily precipitation and average particulate matter of diameter ≥2.5 µm. CMS, Centers for Medicare and Medicaid Services.



**Figure 4** County-level map of the USA 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 USA, 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.

of infections. The absence of SVI being a predictor of COVID-19 outcomes coincided with a sharp increase in cases and deaths within the USA 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 (ie, the ‘third wave’) in the USA, 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 indepth 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 and January 2021. During this period, majority of the white communities demonstrated the highest incidence and death rates. Our nationwide results are congruent with a more indepth analysis completed in Cuyahoga County, Ohio and in Wayne County, Michigan—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 utilisation.<sup>10</sup> In addition, these temporal changes may at least partly also be due to the geographical spread of COVID-19 infections in the USA. 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 midwestern and central states, which 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, marginalisation and poor healthcare access.<sup>23</sup> These factors need to be further studied.

Shortly following reopenings 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.<sup>24 25</sup> During the COVID-19 pandemic, data have suggested that Hispanic communities in the USA are particularly vulnerable to financial insecurities compared with 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.<sup>26</sup> 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 SES communities have a higher burden of pre-existing health conditions,<sup>24</sup> but our findings are independent of community-level comorbidities. Finally, lower income communities are also more likely to live in multifamily crowded environments<sup>27</sup> and studies completed in restricted geographical locales such as New York City show the importance of these vulnerability markers<sup>28</sup> 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.<sup>24 25</sup> However, future studies using individual patient-level information across the USA need to be conducted to further clarify these associations.

### Limitations

One of the major limitations of our study is that it 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 timepoints 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 recognise that

we may not have been able to adjust for all confounders (including vaccinations) driving the associations seen in this analysis. In terms of controlling for county-level comorbidity, we use the HCC risk score, which was designed to reflect healthcare access and hospital admissions in a geographical area, as a proxy which may impact the associations seen. Therefore, studies incorporating individual patient-level data which include more confounders are needed to further delineate the associations seen in this ecological study. Finally, we focused our analysis on county-level proportion of black and Hispanic residents within the USA 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.

### Author affiliations

<sup>1</sup>Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, Georgia, USA

<sup>2</sup>Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

<sup>3</sup>City Operations, Getaround Incorporated, San Francisco, California, USA

<sup>4</sup>Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

**Twitter** Shabatun J Islam @shabatun and Aditi Nayak @AditiNayakMD

**Contributors** SI, AN, AM: concept, drafting and editing of manuscript, statistical analysis. YH, KD, ZA, Y-AK: statistical analysis, review of manuscript. SP, AG, SMS: concept, review of manuscript. TL, VV, AM: concept, drafting and editing of manuscript. AQ: concept, drafting and editing of manuscript, review of manuscript.

**Funding** The study received funding from Abraham J & Phyllis Katz Foundation, Byron Williams Jr, MD Fellowship Fund and the National Institutes of Health (T32 HL130025 and T32 HL007745-26A1).

**Map disclaimer** The inclusion of any map (including the depiction of any boundaries therein), or of any geographical or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** Given that all data are publicly available and no patient-level data are used, this study was exempt from institutional board review (IRB) by Emory University.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data used is publicly available. A table listing data sources is provided in the online supplemental materials.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iD

Shabatun J Islam <http://orcid.org/0000-0002-0886-9021>

## REFERENCES

- 1 Link BG. Epidemiological sociology and the social shaping of population health. *J Health Soc Behav* 2008;49:367–84.
- 2 Drago R, Miller K. *Sick at work: infected employees in the workplace during the H1N1 pandemic*. Institute for Women's Policy Research, 2010.
- 3 Hutchins SS, Fiscella K, Levine RS, *et al*. Protection of racial/ethnic minority populations during an influenza pandemic. *Am J Public Health* 2009;99 Suppl 2:S261–70.
- 4 Leclere FB, Jensen L, Biddlecom AE. Health care utilization, family context, and adaptation among immigrants to the United States. *J Health Soc Behav* 1994;35:370–84.
- 5 Buchmueller TC, Lo Sasso AT, Lurie I, *et al*. Immigrants and employer-sponsored health insurance. *Health Serv Res* 2007;42:286–310.
- 6 Quinn SC, Kumar S, Freimuth VS, *et al*. Racial disparities in exposure, susceptibility, and access to health care in the US H1N1 influenza pandemic. *Am J Public Health* 2011;101:285–93.
- 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:e84380.
- 8 Flanagan BE, Gregory EW, Hallisey EJ, *et al*. A social vulnerability index for disaster management. *J Homel Secur Emerg Manag* 2011;8.
- 9 Flanagan BE, Hallisey EJ, Adams E, *et al*. Measuring community vulnerability to natural and anthropogenic hazards: the centers for disease control and prevention's social vulnerability index. *J Environ Health* 2018;80:34–6.
- 10 Gaynor TS, Wilson ME. Social Vulnerability and Equity: The Disproportionate Impact of COVID -19. *Public Adm Rev* 2020;80:832–8.
- 11 Karaye IM, Horney JA. The impact of social vulnerability on COVID-19 in the U.S.: an analysis of spatially varying relationships. *Am J Prev Med* 2020;59:317–25.
- 12 Khazanchi R, Beiter ER, Gondi S, *et al*. County-Level association of social vulnerability with COVID-19 cases and deaths in the USA. *J Gen Intern Med* 2020;35:2784–7.
- 13 Kim SJ, Bostwick W. Social vulnerability and racial inequality in COVID-19 deaths in Chicago. *Health Educ Behav* 2020;47:509–13.
- 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. doi:10.1101/2020.04.10.20060962. [Epub ahead of print: 14 Apr 2020].
- 15 Neelon B, Mutiso F, Mueller NT, *et al*. Spatial and temporal trends in social vulnerability and COVID-19 incidence and death rates in the United States. *PLoS One* 2021;16:e0248702.
- 16 Muñoz-Price LS, Nattinger AB, Rivera F, *et al*. Racial disparities in incidence and outcomes among patients with COVID-19. *JAMA Netw Open* 2020;3:e2021892.
- 17 Yancy CW. COVID-19 and African Americans. *JAMA* 2020;323:1891–2.
- 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;69:1812–6.
- 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, 2021.
- 21 National centers for environmental information. Available: <https://www.ncei.noaa.gov/>
- 22 Robert Wood Johnson Foundation. County health rankings & roadmaps: rankings data & documentation, 2020. Available: <https://www.countyhealthrankings.org/explore-health-rankings/rankingsdata-documentation>
- 23 Bailey ZD, Krieger N, Agénor M, *et al*. Structural racism and health inequities in the USA: evidence and interventions. *Lancet* 2017;389:1453–63.
- 24 Weill JA, Stigler M, Deschenes O, *et al*. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proc Natl Acad Sci U S A* 2020;117:19658–60.
- 25 Orhun AY, Palazzolo M. Frugality is hard to afford. *J Mark Res* 2019;56:1–17.
- 26 Gonzalez D, Karpman M, Kenney GM, *et al*. *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, *et al*. Household crowding index: a correlate of socioeconomic status and inter-pregnancy spacing in an urban setting. *J Epidemiol Community Health* 2004;58:476–80.
- 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;324:390–2.