10-1-2021

Estimating the effects of race and social vulnerability on hospital admission and mortality from COVID-19

Joshua M Landman

Karen Steger-May

Karen E Joynt Maddox

Gmerice Hammond

Aditi Gupta

See next page for additional authors

Authors

Joshua M Landman, Karen Steger-May, Karen E Joynt Maddox, Gmerice Hammond, Aditi Gupta, Adriana M Rauseo, Min Zhao, and Randi E Foraker

JAMIA Open, 4(4), 2021, 1–11
https://doi.org/10.1093/jamiaopen/ooab111
Research and Applications

Research and Applications

Estimating the effects of race and social vulnerability on hospital admission and mortality from COVID-19

Joshua M. Landman [1,2], Karen Steger-May[3], Karen E. Joynt Maddox[4], Gmerice Hammond[4], Aditi Gupta[1,3], Adriana M. Rauseo [5], Min Zhao[1], and Randi E. Foraker [1,6]

[1]Institute for Informatics, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA, [2]Division of Computational and Data Sciences, Washington University in St. Louis, St. Louis, Missouri, USA, [3]Division of Biostatistics, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA, [4]Cardiovascular Division, Department of Medicine, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA, [5]Division of Infectious Diseases, Department of Medicine, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA, and [6]Division of General Medical Sciences, Department of Medicine, Washington University in St. Louis School of Medicine, St. Louis, Missouri, USA

Corresponding Author: Joshua M. Landman, MS, Campus Box 8102, 600 S. Taylor Avenue, Suite 102, St. Louis, MO 63110, USA; landman@wustl.edu

Received 8 June 2021; Revised 16 November 2021; Editorial Decision 3 December 2021; Accepted 6 December 2021

ABSTRACT

Objective: To estimate the risk of hospital admission and mortality from COVID-19 to patients and measure the association of race and area-level social vulnerability with those outcomes.
Materials and Methods: Using patient records collected at a multisite hospital system from April 2020 to October 2020, the risk of hospital admission and the risk of mortality were estimated for patients who tested positive for COVID-19 and were admitted to the hospital for COVID-19, respectively, using generalized estimating equations while controlling for patient race, patient area-level social vulnerability, and time course of the pandemic.
Results: Black individuals were 3.57 as likely (95% CI, 3.18–4.00) to be hospitalized than White people, and patients living in the most disadvantaged areas were 2.61 times as likely (95% CI, 2.26–3.02) to be hospitalized than those living in the least disadvantaged areas. While Black patients had lower raw mortality than White patients, mortality was similar after controlling for comorbidities and social vulnerability.
Discussion: Our findings point to potent correlates of race and socioeconomic status, including resource distribution, employment, and shared living spaces, that may be associated with inequitable burden of disease across patients of different races.
Conclusions: Public health and policy interventions should address these social factors when responding to the next pandemic.

Key words: COVID-19, social vulnerability, social determinants of health, mortality, hospital admission

Lay Summary

Using records from a hospital system spanning multiple sites and 2 states, we examined patients' risk of hospital admission and death due to COVID-19 over the first 6 months of the pandemic. Specifically, we measured the association of race and social vulnerability with the likelihood of admission and death. Our statistical models showed that Black patients were 3.57 times more likely than White patients to be hospitalized. We also took into account how socially vulnerable a given US Census tract is, and when incorporating that into our models, we estimated that patients living in the most disadvantaged areas were 2.61 times as likely to be hospitalized as those living in the least disadvantaged areas. We found no difference in mortality by race when controlling for underlying conditions and social vulnerability. These findings suggest that public health and policy interventions should address such things as resource distribution, employment, and shared living spaces, and other social factors that may be associated with race, socioeconomic status, and the inequitable burden of disease across patients of different races.

INTRODUCTION

There are striking racial inequities in outcomes among COVID-19 patients in the United States. Compared to Whites, Black individuals in the United States have had higher rates of testing positive for the virus, higher rates of hospitalization, and worse COVID-19-related outcomes such as intensive care unit (ICU) transfer and death.[1–5] In addition to racial differences in COVID-19 outcomes, there also appear to be area-level socioeconomic inequities; less affluent areas of the United States are being hit much harder than more affluent areas.[4,6,7]

Area-level socioeconomic inequities are prevalent and well-documented in the setting of other infectious diseases,[8] and often contribute independently to the variability in disease outcomes in statistical models. Yet relatively few studies to date have attempted to assess for socioeconomic disparities in diverse populations to determine how race and socioeconomic status (SES) may contribute independently to adverse outcomes among our most underserved populations in the setting of COVID-19.

Although both minority race and lower SES appear to have an independent effect on poorer outcomes among COVID-19 patients,[3] the role of SES differences in racially diverse populations has not been sufficiently evaluated at area-level units smaller than a ZIP code. Socioeconomic disparities tend to be difficult to study in the United States, as SES data are not systematically recorded in administrative databases and hospital records,[9] but there exist considerable public health implications for including SES in ongoing surveillance of COVID-19. Doing so may ultimately enhance communication and contact tracing in diverse communities, especially when recorded at more granular levels.

To address this gap, we examined the impact of area-level social vulnerability and patient race on the likelihood of being admitted to the hospital and dying in-hospital among patients with COVID-19. We determined the extent to which these outcomes also varied by age, gender, comorbidity score, and calendar time of diagnosis. We hypothesized that area-level SES defined at the level of the US Census tract (CT) would be associated with adverse outcomes and that model estimates for SES would be larger in magnitude among Black patients as compared to White patients. We used CT-level analysis instead of a less granular analysis at the zip code or county level as other studies have done[10] to better approximate individual-level patient experiences, which we believed would improve our analysis over those of prior works.[11]

MATERIALS AND METHODS

Data

For this study, electronic health record (EHR) data were acquired for all patients who had a COVID-19-related encounter with the Washington University School of Medicine/Barnes Jewish (WUSM/BJC) health system between 7 April 2020 and 31 October 2020. Follow-up outcome data were included for each of these patients through 30 November 2020. COVID-19-related encounters were defined as the set of encounters during which a patient was diagnosed as having COVID-19 or was admitted to the hospital due to COVID-19. We considered a hospitalization to be COVID-19 related if it occurred within 30 days after a positive COVID-19 test. In total, we collected data for 12 884 patients. The data collected for each observation included demographic information, such as age, date of birth, gender, race, ethnicity, and street address; medical information relating to COVID-19 testing and diagnosis, hospital admission codes, ICU transfer records, ventilation status, and hospital discharge codes. Race was categorized using self-reported race. Patients were classified as White or Black race if the patient endorsed the respective race in the absence of any other race, and they were coded as "Other" for race otherwise. We also collected in-hospital mortality records for patients who expired with a diagnosis of COVID-19 in-hospital during the study period. All data were aggregated to permit person-level analysis. This study was approved by the Washington University in St. Louis Institutional Review Board (IRB ID no.: 202003122) as exempt human subjects research with a HIPAA Authorization waiver for protected health information.

In addition to these data, we extracted all patient comorbidity information from our EHR system for each patient in our study data set. Up to 2 sets of comorbidities were used for each patient: comorbidities present at the time of testing positive for COVID-19, and comorbidities present at time of hospital admission due to COVID-19. Each of those sets of comorbidities was evaluated using data entered in the EHR at the time of each of those 2 events, respectively, plus or minus 2 days. Individual comorbidities were recorded as ICD-10 codes and aggregated into the Agency for Healthcare Research and Quality (AHRQ) Elixhauser Mortality Index as modified by Moore et al.[12,13] The AHRQ Index is validated to predict in-hospital death and is based on the presence of 29 conditions and ranges from −32 (lesser disease burden) to +99 (greater disease burden).

Geographic and social vulnerability data

We used ArcGIS to geocode patients' street address of residence and linked the Centers for Disease Control and Prevention's (CDC) social vulnerability index (SVI)[14] data with patient records. Using the *sf* library in R,[15] we matched the coded latitude and longitude coordinates to CTs, geographic data for which were acquired from the US Census shapefile repository. We limited our geographic area of interest to the St. Louis Metropolitan Statistical Area (MSA). The CDC SVI data set includes a percentile ranking of social vulnerabil-

ity for each CT in the United States, with 100% being the most vulnerable and 0% being the least vulnerable. The SVI comprises 4 subscales ("themes") relating to socioeconomic status; household composition; race, ethnicity, and language; and housing and transportation.[14] Both the overall percentile rankings and each theme ranking were included in our data set at the CT level.

Statistical analyses

For patients who tested positive for COVID-19, chi-square tests (for categorical variables) and unpaired *t*-tests (for normally distributed continuous variables) were used to compare patient characteristics by race. For continuous variables that were not normally distributed, Wilcoxon's test was used as a nonparametric alternative. To determine if there were monotonic gradients in outcomes across categories, all categorized variables were analyzed as unordered variables. Characteristics of patients with and without missing comorbidity were compared by unpaired *t*-test or chi-square test.

With the goal to estimate population-averaged (or marginal) effects, generalized estimating equations (GEE) were used to model the probability of COVID-19 hospital admission and the probability of in-hospital mortality, respectively. Initially, the interaction between race and SVI was assessed to test the null hypothesis that the association between race and the probability of each outcome was not significantly different across levels of social vulnerability. When the interaction was not significant ($P > .05$), the interaction was dropped from the model and the focus of analyses was to determine if race and SVI were significantly and independently associated with the probability of each outcome after adjusting for calendar time of diagnosis, age, gender, and AHRQ disease burden. Within the framework of the GEE, null hypotheses were assessed with statistical contrasts that compared the probability of the outcome for each category of the independent variable (IV) as compared to the referent category. Univariable (unadjusted) odds ratios (ORs) and their 95% confidence intervals (CIs) were computed for each independent variable. Adjusted ORs (aORs) and corresponding 95% CIs were reported from multivariable models, adjusted for all IVs. For each outcome, univariable models (ie, 1 model for each IV) and a multivariable model (ie, 1 model with all IVs included) were reported. Due to missing data for AHRQ disease burden, separate GEE models were performed in the subgroup of patients with comorbidity data.

GEE models included a robust variance estimator for US CT to account for the correlation between patients within the same CT with the independent correlation structure, a binomial probability distribution, and logit link function. The quasilikelihood under the independence model criterion (QIC) statistic was used to determine the working correlation structure. QIC and QICu were used to compare models for model selection (ie, goodness-of-fit) when IVs were included as continuous variables compared to categorized. Lack of collinearity was confirmed where the variance inflation factor (VIF) by linear regression was less than 2.[16]

Due to skewed distributions and/or poor model fit, inherently continuous variables were categorized for GEE analyses. We a priori defined tertile categories of *SVI percentile* based on SVI frequencies. The cutpoints for the most socially advantaged, middle, and most socially disadvantaged SVI categories were: SVI $\leq$ 26.36, SVI $>$ 26.36 to $\leq$ 62.91, and SVI $>$ 62.91, respectively. *Pandemic time course* was categorized based on the date of the COVID-positive test. A priori tertile categories for early, middle, and late were based on the number of days between the first and last COVID-positive dates in the data (ie, 69 days, 69 days, and 67 days, respectively), and to avoid cutpoints that fall within the weekend due to known deficiencies in weekend reporting. Categorization of *age* was determined by model fit where certain age groups functioned similarly and combining those groups was biologically plausible. Categorization of the *AHRQ Index* was guided by model fit statistics and from visual inspection seeking values that discriminate between patients with and without hospitalization. The AHRQ Index scores in the least, middle, and most disease burden categories were: AHRQ ≤ 0, AHRQ 1–5, and AHRQ ≥ 6, respectively.

Our data included 310 patients where a hospitalization occurred more than 30 days following the date of the first COVID-positive test. For analyses, these patients were *recoded as not hospitalized* because these hospitalizations are unlikely to be "primary COVID." Sensitivity analyses were performed to assess the robustness of the modeling of hospitalization by race and SVI (for univariable and multivariable models with the entire sample and the subsample with comorbidity data) to including hospitalizations that occurred more than 30 days from the first COVID-positive test.

The data analysis was generated using SAS® software, version 9.4 of the SAS System for Windows (SAS Institute Inc., Cary, NC, USA).

RESULTS

Patient characteristics

After filtering out patients whose addresses were unable to be geocoded and patients whose residences were outside of the St. Louis MSA, we retained observations for 8645 patients who tested positive for COVID-19. Of those, 529 (6.1%) were excluded because the reported race was not Black or White, 1 (<0.1%) was excluded for lack of gender data, and 74 (0.9%) were excluded because there was a discrepancy between hospital admission status and hospitalization outcomes such as ICU admission or in-hospital mortality.

Among 8041 persons who tested positive for COVID-19 in the analytic sample, the average age was 48.2 ± 21.3 years (Table 1). The majority of patients identified as White (55.9% vs 75.5% in the St. Louis MSA[17]), female (59.8%), and had a low comorbidity burden based on the AHRQ Index (73.0%). Forty-four percent of patients identified as Black versus 18.0% in the St. Louis MSA.[17] The most prevalent AHRQ individual comorbidity was hypertension (27.4%). Compared to White patients, Black patients were more likely to live in the most socially disadvantaged areas (54.9% Black; 17.2% White; $P < .0001$), test positive early in the pandemic (39.3% Black; 20.7% White; $P < .0001$), be younger (46.0 ± 20.2 years Black; 50.0 ± 22.1 years White; $P < .0001$), and be female (61.3% Black; 58.7% White; $P = .019$).

Fifteen percent of the analytic sample were missing data for comorbidities at the time of testing positive for COVID-19. Compared to patients with comorbidity data, patients with missing data were more likely to be Black (50.3% vs. 43.0%, $P < .0001$), older (51.5 ± 22.6 years vs. 47.6 ± 21.0 years; $P < .0001$), male (43.2% vs. 39.6%; $P = .02$), test positive early in the pandemic (44.3% vs. 26.1%; $P < .0001$), live in the most socially disadvantaged areas (36.1% vs. 33.4%; $P < .0001$), and less likely to be hospitalized (27.0% vs. 43.4%; $P < .0001$). Among Black patients, missing comorbidity data was not associated with age (47.5 ± 20.3 years vs. 45.7 ± 20.1 years; $P = .05$), gender (40.3% vs. 38.4%; $P = .38$), or living in the most socially disadvantaged areas (51.7% vs. 55.6%, $P = .89$). Black patients with missing data were more likely to test

Table 1. Patient characteristics at first COVID-19 positive test overall and by race for 8041 patients who first tested positive between 7 April 2020 and 31 October 2020

Characteristic	Overall (No. = 8041)		By race				*P*-value
			White (No. = 4494)		Black (No. = 3547)		
	No.	Summary statistic	No.	Summary statistic	No.	Summary statistic	
Demographics							
SVI percentile, median (IQR), range	8041	45.1 (54.9), 0.1 to 99.7	4494	28.6 (40.9), 0.1 to 99.7	3547	68.4 (36.9), 0.1 to 99.7	<.0001[a]
SVI category, no. (%)	8041		4494		3547		<.0001[b]
Most socially advantaged		2657 (33.0%)		2133 (47.5%)		524 (14.8%)	
Middle		2666 (33.2%)		1590 (35.4%)		1076 (30.3%)	
Most socially disadvantaged		2718 (33.8%)		771 (17.2%)		1947 (54.9%)	
Pandemic time course category, no. (%)	8041		4494		3547		<.0001[b]
Early		2322 (28.9%)		929 (20.7%)		1393 (39.3%)	
Middle		3113 (38.7%)		1679 (37.4%)		1434 (40.4%)	
Late		2606 (32.4%)		1886 (42.0%)		720 (20.3%)	
Age (y), mean (SD), range	8041	48.2 (21.3), 0 to 103	4494	50.0 (22.1), 0 to 102	3547	46.0 (20.2), 0 to 103	<.0001[c]
Age category (y), no. (%)	8041		4494		3547		<.0001[b]
0–19		675 (8.4%)		410 (9.1%)		265 (7.5%)	
20–49		3430 (42.7%)		1701 (37.8%)		1729 (48.8%)	
50–64		1931 (24.0%)		1077 (24.0%)		854 (24.1%)	
65–74		1057 (13.2%)		676 (15.0%)		381 (10.7%)	
≥75		948 (11.8%)		630 (14.0%)		318 (9.0%)	
Gender, no. (%)	8041		4494		3547		.019[b]
Female		4810 (59.8%)		2637 (58.7%)		2173 (61.3%)	
Male		3231 (40.2%)		1857 (41.3%)		1374 (38.7%)	
Elixhauser comorbidities[d]							
AHRQ Elixhauser Mortality Index, median (IQR), range	6823	0 (2), -22 to +51	3889	0 (0), -22 to +45	2934	0 (3), -14 to +51	.001[a]
AHRQ Elixhauser Mortality Index category, no. (%)	6823		3889		2934		<.0001[b]
Least disease burden		4979 (73.0%)		3001 (77.2%)		1978 (67.4%)	
Middle		720 (10.6%)		319 (8.2%)		401 (13.7%)	
Most disease burden		1124 (16.5%)		569 (14.6%)		555 (18.9%)	
Presence of the 5 most prevalent AHRQ comorbidities							
Hypertension, uncomplicated or complicated, no. (%)	6823	1867 (27.4%)	3889	866 (22.3%)	2934	1001 (34.1%)	<.0001[b]
Diabetes, uncomplicated, no. (%)	6823	908 (13.3%)	3889	370 (9.5%)	2934	538 (18.3%)	<.0001[b]
Chronic pulmonary disease, no. (%)	6823	874 (12.8%)	3889	393 (10.1%)	2934	481 (16.4%)	<.0001[b]
Depression, no. (%)	6823	474 (6.9%)	3889	291 (7.5%)	2934	183 (6.2%)	.045[b]
Fluid and electrolyte disorders, no. (%)	6823	418 (6.1%)	3889	192 (4.9%)	2934	226 (7.7%)	<.0001[b]

Abbreviations: AHRQ: Agency for Healthcare Research and Quality; IQR: interquartile range defined as the 75th minus the 25th percentile; No./no.: number of patients; SD: standard deviation; SVI: Social Vulnerability Index; y: year; %: percent of the sample represented by the column.

[a]*P*-value compares characteristics for White versus Black patients by Wilcoxon's test.

[b]*P*-value compares characteristics for White versus Black patients by chi-square test.

[c]*P*-value compares characteristics for White versus Black patients by unpaired *t*-test.

[d]Comorbidities reflect data that are available within ±2 days of the first positive COVID-19 test.

positive early in the pandemic (48.3% vs. 37.4%; $P < .0001$) and less likely to be hospitalized (27.7% vs. 64.0%; $P < .0001$) compared to Black patients with comorbidity data.

Likelihood of COVID-related hospital admission

Approximately 40% of patients who tested positive for COVID-19 were hospitalized (3292 of 8041; 40.9%). A greater proportion of COVID-positive Black patients were hospitalized (2048 of 3547; 57.7%) compared to White patients (1244 of 4494; 27.7%) (Table 2). Without adjustment for other characteristics, Black patients were 3.6 times as likely to be hospitalized compared to White patients (95% CI, 3.18–4.00). On average, hospitalized Black patients (48.3 ± 21.1 years) were substantially younger than hospitalized White patients (60.8 ± 21.6 years; $P < .0001$). Patients living in the most disadvantaged areas were 2.6 times as likely to be hospitalized compared to patients living in the most socially advantaged areas (95% CI, 2.26–3.02, respectively). Testing positive early in the pandemic, older age, and male gender were significantly associated with increased likelihood of hospitalization.

After adjusting for SVI, pandemic time course, age, and gender, Black patients remained nearly 4 times as likely to be hospitalized compared to White patients (aOR = 3.92, 95% CI, 3.43–4.49). Similarly, in the fully adjusted model, the risk of hospitalization for patients living in the most disadvantaged areas persisted (aOR = 1.49, 95% CI, 1.28–1.73).

Among the subgroup of 6823 patients with comorbidity data, Black race (aOR = 4.55, 95% CI, 3.95–5.24) and living in high SVI areas (aOR = 1.47, 95% CI, 1.25–1.73) remained independently associated with increased likelihood of hospitalization after adjusting for all measured characteristics including comorbidity disease burden.

The higher likelihood of hospitalization for Black patients did not vary across SVI levels among all COVID-positive patients ($P_{interaction} = .92$) or among the subgroup with comorbidity data ($P = .79$) (see Supplementary Table 1 for patient characteristics stratified by race and hospital admission status).

Sensitivity analyses to assess the robustness of the results for race and SVI when patients whose hospitalization occurred more than 30 days of the first COVID-positive test were analyzed as not hospitalized versus hospitalized revealed no change in conclusions for all race and SVI estimates (data not reported).

Likelihood of in-hospital mortality

Among the patients who had a COVID-related hospital admission (n = 3292), roughly 7% died in the hospital (n = 223; 6.8%). A smaller proportion of admitted Black patients died in the hospital (115 of 2048; 5.6%) compared to admitted White patients (108 of 1244; 8.7%) (Table 3). Without adjustment for other characteristics, the odds of dying in the hospital was 37% lower for Black patients compared to White patients (OR = 0.63, 95% CI, 0.47–0.84). The average age for patients who died in-hospital was similar for Black (74.1 ± 14.1 years) and White patients (76.5 ± 12.9 years; $P = .19$). Living in a high SVI area was not significantly associated with the risk of in-hospital mortality (OR = 0.75, 95% CI, 0.52–1.07). Testing positive early in the pandemic, older age, and male gender were significantly associated with increased likelihood of in-hospital mortality.

After adjusting for SVI, pandemic time course, age, and gender, the association between Black race and risk of mortality was attenuated (aOR = 0.96; 95% CI, 0.68–1.34). In the fully adjusted model, patients who were 75 years or older were more than 40 times as likely to die compared to patients under age 50 (aOR = 43.93, 95% CI, 20.18–95.60).

Among hospitalized patients with comorbidity data (n = 2982), Black race was not independently associated with increased likelihood of in-hospital mortality (aOR = 0.82, 95% CI, 0.57–1.17) after adjusting for measured characteristics including comorbidity burden. In the fully adjusted model, patients who were 75 years or older remained nearly 40 times as likely to die compared to patients under age 50 (aOR = 37.22, 95% CI, 16.44–84.24).

The odds of in-hospital mortality for Black compared to White patients was similar across levels of social vulnerability among all hospitalized patients ($P_{interaction} = .26$) and among the subgroup with comorbidity data ($P = .23$) (see Supplementary Table 2 for patient characteristics stratified by race and in-hospital mortality).

DISCUSSION

In this study, we found that Black patients with COVID-19 were markedly more likely to be admitted to the hospital, conditional on a positive test for COVID-19, than White patients. Similarly, patients from highly socially vulnerable neighborhoods were more likely to be admitted than those from less vulnerable ones. However, among hospitalized patients, mortality was similar for Black and White patients and among those in neighborhoods with high or low social vulnerability. There was no significant interaction between SVI and race such that SVI portended worse prognosis in both Black and White patients with COVID-19.

Our finding of elevated hospitalization rates in Black versus White patients is consistent with prior results.[3,4] The mechanisms underlying these differences is, however, unclear. One possibility is that, due to historic and ongoing structural and systemic racism, poverty and other barriers to access to health are disproportionately concentrated in Black neighborhoods. A lack of access to care may have been associated with presenting for evaluation and testing later in the course of disease, or only when symptoms reach crisis level, and therefore lead to a higher likelihood of hospitalization. A robust literature describes barriers to access for Black individuals and those living in socioeconomically deprived areas, and those issues could underlie the patterns we found.[18,19] It is also possible that COVID manifested more severely among Black patients, such that a higher proportion of patients were hospitalized due to the course of their disease alone. While others have suggested that underlying conditions are an important factor in elevated hospitalization rates in Black populations,[20] the fact that adjusting for comorbidities did not eliminate this association suggests that differential disease severity based on clinical characteristics alone does not likely drive our findings. Our findings echo those of Rentsch et al,[21] who came to similar conclusions in a substantially larger, national study at a less granular geospatial resolution. Even when analyzed through the more granular, CT-level lens we used and having adjusted for social vulnerability, however, Black patients still had elevated hospitalization rates. We suggest addressing spatial autocorrelation in future study designs exploring this and similar questions, which, to our knowledge, has not been taken into account in similar studies.

Another key finding of our study was that in fully adjusted models, neither race nor SVI were predictive of in-hospital mortality, although advanced age (those 75 years and older) was a robust predictor, which is again consistent with prior research.[4] Of note, Black patients who were hospitalized were over 10 years younger on average than White patients. This provides further support for the

Table 2. Association of patient characteristics at first COVID-19 positive test with the likelihood of COVID-related hospital admission

Population	Characteristic	No. (% of category admitted)	Univariable model[a]		Multivariable model[b]	
			OR (95% CI)	*P*-value	aOR (95% CI)	*P*-value
COVID-19 positive (No. = 8041)	Race					
	White	1244 (27.7%)	referent		referent	
	Black	2048 (57.7%)	3.57 (3.18–4.00)	<.0001	3.92 (3.43–4.49)	<.0001
	SVI category					
	Most socially advantaged	808 (30.4%)	referent		referent	
	Middle	1036 (38.9%)	1.45 (1.23–1.72)	<.0001	1.13 (0.97–1.32)	.10
	Most socially disadvantaged	1448 (53.3%)	2.61 (2.26–3.02)	<.0001	1.49 (1.28–1.73)	<.0001
	Pandemic time course category					
	Early	1128 (48.6%)	1.68 (1.47–1.92)	<.0001	0.99 (0.86–1.14)	.89
	Middle	1227 (39.4%)	1.16 (1.04–1.30)	.010	0.96 (0.85–1.09)	.54
	Late	937 (36.0%)	referent		referent	
	Age category (y)					
	0–19	197 (29.2%)	referent		referent	
	20–49	1213 (35.4%)	1.33 (1.11–1.58)	.002	1.21 (1.00–1.45)	.049
	50–64	746 (38.6%)	1.53 (1.26–1.85)	<.0001	1.52 (1.25–1.84)	<.0001
	65–74	519 (49.1%)	2.34 (1.90–2.89)	<.0001	2.79 (2.23–3.48)	<.0001
	≥75	617 (65.1%)	4.52 (3.59–5.69)	<.0001	6.01 (4.72–7.65)	<.0001
	Gender					
	Female	1812 (37.7%)	referent		referent	
	Male	1480 (45.8%)	1.40 (1.27–1.54)	<.0001	1.42 (1.28–1.57)	<.0001
COVID-19 positive, subgroup with nonmissing comorbidity data (No. = 6823)	Race					
	White	1085 (27.9%)	referent		referent	
	Black	1878 (64.0%)	4.60 (4.11–5.14)	<.0001	4.55 (3.95–5.24)	<.0001
	SVI category					
	Most socially advantaged	742 (31.7%)	referent		referent	
	Middle	906 (41.1%)	1.50 (1.24–1.81)	<.0001	1.13 (0.96–1.33)	.13
	Most socially disadvantaged	1315 (57.7%)	2.94 (2.52–3.44)	<.0001	1.47 (1.25–1.73)	<.0001
	Pandemic time course category					
	Early	970 (54.4%)	2.09 (1.81–2.42)	<.0001	1.13 (0.96–1.33)	.13
	Middle	1145 (42.3%)	1.29 (1.14–1.45)	<.0001	1.01 (0.88–1.16)	.88
	Late	848 (36.3%)	referent		referent	
	Age category (y)					
	0–19	197 (32.4%)	referent		referent	
	20–49	1156 (39.4%)	1.36 (1.13–1.63)	.001	1.23 (1.01–1.50)	.038
	50–64	667 (40.3%)	1.40 (1.15–1.72)	.001	1.26 (1.02–1.56)	.031
	65–74	445 (48.4%)	1.95 (1.56–2.44)	<.0001	1.79 (1.40–2.28)	<.0001
	≥75	498 (70.1%)	4.89 (3.84–6.23)	<.0001	4.66 (3.57–6.09)	<.0001
	Gender					
	Female	1641 (39.8%)	referent		referent	
	Male	1322 (48.9%)	1.44 (1.31–1.60)	<.0001	1.46 (1.31–1.63)	<.0001

(continued)

Table 2. continued

Population	Characteristic	No. (% of category admitted)	Univariable model[a]		Multivariable model[b]	
			OR (95% CI)	*P*-value	aOR (95% CI)	*P*-value
	AHRQ Elixhauser Mortality Index category[c]					
	Least disease burden	1769 (35.5%)	referent		referent	
	Middle	444 (61.7%)	2.92 (2.47–3.45)	<.0001	2.26 (1.88–2.73)	<.0001
	Most disease burden	750 (66.7%)	3.64 (3.19–4.15)	<.0001	2.39 (2.03–2.81)	<.0001

Abbreviations: AHRQ: Agency for Healthcare Research and Quality; aOR: adjusted odds ratio; CI: confidence interval for the OR; No./no.: number of patients; OR: unadjusted odds ratio; SVI: Social Vulnerability Index; y: year; %: percent of the row category that were admitted.

[a]For COVID-positive patients, unadjusted odds ratio and *P*-value are calculated from univariable generalized estimating equations modeling the probability of COVID-related hospital admission for each category as compared to the referent.

[b]For COVID-positive patients, adjusted odds ratio and *P*-value are calculated from multivariable generalized estimating equations model of characteristics associated with the probability of COVID-related hospital admission, after adjusting for all other characteristics in the model.

[c]Comorbidities reflect data that are available within ±2 days of the first positive COVID-19 test.

Table 3. Association of patient characteristics with the likelihood of in-hospital mortality among patients with a positive COVID-19 test result and a COVID-related hospital admission

Population	Characteristic	No. (% of category that died)	Univariable model[a]		Multivariable model[b]	
			OR (95% CI)	*P*-value	aOR (95% CI)	*P*-value
COVID-related hospital admission (No. = 3292)	Race					
	White	108 (8.7%)	referent		referent	
	Black	115 (5.6%)	0.63 (0.47, 0.84)	.002	0.96 (0.68, 1.34)	.80
	SVI category					
	Most socially advantaged	66 (8.2%)	referent		referent	
	Middle	67 (6.5%)	0.78 (0.54, 1.13)	.19	0.82 (0.54, 1.24)	.34
	Most socially disadvantaged	90 (6.2%)	0.75 (0.52, 1.07)	.11	0.99 (0.67, 1.45)	.95
	Pandemic time course category					
	Early	133 (11.8%)	2.92 (2.02, 4.22)	<.0001	2.95 (2.02, 4.33)	<.0001
	Middle	49 (4.0%)	0.91 (0.61, 1.35)	.63	1.40 (0.94, 2.09)	.10
	Late	41 (4.4%)	referent		referent	
	Age category (y)					
	0–49[c]	7 (0.5%)	referent		referent	
	50–64	39 (5.2%)	11.06 (4.98, 24.52)	<.0001	9.85 (4.40, 22.04)	<.0001
	65–74	54 (10.4%)	23.28 (10.54, 51.41)	<.0001	21.14 (9.49, 47.12)	<.0001
	≥75	123 (19.9%)	49.90 (23.41, 106.4)	<.0001	43.93 (20.18, 95.60)	<.0001
	Gender					
	Female	98 (5.4%)	referent		referent	
	Male	125 (8.4%)	1.61 (1.26, 2.07)	.0002	1.49 (1.15, 1.93)	.003
COVID-related hospital admission, subgroup with nonmissing comorbidity data *(No. = 2,982)*	Race					
	White	107 (9.8%)	referent		referent	
	Black	114 (6.0%)	0.59 (0.44, 0.79)	.0004	0.82 (0.57, 1.17)	.27
	SVI category					
	Most socially advantaged	66 (8.8%)	referent		referent	
	Middle	67 (7.3%)	0.82 (0.57, 1.17)	.27	0.92 (0.62, 1.37)	.70
	Most socially disadvantaged	88 (6.6%)	0.73 (0.51, 1.05)	.09	1.06 (0.72, 1.56)	.76
	Pandemic time course category:					
	Early	132 (13.4%)	3.06 (2.12, 4.41)	<.0001	3.20 (2.15, 4.77)	<.0001
	Middle	48 (4.2%)	0.87 (0.59, 1.28)	.48	1.39 (0.92, 2.09)	.11
	Late	41 (4.8%)	referent		referent	
	Age category (y)					
	0–49[c]	7 (0.5%)	referent		referent	
	50–64	39 (5.8%)	11.78 (5.31, 26.15)	<.0001	8.71 (3.83, 19.81)	<.0001
	65–74	52 (11.5%)	25.08 (11.33, 55.53)	<.0001	15.61 (6.75, 36.10)	<0.0001
	≥75	123 (24.6%)	62.62 (29.36, 133.5)	<.0001	37.22 (16.44, 84.24)	<.0001
	Gender					
	Female	98 (5.9%)	referent		referent	
	Male	123 (9.2%)	1.61 (1.26, 2.07)	.0002	1.52 (1.16, 2.01)	.003

(continued)

Table 3. continued

Population	Characteristic	No. (% of category that died)	Univariable model[a]		Multivariable model[b]	
			OR (95% CI)	*P*-value	aOR (95% CI)	*P*-value
	AHRQ Elixhauser Mortality Index category[d]					
	Least disease burden	58 (3.3%)	referent	.024	referent	.17
	Middle	27 (5.8%)	1.77 (1.08, 2.89)	<.0001	1.42 (0.86, 2.34)	<.0001
	Most disease burden	136 (17.5%)	6.11 (4.34, 8.60)		2.30 (1.58, 3.36)	

Abbreviations: AHRQ: Agency for Healthcare Research and Quality; aOR: adjusted odds ratio; CI: confidence interval for the OR; No./no.: number of patients; OR: unadjusted odds ratio; SVI: Social Vulnerability Index; y: year; %: percent of the row category that died.

[a]For COVID-positive patients who underwent a COVID-related hospitalization, unadjusted odds ratio and *P*-value are calculated from univariable generalized estimating equations modeling the probability of in-hospital mortality for each category as compared to the referent.

[b]For COVID-positive patients who underwent a COVID-related hospitalization, adjusted odds ratio and *P*-value are calculated from multivariable generalized estimating equations model of characteristics associated with the probability of in-hospital mortality, after adjusting for all other characteristics in the model.

[c]Due to zero deaths in patients less than 20 years old, the 0–19 and 20–49 category were combined for analyses.

[d]Comorbidities reflect data that are available within ±2 days of the COVID-related admission.

idea that the differential outcomes seen at the population level for Black versus White patients and those living in more versus less vulnerable environments are not the result of differential biology or response to treatment, but of differential acquisition of the disease. This finding points to potent correlates of race and SES that remained unmeasured in our analyses, such as racism, less access to personal protective equipment, or shared living spaces among generations which may contribute to a higher burden of COVID-19 in these populations. Goldman et al[22] note that Black workers, while underrepresented in frontline jobs overall, are overrepresented in frontline jobs of lower standing, which may be associated with increased risk. Underlying these differences are both structural racism and classism, which have profound implications for both individual and public health.[23]

Perpetuation of systemic barriers, such as red lining, that have limited the mobility of individuals by race, continue to concentrate barriers to health in minoritized communities. Not surprisingly, Black patients living in high SVI neighborhoods represent a particularly vulnerable group in terms of risk of hospitalization for COVID-19. These findings point out a dire need for targeted investment in better and more equitable public health infrastructure for these groups to improve outcomes in future waves of COVID-19 or for future pandemics, as well as for noncommunicable diseases. Ensuring equitable access to testing and treatment, as well as appropriate workplace and other protections, should be a priority for clinical leaders and policymakers seeking to improve equity in health outcomes.

Limitations

Missing comorbidity data was most prominent among Black patients, older adults, males, those diagnosed early in the pandemic timeframe, those who live in the most socially disadvantaged areas, and those who had less severe disease. Among Black patients, however, missing comorbidity data was not associated with age, gender, or living in the most socially disadvantaged areas. A limitation of these analyses is that we are not able to disentangle the reason for missing comorbidity data from other patient characteristics. Thus, the effect of comorbidity data as reported in our results may not be representative of all patients testing positive for COVID-19 in our data set.

Of note, too, is that our data did not include patient-level social determinants of health (SDoH) information for the patients included in this study. Because area-level SVI may not necessarily capture individual social vulnerability, there may be additional bias in the social vulnerability estimates we computed in our statistical analyses. However, the SVI index captures CT-level housing, SES, household composition, and minority status and language information that may never be captured by the EHR. Especially in the absence of patient-level SDoH data, area-level SDoH data, like the SVI, may allow for better analysis of how a person's environment impacts their health outcomes.

Data for mortality following hospital discharge were not reliably reported in the source data for this study, and as a result, we limited our analyses to the outcomes of admission and in-hospital mortality. Race was self-reported or determined by visual assessment from hospital staff, and consequently, the accuracy of racial assignment is flawed. Race is an imperfect variable; however, it is the best approximation present for racism in a racialized society. It will capture in some imperfect ways the effects of racism. Future research should aim to evaluate the risk of long-term mortality (ie, within 30 or 90 days) and other long-term outcomes by race and SES.

CONCLUSIONS

Both Black race and high SVI were independent predictors of hospitalization among patients diagnosed with COVID-19 during the first 6 months of the pandemic. Even when adjusting for underlying conditions, area-level social vulnerability, and other demographic information, Black race was still predictive of hospitalization. Public health and policy interventions should address these social factors when responding to the next pandemic.

AUTHOR CONTRIBUTIONS

Authors JML, KSM, KEJM, GH, AG, AMR, and REF contributed to study conception and design; author JML contributed to data acquisition, generation, and processing; author KSM contributed to data and statistical analysis; authors JML, KSM, KEJM, GH, and REF contributed to the writing and editing of the manuscript; and author MZ provided critical feedback and final approval of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *JAMIA Open* online.

CONFLICTS OF INTEREST

None declared.

DATA AVAILABILITY STATEMENT

Due to the sensitive and confidential nature of patient records, we are unable to share the data we used in preparing this manuscript.

REFERENCES

1. Millett GA, Jones AT, Benkeser D, *et al.* Assessing differential impacts of COVID-19 on black communities. *Ann Epidemiol* 2020; 47: 37–44.
2. Yehia BR, Winegar A, Fogel R, *et al.* Association of race with mortality among patients hospitalized with coronavirus disease 2019 (COVID-19) at 92 US hospitals. *JAMA Netw Open* 2020; 3 (8): e2018039.
3. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among Black patients and White patients with Covid-19. *N Engl J Med* 2020; 382 (26): 2534–43.
4. Munoz-Price LS, Nattinger AB, Rivera F, *et al.* Racial disparities in incidence and outcomes among patients with COVID-19. *Jama Netw Open* 2020; 3 (9): e2021892.
5. Khanijahani A, Tomassoni L. Socioeconomic and racial segregation and COVID-19: concentrated disadvantage and black concentration in association with COVID-19 deaths in the USA. *J Racial Ethn Heal Disparities* 2021; 1–9. doi:10.1007/s40615-021-00965-1.
6. Fielding-Miller RK, Sundaram ME, Brouwer K. Social determinants of COVID-19 mortality at the county level. *PLoS One* 2020; 15 (10): e0240151.
7. Andrasfay T, Goldman N. Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. *Proc Natl Acad Sci USA* 2021; 118 (5): e2014746118.
8. Harling G, Subramanian S, Barnighausen T, Kawachi I. Socioeconomic disparities in sexually transmitted infections among young adults in the United States. *Sex Transm Dis* 2013; 40 (7): 575–81.

9. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: a systematic review. *J Am Med Inform Assoc* 2020; 27 (11): 1764–73.
10. Dalsania AK, Fastiggi MJ, Kahlam A, *et al.* The relationship between social determinants of health and racial disparities in COVID-19 mortality. *J Racial Ethn Heal Disparities* 2021; 1–8. doi:10.1007/s40615-020-00952-y.
11. Roux AVD, Mair C. Neighborhoods and health. *Ann N Y Acad Sci* 2010; 1186: 125–45.
12. Gasparini A. Comorbidity: an R package for computing comorbidity scores. *J Open Sour Softw* 2018; 3 (23): 648.
13. Moore BJ, White S, Washington R, Coenen N, Elixhauser A. Identifying increased risk of readmission and in-hospital mortality using hospital administrative data. *Med Care* 2017; 55 (7): 698–705.
14. Flanagan BE, Hallisey EJ, Adams E, Lavery A. Measuring community vulnerability to natural and anthropogenic hazards: the Centers for Disease Control and Prevention's Social Vulnerability Index. *J Environ Health* 2018; 80 (10): 34–6.
15. Pebesma E. Simple features for R: standardized support for spatial vector data. *R J* 2018; 10 (1): 439–46.
16. Fox J, Monette G. Generalized collinearity diagnostics. *J Am Stat Assoc* 1992; 87 (417): 178–83.
17. U.S. Census Bureau. Race, American Community Survey 2019 1-Year Estimates Detailed Tables. https://data.census.gov. Accessed September 25, 2021.
18. Adebayo CT, Walker K, Hawkins M, *et al.* Race and Blackness: a thematic review of communication challenges confronting the black community within the U.S. Health Care System. *J Transcult Nurs* 2020; 31 (4): 397–405.
19. Lopez L, Hart LH, Katz MH. Racial and ethnic health disparities related to COVID-19. *JAMA* 2021; 325 (8): 719–20.
20. Krishnamoorthy G, Arsene C, Jena N, *et al.* Racial disparities in COVID-19 hospitalizations do not lead to disparities in outcomes. *Public Health* 2021; 190: 93–8.
21. Rentsch CT, Kidwai-Khan F, Tate JP, *et al.* Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: a nationwide cohort study. *PLoS Med* 2020; 17 (9): e1003379.
22. Goldman N, Pebley AR, Lee K, *et al.* Racial and ethnic differentials in COVID-19-related job exposures by occupational standing in the US. *PLoS One* 2021; 16 (9): e0256085.
23. Kim SJ, Bostwick W. Social vulnerability and racial inequality in COVID-19 deaths in Chicago. *Health Educ Behav* 2020; 47 (4): 509–13.