medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

COVID-19 outcomes, risk factors and associations by race: a comprehensive analysis using electronic health records data in Michigan Medicine

Tian Gu, MS[1]; Jasmine A. Mack, MPH[1]; Maxwell Salvatore, MPH[1]; Swaraaj Prabhu Sankar, MS[2,3]; Thomas S. Valley, MD, MSc[4,5]; Karandeep Singh, MD, MMSc[5,6]; Brahmajee K. Nallamothu, MD, MPH[7]; Sachin Kheterpal, MD, MBA[5,8]; Lynda Lisabeth, PhD[9]; Lars G. Fritsche, PhD[1,2,10]; Bhramar Mukherjee, PhD[1,2,9*]

Author Affiliations

[1] Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI
[2] Rogel Cancer Center, University of Michigan Medicine, Ann Arbor, MI
[3] Data Office for Clinical and Translational Research, University of Michigan, Ann Arbor, MI
[4] Division of Pulmonary and Critical Care Medicine and Department of Internal Medicine, University of Michigan, Ann Arbor, MI
[5] Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI
[6] Department of Learning Health Sciences, University of Michigan, Ann Arbor, MI
[7] Division of Cardiovascular Medicine and Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI
[8] Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, MI
[9] Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI
[10] Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI

Corresponding author:
Bhramar Mukherjee
Department of Biostatistics
University of Michigan School of Public Health
1415 Washington Heights
Ann Arbor, MI 48109
bhramar@umich.edu

Word count: 3174

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Key Points

Question: What are the sociodemographic and pre-existing health conditions associated with COVID-19 outcomes and how do they differ by race/ethnicity?

Findings: In this retrospective cohort of 5,698 patients tested for COVID-19, high population density and comorbidities such as type II diabetes/kidney disease were associated with hospitalization, in addition to older age, male sex and obesity. Adjusting for covariates, non-Hispanic Blacks were 1.66 times more likely to be hospitalized and 1.52 times more likely to be admitted to ICUs than non-Hispanic Whites.

Meaning: Targeted interventions to support vulnerable populations are needed. Racial disparities existed in COVID-19 outcomes that cannot be explained after controlling for age, sex, and socioeconomic status.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Structured Abstract

Importance: Blacks/African-Americans are overrepresented in the number of COVID-19 infections, hospitalizations and deaths. Reasons for this disparity have not been well-characterized but may be due to underlying comorbidities or sociodemographic factors.

Objective: To systematically determine patient characteristics associated with racial/ethnic disparities in COVID-19 outcomes.

Design: A retrospective cohort study with comparative control groups.

Setting: Patients tested for COVID-19 at University of Michigan Medicine from March 10, 2020 to April 22, 2020.

Participants: 5,698 tested patients and two sets of comparison groups who were not tested for COVID-19: randomly selected unmatched controls (n = 7,211) and frequency-matched controls by race, age, and sex (n = 13,351).

Main Outcomes and Measures: We identified factors associated with testing and testing positive for COVID-19, being hospitalized, requiring intensive care unit (ICU) admission, and mortality (in/out-patient during the time frame). Factors included race/ethnicity, age, smoking, alcohol consumption, healthcare utilization, and residential-level socioeconomic characteristics (SES; i.e., education, unemployment, population density, and poverty rate). Medical comorbidities were defined from the International Classification of Diseases (ICD) codes, and were aggregated into a comorbidity score.

Results: Of 5,698 patients, (median age, 47 years; 38% male; mean BMI, 30.1), the majority were non-Hispanic Whites (NHW, 59.2%) and non-Hispanic Black/African-Americans (NHAA, 17.2%). Among 1,119 diagnosed, there were 41.2% NHW and 37.4% NHAA; 44.8% hospitalized, 20.6% admitted to ICU, and 3.8% died. Adjusting for age, sex, and SES, NHAA were 1.66 times more likely to be hospitalized (95% CI, 1.09-2.52; *P*=.02), 1.52 times more likely to enter ICU (95% CI, 0.92-2.52; *P*=.10). In addition to older age, male sex and obesity, high population density neighborhood (OR, 1.27 associated with one SD change [95% CI, 1.20-1.76]; *P*=.02) was associated with hospitalization. Pre-existing kidney disease led to 2.55 times higher risk of hospitalization (95% CI, 1.62-4.02; *P*<.001) in the overall population and 11.9 times higher mortality risk in NHAA (95% CI, 2.2-64.7, *P*=.004).

Conclusions and Relevance: Pre-existing type II diabetes/kidney diseases and living in high population density areas were associated with high risk for COVID-19 susceptibility and poor prognosis. Association of risk factors with COVID-19 outcomes differed by race. NHAA patients were disproportionately affected by obesity and kidney disease.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

The COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus SARS-CoV-2, has demonstrated racial disparities in those affected in the United States (US)[1–13]. In the state of Michigan in particular, there have been 64,998 confirmed COVID-19 cases and 5,943 deaths as of June 11, 2020, which makes Michigan one of the most affected states in the US[14]. While Blacks/African-Americans represent 14% of the Michigan population[15], they account for 31% of COVID-19 cases and 40% of deaths attributed to COVID-19[14]. Similar trends are observed in New York[9] and Illinois, where there is an overrepresentation of African-Americans and Latinos in COVID-19 cases and deaths[16].

Overrepresentation of minority populations in poorer COVID-19 outcomes may be explained by a myriad of factors, such as by weathering, or early health deterioration due to cumulative impact of socioeconomic disparity [17,18], higher comorbidity burden[19], inadequate healthcare[19], and socioeconomic differences related to unemployment, food insecurity, and housing instability[17]. Several studies have reported non-White, male, older age, current smoking, and comorbid conditions as high risk factors of COVID-19 susceptibility and hospitalization[2,13,20–24]. Racial/ethnic minorities who maintain livelihood as essential workers are more likely to be exposed to the virus[16], whereas living in high density areas[1], high proportion of homelessness[25] and incarceration[26] adds to the barriers to social distancing[16].

Although studies had reported many possible reasons for the overrepresentation of minority populations in poorer health outcomes, the evidence supporting the observed disparity in COVID-19 outcomes remains limited, and more data from diverse

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

communities need to be analyzed. In addition, experiences from COVID-19 highlight the need to not only identify risk factors but also to avoid spurious conclusions of racial/ethnic differences being explained by biology, which could further perpetuate racial/ethnic stereotypes[17]. Data on holistic clinical and sociodemographic factors contributing to racial/ethnic differences in COVID-19 outcomes is limited. Some previous studies have also compared those who tested positive for COVID-19 to those who are negative, instead of population-based controls where selection bias is potentially observed[27,28].

The objective of this study is to determine sociodemographic and comorbid conditions that are associated with COVID-19 outcomes (e.g., testing positive, hospitalization, admission to ICU, and mortality), utilizing electronic health records (EHR) from the University of Michigan, which serves a large patient population in the US Midwest.

Subjects and Methods

Evaluation cohorts

COVID-19 cohort: We extracted the EHR data for patients tested for COVID-19 at the University of Michigan Medicine Health System, also known as Michigan Medicine (MM), from March 10, 2020 to April 22, 2020. Our study cohort of 5,698 patients comprises 5,500 patients (96.5%) who were tested at MM and 198 patients (3.5%) who were treated for COVID-19 in MM but tested elsewhere, of which 1,119 were COVID-19 positive. For ease of notation, we refer to them as **the tested cohort (n=5,698)** and **the positive cohort (n=1,119)**. The tested cohort is a non-random sample of the population, since the testing protocol at MM focused on prioritized testing[29] (e.g., testing

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

symptomatic patients and those at the highest risk of exposure). This cohort also contained transfer patients from other hospitals.

Control selection: To understand how selection bias factored into our sample, in addition to comparing COVID-19 positive patients with those testing negative, we created two sets of controls from the MM database. The first **unmatched control group (n=7,211)** is a similar-sized random sample of contemporaneous patients**.** The second 1:3 **frequency-matched control group (n=13,351)** is matched by race, sex and age (above or below 50). All controls were alive at the time of data extraction. Study protocols were reviewed and approved by the University of Michigan Medical School Institutional Review Board (IRB ID HUM00180294 and HUM00155849).

Description of variables: A summary data dictionary, eTable S2A, is available in Supplement with source and definition of each variable used in our analysis.

COVID-19 prognosis outcomes: Among the patients diagnosed with COVID-19, we considered various stages of progression of the disease that included hospitalization, admission to the ICU and death. Hospitalizations were defined by inpatients with a COVID-19 diagnosis where the admission date was within the time frame of the data extraction. ICU patients were defined as patients who were admitted to ICU units any time during their COVID-19 related hospitalization. Mortality data including inpatient and non-hospitalized deaths was extracted from EHR.

Classifying patients who were still in hospital and ICU: We categorized patients into non-hospitalized, hospitalized (includes ICU stays), and hospitalized with ICU stay based on the admission and discharge data. Several patients were still admitted in the hospital (non-ICU, n=53) or were still in an ICU (n=113) at the time of the data

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

extraction. We performed a sensitivity analysis by excluding these patients whose final prognostic outcome is unclear from the analysis (eTable S4 in Supplement).

Generation of comorbidities from electronic health records

We constructed the comorbid conditions using available International Classification of Diseases (ICD; ninth and tenth editions) code for 23,769 individuals (n_{tested}: 5,225, $n_{unmatched}$: 6,811, $n_{matched}$: 11,733) from EHR. Longitudinal time-stamped diagnoses were recoded to indicator variables for whether a patient ever had a given diagnosis code recorded by MM. To differentiate *pre-existing* conditions from diagnoses related to COVID-19 testing/treatment, we applied a 14-day-prior restriction on the tested cohort by removing diagnoses that first appeared within the 14 days before the first test or diagnosis date, whichever was earlier (4,622 of the 5,225 tested individuals had diagnoses data after the 14-day-prior restriction). We focused on seven binary disease indicators that have been specifically mentioned in relation to COVID-19 outcomes: respiratory, circulatory, any cancer, type II diabetes, kidney, liver, and autoimmune diseases (ICD codes in eTable S2A in Supplement). We calculated a comorbidity score as the sum of these seven that ranges from 0-7. For exploratory analysis, we defined prior medication use as at least one appearance of a given class of medication in the patient's EHR.

Defining race/ethnicity groups, SES and other adjustment covariates: Variables such as self-reported sex, race/ethnicity, smoking status, alcohol consumption, body mass index (BMI), and age were extracted from the EHR. We classified a patient to be seeking primary care in MM if they have had an encounter in any of the primary care locations in MM since January 1, 2018. Measures of socioeconomic characteristics are

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

defined by US census tract (based on residential address available in each patient's EHR) for the year 2010. The boundaries for the census tracts were normalized to 2010 tract boundaries using the Longitudinal Tract Data Base [30]. We chose three SES indicators included in the National Neighborhood Data Archive (NaNDA)[31]: percentage of population with below high school (<HS) education, percentage unemployed and percentage with annual income below federal poverty level (FPL). We also used population density per square mile as a potential predictor[31].

Statistical analysis

Since all outcomes were binary, we performed logistic regression to assess the risk factors of COVID-19 outcomes, by reporting Firth bias-corrected estimate of the odds ratio to address potential separation issues, as well as 95% Wald-type confidence interval and *P-value*. Continuous predictors were standardized before modeling. Three sets of adjustment covariates were used to check the robustness of inference to the choice of potential confounders: (i) age, race/ethnicity, sex, (ii) adjustment (i)+SES, and (iii) adjustment (ii)+comorbidity score. We present the results of adjustment (ii) in the main text, since adjustment (iii) led to loss of sample size as many of the patients did not have EHR data to construct the pre-existing conditions. Sensitivity analysis after removing those still in ICU or hospital at the time of data extraction were also performed (eTable S4 in Supplement).

Three different types of association models were created (eTable S2B in Supplement):

(a) Predictors of being tested: comparing the tested cohort with those who were not tested for COVID-19 (unmatched controls).

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(b) Predictors of COVID-19 susceptibility: comparing the positive cohort with those who were not diagnosed with COVID-19 (unmatched controls)

(c) Predictors of three COVID-19 prognostic outcomes: among the positive cohort, (i) comparing those who were hospitalized with those who were not, (ii) those who were admitted to ICU with those who were not, and (iii) those who died with those who did not. (this analysis does not use untested controls)

After analyzing the full cohort including all races, we carried out association analysis stratified by race/ethnicity (NHW or NHAA, given limited sample size for other racial/ethnic groups). Frequency-matched, instead of unmatched controls, were used to assess the COVID-19 susceptibility as the proportion of race/ethnicity in unmatched controls are not comparable to the stratified study population.

Analyses were performed in R version 3.6.2 (R Foundation for Statistical Computing). Statistical significance was defined using a 2-sided significance level of $\alpha = .05$. All unadjusted models and results using other adjustments can be found in eTable S3A for all cohort, S3B and S3C for race/ethnicity-stratified in Supplement.

Results

Descriptive Statistics

A total of 5,698 patients were tested for COVID-19 (median age, 47 years [IQR, 31]; 38% male [2030/5336]; mean BMI 30.1 [SD, 8.1]), among whom 1,119 (19.6%) tested positive (**Table 1**). Over half of the tested (3,026, 53.1%) had primary care at MM. A majority of those tested were either NHW (3,374, 59.2%) or NHAA (981, 17.2%). Among the 1,119 positive patients, 44.8% (501) were hospitalized, 20.6% (231) were in an ICU, and 3.8% (43) died. As the disease progressed among patients testing positive

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

(from non-hospitalized to hospitalized, ICU, and deceased), the proportion of older age (65 and above), male sex, and ever-smoker, consistently increased (**Table 1**). BMI showed an increasing trend from non-hospitalized to hospitalized and ICU but not in those who died (pairwise comparison results in eTable S1B in Supplement).

The descriptive trends of all three residential SES variables indicate lower SES is associated with poorer COVID outcomes (**Table 1**). Both the tested cohort and the positive cohort had a higher enrichment of diseases. In particular, type II diabetes and kidney diseases were more common in groups with more severe outcomes. The tested population were also more frequent users of medication (e.g. NSAIDs, statins, and anticoagulants) compared with the unmatched controls (eTable S1A in Supplement).

Descriptive statistics stratified by NHW and NHAA (eTable S1A and S1B in Supplement) show a stark contrast in COVID-19 outcomes across these two groups that we capture in **Figure 1**. The test-positive rate was significantly higher in NHAA compared with NHW (42.6% vs. 13.7%, $P<.001$), hospital admission (52.2% vs. 39%, $P<.001$), ICU (27% vs. 14.8%, $P<.001$), and death (5.3% vs. 3.0%, $P=.12$). Moreover, racial disparities also exist across demographics such as BMI, lower SES, and comorbidity conditions, which were consistently more prevalent in NHAA compared with NHW across subpopulations defined by COVID-19 outcomes (pairwise $P<.001$, eTable S1C in Supplement).

Association Analysis of COVID-19 Outcomes Using Multivariable Logistic Regression

Factors Associated with Getting Tested

Due to dependence on test availability, the testing guidelines varied during the time of

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the study. Comparing with the random controls from MM, overall, we noticed that male sex, current smoking, percentage of population with <HS education and percentage below FPL were inversely associated with the chance of getting tested whereas NHAA race/ethnicity, past smoking, age, BMI, and all comorbidities were positively associated with getting tested (eTable S5 in Supplement).

Factors Associated with being Positive or with COVID19 Susceptibility

Full cohort: We identified the following factors that differed between testing positive for COVID-19 and the randomly selected untested population from MM: NHAA vs NHW odds ratio (OR) was 7.11 (95% CI, 5.71-8.87; *P*<.001); older age (OR, 1.66 [95% CI, 1.52-1.82]; *P*<.001), higher BMI (OR, 1.35 [95% CI, 1.23-1.47]; *P*<.001), and alcohol consumption (ever vs. never; OR, 1.42 [95% CI, 1.17-1.73]; *P*<.001) (**Table 2**). Current smoking (self-reported in the latest EHR record) was observed as being protective against COVID-19 outcomes (OR, 0.34 [95% CI, 0.22-0.51]; *P*<.001), with a positive but insignificant association with past smoking (possibly due to quitting smoking for health reasons). In addition, population density was associated with testing positive (OR, 1.46 [95% CI, 1.33, 1.60]; *P*<.001). Percentage with income below FPL showed consistent association in protective direction, perhaps due to high correlation and overlap with other SES variables. **Table 2** indicates that pre-existing comorbidities were associated with testing positive. Conversely, a comparison between those **who testing positive versus those testing negative** showed *protective effects of existing comorbidity conditions*, a reflection of the bias due to targeted testing criteria.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Race/ethnicity-stratified: Former smokers were strongly associated with testing positive in both NHW (OR, 1.54 [95% CI, 1.22-1.95]; *P*<.001) and NHAA (OR, 1.74 [95% CI, 1.25-2.42]; *P*=.001). Overweight/obesity and having any cancer showed stronger effect in the race-stratified analysis, especially obesity in NHAA (OR, 6.15 [95% CI, 3.50-10.82]; *P*<.001). The stronger association was most likely due to stratification which addresses potential presence of interaction by race. All comorbidities and SES variables show similar directionalities as the full cohort.

Factors Associated with Prognosis among COVID-19 Diagnosed Patients

Full Cohort: Among the COVID-19 positive cohort, NHAA were 1.66 (95% CI, 1.09-2.52; *P*=.02) times more likely to be hospitalized, 1.52 (95% CI, 0.92-2.52; *P*=.1) times more likely of ICU admission, and 1.17 (95% CI, 0.4-3.45; *P*=.77) times more likely of death than NHW (**Table 3**), shown in Figure 1. Higher population density was associated with higher chance of hospitalization (OR, 1.27 [95% CI, 1.20-1.76]; *P*=.02), whereas older age, male sex and obesity consistently showed strong association with poor prognosis (**Table 3**). Type II diabetes (OR, 1.62 [95% CI 1.09-2.41]; *P*=.02) and kidney disease (OR, 2.55 [95% CI, 1.62-4.02]; *P*<.001) stood out as the strongest risk factors of hospitalization amongst the seven comorbidities. In spite of limited power due to sample size when comparing deceased versus alive, composite comorbidity score (OR, 1.72 [95%CI, 1.08-2.75]; *P*=.02) and kidney disease (OR, 2.82 [95%CI, 1.09-7.28]; *P*=.03) appear to be the strongest risk factors, in addition to age and male sex.

Race/ethnicity-stratified: In NHW, patients with obesity and type II diabetes were 2.89 (95% CI, 1.31-6.35; *P*=.009) and 2.53 (95% CI, 1.41-4.57; *P*=.002) times more likely to be hospitalized compared to those without, respectively. In NHAA, males were at higher

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

risk of being hospitalized (OR, 2.96 [95% CI, 1.68-5.2]; *P*<.001) than females, as well as being admitted to ICU (OR, 2.65 [95% CI, 1.37-5.11]; *P*=.004). In general, in addition to older age, patients with pre-existing kidney diseases had worse prognosis, especially for mortality in NHAA (OR, 11.9 [95% CI, 2.2-43.5]; *P*=.004).

Discussion

This is one of the first studies to compare patient neighborhood SES, sociodemographic factors and health conditions from EHR data to uncover potential risk factors for observed racial disparities in COVID-19 susceptibility and prognosis. We identified racial and socioeconomic disparity in all COVID-19 outcomes (statistically significant for testing positive, hospitalization and suggestive for ICU admission and mortality due to small sample size). Pre-existing health conditions, and higher comorbidity burden were associated with poor prognosis.

This work contributes to a new area of COVID-19 research that incorporates both EHR data and SES data to rigorously examine racial/ethnic differences in disease susceptibility and prognosis. Other novel contributions include: (i) we compare the COVID-19 tested positive population with a random subset of the MM population to avoid the biased sample of who gets tested for COVID-19. (ii) We consider outcomes related to both susceptibility and prognosis. (iii) While most of the early studies present unadjusted analysis of various risk factors, we present a comprehensive analytic framework that attempts to adjust for an expanded set of potential confounders with suitably chosen comparison groups.

In general, our findings are consistent with existing studies. The observed high prevalence of kidney disease in COVID-19 mortality in NHAA can be possibly explained

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

by the racial/ethnic differences in kidney disease[32]; males were at a higher risk of hospitalization and death, especially among those 50 years and older[33]; health conditions such as obesity, cancer, diabetes mellitus and renal conditions were prevalent in severe prognosis of COVID-19[22,34–36]. Notably, our findings largely agree with another recently published work looking at racial/ethnic differences in COVID-19 outcomes in an integrated-delivery health system in Louisiana[13], which did not identify type II diabetes and kidney disease as risk factors of hospitalization as we do. Similar directional results but different strength of association with SES variables is likely because we used a continuous metric as opposed to the categorical measures used in Price-Haywood et al.[13]

Due to the prioritized testing scenario, there may be many asymptomatic or mild-symptomatic patients in the control population. Therefore, the comparison results between the positive and the unmatched control reveals the fact that in general, people with low SES and pre-existing health conditions had higher risk of developing severe disease outcomes after being infected with COVID-19. In contrast, a naïve comparison between the positive and the negative in the tested population leads to counterintuitive findings (protective effect of smoking, medication use, and worse comorbidity condition) contradicting findings in other COVID-19 studies[27,28]. This amplifies the need for choosing an appropriate control group.

This study has several limitations. First, we may not have captured all hospitalized patients given that only about half (48.3%) of the positive cohort had primary care at MM. It is possible that some of the “not hospitalized” patients actually were hospitalized elsewhere. Second, we did not consider the transfer patients from other hospitals as a

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

special sub-group, who often had more severe outcomes. Third, early in COVID-19 course at MM, all COVID-19 patients were placed in regional infectious containment unit (RICU), some of which did not require ICU-level care. We suggest future study to define ICU patients using requirement of mechanical ventilators. Fourth, one may argue that the untested MM controls are intrinsically different than the tested cohort and do not serve as proper controls and may impact the estimation of the ORs observed in the susceptibility models. In the future, when testing is abundantly available, we may be able to restrict the patient population only to the catchment of MM or those who seek primary care at MM and exclude transfer patients. However, this is only a limitation in susceptibility models, and, the prognosis-outcome models only focused on the tested positive cohort and did not use the untested controls and thus, are not subject to the same selection issues. Lastly, the study results should be interpreted cautiously for generalizability given the demographics varies by region and by countries.

Conclusion

The COVID-19 pandemic has reached every corner of the globe affecting millions and tens of millions of their loved ones. Our findings highlight that COVID-19 disproportionately affects those who are vulnerable: the elderly, those with pre-existing conditions, and those in population dense communities. Specifically, NHAAs have a greater risk burden – more likely to test positive, more likely to be hospitalized and more likely to be admitted to the ICU. Our results support targeted screening for the elderly and those with type II diabetes and kidney disease. Moreover, we call for increased investments in testing and prevention efforts in low SES, densely populated communities and racially diverse communities. It is these same communities that

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

employ a greater proportion of essential workers who have been rightly called heroes but have still not been treated and protected as such. Genetics could also play a role in determining susceptibility and prognosis. Thus, it is important to expand this study to consider the multifactorial interplay of genetics and environment in the future.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Vahidy FS, Nicolas JC, Meeks JR, et al. Racial and Ethnic Disparities in SARS-CoV-2 Pandemic: Analysis of a COVID-19 Observational Registry for a Diverse U.S. Metropolitan Population. *medRxiv*. Published online April 28, 2020:2020.04.24.20073148. doi:10.1101/2020.04.24.20073148

2. Patel AP, Paranjpe MD, Kathiresan NP, Rivas MA, Khera AV. Race, Socioeconomic Deprivation, and Hospitalization for COVID-19 in English participants of a National Biobank. *medRxiv*. Published online May 2, 2020:2020.04.27.20082107. doi:10.1101/2020.04.27.20082107

3. Henning-Smith C, Tuttle M, Kozhimannil KB. Unequal Distribution of COVID-19 Risk among Rural Residents by Race and Ethnicity. *J Rural Health*. n/a(n/a). doi:10.1111/jrh.12463

4. Li AY, Hannah TC, Durbin J, et al. Multivariate Analysis of Factors Affecting COVID-19 Case and Death Rate in U.S. Counties: The Significant Effects of Black Race and Temperature. *medRxiv*. Published online April 24, 2020:2020.04.17.20069708. doi:10.1101/2020.04.17.20069708

5. Gross CP, Essien UR, Pasha S, Gross JR, Wang S, Nunez-Smith M. Racial and Ethnic Disparities in Population Level Covid-19 Mortality. *medRxiv*. Published online May 11, 2020:2020.05.07.20094250. doi:10.1101/2020.05.07.20094250

6. Abedi V, Olulana O, Avula V, et al. Racial, Economic and Health Inequality and COVID-19 Infection in the United States. *medRxiv*. Published online May 1, 2020:2020.04.26.20079756. doi:10.1101/2020.04.26.20079756

7. Shah M, Sachdeva M, Dodiuk-Gad RP. COVID-19 and Racial Disparities. *J Am Acad Dermatol*. Published online April 17, 2020. doi:10.1016/j.jaad.2020.04.046

8. Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California. *Health Aff (Millwood)*. Published online May 21, 2020:10.1377/hlthaff.2020.00598. doi:10.1377/hlthaff.2020.00598

9. Marcello RK, Dolle J, Grami S, et al. Characteristics and Outcomes of COVID-19 Patients in New York City's Public Hospital System. *medRxiv*. Published online June 1, 2020:2020.05.29.20086645. doi:10.1101/2020.05.29.20086645

10. Ahmed SM, Shah RU, Bale M, et al. Comprehensive Testing Highlights Racial, Ethnic, and Age Disparities in the COVID-19 Outbreak. *medRxiv*. Published online May 9, 2020:2020.05.05.20092031. doi:10.1101/2020.05.05.20092031

11. Millett G, Jones AT, Benkeser D, et al. Assessing Differential Impacts of COVID-19 on Black Communities. *medRxiv*. Published online May 8, 2020:2020.05.04.20090274. doi:10.1101/2020.05.04.20090274

12. Rentsch CT, Kidwai-Khan F, Tate JP, et al. Covid-19 by Race and Ethnicity: A National Cohort Study of 6 Million United States Veterans. *medRxiv*. Published online May 18, 2020:2020.05.12.20099135. doi:10.1101/2020.05.12.20099135

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

13. 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;0(0):null. doi:10.1056/NEJMsa2011686

14. Coronavirus - Michigan Data. Accessed June 10, 2020. https://www.michigan.gov/coronavirus/0,9753,7-406-98163_98173---,00.html

15. U.S. Census Bureau QuickFacts: Michigan. Accessed May 11, 2020. https://www.census.gov/quickfacts/MI

16. Bibbins-Domingo K. This Time Must Be Different: Disparities During the COVID-19 Pandemic. *Ann Intern Med*. Published online April 28, 2020. doi:10.7326/M20-2247

17. Chowkwanyun M, Reed AL. Racial Health Disparities and Covid-19 — Caution and Context. *N Engl J Med*. 2020;0(0):null. doi:10.1056/NEJMp2012910

18. Geronimus AT, Hicken M, Keene D, Bound J. “Weathering” and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *Am J Public Health*. 2006;96(5):826-833. doi:10.2105/AJPH.2004.060749

19. Milam AJ, Furr-Holden D, Edwards-Johnson J, et al. Are Clinicians Contributing to Excess African American COVID-19 Deaths? Unbeknownst to Them, They May Be. *Health Equity*. 2020;4(1):139-141. doi:10.1089/heq.2020.0015

20. Jackson SE, Brown J, Shahab L, Steptoe A, Fancourt D. COVID-19, smoking, and inequalities: a cross-sectional survey of adults in the UK. *medRxiv*. Published online May 5, 2020:2020.04.30.20086074. doi:10.1101/2020.04.30.20086074

21. Niedzwiedz CL, O’Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *medRxiv*. Published online April 30, 2020:2020.04.22.20075663. doi:10.1101/2020.04.22.20075663

22. Ebinger JE, Achamallah N, Ji H, et al. Pre-Existing Characteristics Associated with Covid-19 Illness Severity. *medRxiv*. Published online May 5, 2020:2020.04.29.20084533. doi:10.1101/2020.04.29.20084533

23. Collaborative TO, Williamson E, Walker AJ, et al. OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *medRxiv*. Published online May 7, 2020:2020.05.06.20092999. doi:10.1101/2020.05.06.20092999

24. Alqahtani JS, Oyelade T, Aldhahir AM, et al. Prevalence, Severity and Mortality associated with COPD and Smoking in patients with COVID-19: A Rapid Systematic Review and Meta-Analysis. *PLOS ONE*. 2020;15(5):e0233147. doi:10.1371/journal.pone.0233147

25. Fusaro VA, Levy HG, Shaefer HL. Racial and Ethnic Disparities in the Lifetime Prevalence of Homelessness in the United States. *Demography*. 2018;55(6):2119-2128. doi:10.1007/s13524-018-0717-0

26. Wildeman C, Wang EA. Mass incarceration, public health, and widening inequality in the USA. *The Lancet*. 2017;389(10077):1464-1474. doi:10.1016/S0140-6736(17)30259-3

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

27. de Lusignan S, Dorward J, Correa A, et al. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. *Lancet Infect Dis*. Published online May 15, 2020. doi:10.1016/S1473-3099(20)30371-6

28. Miyara M, Tubach F, Pourcher V, et al. Low rate of daily active tobacco smoking in patients with symptomatic COVID-19. *Qeios*. Published online May 9, 2020. doi:10.32388/WPP19W.4

29. Indications for COVID-19 diagnostic testing for adult patients in all clinical settings. Accessed June 9, 2020. http://www.med.umich.edu/asp/pdf/adult_guidelines/COVID-19-testing.pdf

30. Logan JR, Xu Z, Stults B. Interpolating U.S. Decennial Census Tract Data from as Early as 1970 to 2010: A Longtitudinal Tract Database. *Prof Geogr J Assoc Am Geogr*. 2014;66(3):412-420. doi:10.1080/00330124.2014.905156

31. NaNDA | Social Environment and Health Program. Accessed June 10, 2020. https://seh.isr.umich.edu/signature-projects/nanda/

32. Nicholas SB, Kalantar-Zadeh K, Norris KC. Racial Disparities in Kidney Disease Outcomes. *Semin Nephrol*. 2013;33(5):409-415. doi:10.1016/j.semnephrol.2013.07.002

33. Singh S, Chowdhry M, Chatterjee A, Khan A. Gender-Based Disparities in COVID-19: Clinical Characteristics and Propensity-matched Analysis of Outcomes. *medRxiv*. Published online April 29, 2020:2020.04.24.20079046. doi:10.1101/2020.04.24.20079046

34. Garg S. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1–30, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69. doi:10.15585/mmwr.mm6915e3

35. Ssentongo P, Ssentongo AE, Heilbrunn ES, Chinchilli VM. The association of cardiovascular disease and other pre-existing comorbidities with COVID-19 mortality: A systematic review and meta-analysis. *medRxiv*. Published online May 14, 2020:2020.05.10.20097253. doi:10.1101/2020.05.10.20097253

36. Khan M, Khan MN, Mustagir MG, Rana J, Islam MS, Kabir MI. Effects of pre-existing morbidities on occurrence of death among COVID-19 disease patients: A systematic review and meta-analysis. *medRxiv*. Published online May 13, 2020:2020.05.08.20095968. doi:10.1101/2020.05.08.20095968

Table 1. Descriptive Characteristics of the COVID-19 Tested/Diagnosed cohort

	COVID-19 Tested					
			COVID-19 Positive			
	Overall (n=5698)	Negative (n=4579)	Overall (n=1119)	Hospitalized (n=501)	ICU (n=231)	Deceased (n=43)
Variable	**No./No. (%)**					
Age (in years)						
mean (SD); n	47.3 (20.9); 5698	45.9 (21.3); 4579	52.8 (17.9); 1119	60.9 (16.3); 501	59.2 (15.8); 231	68.3 (12.8); 43
median (IQR)	47 (31)	45 (31)	53 (28)	63 (22)	62 (19)	70 (17)
Age Range, y						
[0,18)	374 (6.56)	365 (7.97)	9 (0.8)	5 (0.81)	4 (0.8)	3 (1.3)
[18,35)	1313 (23.04)	1109 (24.22)	204 (18.23)	171 (27.67)	33 (6.59)	19 (8.23)
[35,50)	1358 (23.83)	1101 (24.04)	257 (22.97)	177 (28.64)	80 (15.97)	33 (14.29)
[50,65)	1377 (24.17)	1040 (22.71)	337 (30.12)	172 (27.83)	165 (32.93)	82 (35.5)
[65,80)	923 (16.2)	699 (15.27)	224 (20.02)	77 (12.46)	147 (29.34)	75 (32.47)
[80,100)	353 (6.2)	265 (5.79)	88 (7.86)	16 (2.59)	72 (14.37)	19 (8.23)
Male Sex	2030/5336 (38)	1524/4246 (35.9)	506/1090 (46.4)	280/496 (56.5)	138/227 (60.8)	28/43 (65.1)
Primary Care in MM	3026/5698 (53.1)	2486/4579 (54.3)	540/1119 (48.3)	127/501 (25.3)	45/231 (19.5)	7/43 (16.3)
BMI, mean (SD); No.	30.1 (8.1); 4608	29.4 (7.5); 3609	32.4 (9.5); 999	33.6 (11.1); 467	35.3 (13.3); 217	31.5 (8.0); 42
BMI Range, y						
<18.5	76/4608 (1.3)	68/3609 (1.5)	8/999 (0.7)	4/467 (0.8)	1/217 (0.4)	1/42 (2.3)
[18.5, 25)	196/4608 (21)	1033/3609 (22.6)	163/999 (14.6)	55/467 (11)	23/217 (10)	6/42 (14)
[25, 30)	1394/4608 (24.5)	1091/3609 (23.8)	303/999 (46.9)	150/467 (29.9)	60/217 (26)	15/42 (34.9)
>=30	1942/4608 (34.1)	1417/3609 (30.9)	525/999 (10.7)	258/467 (51.5)	133/217 (57.6)	20/42 (46.5)
Ever-Smoker	1806/4699 (38.4)	1502/3815 (39.4)	304/884 (34.4)	148/373 (39.7)	64/134 (47.8)	14/20 (70)
Smoking Status, No. (%)						
Never-Smoker	2893/4699 (61.6)	2313/3815 (60.6)	580/884 (65.6)	225/373 (60.3)	70/134 (52.2)	6/20 (30)
Past-Smoker	1406/4699 (29.9)	1142/3815 (29.9)	264/884 (29.9)	138/373 (37)	61/134 (45.5)	14/20 (70)
Current-Smoker	400/4699 (8.5)	360/3815 (9.4)	40/884 (4.52)	10/373 (2.7)	3/134 (2.2)	0/20 (0)
Alcohol Consumption	2573/3879 (66.3)	2134/3198 (66.7)	439/681 (64.5)	151/243 (62.1)	61/91 (67)	13/16 (81.2)
Race Ethnicity, No. (%)						
NHW	3374 (59.2)	2913 (63.6)	461 (41.2)	180 (35.9)	68 (29.4)	14 (32.6)
NHAA	981 (17.2)	563 (12.3)	418 (37.4)	218 (43.5)	113 (48.9)	22 (51.2)
Other Ethnicity	486 (8.5)	378 (8.3)	108 (9.7)	45 (9.0)	14 (6.1)	1 (2.3)
Unknown Ethnicity	857 (15.0)	725 (15.8)	132 (11.8)	58 (11.6)	36 (15.6)	6 (14.0)
SES, mean (SD); No.						
< HS Education	0.08 (0.07); 4593	0.08 (0.06); 3807	0.09 (0.08); 786	0.10 (0.09); 265	0.12 (0.10); 104	0.14 (0.11); 18
Unemployed	0.07 (0.04); 4593	0.07 (0.03); 3807	0.08 (0.05); 786	0.08 (0.05); 265	0.09 (0.05); 104	0.09 (0.03); 18

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Annual Income < FPL	0.13 (0.12); 4592	0.12 (0.114); 3806	0.14 (0.13); 786	0.16 (0.14); 265	0.17 (0.15); 104	0.22 (0.17); 18
Persons per Mile^2	2650 (2340); 4593	2550 (2330); 3807	3150 (2310); 786	3600 (2580); 265	3780 (2500); 104	4490 (3220); 18
Comorbidity Score, mean (SD); No.	2.53 (1.59); 4615	2.53 (1.58); 3838	2.53 (1.64); 777	3.02 (1.77); 272	2.78 (1.73); 107	3.89 (1.75); 18

Abbreviations: MM, Michigan Medicine; BMI, body mass index; NHW, non-Hispanic Whites; NHAA, non-Hispanic African Americans; SES, residential-level socioeconomic characteristics; HS, high school; FPL, federal poverty level.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. Multivariable Logistic Regression Comparing COVID-19 Positive Patients with Untested Controls and Tested Negatives.

	Positive (1) versus Control (0)			Positive (1) vs. Negative (0)
	Unmatched Control	Matched Control		
	All (n_0=5942, n_1=766)	NHW (n_0=8521, n_1=371)	NHAA (n_0=1029, n_1=271)	All (n_0=3533, n_1=766)
	OR (95% CI)			
Age* (y)	**1.66 (1.52, 1.82)**	**1.43 (1.28, 1.6)**	**1.75 (1.51, 2.02)**	**1.37 (1.26, 1.49)**
<50	REF	REF	REF	REF
[50, 65)	**2.20 (1.80, 2.68)**	**2.13 (1.65, 2.73)**	**2.48 (1.79, 3.42)**	**1.80 (1.48, 2.19)**
[65, 80)	**1.88 (1.50, 2.37)**	**1.55 (1.16, 1.07)**	**2.19 (1.49, 3.23)**	**1.65 (1.31, 2.08)**
>=80	**2.21 (1.59, 3.06)**	**1.50 (1.02, 2.22)**	**2.08 (1.10, 3.96)**	**2.06 (1.48, 2.86)**
Male Sex	0.91 (0.77, 1.08)	**1.45 (1.18, 1.79)**	1.27 (0.96, 1.67)	**1.5 (1.27, 1.77)**
BMI*	**1.35 (1.23, 1.47)**	**1.34 (1.21, 1.49)**	**1.33 (1.14, 1.54)**	**1.24 (1.14, 1.35)**
<18.5	0.67 (0.28, 1.57)	0.77 (0.26, 2.30)	1.47 (0.33, 6.49)	0.705 (0.30,1.66)
[18.5, 25)	REF	REF	REF	REF
[25, 30)	**1.75 (1.36, 2.25)**	**1.80 (1.34, 2.42)**	**3.16 (1.71, 5.84)**	**1.41 (1.10, 1.81)**
>=30	**2.04 (1.60, 2.60)**	**2.20 (1.64, 2.94)**	**6.15 (3.50, 10.82)**	**1.63 (1.29, 2.06)**
Ever-Smoker	0.87 (0.71, 1.05)	1.19 (0.95, 1.5)	1.24 (0.91, 1.68)	**0.69 (0.57, 0.83)**
Never-Smoker	REF	REF	REF	REF
Past-Smoker	1.12 (0.91, 1.38)	**1.54 (1.22, 1.95)**	**1.74 (1.25, 2.42)**	**0.81 (0.66, 0.98)**
Current-Smoker	**0.34 (0.22, 0.51)**	**0.36 (0.2, 0.65)**	**0.45 (0.24, 0.84)**	**0.34 (0.22, 0.51)**
Alcohol Consumption	**1.42 (1.16, 1.73)**	1.65 (1.28, 2.12)	1.96 (1.41, 2.71)	0.98 (0.81, 1.19)
Race/Ethnicity				
NHW	REF	NA	NA	REF
NHAA	**7.11 (5.71, 8.87)**	NA	NA	**3.53 (2.88, 4.33)**
Other	**1.69 (1.28, 2.22)**	NA	NA	**1.69 (1.29, 2.22)**
Unknown	**0.16 (0.12, 0.22)**	NA	NA	0.86 (0.61, 1.21)
SES				
< HS Education	0.93 (0.83, 1.05)	0.85 (0.71, 1.01)	1.11 (0.94, 1.31)	1.08 (0.96, 1.21)
Unemployed	**1.16 (1.04, 1.31)**	1.01 (0.86, 1.19)	1.16 (1, 1.34)	**1.2 (1.07, 1.33)**
Annual Income < FPL	**0.67 (0.59, 0.76)**	**0.75 (0.64, 0.89)**	**0.64 (0.53, 0.77)**	**0.75 (0.65, 0.85)**
Persons per Mile^2	**1.46 (1.33, 1.60)**	**1.48 (1.31, 1.66)**	1.17 (1, 1.37)	**1.28 (1.16, 1.41)**
Comorbidity Score*	**1.87 (1.71, 2.05)**	**2.09 (1.91, 2.29)**	**2.43 (2.08, 2.82)**	**0.81 (0.74, 0.88)**
Respiratory	**3.28 (2.71, 3.97)**	**4.00 (3.12, 5.14)**	**3.46 (2.51, 4.77)**	**0.69 (0.57, 0.85)**
Circulatory	**2.46 (2.02, 2.99)**	**3.15 (2.49, 3.98)**	**4.86 (3.47, 6.79)**	**0.62 (0.5, 0.76)**
Any Cancer	1.18 (0.95, 1.45)	**1.78 (1.41, 2.25)**	**2.65 (1.88, 3.72)**	**0.75 (0.61, 0.91)**
Type II Diabetes	**1.92 (1.53, 2.41)**	**3.07 (2.36, 4)**	**3.05 (2.2, 4.22)**	0.97 (0.79, 1.2)
Kidney	**2.58 (1.98, 3.37)**	**4.28 (3.19, 5.75)**	**3.08 (2.13, 4.46)**	**0.55 (0.44, 0.69)**
Liver	**3.02 (2.18, 4.17)**	**4.56 (3.18, 6.53)**	**5.78 (3.07, 10.9)**	0.76 (0.58, 1.01)
Autoimmune	**2.55 (2.02, 3.21)**	**3.84 (2.96, 4.96)**	**3.42 (2.36, 4.95)**	0.95 (0.78, 1.17)

Abbreviations: n_1, sample size for COVID positive patients; n_0, sample size for comparative group; OR, odds ratio; NHW, non-Hispanic Whites; NHAA, non-Hispanic African Americans; BMI, body mass index; NA, not applicable; REF, reference group; SES, residential-level socioeconomic characteristics; HS, high school; FPL, federal poverty level.

*Odds ratios associated with one standard deviation (SD) change, with SD defined as the SD of the sample under consideration in each model.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Multivariable Logistic Regression Odds Ratio of Prognostic Outcomes Among Diagnosed Patients.

Hospitalized (1) vs. Non-Hospitalized (0)	**All (n_0=501, n_1=265)**	**NHW (n_0=252, n_1=119)**	**NHAA (n_0=160, n_1=111)**
	OR (95% CI)		
Age* (y)	**3.25 (2.64, 4.01)**	**3.38 (2.5, 4.56)**	**2.76 (1.99, 3.84)**
<50	REF	REF	REF
[50, 65)	**3.61 (2.39, 5.46)**	**3.84 (2.03, 7.26)**	**3.39 (1.80, 6.38)**
[65, 80)	**7.65 (4.76, 12.30)**	**11.16 (5.58, 22.34)**	**4.35 (2.05, 9.23)**
>=80	**28.57 (13.52, 60.38)**	**18.93 (7.56, 47.41)**	**123.36 (6.32, 2406.84)**
Male sex	**2.32 (1.64, 3.29)**	1.51 (0.92, 2.49)	**2.96 (1.68, 5.2)**
BMI*	**1.45 (1.2, 1.76)**	**1.48 (1.13, 1.93)**	1.26 (0.94, 1.7)
<18.5	1.61 (0.18, 14.14)	0.58 (0.03, 9.77)	17.84 (0.33, 971.58)
[18.5, 25)	REF	REF	REF
[25, 30)	**2.26 (1.20, 4.25)**	1.81 (0.81, 4.01)	7.16 (0.88, 57.92)
>=30	**3.03 (1.63, 5.64)**	**2.89 (1.31, 6.35)**	6.85 (0.90, 51.99)
Ever-smoker	1.14 (0.78, 1.66)	1.07 (0.63, 1.82)	1.12 (0.61, 2.05)
Never-Smoker	REF	REF	REF
Past-Smoker	1.21 (0.82, 1.79)	1.09 (0.63, 1.89)	1.28 (0.679, 2.43)
Current-Smoker	0.72 (0.26, 2.03)	0.89 (0.18, 4.39)	0.48 (0.10, 2.23)
Alcohol consumption	0.82 (0.54, 1.24)	1.28 (0.68, 2.4)	0.59 (0.31, 1.11)
Race/Ethnicity			
NHW	REF	NA	NA
NHAA	**1.66 (1.09, 2.52)**	NA	NA
Other Ethnicity	1.31 (0.73, 2.37)	NA	NA
Unknown Ethnicity	**0.33 (0.14, 0.77)**	NA	NA
SES			
%< HS Education	1.10 (0.83, 1.45)	0.90 (0.63, 1.29)	1.07 (0.66, 1.71)
%Unemployed	1.08 (0.84, 1.40)	1.17 (0.87, 1.58)	0.995 (0.64, 1.54)
%Annual Income < FPL	1.10 (0.85, 1.42)	1.17 (0.87, 1.57)	1.19 (0.74, 1.91)
Persons per Mile^2	**1.27 (1.04, 1.54)**	1.19 (0.89, 1.59)	1.31 (0.96, 1.81)
Comorbidity score*	1.08 (0.96, 1.20)	**1.40 (1.07, 1.83)**	0.93 (0.69, 1.25)
Respiratory	0.74 (0.49, 1.12)	0.76 (0.41, 1.41)	0.62 (0.31, 1.25)
Circulatory	1.01 (0.66, 1.56)	1.47 (0.78, 2.78)	0.68 (0.32, 1.46)
Any Cancer	0.85 (0.56, 1.27)	1.07 (0.61, 1.88)	0.76 (0.39, 1.46)
Type II Diabetes	**1.62 (1.09, 2.41)**	**2.53 (1.41, 4.57)**	0.91 (0.48, 1.72)
Kidney	**2.55 (1.62, 4.02)**	**2.82 (1.47, 5.39)**	**2.91 (1.41, 5.97)**
Liver	0.93 (0.53, 1.63)	1.53 (0.72, 3.24)	0.47 (0.18, 1.21)
Autoimmune	1.18 (0.77, 1.80)	1.63 (0.90, 2.96)	0.94 (0.49, 1.81)
ICU (1) vs. Not-ICU (0)	**All (n_0=657, n_1=109)**	**NHW (n_0=325, n_1=46)**	**NHAA (n_0=220, n_1=51)**
	OR (95% CI)		
Age* (y)	**2.16 (1.69, 2.75)**	**2.52 (1.76, 3.62)**	**1.88 (1.30, 2.72)**
<50	REF	REF	REF
[50, 65)	3.40 (1.89, 6.11)	**6.38 (2.24, 18.18)**	**2.99 (1.30, 6.86)**
[65, 80)	**5.75 (3.09, 10.70)**	**7.98 (2.68, 23.73)**	**5.09 (2.05, 12.65)**
>=80	**6.73 (3.04, 14.90)**	**15.27 (4.64, 50.21)**	**4.80 (1.17, 19.62)**
Male sex	**2.53 (1.63, 3.93)**	1.65 (0.86, 3.16)	**2.65 (1.37, 5.11)**
BMI*	**1.38 (1.10, 1.73)**	1.20 (0.88, 1.66)	**1.46 (1.02, 2.07)**
<18.5	2.23 (0.20, 25.36)	2.72 (0.15, 47.64)	0.00 (0.00, 0.00)
[18.5, 25)	REF	REF	REF
[25, 30)	1.48 (0.62, 3.54)	1.46 (0.50, 4.21)	4.79 (0.22, 105.0)
>=30	**2.71 (1.17, 6.29)**	2.72 (0.96, 7.71)	5.68 (0.27, 118.1)
Ever-smoker	1.50 (0.94, 2.38)	1.40 (0.71, 2.77)	1.19 (0.59, 2.39)

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Never-Smoker	REF	REF	REF
Past-Smoker	1.63 (1.02, 2.61)	1.40 (0.70, 2.81)	1.39 (0.68, 2.85)
Current-Smoker	0.59 (0.11, 3.23)	1.94 (0.29, 13.0)	0.24 (0.01, 4.62)
Alcohol consumption	1.26 (0.74, 2.13)	1.16 (0.53, 2.56)	0.998 (0.46, 2.16)
Race/Ethnicity			
NHW	REF	NA	NA
NHAA	1.52 (0.92, 2.52)	NA	NA
Other	0.93 (0.41, 2.11)	NA	NA
Unknown	0.62 (0.21, 1.82)	NA	NA
SES			
< HS Education	1.35 (0.98, 1.86)	1.02 (0.65, 1.59)	1.25 (0.75, 2.07)
Unemployed	1.17 (0.87, 1.58)	1.15 (0.78, 1.70)	1.17 (0.73, 1.86)
Annual Income < FPL	0.90 (0.65, 1.23)	0.90 (0.60, 1.36)	1.13 (0.67, 1.92)
Persons per Mile^2	1.16 (0.92, 1.46)	1.32 (0.92, 1.90)	1.07 (0.76, 1.50)
Comorbidity score*	0.93 (0.74, 1.17)	1.10 (0.79, 1.53)	0.89 (0.64, 1.25)
Respiratory	0.72 (0.44, 1.18)	1.05 (0.47, 2.35)	0.58 (0.28, 1.21)
Circulatory	0.90 (0.52, 1.55)	0.97 (0.43, 2.20)	1.07 (0.43, 2.65)
Any Cancer	0.75 (0.45, 1.24)	0.99 (0.49, 1.97)	0.56 (0.25, 1.23)
Type II Diabetes	1.17 (0.73, 1.88)	1.81 (0.91, 3.62)	1.02 (0.50, 2.09)
Kidney	1.49 (0.90, 2.46)	1.29 (0.61, 2.75)	1.91 (0.92, 3.97)
Liver	0.57 (0.26, 1.24)	1.02 (0.38, 2.78)	0.28 (0.07, 1.14)
Autoimmune	0.87 (0.50, 1.48)	0.89 (0.40, 1.97)	0.98 (0.46, 2.10)

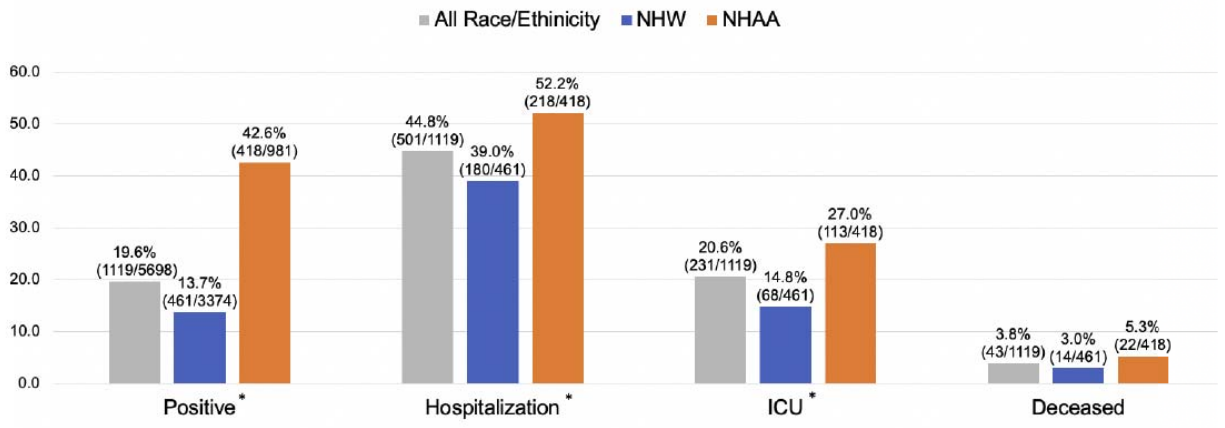
Deceased (1) vs. Alive (0)	**All (n_0=748, n_1=18)**	**NHW (n_0=363, n_1=8)**	**NHAA (n_0=263, n_1=8)**
	OR (95% CI)		
Age*	**3.15 (1.79, 5.57)**	**7.65 (2.63, 22.2)**	**1.77 (0.85, 3.69)**
Male sex	**2.99 (1.12, 7.98)**	2.67 (0.60, 11.8)	1.83 (0.48, 6.95)
BMI*	1.22 (0.74, 2.01)	0.73 (0.28, 1.91)	1.04 (0.52, 2.09)
Ever-smoker	1.94 (0.70, 5.39)	1.58 (0.35, 7.16)	2.10 (0.54, 8.12)
Never-Smoker	REF	REF	REF
Past-Smoker	2.06 (0.73, 5.77)	1.61 (0.36, 7.32)	2.36 (0.60, 9.28)
Current-Smoker	2.40 (0.15, 39.6)	9.04 (0.28, 295.2)	1.77 (0.08, 37.81)
Alcohol consumption	2.60 (0.79, 8.52)	2.39 (0.37, 15.3)	2.01 (0.47, 8.53)
Race/Ethnicity			
NHW	REF	NA	NA
NHAA	1.17 (0.4, 3.45)	NA	NA
Other Ethnicity	0.89 (0.15, 5.41)	NA	NA
Unknown Ethnicity	1.28 (0.21, 7.75)	NA	NA
SES			
< HS Education	1.63 (0.95, 2.8)	1.03 (0.50, 2.15)	1.47 (0.65, 3.29)
Unemployed	0.86 (0.49, 1.51)	1.55 (0.72, 3.31)	0.60 (0.23, 1.35)
Annual Income < FPL	1.26 (0.72, 2.20)	1.20 (0.57, 2.53)	1.95 (0.78, 4.89)
Persons per Mile^2	1.23 (0.81, 1.86)	1.52 (0.73, 3.16)	1.14 (0.60, 2.15)
Comorbidity score*	**1.72 (1.08, 2.75)**	1.43 (0.69, 2.97)	**2.95 (1.33, 6.52)**
Respiratory	2.61 (0.71, 9.64)	1.74 (0.26, 11.8)	6.61 (0.47, 93.7)
Circulatory	1.49 (0.42, 5.33)	0.70 (0.11, 4.3)	3.30 (0.25, 43.5)
Any Cancer	1.99 (0.77, 5.15)	1.83 (0.46, 7.27)	3.20 (0.80, 12.8)
Type II Diabetes	2.16 (0.85, 5.49)	2.64 (0.67, 10.4)	1.88 (0.48, 7.45)
Kidney	**2.82 (1.09, 7.28)**	1.23 (0.28, 5.39)	**11.90 (2.20, 64.7)**
Liver	1.16 (0.31, 4.37)	2.11 (0.28, 15.9)	1.50 (0.26, 8.67)
Autoimmune	1.69 (0.61, 4.70)	0.79 (0.13, 4.87)	3.58 (0.95, 13.6)

Abbreviations: n_1, sample size for COVID outcome patients; n_0, sample size for contrast group; OR, odds

medRxiv preprint doi: https://doi.org/10.1101/2020.06.16.20133140; this version posted June 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ratio; NHW, non-Hispanic Whites; NHAA, non-Hispanic African Americans; BMI, body mass index; NA, not applicable; REF, reference group; SES, residential-level socioeconomic characteristics; HS, high school; FPL, federal poverty level.
*Odds ratios associated with one standard deviation (SD) change, with SD defined as the SD of the sample under consideration in each model.

Figure 1. COVID-19 Outcomes by Race/Ethnicity



OR (95% CI)				
Unadjusted	4.69 (4.00, 5.50)	1.70 (1.30, 2.22)	2.14 (1.53, 2.99)	1.77 (0.89, 3.51)
Unadjusted †	4.69 (3.99, 5.50)	1.70 (1.30, 2.22)	2.13 (1.53, 2.98)	1.75 (0.89, 3.44)
Adjustment 1[a] †	3.56 (2.94, 4.31)	2.35 (1.73, 3.20)	2.39 (1.20, 3.14)	2.24 (1.12, 4.48)
Adjustment 2[b] †	3.53 (2.88, 4.33)	1.66 (1.09, 2.52)	1.52 (0.92, 2.52)	1.17 (0.40, 3.45)
Adjustment 3[c] †	3.51 (2.84, 4.33)	1.73 (1.13, 2.65)	1.49 (0.89, 2.50)	0.96 (0.32, 2.89)

Abbreviations: NHW, non-Hispanic Whites; NHAA, non-Hispanic African Americans; ICU, intensive care unit; OR, odds ratio.
* Chi-squared test P<.001, comparing the proportion between NHW and NHAA.
† Firth correction applied logistic regression.
[a] Multivariable logistic regression, adjusting for age, gender, race/ethnicity, persons per Mile^2.
[b] Multivariable logistic regression, adjusting for adjustment 1 + less than high school education, unemployed, annual income below federal poverty level (FPL).
[c] Multivariable logistic regression, adjusting for adjustment 2 + the comorbidity score.