Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California

ABSTRACT As the coronavirus disease (COVID-19) pandemic spreads throughout the United States, evidence is mounting that racial and ethnic minorities and socioeconomically disadvantaged groups are bearing a disproportionate burden of illness and death. We conducted a retrospective cohort analysis of COVID-19 patients at Sutter Health, a large integrated health care system in northern California, to measure potential disparities. We used Sutter’s integrated electronic health record to identify adults with suspected and confirmed COVID-19, and used multivariable logistic regression to assess risk of hospitalization, adjusting for known risk factors, such as race/ethnicity, sex, age, health, and socioeconomic variables. We analyzed 1,052 confirmed cases of COVID-19 from January 1–April 8, 2020. Among our findings, we observed that, compared with non-Hispanic white patients, African Americans had 2.7 times the odds of hospitalization, after adjusting for age, sex, comorbidities, and income. We explore possible explanations for this, including societal factors that either result in barriers to timely access to care or create circumstances in which patients view delaying care as the most sensible option. Our study provides real-world evidence that there are racial and ethnic disparities in the presentation of COVID-19. [Editor’s Note: This Fast Track Ahead Of Print article is the accepted version of the peer-reviewed manuscript. The final edited version will appear in an upcoming issue of Health Affairs.]

In late December 2019, the severe acute respiratory syndrome coronavirus (SARS-CoV-2) was first detected in Wuhan, China, and quickly developed into a devastating international outbreak, the likes of which have not been seen since the influenza pandemic of 1918. According to the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO), the disease caused by the novel coronavirus, COVID-19, has now been detected in more than 100 countries, including in the United States. On March 11, 2020, COVID-19 was officially declared a global pandemic. Since February, when the first US community-acquired case was detected in California, the disease quickly spread across the nation, and at the time of this writing, the US had an estimated 1.2 million confirmed cases and more than 70,000 COVID-19-related deaths. However, these numbers are suspected to vastly underestimate true disease prevalence due to a widespread shortage of testing kits and an unknown number of asymptomatic cases. Thus, accurate rates of infection and mortality remain elusive.

As the pandemic spreads throughout the US, alarming evidence is emerging to suggest that some racial and ethnic minorities, as well as...
socioeconomically disadvantaged groups, are bearing a disproportionate burden of illness and death. This is especially concerning in California, a state that has embraced the Affordable Care Act and Medicaid expansion in an attempt to provide coverage and access to the vast majority of its residents. On April 8, California Governor Gavin Newsom presented a statewide analysis of COVID-19, using partial data, asserting that there were no apparent disparities in rates of testing by race or ethnicity. This was based on the observation that the demographic distribution of testing mirrored the underlying distribution of race and ethnicity across the state. However, the analysis was limited to 54 percent of the population, and race and ethnicity information was available for only 40 percent of the 16,957 people who tested positive for the novel coronavirus in California at the time. As of May 12, race and ethnicity information was available for only 67 percent of the 74,141 people who tested positive.

Testing rates aside, even less is known about how disparities are reflected in severity of illness and outcomes. COVID-19 infections have caused a range of symptoms, from mild (such as fever and malaise) to life-threatening (such as acute respiratory distress), while some infected individuals remain asymptomatic. Although efforts in California to expand no-cost COVID-19 testing to all state residents have likely contributed to more equitable testing rates, African Americans comprise 6.0 percent of the population of California and as of May 12 6.1 percent of COVID-19 cases, but by contrast represent 10.3 percent of COVID-19-related deaths (to the extent race and ethnicity is documented). In late March, the CDC published a report examining sociodemographic and clinical characteristics of cases using hospitalization data from a surveillance network tracking COVID-19 in 14 states during the month of March, stratified by age, race, and sex. Among those included in the analysis (n = 1,482 patients hospitalized with COVID-19), data on race and ethnicity was available in less than half (n = 580). Nevertheless, the CDC’s findings indicated that despite representing 18 percent of cases in the analysis, African American patients comprised 33 percent of those hospitalized, raising questions about whether the clinical course for African Americans affected by COVID-19 may differ from other races and ethnicities.

Health systems are the focal point of the COVID-19 pandemic, especially in the absence of widespread community testing, and are vital to understanding the extent of this pandemic and identifying groups at highest risk. In this study, we used electronic health record (EHR) data from Sutter Health, a large not-for-profit integrated health care system in northern California, to characterize COVID-19 tested and confirmed cases by key sociodemographic and clinical characteristics, including self-reported race and ethnicity.

Study Data And Methods

**SETTING** This study was conducted at Sutter Health, a large mixed-payer, integrated health care delivery system in northern California. Sutter delivers comprehensive medical services in 100+ ambulatory clinics and 24 acute-care hospitals, caring for approximately 3.5 million people each year, across 22 counties in California, in both urban and rural settings. Ten of those counties are in the San Francisco Bay Area, a highly populated and racially diverse region. Sutter’s Epic EHR (Epic Systems Corporation) is fully integrated across all hospital and ambulatory sites. Sutter has collected patient self-reported race, ethnicity, ancestry, and language (REAL) data since 2010. As of 2019, Sutter patients self-identified as 45.9 percent white, 15.6 percent Hispanic, 15.8 percent Asian, 4.9 percent Black/African American, and 17.8 percent Other (American Indian/Alaskan Native, mixed race, declined-to-state, and unknown). Data were retrospectively extracted from the Sutter EHR for the study period between January 1 and April 8, 2020. This study was approved by Sutter’s Institutional Review Board and was conducted according to Health Insurance Portability and Accountability Act standards.

**COHORT IDENTIFICATION** We identified patients ages 18 or older, with at least one encounter at a Sutter facility during the study period. We defined index encounter as the date of the first encounter during the study period in which a patient satisfied criteria for one of the following groups:

- **Group 1** (suspected cases with evidence of testing): This group includes patients with a record of a COVID-19 test in their EHR laboratory records, regardless of the test result or testing location; testing could have occurred at a different institution.
- **Group 2** (confirmed cases): This group includes patients with evidence of a positive test result in the EHR laboratory records, regardless of testing location, or patients who had a documented ICD-10 diagnosis of confirmed COVID-19 in the EHR without a positive COVID-19 test result. A random sample chart review was performed to confirm evidence of COVID-19 positive status (in unstructured notes) for the latter.

**DATA EXTRACTION AND MANAGEMENT** For all patients we extracted demographic information from the EHR, including patients’ dates of birth,
sex, self-reported race and ethnicity, and primary insurance. Age was classified as 18–29, and in 10-year age categories thereafter with 80+ as the oldest category. Race and ethnicity were defined by Hispanic identity, followed by racial group. If a patient did not self-identify as Hispanic, we classified them based on their race. Insurance status was identified by the active primary payer documented in the EHR, at the most recent encounter, and classified as commercial, Medicaid, Medicare (Part A/B or Part C), other insurance, or self-pay/not reported. Homelessness status was assessed using documentation on patient registration and patient address associated with the encounter. Homelessness status is assessed at every inpatient and ED encounter with standardized system-wide protocols.

From the EHR, we extracted patients’ comorbidities, using ICD-10 diagnoses as of the index encounter, and smoking status reported at the index or most recent encounter during the study period. We extracted post-index hospital admission data and mortality data for all patients. By chart review, we confirmed mortality for those who died within the hospital and for hospitalized patients who were discharged.

For Group 2, using patients’ last known address, we geocoded each patient and acquired median income level by Census ZIP Code Tabulation Areas (ZCTAs). Household income was categorized by quartiles.

Among patients who were tested for COVID-19 at a Sutter facility, we extracted information on the testing location (ambulatory/outpatient, ED, or inpatient).

**STATISTICAL ANALYSES** Patient sociodemographic and clinical characteristics for each of the two groups were summarized using descriptive statistics. We also summarized the sociodemographic and clinical characteristics of hospital-admitted COVID-19 cases and patients who died. Standard univariate tests were used: t-tests for continuous variables and chi-squared tests for categorical variables. We described testing site stratified by racial and ethnic group.

Among confirmed COVID-19 cases (Group 2), we used logistic regression to examine the association between clinical and sociodemographic factors and hospital admission. We built a series of step-wise models that incrementally included more covariates as follows: univariate models for all covariates (unadjusted models); demographics (Adjusted Model 1); demographics plus clinical characteristics (Adjusted Model 2); demographics, clinical characteristics, and socio-demographic characteristics (Adjusted Model 3; full model). For these models, for age, we combined the two youngest age groups (18–39) as the reference group. Odds ratios (OR) and 95% confidence intervals (95% CI) were generated. A p-value <0.05 was considered statistically significant.

We geocoded the current residential addresses of patients in Group 2 using Address Verification by Informatica. Of 1052 confirmed COVID-19 patients, 28 were not geocoded because of missing addresses (n = 18) or because their addresses were outside of California (n = 10). We overlaid ZCTAs and attributed median household income based on the 2013–2018 American Community Survey 5-year estimates. One additional patient was removed because they lived outside the boundary of the ZCTAs.

We conducted additional analyses to test the robustness of our findings. To examine possible racial and ethnic disparities in COVID-19 testing, given the limited supply of test-kits, we compared the race and ethnicity of patients who presented at the index encounter with symptoms consistent with COVID-19 clinical presentation, but were not tested. Details of the ICD-10 codes and definition used to identify this group are provided in the online appendix. Additionally, given the rapidly changing nature of the pandemic, we confirmed that our findings were consistent at two different points in time, one week apart within the study period.

Data extraction and statistical analyses were performed in R v3.6.3, SAS9.4, and STATA. Geographic data processing and visualizations were designed using R version 3.6.3 in RStudio.

**LIMITATIONS** Our study has several limitations. First, the pandemic is very fast moving, and the count of COVID-19 patients is accumulating very quickly, but we had to freeze the cohort in order to process the data and summarize our findings. We addressed this with the sensitivity analysis described above. Second, we did not consider severity of comorbid conditions. Instead, we relied only on diagnosis codes, so patients with severe and mild disease are classified together. Third, we assigned income at the ZIP code level due to institutional constraints, though income would have been more accurate at the census tract level. Fourth, we examined differential patterns of missing values for smoking status by race and ethnicity and found some differences. However, African Americans had the lowest percentage of missing values, so we expect any bias would be toward no effect. Fifth, statistical analysis of interactions and risk of mortality was limited by power. Finally, we classified 158 patients as confirmed COVID-19 cases with only a diagnostic code, but no lab record of a positive test. However, chart review demonstrated that most did have a documented positive lab result, and in the small subset with documentation of a negative test result, some may have recovered or
have been false negatives. We therefore believe that this would not have a large effect on our conclusions as, at most, falsely classified Group 2 patients would represent a very small percentage of the overall confirmed sample.

The study also has several strengths. Sutter collects patient self-reported REAL data. This is important given the limitations of state and national estimates, which have large proportions of missing race and ethnicity data. Further, Sutter’s patient population is representative of the state of California according to underlying racial/ethnic and economic distributions. Finally, Sutter has an integrated system-wide EHR that allows us to analyze data across our network of clinics and hospitals.

**Study Results**

**CHARACTERISTICS OF SUSPECTED COVID-19 CASES (GROUP 1)** We identified 14,036 patients who were tested for COVID-19 from January 1–April 8, 2020 (Group 1). The average age was 50.7 years and 68.0 percent of patients were younger than 60 years old (supplemental exhibit 1). Females comprised 60.7 percent of patients in this group. Almost half (48.3 percent) of Group 1 patients were of non-Hispanic white race and ethnicity, 6.7 percent identified as African American, and 19.1 percent as Hispanic; 12.2 percent and 16.9 percent of patients were insured by Medicaid and Medicare, respectively. Approximately one-fourth (27.4 percent) of patients were identified as current or past smokers and 29.8 percent had diagnosed hypertension.

A smaller percentage of African Americans (29.9 percent) were COVID-19-tested in an ambulatory setting compared to non-Hispanic whites (56.0 percent), Asians (60.0 percent) and Hispanics (53.8 percent) (supplemental exhibit 2). The majority of African Americans were tested in hospitals, either in the ED (37.8 percent) or as inpatients (32.3 percent).

In a sensitivity analysis, we identified 2,648 patients who presented at the index encounter with symptoms consistent with COVID-19 clinical presentation, but were not tested (see the appendix for the list of symptoms included). Compared with those who were tested, these individuals had similar distributions of sex, and race and ethnicity (data not shown).

**CHARACTERISTICS OF CONFIRMED COVID-19 CASES (GROUP 2)** By April 8, Sutter had treated 1,052 individual patients with confirmed COVID-19, representing 5.7 percent of California COVID-19 cases at that time. Most had positive test results, while some had a formal diagnosis (ICD-10 code) of COVID-19 in the EHR, but no laboratory record of a positive result (n = 158).

We reviewed charts for a 20 percent random sample of these (32 charts), of which 28 (87.5 percent) had positive test results documented in their progress notes. The remaining 4 (12.5 percent) had documentation of a negative test result after the formal diagnosis was given. Given this, and the potentially high false-negative rates of the tests, we estimate that at most a very low percentage (12.5 percent of 158, or about 20 individuals) of the total confirmed cases were potentially misclassified.

The average age among confirmed cases was 53.0 years and the majority of these patients (62.7 percent) were younger than 60 years of age (supplemental exhibit 1). Approximately half of all confirmed cases were female, 5.8 percent were African American and more than one-quarter (25.8 percent) were Hispanic. Further, 31.8 percent had a diagnosis of hypertension, and 14.3 percent had type 2 diabetes.

Sociodemographic and clinical characteristics of confirmed cases who were admitted to an inpatient setting (n = 256; 24.3 percent), and further, transferred to an intensive care unit (ICU) (n = 110; 10.5 percent) are presented in supplemental exhibits 1 and 3. Overall, 66.0 percent of hospital-admitted cases were 60 years old or older and the mean age was 65.6 years (supplemental exhibit 1). Among African American patients with confirmed COVID-19, 52.5 percent (n = 32) were hospitalized compared with 25.7 percent (n = 110) of non-Hispanic white patients (supplemental exhibit 3), and a higher proportion of African American patients were transferred to the ICU than their non-Hispanic white counterparts (24.6 percent versus 10.7 percent, respectively). A greater proportion of male than female patients were hospitalized, and a higher proportion of males ultimately were transferred to the ICU than females.

Of the total confirmed cases, we geocoded and mapped 97.2 percent (n = 1793) (see the appendix for visualizations of the income disparities for COVID-19 confirmed cases by race and ethnicity). Median annual household income values ranged from $31,379 to $250,001. Overall, African Americans with COVID-19 lived in ZIP codes with lower income compared to all other racial and ethnic groups (p < 0.001).

Among confirmed COVID-19 cases, 51 patients died (mean age: 79.5 years old). Among confirmed cases, 3.4 percent of females died compared with 6.4 percent of males. By race and ethnicity, 6.1 percent of non-Hispanic White patients died compared with 5.6 percent of Asian patients, 4.9 percent of African American patients, and 3.7 percent of Hispanic patients.

**RESULTS OF MULTIVARIABLE ANALYSIS** Several sociodemographic and clinical characteristics
Discussion

Health systems are the front line in the current COVID-19 pandemic, as symptomatic individuals surge to hospitals and clinics in search of testing and treatment. Data from integrated health systems are vital for understanding the impact of the epidemic, especially in the absence of widespread testing and community-based sampling. On April 9, the California Department of Public Health reported that approximately 177,600 people in California have been tested for COVID-19. Of these, 7.9 percent were Sutter patients \( (n = 14,036) \). At that time, Sutter had treated about 5.7 percent \( (n = 1,052) \) of the estimated 18,309 confirmed cases in the state. The breakdown of COVID-19 cases by age and sex were similar for Sutter and the state.

Despite predictions that the COVID-19 pandemic would be the “great equalizer,” sparing no segment of the population from risk, reports of disparities in testing, treatment, and outcomes are emerging. California’s death rate among African Americans is higher than that group’s representation in the population (10 percent mortality versus 6 percent population) and even more disproportionate in some counties.6 Recent data from Los Angeles County show a 14 percent mortality rate among African Americans, who make up 9 percent of the county’s population.13,24 The California Health Care Foundation has identified the elevated risk among African Americans in the context of this pandemic as “a perfect storm of irrefutable evidence that people of color are caught in a web of social inequality.”15 While we did not find mortality differences by race and ethnicity in our study, the study was not powered to detect such differences.

We found that African Americans with confirmed COVID-19 are significantly more likely to be admitted to the hospital than their non-Hispanic white counterparts, after controlling for sociodemographics, clinical factors, and income. There are several possible explanations for these observed disparities, which likely result from a constellation of factors.

The greater odds of hospital admission may indicate that African Americans have more advanced or severe illness at the time of presenting for COVID-19 testing and medical care. One hypothesis is that there may be some unknown or unmeasured genetic or biological factors that increase the severity of this illness for African Americans. Another possibility is that our results are explained by societal factors that either result in barriers to timely access to care or create circumstances in which patients view delaying care as the most sensible option.16 This is supported by our observation that, although there were not apparent racial disparities in testing, African American patients were more likely to have been tested at a hospital than in the ambulatory environment. Our sensitivity analysis did not reveal any racial or ethnic differences among those presenting with symptoms, but not tested. The disparity therefore may not be in who is tested, but when.

Our prior research has found that our African American patients are more likely to access care later, in the acute setting. Even those with ambulatory-care-sensitive conditions, such as asthma, have demonstrated a propensity to present through the ED and hospital rather than ambulatory settings.17 The pattern we observed with COVID-19 testing is similar. For African Americans, this is likely not due to lack of insurance coverage. Compared with other states, support in California for the Affordable Care Act and expansion of Medicaid has resulted in a high level
of insurance coverage among African Americans. But, having insurance does not guarantee access to primary care, which can be affected by structural inequities as well as personal agency.

In addition, unconscious biases on the part of providers and patients’ prior negative experiences with health care can lead to distrust and the decision to seek care only in the most extreme circumstances. This has been documented in health care settings across the US and remains a major threat to health equity. Presentation and testing in the ED and hospital may indicate that African Americans seek care at a later stage, leading to the higher rates of admissions. Policies that support community-based outreach, testing, and access to culturally competent care within the African American community hold the promise of earlier testing, diagnosis, and the potential to have a positive impact on some of the disparities we have observed. In the case of COVID-19, early identification is especially important as this can reduce the community spread of the disease.

Given that financial stress may contribute to a need to continue employment and delay medical care despite symptoms, we also explored the potential for underlying socioeconomic disparities to increase odds of hospital admission. We found disparities among Medicaid versus commercially insured beneficiaries and among lower-income (25th percentile of ZIP code median household income) versus higher-income (51st percentile and above) patients. As a group, African American patients tended to live in areas in the lower-income brackets. While the adjustment for income level attenuated the effect of race and ethnicity, it did not explain the entire effect. However, as we acknowledge in the limitations of this study, measurement of median household income at a more granular level may provide additional clarity.

Our findings are in agreement with other studies that suggest that advanced age and severe underlying health conditions appear to increase the risk of developing serious COVID-19 illness. According to the CDC, the most commonly reported conditions co-occurring with COVID-19 were diabetes mellitus, chronic lung disease, and cardiovascular disease. Because African Americans continue to bear a disproportionate burden of chronic disease, with pronounced disparities in type 2 diabetes, congestive heart failure, and hypertension, there are concerns that existing racial and ethnic disparities will be exacerbated and compounded by COVID-19. We found type 2 diabetes and congestive heart failure were independently associated with an increased risk of hospital admission in the fully adjusted model. We did not find that hypertension, cardiovascular disease, chronic obstructive pulmonary disease, or asthma were independent risk factors for admission after adjusting for age, gender, and race and ethnicity. While some of these conditions have been associated with an increased risk for COVID-19 infection, our focus was hospitalization. Of note, the increased risk of hospitalization among African Americans persisted, even after adjusting for all of these co-morbidities.

Our finding that males with confirmed COVID-19 infection had twice the odds of hospital admission compared with females also supports other recent findings. The CDC, for example, reports that despite the fact that 49 percent of COVID-19 cases were male, males made up 54 percent of the hospitalizations, reflecting earlier case reports from China, Italy, and South Korea. Studies have shown that females tend to access health care and preventive care more frequently than males, which may help to explain why more females are tested for COVID-19, but both sexes in our study appear to be infected in comparable proportions. More investigation of these differences is needed.

Conclusion
The COVID-19 pandemic has presented an enormous and unprecedented test of the adequacy of health systems around the world to provide appropriate care with high-quality outcomes to large numbers of people. Despite the fact that California, one of the most populous and diverse states in our nation, has embraced policies designed to provide extended health care coverage for all, challenges remain to providing equitable care, with comparable outcomes for all. The experience of Sutter Health highlights the fact that race and ethnicity play a pivotal role in determining how and when care is accessed, and the outcome. Our findings suggest that the greatest risk, in terms of hospitalization, is borne by the African American community. This pandemic offers the opportunity to identify and quantify these inequities, and to seek solutions. Health care systems have an ethical obligation to ensure that all patients receive the right care at the right time, especially in times of crisis.
NOTES


10 To access the appendix, click on the Details tab of the article online.


Downloaded from HealthAffairs.org on June 10, 2020. Copyright Project HOPE—The People-to-People Health Foundation, Inc. For personal use only. All rights reserved. Reuse permissions at HealthAffairs.org.