Comorbidities and Disparities in Outcomes of COVID-19 Among Black and White Patients

Ahmad Khan, Arka Chatterjee, *Shailendra Singh

Authors

Ahmad Khan MD
Department of Medicine, West Virginia University Health Sciences Center Charleston Division, Charleston, WV.

Arka Chatterjee MD
Division of Cardiovascular Disease, University of Alabama at Birmingham.

*Shailendra Singh MD
1. Division of Gastroenterology, West Virginia University Health Sciences Center Charleston Division, Charleston, WV.
2. Charleston Area Medical Center Health System, Charleston, WV.

Corresponding Author

Shailendra Singh M.D
Charleston Area Medical Center Health System, Charleston, WV

Email: Shail121@gmail.com
Phone no: 304-342-0821
Fax: 304-345-6679
Abstract

Initial surveillance data suggests a disproportionately high number of deaths among Black patients with COVID-19. However, high-risk comorbidities are often over-represented in the Black population, and understanding whether the disparity is entirely secondary to them is essential. We performed a retrospective cohort study using real-time analysis of electronic medical records (EMR) of patients from multiple healthcare organizations in the United States. Our results showed that Black patients with COVID-19 have a significantly higher risk of mortality, hospitalization, and invasive mechanical ventilation compared to White patients. The incremental risk of poor outcomes in Blacks persists despite accounting for a higher prevalence of comorbidities. This may point to the disparities in socioeconomic determinants of health affecting Blacks and the need for an improvement in the care of this vulnerable population.

Main

Background

Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) infection has emerged as unprecedented global health, economic, and social crisis. More than 2 million COVID-19 cases and 160,000 deaths have been reported worldwide as of April 19, 2020 [1]. Initial surveillance data from 14 states suggests a disproportionately high number of cases and deaths for Black or African Americans [2]. However, it has also been observed that comorbid conditions such as hypertension (HTN), diabetes mellitus (DM), and cardiovascular diseases are associated with worse outcomes with COVID-19 infection[3, 4]. These comorbidities are often over-represented in Black communities[5]. Thus, an understanding of the outcomes among Black patients with COVID-19 in the milieu of these comorbidities is required. We attempt to provide insight into this with a propensity-matched analysis of a large cohort of Black and white COVID 19 patients.

Methods
**Study Design and Data Source**

This was a retrospective cohort study conducted using TriNetX (Cambridge, MA, USA), a global federated health research network providing real-time access to electronic health records (EHR) of more than 40 million patients from 34 member healthcare organizations (HCOs). TriNetX provides access to HIPPA compliant de-identified longitudinal clinical data to member Health Care Organizations (HCOs) in a real-time fashion on its cloud-based platform. The de-identified clinical data such as diagnoses, procedures, medications, laboratory values, and genomic information are aggregated directly from the EHR systems of the member HCOs continuously. Both the patients and HCO's as data sources stay anonymous. Member HCO's include a mix of inpatient, outpatient, and specialty care services and provide care to a diverse patient population from different age groups, ethnicity, income levels, and geographical region. TriNetX recently fast-tracked data inflow to incorporate COVID-19 specific diagnosis and terminology following the World Health Organization (WHO) and Centers for Disease Control (CDC) COVID-19 criteria for coronavirus research. TriNetX data has also been granted a waiver from the Western IRB because it is a federated network, and only aggregated counts and statistical summaries of the de-identified information without any protected health information is received from participating HCO's. Further details of the data source, data quality checks, mapping, and coding systems to present data are available in the Supplementary.

**Patient Selection**

*Selection of COVID-19 patients*

We performed a real-time search query (April 21, 2020) on the research network to identify patients > 10 years with a diagnosis of COVID-19 between January 20, 2020, and April 21, 2020. The search criteria to identify potential COVID-19 patients was based on specific diagnosis and terminology recommended by the WHO and CDC. These codes included International Classification of Diseases, Ninth Revision and tenth Revision, Clinical Modification (ICD-10-CM) codes U07.1 (COVID-19, virus identified), B34.2 (Coronavirus infection, unspecified), B97.29 (Other coronavirus as the cause of diseases classified elsewhere) and J12.81 (Pneumonia due to SARS-associated coronavirus). Patients identified with diagnosis code 079.89 (Other specified viral infection) and Z03.818 (Encounter for
observation for suspected exposure to other biological agents ruled out) were excluded. Only patients diagnosed with the codes, as mentioned earlier after January 20, 2020 (first confirmed case in the USA) were included. The B97.29 code was specifically included based on the recommendation from the general guidance of the ICD-10-CM Official Coding Guidelines released by the CDC on February 20, 2020. Similarly, U07.1 is the new specific code for a confirmed diagnosis of the COVID-19 with a positive COVID-19 test result starting April 1, 2020, as per the new CDC guidelines. The codes B34.2 and J12.81 were used more often before the CDC guidelines. Patients with ICD-9 code 079.89 (mapped to ICD-10 code B34.2 and B97.2) were excluded to reduce any false positive COVID-19 patients because this ICD-9 code can still be used occasionally as "catch-all" code for more than 50 viral infections.

**Selection of Black and White Patient Cohort**

Patient demographics, including race, and ethnicity, are coded to HL7 version 3 administrative standards. Races and Ethnicity are self-reported and categorized as per the United States (U.S) Census Bureau [6]. Federal standards acknowledge that race is a social-political construct, and self-identification with one more race categories is preferred to observer identification. The most recent United States Census officially recognized five racial categories (White, Black or African American, American and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander). We only included COVID-19 patients with races identified as Black or African American (Black cohort) and White American (White cohort) (Figure 1). We excluded patients with race identified as Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, or with unknown race. The reporting of Ethnicity is separate from the race and can be categorized as Non-Hispanic or Latino and Hispanic or Latino as per the U.S census. Hispanics could report as any race [6]. We excluded patients identified as Hispanics from the white cohort for a more generalizable comparison between white and black patients.

**Definition of Variables and Outcomes**

TriNetX has the capability of analyzing data based on a temporal relationship to the index event. The index event in our study was defined as the diagnosis of COVID-19. Baseline characteristics were estimated from any time before the index event. Presenting signs and symptoms were recorded from the time of index
event up to two weeks before the index event. The outcomes were assessed from the index event up to 30 days after the index event. The primary outcome was the mortality in Black and white COVID-19 patients. Other outcomes of interest were mechanical ventilation, hospitalization, critical care services, development of acute respiratory distress syndrome (ARDS), septic shock, acute kidney injury (AKI), and continuous renal replacement therapy (CRRT) requirement. We also performed a subgroup analysis of mortality stratified by age into three sub-groups ≤49, 50-64, and ≥65 years. Subgroups were divided based on the mean age of the cohort and the risk of poor outcomes while obtaining a comparable sample size in each group. We also compared the laboratory findings after the diagnosis of COVID-19. Codes used to define outcomes (Supplementary Table 1), and laboratory values (Supplementary Table 2) can be found in the Supplementary.

Statistical Analysis

We conducted all the statistical analyses using TriNetX, which has inbuilt analytics features. Descriptive statistics were reported as means, standard deviations for continuous variables, and total counts with frequencies for categorical variables. However, patient counts were obfuscated by rounding up to the nearest 10 in the analysis to safeguard the protected health information. We performed the inferential statistics by calculating the risk difference and risk ratio for the comparison of each outcome in both the groups. However, to adjust for the confounding factors, we performed 1:1 propensity score matching using a greedy nearest-neighbor matching algorithm with a caliper of 0.1 standard deviations. The outcomes are reported with and without propensity matching. An alpha of less than 0.05 was considered as the cutoff for statistical significance.

Results:

A total of 7,082 COVID-19 patients from 18 HCOs were identified. The mean age for these COVID-19 patients was 51.3 ±18.2 years, and the majority of them were females (60%). Out of these, 3,435 white patients and 2,621 black patients were included in the study (Figure 1).
Table 1 compares the characteristics between the Black and white COVID-19 patient cohorts. Patients in the Black cohort were older than as compared to the white cohort (mean age 54.1±16.8 vs. 50.9 ± 18.7, P<0.001). Black patients had higher BMI and high rates of hypertension, diabetes, chronic kidney disease, and heart failure. The prevalence of chronic respiratory disease (Asthma and COPD combined) was higher in white patients, while the rate of Ischemic heart disease was similar. Propensity matching of cohorts was performed accounting for age, gender, BMI, and major comorbidities described in Table 1. After propensity score matching, there were 1862 patients in each group, and the cohorts were well-balanced (Figure 2 and Table 1). Presenting clinical signs and symptoms are also described in Table 1. Laboratory findings in each group after the diagnosis of COVID-19 are described in Supplementary Table 3. Significantly higher creatinine (mg/dl) (2.13 ±2.51 Vs. 1.25 ±1.27, P <0.001), erythrocyte sedimentation rate (mm/hr) (81.04 ±34.95 Vs. 53.96 ±29.82, P <0.001), and C-Reactive Protein (mg/L) (101.31 ±107.22 Vs. 85.27 ±90.16, P: 0.004) levels were seen in black patients.

The risk of 30-day mortality was significantly higher in both the unmatched (8.08% vs. 4.22%, RR 1.91, 95 % CI: 1.56,2.35, P <0.001) and matched black patients (6.92% vs. 5.20%, RR 1.33, 95 % CI: 1.03,1.72, P 0.02) compared to whites. Overall and age-wise subgroup analysis of mortality outcomes is described in Table 2 and Figure 3. A higher proportion of patients in the black cohort required hospitalization, invasive mechanical ventilation, and critical care services compared to the White cohort. Similarly, the prevalence of ARDS, AKI, requirement for CRRT, and septic shock were also higher in the black cohort (Table 2). Characteristics of non-survivors, patients requiring mechanical ventilation, and hospitalizations among black and White patients are compared in Supplementary Table 4.

**Discussion**

Current data suggests worse COVID-19 outcomes in older patients with pre-existing HTN, DM, and cardiovascular diseases[3, 4]. This is the first attempt to examine and understand the racial disparity in COVID-19 after disaggregating the effect of comorbidities. In the unmatched sample, we found a 91% increased risk of death in Blacks and a 137% higher need for hospitalization along with a greater than threefold higher need for mechanical ventilation. These findings go
along with news media reports of drastically higher COVID-19 infection rates and mortality in Black patients in Chicago, Louisiana, and Michigan with up to 6-fold higher mortality in predominantly black counties compared to predominantly white ones[7-9].

A large part of these disparities in outcomes can be explained by the differences in high-risk characteristics between black and white patients (Unmatched RR for mortality 1.91 vs. matched RR 1.33). The cohorts derived from our data source are reasonably representative of the differences in these comorbidities vis-à-vis Black and White populations in the United States. It is well known that the incidence of HTN is exceptionally high amongst Black adults (>40%), and even if recognized, it is much less likely to be controlled[10-12]. Similarly, the prevalence of diabetes in our study cohort also mirrors known data about racial differences [13]. Obesity also disproportionately affects minorities and individuals from lower socioeconomic groups, particularly Blacks[14]. Additionally, these comorbidities are not independent but interrelated, and having one disease also directly raises the hazard of developing others. A higher incidence of HTN and DM is associated with chronic kidney disease and, most likely, the higher prevalence of heart failure in the Black population[15, 16]. Moreover, the mortality from certain diseases is higher in Blacks than whites, despite being less or equally prevalent such as coronary heart disease and smoking[17, 18].

A sizeable incremental risk in Blacks persists even after accounting for an exhaustive list of comorbidities. Multiple factors likely interact in intricate ways to contribute to this disparity. It is possible that socioeconomic environments that promote poor health in the Black population likely have a critical influence on COVID-19 outcomes. Black communities still live in more segregated, closely spaced urban housing [19], and this may negatively impact the ability to practice social distancing. A disproportionate number of Blacks are employed in the service industry and in low paying jobs, which cannot be transitioned to working remotely during the COVID-19 pandemic[20]. These jobs also may lack protection from furloughs and lay-offs, requiring families to be forced to find work, which increases the chances of exposure. Blacks also face challenges in access to health care in the United States. Moreover, when they receive health care, it may not be equivalent to that for the white population[21]. A greater number of Blacks lack health insurance and have a much higher rate of loss of insurance due to
job/life circumstances[10, 22]. Lack of access to health care can delay diagnosis and medical care in Blacks resulting in poor COVID-19 outcomes. Moreover, patient self-care, a behavior ranging from compliance with preventive measures to seeking health care, is perceived as lacking in Blacks [21].

Other potential etiologies also need to be explored – the angiotensin-converting enzyme 2 (ACE-2) receptor is responsible for COVID 19 virus entry into host cells and also has been hypothesized to play a part in the lung injury. The frequency of ACE-2 variants and their expression may explain susceptibility and severity of disease in different races, and a wide variety of allele frequency differences have already been shown in Asians, Eurasians and African populations[23]. There are also genetic polymorphisms for clotting factors, especially Fibrinogen in the Black population, which may increase thrombosis[24]. There are signals that thromboembolic complications are seen with sicker COVID-19 patients, and a link between this and increased mortality in Black patients may be a potential area for investigation.

Our analysis carries the limitations of using an EMR database despite the ability of TriNetX to aggregate the data directly from the EMRs in a real-time fashion minimizing the risk of data collection errors. Patient counts were rounded up to the nearest 10 in our analysis to guard Protected Health Information (PHI). Rounding may influence measures of association results for small cohorts and infrequent outcomes; however, most of our outcomes have a relatively large number of patients. Exposure history, incubation time, and dynamic changes in patients’ clinical condition could not be estimated. Socioeconomic factors, geographical variations, delivery, and access to health care during COVID-19 were beyond the scope of our study; however, they have a significant influence on differences in high-risk characteristics and outcomes for COVID-19. We did not include other minorities in our study because of the relatively small sample size.

In conclusion, our study provides evidence of poor outcomes among Black patients affected by COVID-19. Our results indicate the role of disproportionate comorbidities in the Black population but also shows an unexplained significant residual difference that persists after careful propensity matching. This highlights
the role socioeconomic factors and healthcare disparities have played in the COVID-19 pandemic and the need for a vast improvement in the care of Black population.

Acknowledgment

We acknowledge the West Virginia Clinical and Translational Science Institute to provide us access, and training to the TriNETX global healthcare network.

We also acknowledge the TriNETX (Cambridge, MA, USA) healthcare network for design assistance to complete this project.

The authors received no financial support or grants for the research, authorship, and publication of this article.

Ahmad Khan, Arka Chatterjee, and Shailendra Singh declare that they have no conflict of interest.

Data availability

Data is available to member Health Care Organizations on the TriNetX https://www.trinetx.com/ research network platform. Data aggregated directly from the electronic health records systems of the participating HCOs is provided a real-time fashion on the TriNetX cloud-based platform. Datasets for individual patient level data can be requested from TriNetX by participating Health Care Organizations. Codes for creating cohort of COVID-19 patients races are included in methods. Race is coded to HL7 version 3 administrative standards. Codes for other characteristics, laboratory values, and outcomes are included in the supplementary.

References:


8. Reyes C., H.N., Gutowski C, St Clair S, Pratt G, Chicago’s coronavirus disparity: black Chicagoans are dying at nearly six times the rate of white residents, data show, in Chicago Tribune. 2020.


Table Legends

Table 1: Comparison of patient characteristics and presenting signs and symptoms among COVID-19 African American and White patients. Demographics and comorbidities are compared before and after propensity matching of cohorts.

Table 2: Outcomes among COVID-19 African American and White patients. Outcomes are compared before and after propensity score matching of cohorts.

Figure Legends

Figure 1: Flow chart showing the selection of African American and White COVID-19 patient cohorts.

Figure 2: Propensity Score Density Graph before and after propensity score matching of the African-American and White cohort of patients.

Figure 3: Comparison of Outcomes among African American and White patients after propensity score matching of cohorts. Mortality is compared in the overall cohort and in subgroups based on age. Secondary outcomes described in the overall cohort of patients.

Supplementary

Supplementary Methods: Data Source, data quality checks, mapping, and coding system to present data.

Supplementary Table 1: Search query using ICD-10 and CPT codes used to define outcomes of the study.

Supplementary Table 2: TriNetX (TNX) laboratory or Logical Observation Identifiers Names and Codes (LOINC) to capture laboratory values.

Supplementary Table 3: Comparison of Laboratory findings in African American and White patients after diagnosis of COVID-19.

Supplementary Table 4: Comparison of characteristics in patients with mortality, mechanical ventilation, and hospitalizations among African Americans and Whites.
Table 1: Comparison of patient characteristics and presenting signs and symptoms among COVID-19 African American and White patients. Demographics and comorbidities are compared before and after propensity matching of cohorts.

<table>
<thead>
<tr>
<th></th>
<th>Before Propensity Score Matching</th>
<th>After Propensity Score Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>African Americans (n or %)</td>
<td>White (n or %)</td>
</tr>
<tr>
<td>No. of patients</td>
<td>2,621</td>
<td>3,435</td>
</tr>
<tr>
<td>Age at Index</td>
<td>54.1±16.8</td>
<td>50.9±18.7</td>
</tr>
<tr>
<td>Female</td>
<td>1,614 (61.58%)</td>
<td>2,057 (59.88%)</td>
</tr>
<tr>
<td>Male</td>
<td>1,007 (38.42%)</td>
<td>1,376 (40.06%)</td>
</tr>
<tr>
<td>Basal metabolic index (BMI)</td>
<td>33.9 ± 9.29</td>
<td>29.1 ± 7.5</td>
</tr>
<tr>
<td><strong>Comorbid Conditions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Hypertension</td>
<td>1,365 (52.08%)</td>
<td>1,159 (33.74%)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>779 (29.72%)</td>
<td>562 (16.36%)</td>
</tr>
<tr>
<td>Chronic lower respiratory diseases (Asthma and COPD)</td>
<td>596 (22.74%)</td>
<td>915 (26.64%)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>425 (16.25%)</td>
<td>290 (8.44%)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>305 (11.64%)</td>
<td>277 (8.06%)</td>
</tr>
<tr>
<td>Ischemic heart diseases</td>
<td>304 (11.60%)</td>
<td>434 (12.63%)</td>
</tr>
<tr>
<td>Cerebrovascular diseases</td>
<td>242 (9.23%)</td>
<td>344 (10.01%)</td>
</tr>
<tr>
<td>Nicotine dependence</td>
<td>183 (6.98%)</td>
<td>616 (17.93%)</td>
</tr>
<tr>
<td>Alcohol related disorders</td>
<td>61 (2.33%)</td>
<td>352 (10.25%)</td>
</tr>
<tr>
<td><strong>Signs and symptoms</strong></td>
<td><strong>African Americans</strong> Patients (% of the cohort)</td>
<td><strong>White</strong> Patients (% of the cohort)</td>
</tr>
<tr>
<td>Dyspnea</td>
<td>859 (32.72%)</td>
<td>1,061 (30.84%)</td>
</tr>
<tr>
<td>Cough</td>
<td>779 (29.68%)</td>
<td>1,221 (35.49%)</td>
</tr>
<tr>
<td>Fever</td>
<td>490 (18.66%)</td>
<td>1,037 (30.14%)</td>
</tr>
<tr>
<td>Condition</td>
<td>Group A</td>
<td>Group B</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Hypoxemia</td>
<td>285 (10.85%)</td>
<td>228 (6.62%)</td>
</tr>
<tr>
<td>Chest pain</td>
<td>211 (8.038%)</td>
<td>369 (10.72%)</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>116 (4.42%)</td>
<td>245 (7.12%)</td>
</tr>
<tr>
<td>Nausea and vomiting</td>
<td>98 (3.73%)</td>
<td>315 (9.16%)</td>
</tr>
</tbody>
</table>
Table 2: Outcomes among COVID-19 African American and White patients. Outcomes are compared before and after propensity score matching of cohorts.

<table>
<thead>
<tr>
<th></th>
<th>Before Propensity Score Matching</th>
<th>After Propensity Score Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>African Americans</td>
<td>White</td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>212/2,621 (8.08%)</td>
<td>145/3,435 (4.22%)</td>
</tr>
<tr>
<td>Age ≤ 49</td>
<td>22/1,028 (2.14%)</td>
<td>16/1,674 (0.96%)</td>
</tr>
<tr>
<td>Age 50-64</td>
<td>58/838 (6.92%)</td>
<td>26/900 (2.89%)</td>
</tr>
<tr>
<td>Age ≥ 65</td>
<td>132/755 (17.48%)</td>
<td>103/862 (11.94%)</td>
</tr>
<tr>
<td>Mechanical Ventilation</td>
<td>179/2,621 (6.82%)</td>
<td>78/3,435 (2.27%)</td>
</tr>
<tr>
<td>Critical Care Services</td>
<td>419/2,621 (15.98%)</td>
<td>179/3,435 (5.21%)</td>
</tr>
<tr>
<td>Hospitalization</td>
<td>1,023/2,621 (39.03%)</td>
<td>565/3,435 (16.45%)</td>
</tr>
<tr>
<td>Acute Respiratory Distress Syndrome</td>
<td>236/2,621 (9.01%)</td>
<td>108/3,435 (3.14%)</td>
</tr>
<tr>
<td>Acute Kidney Injury</td>
<td>361/2,621 (13.73%)</td>
<td>167/3,435 (4.86%)</td>
</tr>
<tr>
<td>Condition</td>
<td>Incidence</td>
<td>Reference Incidence</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Septic Shock</td>
<td>82/2,621 (3.12%)</td>
<td>52/3,435 (1.51%)</td>
</tr>
<tr>
<td>Continuous Renal Replacement Therapy</td>
<td>50/2,621 (1.91%)</td>
<td>21/3,435 (0.61%)</td>
</tr>
</tbody>
</table>
Figure 1: Flow chart showing the selection of African American and White COVID-19 patient cohorts.

Research Network
Patients: 49,333,623
HCOs: 34

Excluded

Age < 10 years
4,303,869

Patients: 45,029,754
HCOs: 34

MUST Have
B34.2: Coronavirus Infection, unspecified
OR
J12.81: Pneumonia due to SARS-associated coronavirus
OR
U07.1: 2019-nCoV acute respiratory disease (WHO)
OR
B97.29: Other coronavirus as the cause of diseases classified elsewhere

Excluded

Patients: 7,082
HCOs: 18

CANNOT Have
079.89: Other specified viral infection
OR
Z03.818: Encounter for observation for suspected exposure to other biological agents ruled out

Unknown: 739
Asian: 164
American Indian or Alaska Native: 24
Hawaiian or Other Pacific Islander: 14
Hispanic Whites: 85

Non-Hispanic Whites
3,435 (47%)
African Americans
2,621 (37%)

Northeast: 1012 (14%), Midwest 1803 (25%), South 2967 (42%), West 1300 (18%)
Figure 2: Propensity Score Density Graph before and after propensity score matching of the African-American and White cohort of patients.
Figure 3: Comparison of Outcomes among African American and White patients after propensity score matching of cohorts. Mortality is compared in the overall cohort and subgroups based on age. Secondary outcomes described in the overall cohort of patients.