Research Paper

# COVID-19 risk, disparities and outcomes in patients with chronic liver disease in the United States

QuanQiu Wang[a], Pamela B. Davis[b], Rong Xu[a,*]

[a] Center for Artificial Intelligence in Drug Discovery, School of Medicine, Case Western Reserve University, 2103 Cornell Rd, Cleveland 44106, OH, USA
[b] Center for Clinical Investigation, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

ARTICLE INFO

ABSTRACT

*Background:* Scientific evidence is lacking regarding the risk of patients with chronic liver disease (CLD) for COVID-19, and how these risks are affected by age, gender and race.

*Methods:* We performed a case-control study of electronic health records of 62.2 million patients (age >18 years) in the US up to October 1st, 2020, including 1,034,270 patients with CLD, 16,530 with COVID-19, and 820 with both COVID-19 and CLD. We assessed the risk, disparities, and outcomes of COVID-19 in patients with six major CLDs.

*Findings:* Patients with a recent medical encounter for CLD were at significantly increased risk for COVID-19 compared with patients without CLD, with the strongest effect in patients with chronic non-alcoholic liver disease [adjusted odd ratio (AOR)=13.11, 95% CI: 12.49−13.76, $p < 0.001$] and non-alcoholic cirrhosis (AOR=11.53, 95% CI: 10.69−12.43, $p < 0.001$), followed by chronic hepatitis C (AOR=8.93, 95% CI:8.25−9.66, $p < 0.001$), alcoholic liver damage (AOR=7.05, 95% CI:6.30−7.88, $p < 0.001$), alcoholic liver cirrhosis (AOR=7.00, 95% CI:6.15−7.97, $p < 0.001$) and chronic hepatitis B (AOR=4.37, 95% CI:3.35−5.69, $p < 0.001$). African Americans with CLD were twice more likely to develop COVID-19 than Caucasians. Patients with COVID-19 and a recent encounter for CLD had a death rate of 10.3% (vs. 5.5% among COVID-19 patients without CLD, $p < 0.001$) and a hospitalization rate of 41.0% (vs. 23.9% among COVID-19 patients without CLD, $p < 0.001$).

*Interpretation:* Patients with CLD, especially African Americans, were at increased risk for COVID-19, highlighting the need to protect these patients from exposure to virus infection.

## 1. Introduction

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and has rapidly escalated into a global pandemic [1,2].

Severe illness of COVID-19 predominantly occurs in older people and in individuals with underlying medical comorbidities, including asthma, cardiovascular diseases, diabetes, hypertension, chronic lung disease, chronic kidney disease, cancers, obesity, substance use disorders and mental disorders [3−7]. Currently, there is limited information about whether people with a chronic liver disease (CLD), including liver cirrhosis, hepatitis B, hepatitis C, and chronic non-alcoholic liver diseases are at increased risk for getting COVID-19 or having severe COVID-19 [8−11].

SARS-CoV-2 virus can damage liver and patients with COVID-19 and pre-existing chronic liver disease often had high mortality rates [12−17]. However, it remains unknown whether patients with a CLD were at increased risk for SARS-CoV-2 virus infection and how these risks were further affected by age, gender and race. Patients with CLD, including liver cirrhosis, chronic hepatitis B, chronic hepatitis C, and chronic non-alcoholic liver diseases often have multiple comorbid conditions, including obesity, type 2 diabetes, hypertension, cardiovascular diseases, cancers, infections and malnutrition [18−22], many of which overlap with known risk factors for severe COVID-19 illness [3−7]. In addition, abnormalities of immune function, immunodeficiency and systemic inflammation are often present in individuals with CLD, including liver cirrhosis and chronic hepatitis [23−25]. Because of the overlaps between known COVID-19 risk factors and risk factors for CLD and compromised immune functions in patients with CLD, we hypothesized that patients with a CLD were at increased risk for morbidity and mortality from COVID-19. The large database available to us permitted us also to test the impact of gender, race, and age on the association of CLD with COVID-19 in the US.

### Research in context

*Evidence before this study*

People who have serious underlying medical conditions, including chronic liver disease (CLD), might be at higher risk for severe illness from COVID-19. Evidence showed that SARS-CoV-2 virus damages the liver in infected patients and that pre-existing liver disease was associated with increased mortality in patients with COVID-19. However, it remains unknown whether patients with CLD are at increased risk for getting COVID-19 compared with individuals without CLD and how the risk for COVID-19 is further affected by age, gender and race.

*Added value of this study*

This analysis of patient electronic health records (EHRs) quantifies for the first time the risks, racial disparities, and outcomes for COVID-19 in individuals with a CLD, including alcoholic cirrhosis, non-alcoholic cirrhosis, alcoholic liver damage, chronic hepatitis B, chronic hepatitis C, and chronic non-alcoholic liver disease.

*Implications of all the available evidence*

Based on this analysis, patients with a CLD, especially African Americans with CLD, were at increased risk for both COVID-19 and its adverse outcomes, highlighting the need to protect these patients from exposure to virus infection.

## 2. Methods

### 2.1. Database description

We performed a case-control study using de-identified patient electronic health record (EHR) data collected by the IBM Watson Health Explorys from 360 hospitals and 317,000 providers across 50 states from 1999 up to October 1st, 2020, representing 20% of US population [26]. The EHR data are de-identified according to the Health Insurance Portability and Accountability Act and the Health Information Technology for Economic and Clinical Health Act standards as described in an early study of this dataset as described in an early study using this database [27]. After the de-identification process, curation process normalizes the data through mapping key elements to widely-accepted standards [28]. Specifically, disease terms are coded using the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT), a global standard for health terms that provides the core general terminology for EHRs [29]. More than 160 studies have shown that with this large-scale and standardized EHR database and the cloud-based Explorys Cohort Discovery informatics tools, large case-control studies can be undertaken efficiently [30], We have recently analyzed Explorys EHR database to examine COVID-19 in patients with substance use disorders, mental disorders and cancers [4−7]. The EHR data are de-identified and aggregated (not patient-level) and institutional review board (IRB) review was exempt. Q.W and R.X had access to the EHR data through Explorys web-based informatics tools. The data was accessed from June to October 2020. The EHR database is updated on weekly basis. The final results for this study was based on data on October 1, 2020.

### 2.2. Study population

At the time of this study (October 1st, 2020), the study population consisted of 62,266,410 adult and senior patients (age >18 years old), including 1034,270 patients who had encounters with healthcare

systems for their diagnosis of chronic liver disease (CLD), 16,530 with COVID-19, and 820 with both COVID-19 and CLD. The status of COVID-19 was based on the concept "Coronavirus infection (disorder)" (Concept Code 186,747,009). Based on Centers for Disease Control and Prevention (CDC) and The American Association for the Study of Liver Diseases (AASLD) Expert Panel Consensus Statement, people with chronic liver disease (CLD), including hepatitis B, hepatitis C, alcohol-related liver disease, non-alcoholic fatty liver disease, and cirrhosis may be at increased risk for COVID-19 susceptibility and its adverse outcomes [8−10]. Six types of CLD that had sufficient sample sizes in the EHR database were examined in this study: alcoholic liver damage, alcoholic cirrhosis, non-alcoholic cirrhosis, chronic non-alcoholic liver disease, chronic hepatitis C, and chronic hepatitis B. The status of alcoholic cirrhosis was based on the diagnosis of "Alcoholic cirrhosis (disorder)" (SNOMED-CT code 420,054,005), non-alcoholic cirrhosis on the diagnosis of "Cirrhosis - non-alcoholic (disorder)" (code 266,468,003), chronic non-alcoholic liver disease on the diagnosis of "chronic nonalcoholic liver disease (disorder)" (c79720007), chronic hepatitis C on the diagnosis of "Chronic hepatitis C (disorder)" (code 128,302,006), chronic hepatitis B on the diagnosis of "Viral hepatitis type B (disorder)" (code 66,071,002), and alcoholic liver damage on the diagnosis of "Alcoholic liver damage (disorder)" (code 41,309,000).

Among six types of CLD we examined, alcoholic cirrhosis, non-alcoholic cirrhosis, chronic non-alcoholic liver disease, chronic hepatitis C and chronic hepatitis B are non-overlapping diagnosis codes based on the SNOMED-CT disease classification scheme. Alcoholic liver damage is a parent term of alcoholic cirrhosis and also includes alcoholic fatty liver, alcoholic hepatic failure and alcoholic hepatitis. Alcoholic cirrhosis is a major term under alcoholic liver damage and constituted of 72.6% of patients with alcoholic damage. Since there was insufficient samples size for other specific types of alcoholic liver damage, we used both the parent term "alcoholic liver damage" and the major child term "alcoholic cirrhosis" in order to deduce the risk of COVID-19 in patients with other types of alcoholic liver damage.

The outcome measures were COVID-19 diagnosis, rates of death, and hospitalization. The SNOMED-CT concept "Hospital admission (procedure)" (ID 32,485,007) was used to obtain hospital admission status from patient EHRs. Explorys regularly imports from the Social Security Death index for the "deceased" status.

The following analyses were performed: First, the odds of COVID-19 diagnosis in patients with CLD compared with patients without CLD were calculated, adjusted for age, gender, race, and known COVID-19 risk factors, including asthma, cardiovascular diseases, cancers, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease, substance use disorders and mental disorders, viral treatments, transplantation procedures and nursing home stay [3−7]. The status of viral treatments was based on the Pharmacology Class code "Antiviral agent", transplantation on the Procedure code "Transplantation of bone marrow" and "Solid organ transplant", and nursing home stay on the Encounter code "Skilled Nursing Facility (encounter)".

The status of other COVID-19 risk factors were based on SNOMED-CT disease diagnosis codes. The exposure group consisted of patients who had any medical encounter for their diagnosis of CLD. The unexposed group consisted of patients who had a medical encounter with healthcare systems but had no diagnosis of CLD. The outcome measure was COVID-19 diagnosis. Separate analysis was done for six types of CLD (alcoholic liver cirrhosis, non-alcoholic liver cirrhosis, chronic non-alcoholic liver disease, chronic hepatitis C, chronic hepatitis B, and alcoholic liver damage). Separate analysis was done for patients with any CLD encounter (encounter for their CLD happened within past year or at any time prior) and patients with recent CLD encounter (at least one encounter for their CLD within past year). In the EHR database, three encounter cutoffs are available: Ever, Last 3 years, Last year. In our study, we used "Ever" and "Last year" encounters for CLD. Patients with an encounter in the last year are a subset of those who have ever had an encounter for that diagnosis.

**Table 1**

Patient characteristics. Number of cases and percentage (%) are shown. There are 15 race categories in the database and only the largest race categories were shown in the table. CLD − chronic liver disease, which in our study included alcoholic cirrhosis, non-alcoholic cirrhosis, alcoholic liver damage, chronic hepatitis B, chronic hepatitis C, and chronic non-alcoholic liver disease.

| Patient | Study population | CLD (any encounter) | CLD (recent encounter) | COVID-19 | COVID-19 + CLD (any encounter) | COVID-19 + CLD (recent encounter) |
|---|---|---|---|---|---|---|
| Total | 62,266,410 | 1034,270 | 81,360 | 16,530 | 820 | 390 |
| Gender | | | | | | |
| Female | 33,932,090 (54%) | 497,790 (48%) | 38,170 (47%) | 9840 (60%) | 450 (55%) | 200 (51%) |
| Male | 27,972,660 (45%) | 535,410 (52%) | 42,700 (52%) | 6680 (40%) | 370 (45%) | 190 (49%) |
| Unknown | 378,360 (<1%) | 1140 (<1%) | 560 (<1%) | 30 (<1%) | 0 (0%) | 0 (0%) |
| Age | | | | | | |
| Adult (18 - 65) | 44,208,280 (71%) | 666,710 (64%) | 58,290 (72%) | 12,340 (75%) | 580 (71%) | 280 (72%) |
| Senior (> 65) | 18,101,280 (29%) | 369,340 (36%) | 23,210 (28%) | 4230 (25%) | 240 (29%) | 110 (28%) |
| Race | | | | | | |
| Caucasian | 35,255,240 (57%) | 771,770 (75%) | 58,700 (72%) | 8370 (51%) | 460 (56%) | 210 (54%) |
| African American | 6438,760 (10%) | 130,770 (13%) | 14,670 (18%) | 6610 (40%) | 330 (40%) | 160 (41%) |
| Asian | 1012,900 (2%) | 25,240 (2%) | 2120 (3%) | 160 (1%) | 0 (0%) | 0 (0%) |
| Hispanic/Latino | 860,490 (1%) | 15,890 (2%) | 570 (1%) | 20 (<1%) | 0 (0%) | 0 (0%) |
| Unknown | 7984,920 (13%) | 139,940 (14%) | 6040 (7%) | 900 (5%) | 60 (7%) | 30 (8%) |

Second, we examined how age, gender and race further differentially affected COVID-19 risk in patients with CLD. The case groups were patients with CLD and one of the following demographic factors: Female, Senior (age >65), African American. Although other racial and ethnic groups are specified in the database, none had sufficient numbers of patients with CLD or COVID for this analysis. The comparison groups were also patients with chronic liver diseases but with one of the following corresponding demographic factors (Male, Adult (age 18−65), Caucasian). The outcome measure was COVID-19 diagnosis. For example, we examined whether African Americans with CLD were more likely to get COVID-19 compared with Caucasians with CLD, adjusting for age, gender and known COVID-19 risk factors. Similarly we examine whether women with CLD were more likely to get COVID-19 compared with men with CLD, adjusting for age, race, and known COVID-19 risk factors; Third, the 8-month (February−October, 2020) death rate and hospitalization rate in patients with COVID-19 and recent encounter for CLD were compared with those for patients with COVID-19 but no CLD, and patients with recent encounter for CLD but no COVID-19.

## 2.3. Statistical analysis

The EHR data are de-identified population-level (not patient-level) data, therefore we used odds ratios instead of performing true regression analyses, as was done in previous studies using Explorys EHR database [4−6, 24−25] [4−6,24,25]. For a given input set of patient characteristics (e.g., age, gender, race, diagnosis, comorbidities), the Explorys Explore Cohort Discovery tool built a patient cohort by querying the EHR database for patients matching the inputs. Patients with missing values for the input queries were not included in the returned cohorts. The adjusted odds ratio (AOR), 95% CI and $P$-values were calculated using the Cochran-Mantel-Haenszel (CMH) method [31], by controlling for age groups in 5-year categories (e.g., 20−24, 25−29, ...65−69, 70−74, ..., 85−89), gender (Female, Male), race (Caucasian, African American) and risk factors for COVID-19 [3−7]. Two-sided, 2-sample test for equality of proportions with continuity correction were used to compare death and hospitalization rates. Multiple comparisons were corrected by Bonferroni correction. Statistical tests were conducted with significance set at $P$-value < 0.05 (two-sided). All analyses were done using R, version 3.6.3.

## 3. Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, and writing of the report. The corresponding author (R.X) had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## 4. Results

### 4.1. Patient characteristics

The baseline characteristics of the study population (as of October 1st, 2020) are presented in Table 1. Among 62,266,410 adult patients (age >18 years old), 1034,270 patients have encountered with healthcare systems for their diagnosis of CLD in the past year or prior ("any encounter"), among whom 81,360 had encounter within the past year ("recent encounter").
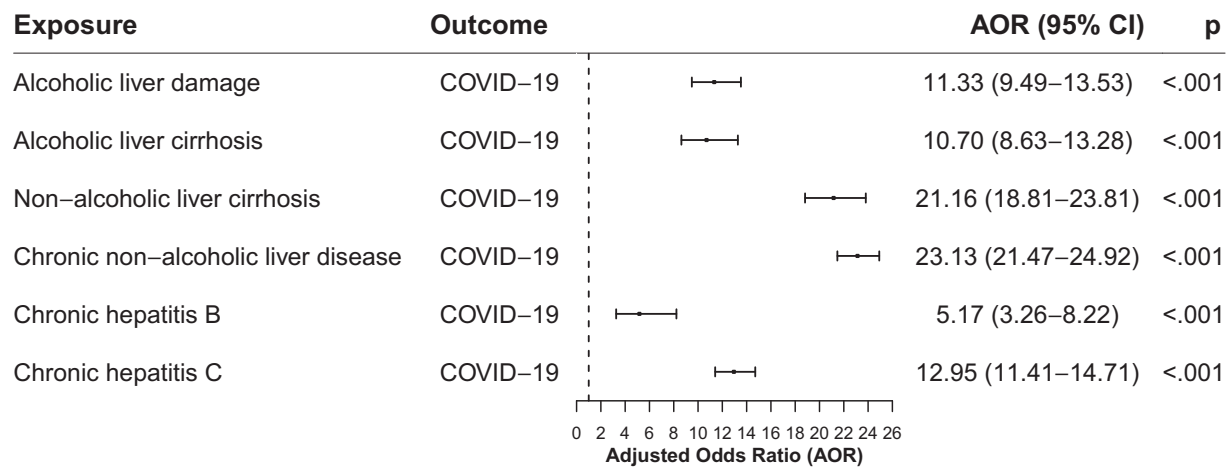
The specifics for all and recent encounters were as follows: any 108,760, recent 9850 for alcoholic cirrhosis; any 213,170, recent 16,300 for non-alcoholic cirrhosis; any 576,770, recent 39,300 for chronic non-alcoholic liver disease; any 242,260, recent 20,410 for chronic hepatitis C; any 41,860, recent 4490 for chronic hepatitis B; any 173,850, recent 13,570 for alcoholic liver damage.

Among 16,530 adult COVID-19 patients in the database, 820 had some encounter with healthcare systems for their CLD (past year or prior, but prior to their COVID-19 encounter), and 390 had a recent encounter (past year, but prior to their COVID-19 encounter). Any and recent encounter for specific disorders in the COVID-19 population: alcoholic cirrhosis (any: 50; recent: 30), alcoholic cirrhosis (any: 50; recent: 30), non-alcoholic cirrhosis (any: 150; recent: 90 chronic non-alcoholic liver disease (any: 570; recent: 240), chronic hepatitis C (any: 180; recent: 80), chronic hepatitis B (any: 30; recent: 10), and alcoholic liver damage (any: 70; recent: 40).

## 5. COVID-19 risk in patients with CLD

Patients with a recent encounter for CLD had significantly higher odds of acquiring COVID-19 compared with patients without recent encounter for CLD, after adjusting for age, gender and race, with the

## Odds of COVID−19 in patients with recent encounter for CLD
## (adjusted for demographics only)

| Exposure | Outcome | | AOR (95% CI) | p |
|---|---|---|---|---|
| Alcoholic liver damage | COVID−19 | | 11.33 (9.49−13.53) | <.001 |
| Alcoholic liver cirrhosis | COVID−19 | | 10.70 (8.63−13.28) | <.001 |
| Non−alcoholic liver cirrhosis | COVID−19 | | 21.16 (18.81−23.81) | <.001 |
| Chronic non−alcoholic liver disease | COVID−19 | | 23.13 (21.47−24.92) | <.001 |
| Chronic hepatitis B | COVID−19 | | 5.17 (3.26−8.22) | <.001 |
| Chronic hepatitis C | COVID−19 | | 12.95 (11.41−14.71) | <.001 |

0 2 4 6 8 10 12 14 16 18 20 22 24 26
**Adjusted Odds Ratio (AOR)**

## Odds of COVID−19 in patients with any encounter for CLD
## (adjusted for demographics only)

| Exposure | Outcome | | AOR (95% CI) | p |
|---|---|---|---|---|
| Alcoholic liver damage | COVID−19 | | 1.64 (1.44−1.87) | <.001 |
| Alcoholic liver cirrhosis | COVID−19 | | 1.68 (1.43−1.99) | <.001 |
| Non−alcoholic liver cirrhosis | COVID−19 | | 2.80 (2.55−3.06) | <.001 |
| Chronic non−alcoholic liver disease | COVID−19 | | 3.92 (3.74−4.12) | <.001 |
| Chronic hepatitis B | COVID−19 | | 1.70 (1.31−2.21) | <.001 |
| Chronic hepatitis C | COVID−19 | | 2.59 (2.38−2.82) | <.001 |

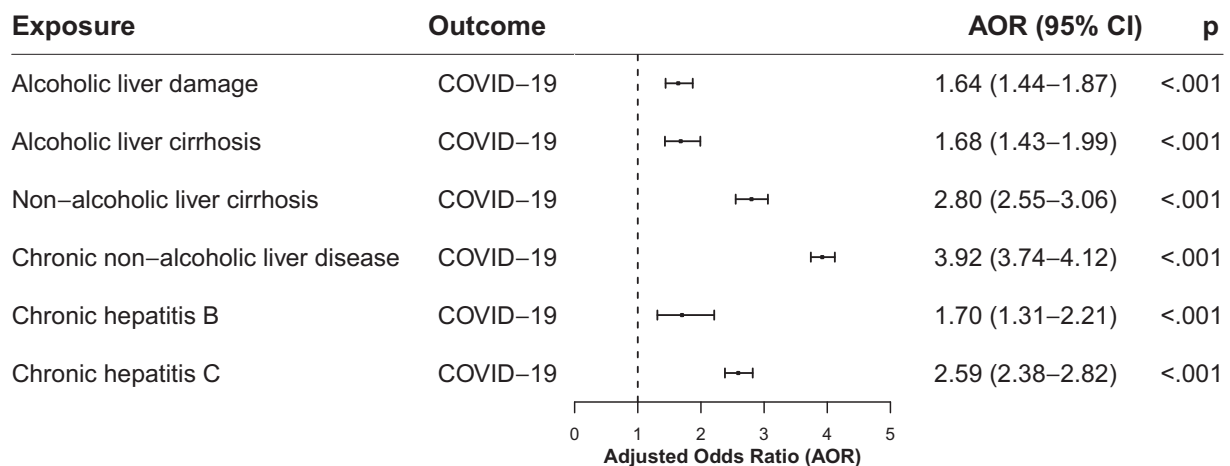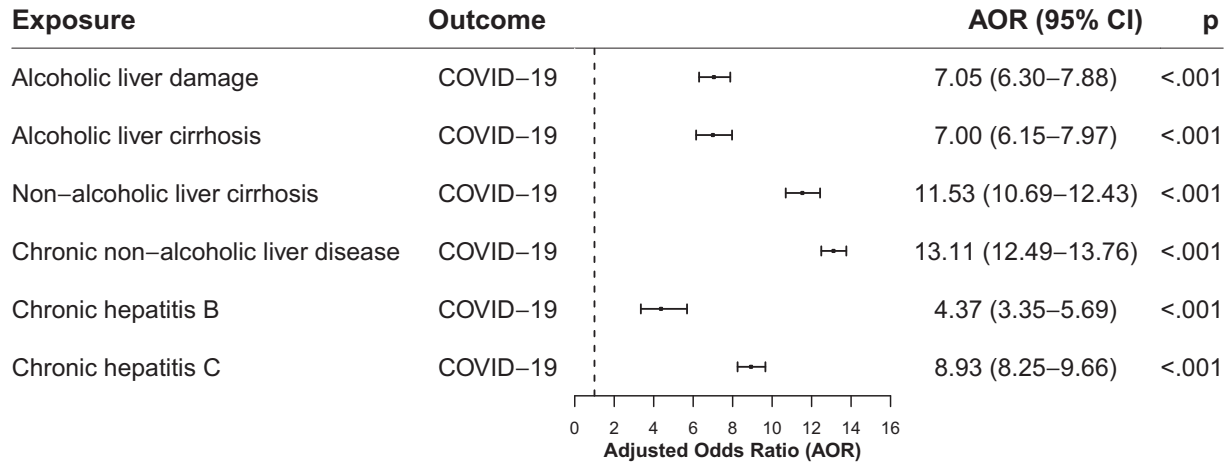0 1 2 3 4 5
**Adjusted Odds Ratio (AOR)**

**Fig. 1.** COVID-19 risk in patients with CLD (recent vs any encounter for CLD) after adjusting for age, gender and race (but not for other known COVID-19 risk factors). Recent CLD encounter − patients have encountered with healthcare systems for their CLD within the past year. Any CLD encounter − patients have encountered with healthcare systems for their CLD at any time in the past (including the last year).

strongest effect for patients with chronic non-alcoholic liver disease (AOR=23.13, 95% CI: 21.47−24.92, $p < 0.001$) and non-alcoholic cirrhosis (AOR=21.16, 95% CI: 18.81−23.81, $p < 0.001$), followed by patients with chronic hepatitis C (AOR=12.95, 95% CI: 11.41−14.71, $p < 0.001$), alcoholic liver damage (AOR=11.33, 95% CI: 9.49−13.53, $p < 0.001$), alcoholic liver cirrhosis (AOR=10.70, 95% CI: 8.63−13.28, $p < 0.001$), and chronic hepatitis B (AOR=5.17, 95% CI: 3.26−8.22, $p < 0.001$) (Fig. 1, top). The trend was similar for patients with any encounter ever for CLD, but the risk associations were lower (Fig. 1, bottom). The parent term "Alcoholic liver damage" includes the child term "alcoholic liver cirrhosis" and other child terms such as alcoholic fatty liver, alcoholic hepatic failure and alcoholic hepatitis. The 13,570 patients with recent encounter for alcoholic liver damage included 9850 individuals with recent encounter for alcoholic liver cirrhosis (72.6%). The risk for COVID-19 in patients with alcoholic liver cirrhosis was similar to that for patients with alcoholic liver damage, suggesting that patients with other alcoholic liver damage including

alcoholic fatty liver, alcoholic hepatitis, alcoholic hepatic failure had similar risk for COVID-19 as patients with alcoholic liver cirrhosis. However, due to the limited sample sizes for other specific types of alcoholic liver damage, we could not directly assess the risk of COVID-19 for each of the alcoholic liver damage types. Instead, we used the parent term alcoholic liver damage and the child term alcoholic liver cirrhosis to infer their COVID-19 risks.

Patients with non-alcoholic liver cirrhosis had similar risk for COVID-19 as patients with chronic non-alcoholic liver diseases (AOR=21.16 vs 23.13). Based on SNOMED-CT disease classification scheme, non-alcoholic liver cirrhosis and non-alcoholic liver diseases are two distinct non-overlapping disease diagnosis. Among 576,770 patients with a diagnosis code of chronic non-alcoholic liver disease, 8.7% also had a diagnosis code of non-alcoholic liver cirrhosis. Among 213,170 patients with a diagnosis code of non-alcoholic liver cirrhosis, 23.6% had a diagnosis code of chronic non-alcoholic liver disease.

## Odds of COVID−19 in patients with recent encounter for CLD
## (adjusted for demographics and known COVID−19 risk factors)

| Exposure | Outcome | | AOR (95% CI) | p |
|---|---|---|---|---|
| Alcoholic liver damage | COVID−19 | | 7.05 (6.30−7.88) | <.001 |
| Alcoholic liver cirrhosis | COVID−19 | | 7.00 (6.15−7.97) | <.001 |
| Non−alcoholic liver cirrhosis | COVID−19 | | 11.53 (10.69−12.43) | <.001 |
| Chronic non−alcoholic liver disease | COVID−19 | | 13.11 (12.49−13.76) | <.001 |
| Chronic hepatitis B | COVID−19 | | 4.37 (3.35−5.69) | <.001 |
| Chronic hepatitis C | COVID−19 | | 8.93 (8.25−9.66) | <.001 |

Adjusted Odds Ratio (AOR)

## Odds of COVID−19 in patients with any encounter for CLD
## (adjusted for demographics and known COVID−19 risk factors)

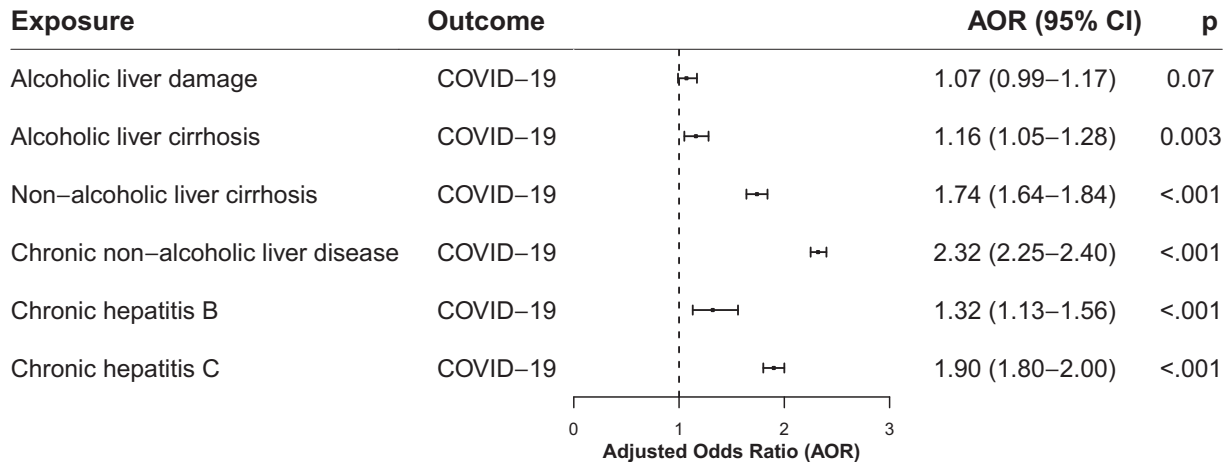| Exposure | Outcome | | AOR (95% CI) | p |
|---|---|---|---|---|
| Alcoholic liver damage | COVID−19 | | 1.07 (0.99−1.17) | 0.07 |
| Alcoholic liver cirrhosis | COVID−19 | | 1.16 (1.05−1.28) | 0.003 |
| Non−alcoholic liver cirrhosis | COVID−19 | | 1.74 (1.64−1.84) | <.001 |
| Chronic non−alcoholic liver disease | COVID−19 | | 2.32 (2.25−2.40) | <.001 |
| Chronic hepatitis B | COVID−19 | | 1.32 (1.13−1.56) | <.001 |
| Chronic hepatitis C | COVID−19 | | 1.90 (1.80−2.00) | <.001 |

Adjusted Odds Ratio (AOR)

**Fig. 2.** COVID-19 risk in patients with CLD (recent vs any CLD encounter) after adjusting for age, gender and race and known COVID-19 risk factors. Recent CLD encounter - patients have encountered the healthcare systems for CLD within the past year. Any CLD encounter − patients have encountered the healthcare system for CLD at any time in the past (including the last year).

After adjusting for known COVID-19 risk factors, in addition to age, gender and race, the odds of COVID-19 in patients with recent or any encounter for CLD decreased, but remained highly significant (Fig. 2). The strongest effects were observed in patients with recent encounter for chronic non-alcoholic liver disease (AOR=13.11, 95% CI: 12.49−13.76, $p < 0.001$), non-alcoholic cirrhosis (AOR=11.53, 95% CI: 10.69−12.43, $p < 0.001$) followed by patients with chronic hepatitis C (AOR = 8.93, 95% CI: 8.25−9.66, $p < 0.001$), alcoholic liver damage (AOR=7.05, 95% CI: 6.30−7.88, $p < 0.001$), alcoholic liver cirrhosis (AOR=7.00, 95% CI: 6.15−7.97, $p < 0.001$), and chronic hepatitis B (AOR=4.37, 95% CI: 3.35−5.69, $p < 0.001$).

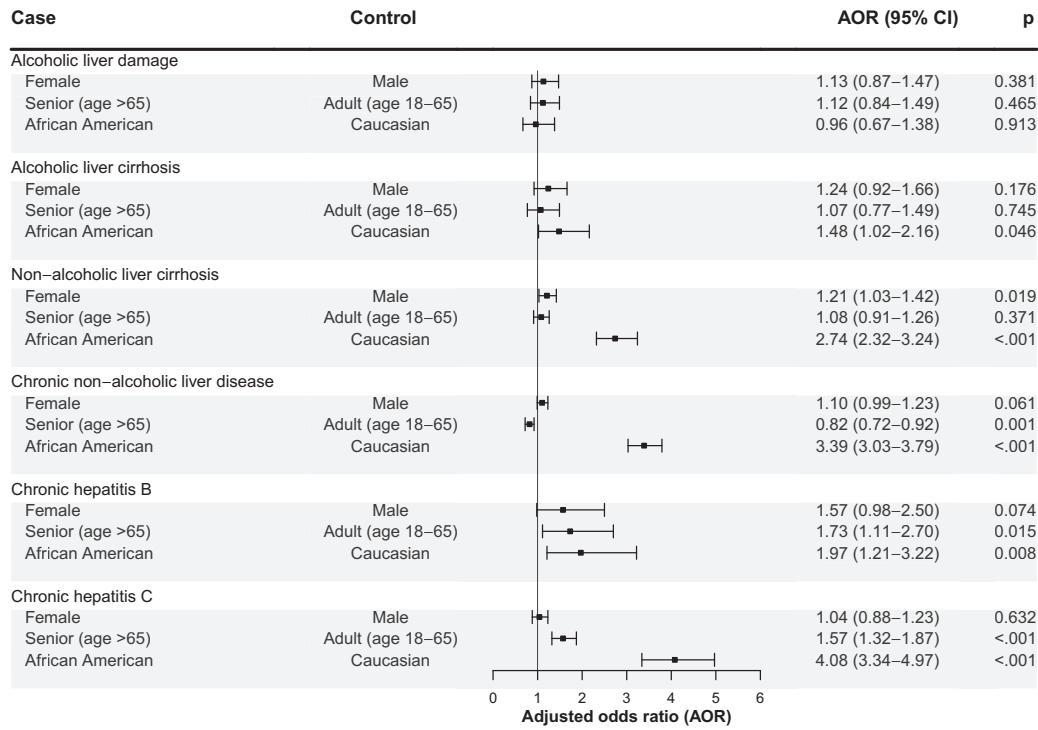### 5.1. Demographic disparity of COVID-19 risk among patients with CLD

We examined how age, gender and race differentially affected COVID-19 risk among patients with CLD after adjusting for COVID-19 risk factors. African Americans with CLD were twice more likely to get COVID-19 than Caucasians with CLD after adjusting for age, gender, and COVID-19 risk factors: non-alcoholic liver cirrhosis (AOR=2.74 and 3.41 for recent and any encounter, respectively), chronic non-alcoholic liver disease (AOR=3.39 and 4.79 for recent and any encounter, respectively), chronic hepatitis B (AOR=1.97 and 2.08 for recent and any encounter, respectively), and chronic hepatitis C (AOR=4.08 and 4.42 for recent and any encounter, respectively). No marked age or gender disparities were observed: women with CLD had similar risk for COVID-19 as men with CLD. Older patients (age >65 years) with CLD had similar risk for COVID-19 as younger patients (age 18−65 years) with CLD (Fig. 3).

### 5.2. Hospitalization and death rates in patients with CLD and COVID-19

The overall hospitalization rate over 8-month period (from the start of the pandemic in February up to October 1, 2020) for 16,530 adult COVID-19 patients was 24.3%, higher for African Americans

## Demographic disparities of COVID−19 in patients with recent encounter for CLD

| Case | Control | | AOR (95% CI) | p |
|---|---|---|---|---|
| **Alcoholic liver damage** | | | | |
| Female | Male | | 1.13 (0.87–1.47) | 0.381 |
| Senior (age >65) | Adult (age 18–65) | | 1.12 (0.84–1.49) | 0.465 |
| African American | Caucasian | | 0.96 (0.67–1.38) | 0.913 |
| **Alcoholic liver cirrhosis** | | | | |
| Female | Male | | 1.24 (0.92–1.66) | 0.176 |
| Senior (age >65) | Adult (age 18–65) | | 1.07 (0.77–1.49) | 0.745 |
| African American | Caucasian | | 1.48 (1.02–2.16) | 0.046 |
| **Non−alcoholic liver cirrhosis** | | | | |
| Female | Male | | 1.21 (1.03–1.42) | 0.019 |
| Senior (age >65) | Adult (age 18–65) | | 1.08 (0.91–1.26) | 0.371 |
| African American | Caucasian | | 2.74 (2.32–3.24) | <.001 |
| **Chronic non−alcoholic liver disease** | | | | |
| Female | Male | | 1.10 (0.99–1.23) | 0.061 |
| Senior (age >65) | Adult (age 18–65) | | 0.82 (0.72–0.92) | 0.001 |
| African American | Caucasian | | 3.39 (3.03–3.79) | <.001 |
| **Chronic hepatitis B** | | | | |
| Female | Male | | 1.57 (0.98–2.50) | 0.074 |
| Senior (age >65) | Adult (age 18–65) | | 1.73 (1.11–2.70) | 0.015 |
| African American | Caucasian | | 1.97 (1.21–3.22) | 0.008 |
| **Chronic hepatitis C** | | | | |
| Female | Male | | 1.04 (0.88–1.23) | 0.632 |
| Senior (age >65) | Adult (age 18–65) | | 1.57 (1.32–1.87) | <.001 |
| African American | Caucasian | | 4.08 (3.34–4.97) | <.001 |

Adjusted odds ratio (AOR)   0 1 2 3 4 5 6

## Demographic disparities of COVID−19 in patients with any encounter for CLD

| Case | Control | | AOR (95% CI) | p |
|---|---|---|---|---|
| **Alcoholic liver damage** | | | | |
| Female | Male | | 1.36 (1.12–1.66) | 0.002 |
| Senior (age >65) | Adult (age 18–65) | | 0.79 (0.64–0.97) | 0.033 |
| African American | Caucasian | | 1.14 (0.89–1.47) | 0.324 |
| **Alcoholic liver cirrhosis** | | | | |
| Female | Male | | 1.35 (1.06–1.72) | 0.017 |
| Senior (age >65) | Adult (age 18–65) | | 0.87 (0.68–1.11) | 0.313 |
| African American | Caucasian | | 1.34 (0.98–1.83) | 0.076 |
| **Non−alcoholic liver cirrhosis** | | | | |
| Female | Male | | 1.20 (1.06–1.36) | <.001 |
| Senior (age >65) | Adult (age 18–65) | | 0.83 (0.74–0.95) | <.001 |
| African American | Caucasian | | 3.41 (3.00–3.88) | <.001 |
| **Chronic non−alcoholic liver disease** | | | | |
| Female | Male | | 1.18 (1.10–1.26) | <.001 |
| Senior (age >65) | Adult (age 18–65) | | 0.62 (0.57–0.67) | <.001 |
| African American | Caucasian | | 4.79 (4.47–5.14) | <.001 |
| **Chronic hepatitis B** | | | | |
| Female | Male | | 1.68 (1.15–2.46) | 0.008 |
| Senior (age >65) | Adult (age 18–65) | | 1.61 (1.11–2.34) | 0.013 |
| African American | Caucasian | | 2.08 (1.40–3.09) | <.001 |
| **Chronic hepatitis C** | | | | |
| Female | Male | | 1.27 (1.14–1.43) | <.001 |
| Senior (age >65) | Adult (age 18–65) | | 1.27 (1.13–1.43) | <.001 |
| African American | Caucasian | | 4.42 (3.91–5.00) | <.001 |

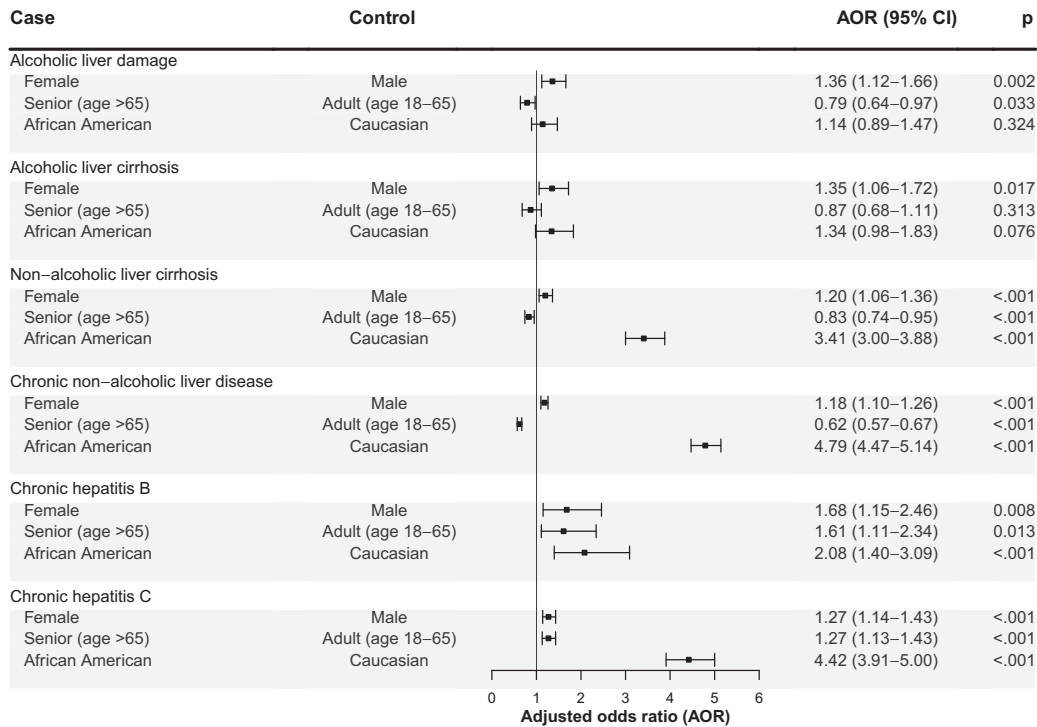Adjusted odds ratio (AOR)   0 1 2 3 4 5 6

**Fig. 3.** Demographic disparity of COVID-19 among patients with recent versus any encounter for CLD after adjusting for potential COVID-19 risk factors. Recent CLD encounter − patients have encountered the healthcare system for CLD within the past year. Any CLD encounter − patients have encountered the healthcare systems for CLD at any time in the past (including the last year).

## Hospitalization rates in patients with COVID-19 and recent CLD encounter
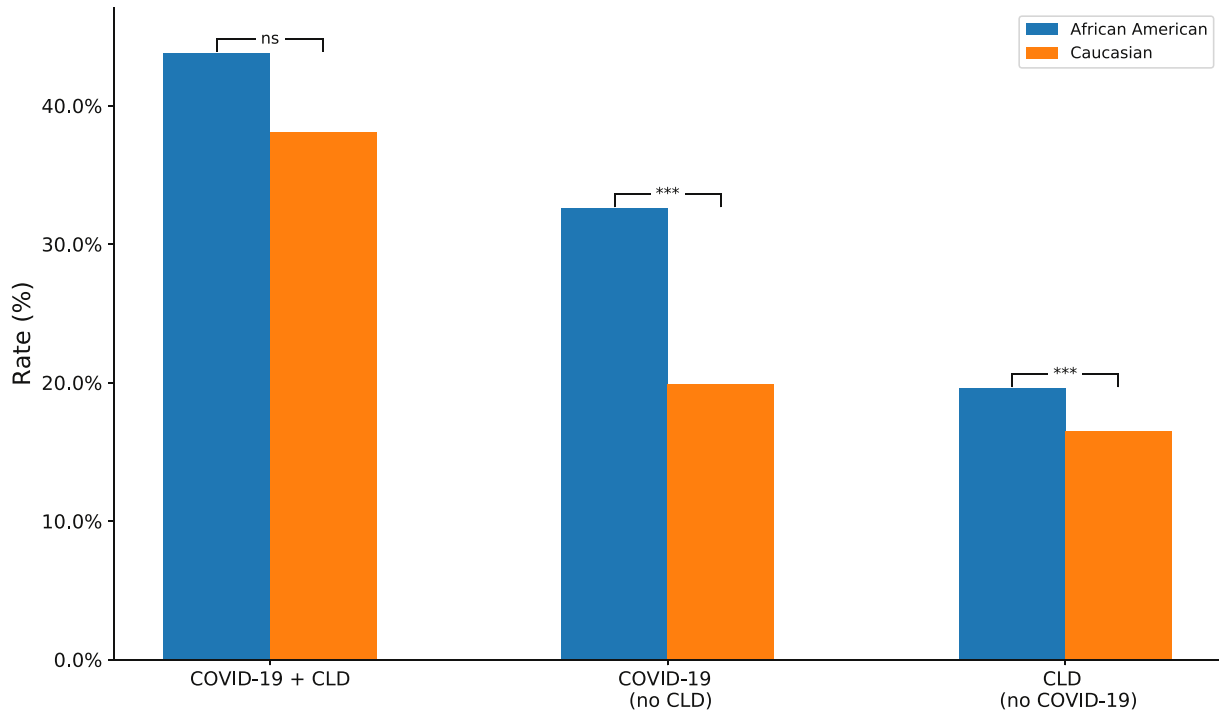


## Hospitalization rates in patients with COVID-19 and any CLD encounter
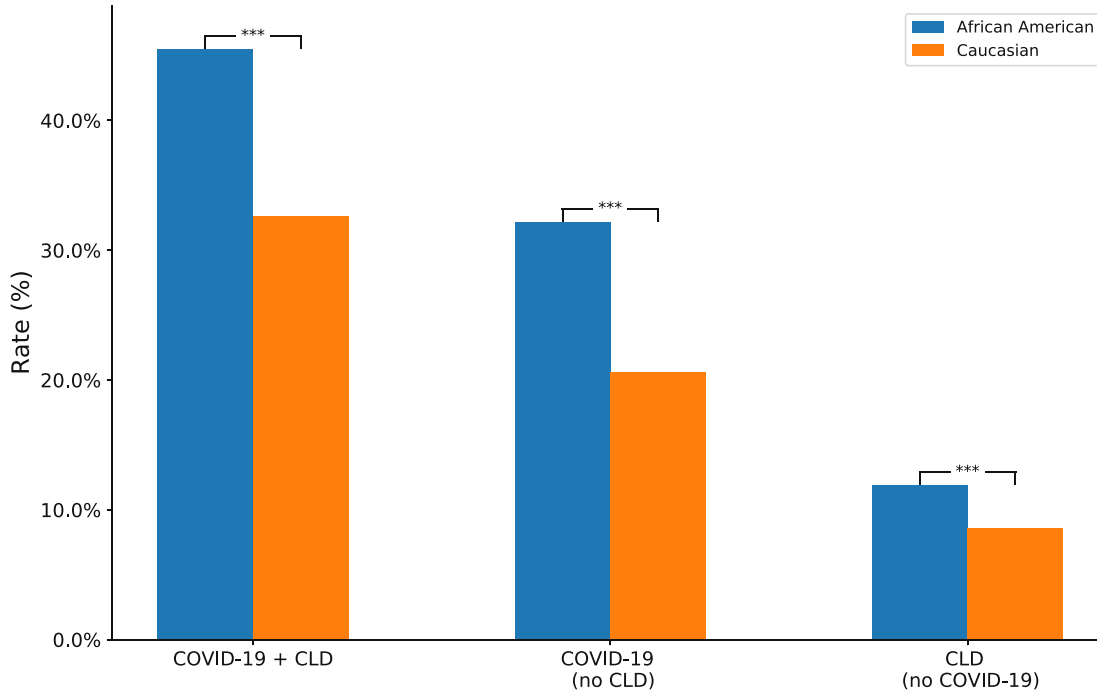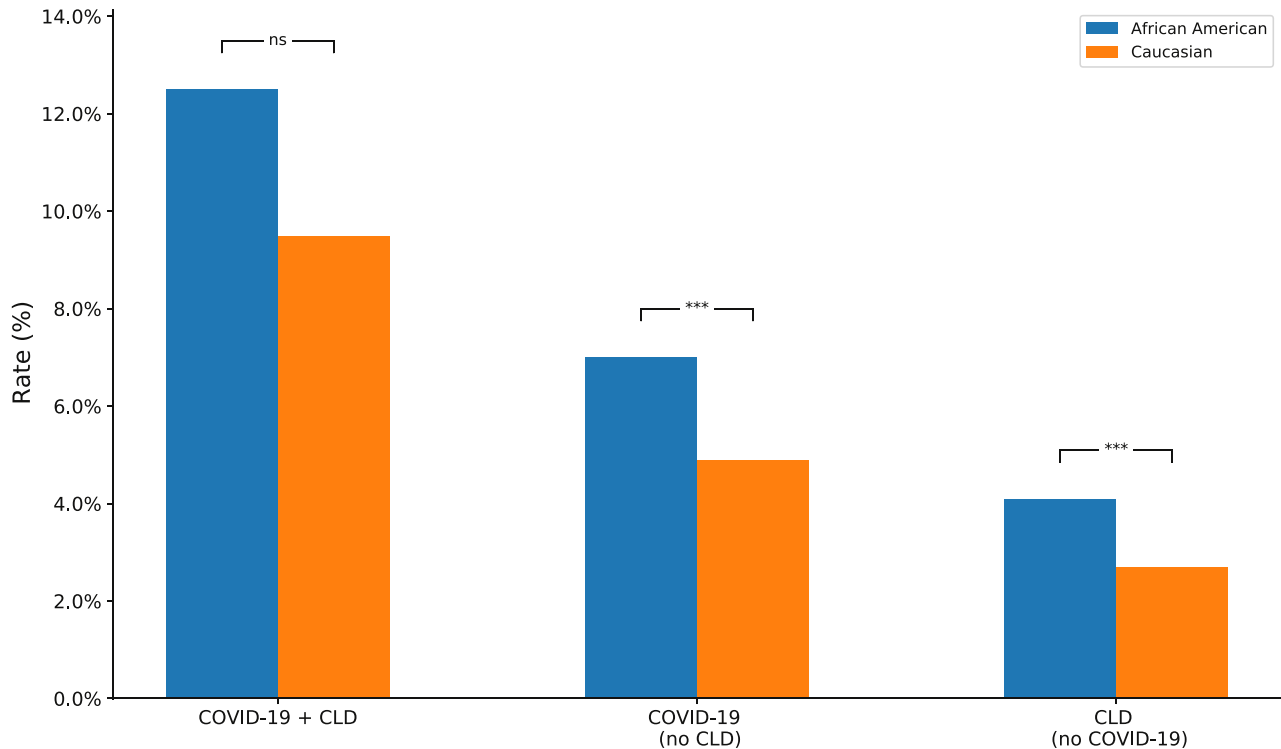


**Fig. 4.** Hospitalization rates among three adult and senior populations (age >18 years): COVID-19 patients with recent or all counter with healthcare systems for CLD, COVID-19 negative patients with recent or any encounter for CLD, and COVID-19 patients without recent or any encounter for CLD. ***: $p < 0.001$; ns: not significant.

(32.9%) than Caucasians (20.3%) ($p < 0.001$). Among 390 patients with COVID-19 and recent encounter for CLD, 160 were hospitalized (41.0%), similar for African Americans (43.8%) and Caucasians (38.1%) ($p$ = 0.321). Among 16,140 COVID-19 patients without CLD, 3850 were hospitalized (23.9%), higher for African Americans (32.6%) than

Caucasians (19.9%) ($p < 0.001$). The hospitalization rate for COVID-19 negative patients with recent encounter for CLD was 16.2%, higher for African Americans (19.6%) than Caucasians (16.5%) ($p < 0.001$) (Fig. 4). The trend is similar for patients with any encounter for CLD, except that among COVID-10 patients with all encounter for CLD,

## Death rates in patients with COVID-19 and recent CLD encounter



## Death rates in patients with COVID-19 and any CLD encounter
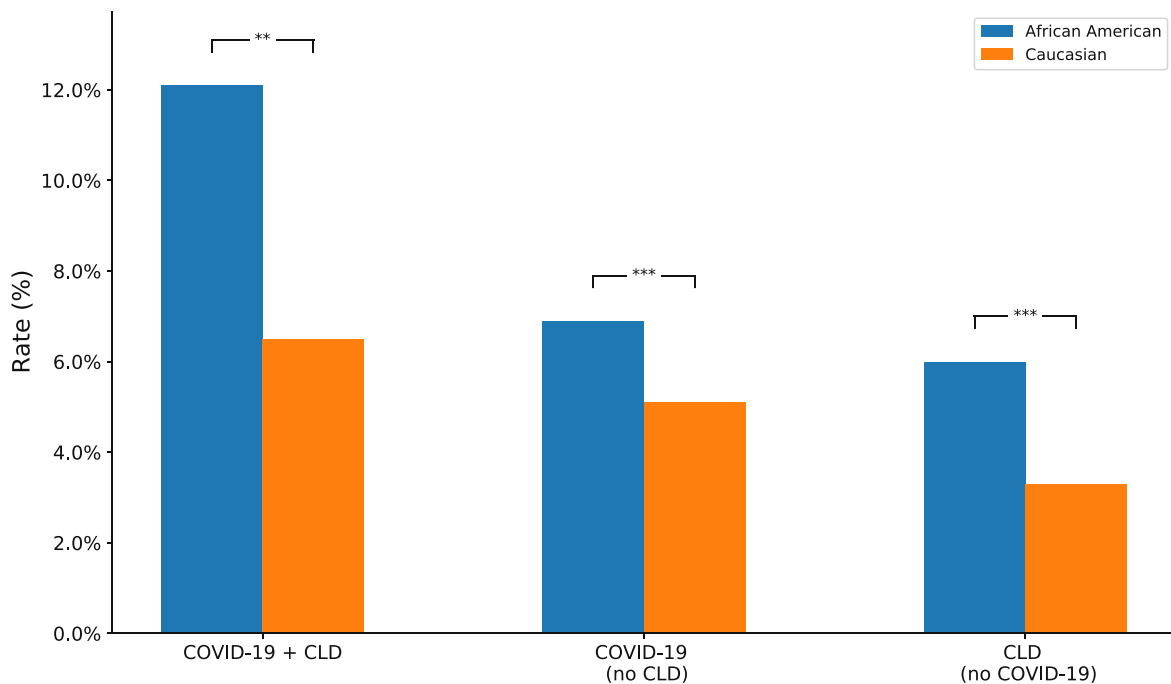


**Fig. 5.** Death rates among three adult populations: COVID-19 patients with recent or all counter with healthcare systems for CLD, COVID-19 negative patients with recent or all encounter for CLD, and COVID-19 patients without recent or all encounter for CLD. ***: $p < 0.001$; **: $p < 0.01$; ns: not significant.

African Americans had higher hospitalization rate than Caucasians (45.5% vs 32.6%, $p < 0.001$) (Fig. 4). Overall the hospitalization rate for COVID-19 patients CLD (41.0% and 39.2% for recent and any encounter, respectively) was higher than for COVID-19 patients without CLD (23.9% vs 23.5% for recent and all encounter, respectively)

($p < 0.001$) and that for COVID-19 negative patients with CLD (16.2% and 8.8% for recent and all encounter, respectively) ($p < 0.001$).

The overall death rate over 8-month period for 16,530 adult COVID-19 patients was 5.6%, higher for African Americans (7.1%) than Caucasians (5.0%) ($p < 0.001$). Among 390 patients with COVID-

19 and recent encounter for CLD, 40 died (10.3%), similar for African Americans (12.5%) and Caucasians (9.5%) (p = 0.457). Among 16,140 COVID-19 patients without CLD, 890 died (5.5%), higher for African Americans (7.0%) than Caucasians (4.9%) (p < 0.001). The death rate for COVID-19 negative patients with recent encounter for CLD was 2.9%, higher for African Americans (4.1%) than Caucasians (2.7%) (p < 0.001) (Fig. 5). The trend is similar for patients with all encounter for CLD, except that among COVID-10 patients with all encounter for CLD, African Americans had higher death rate than Caucasians (12.1% vs 6.5%, p < 0.01) (Fig. 5). Overall the death rate for COVID-19 patients CLD (10.3% and 8.5% for recent and all encounter, respectively) was higher than for COVID-19 patients without CLD (5.5% vs 5.5% for recent and all encounter, respectively) (p < 0.001) and that for COVID-19 negative patients with CLD (2.9% and 3.6% for recent and all encounter, respectively) (p < 0.001).

## 6. Discussion

Based on analyses of a nation-wide EHR database in the US we show that patients with CLD, especially those who had recent medical encounter for CLD, were at significantly increased risk for COVID-19 acquisition compared with patients without CLD. African Americans with CLD were twice more likely to get COVID-19 than Caucasians with CLD. COVID-19 patients with CLD had higher rates of hospitalization and death than COVID-19 negative patients with CLD and COVID-19 patients without CLD.

Our study shows that patients who had recent medical encounter for their CLD had significantly increased risk for COVID-19 compared with patients without recent medical encounter for CLD after demographics and known COVID-19 risk factors. Comparing the odds of COVID-19 in patients with recent medical encounter for CLD before and after adjusting known COVID-19 risk factors, it is clear these factors contributed to the high risk for COVID-19 in patients with CLD. For example, the odds for COVID-19 in patients with recent medical encounter for chronic non-alcoholic liver disease decreased from 23.13 to 13.11 after adjusting for COVID-19 risk factors. Yet, even after adjusting for these risk factors, patients with CLD still had high risk for COVID-19 compared with patients without CLD. Several reasons might have counted for this observed high risk for COVID-19: first, certain residual and unmeasured confounding factors (e.g., socio-economic determinants, behavioral factors, life-style) may have contributed to the increased risk for COVID-19 in patients with CLD. Second, CLD as a disease entity may have effects on patients' increased susceptibility to SARS-CoV2 infection because of abnormalities of immune function, immunodeficiency and systemic inflammation present in individuals with CLD [22−24]. However, due to the limited number of patients with CLD and COVID-19 in our database and due to limited information for socio-economic determinants, behavioral factors, and life-style factors available in the EHR database, we were unable to assess which and how these factors, alone and together, contributed to the increased risk for COVID-19 in patients with CLD.

In our study, patients who had a recent encounter (within the past year but before COVID-19 diagnosis) with healthcare systems for their CLD had higher odds of COVID-19 than those with all encounter (AOR=13.11 vs 2.32 for patients with recent vs all encounter for chronic non-alcoholic liver disease, respectively). Furthermore, we observed markedly different odds of COVID-19 for patients with different CLD, ranging from AOR of 4.37 for chronic hepatitis B, 8.93 for chronic hepatitis C, to 13.11 to chronic non-alcoholic liver disease. There is no obvious trend that patients with more severe CLD (e.g., chronic hepatitis B) had higher odds of COVID-19 infection than those with less severe CLD (e.g., chronic hepatitis C). Reasons for substantially different risks for COVID-19 in patients with recent vs ever CLD encounter and in patients with different CLD warrant further investigation.

An important finding of our study is that African Americans with CLD were more twice more likely to get COVID-19 than Caucasians with CLD,

after adjusting for age, gender and COVID-19 risk factors. This is consistent with data showing that COVID-19 affects African Americans at a disproportionately high rate [32]. This profound racial disparity suggests that factors other than medical conditions such as access to healthcare, socioeconomic status and other social adversity components may have contributed to the increased risk for COVID-19 among African Americans with CLD. However, due to limited socioeconomic information available in the patient EHRs we were unable to assess how these factors further contributed to the high risk for COVID-19 infection in African Americans with CLD compared with Caucasians with CLD. We were unable to examine how COVID-19 risks were affected by other races due to their limited sample sizes for patients with CLD and COVID-19 in the EHR database. There are more than 15 race categories in the EHR database for the 62 million adult and senior patients. For example, though Asians count for 2% of CLD patients, but 0% for patients with COVID-19 and CLD (Table 1). On the other hand, among 820 COVID-19 patients with CLD, 56% are Caucasians, 40% are African Americans. Previous studies showing that advanced age and being female are risk factors for COVID-19 infection where age-and gender-related medical conditions were not controlled [33,34], our study demonstrates that age and gender had no additional effect on the risk of developing COVID-19 among patients with CLD when potential COVID-19 risk factors such as cardiovascular diseases, obesity, diabetes, cancers, and mental disorder were controlled. Our study showed that women with CLD had similar risk for COVID-19 as men with CLD. Older patients (age >65 years) with CLD had similar risk for COVID-19 as younger patients (age 18−65 years) with CLD.

Our study has several limitations. First, this is an analysis of patient EHR data that were collected for clinical convenience and billing, not for research purposes. Though EHR data have been widely used for research purposes, they have inherent limitations including under-diagnosis, over-diagnosis, or mis-diagnosis, limited information on time-series, timing and adherence of medications and treatments, socio-economic and life style determinants, among others [35−38]. Second, the Explorys EHR database included 16,530 cases of adult and senior patients based on COVID-19 diagnosis code at the time of our study, which is substantially lower, proportionately to sample size, than the total confirmed cases in US reported by CDC [1]. COVID-19 is regularly tested at drive-ups and popup testing locations, many of which may have not been captured by EHRs. It is also likely that during the process of Explorys importing and standardizing EHR data from healthcare systems, its disease coding system did not capture all cases of COVID-19. Third, no detailed information was provided regarding rural vs urban, socioeconomic and geographic composition of the patient population as well as healthcare facility venue in the Explorys database. Fourth, we used the SNOMED-CT concept "Hospital admission (procedure)" (ID 32,485,007) to obtain hospital admission status from patient EHRs. Explorys imports information from the Social Security Death index for the "deceased" status. However, we were unable to determine contributing causes of death and hospitalization of patients. Fifth, though patients in this EHR database represent 20% of US population, they represent patients who had encounters with healthcare systems and are not necessarily representative of the general US population or patients in other countries. For example, least sick COVID-19 patients are likely to be missed from the EHR systems since they may never encounter the health system. Our reported rate of all-cause hospitalization rate of 24.3% for COVID-19 patients in the EHR database are higher than the overall cumulative COVID-19-associated hospitalization rate of 207.1 per 100,000 reported by the Centers for Disease Control and Prevention [39]. Last but not least, findings from this EHR-based study are associational not causal. Due to these limitations, this study serves as a baseline study of COVID-19 risk, disparities and outcomes in patients with CLD. Because of the markedly increased risks associated with acquisition of COVID in the chronic liver disease population that we demonstrated here, these initial findings need to be replicated in and compared to other EHR databases and in other

populations and investigated in prospective manner in which additional confounding factors can be identified and controlled.

## Data sharing statement

IBM Watson Health is a for-profit company and access to its Explorys patient EHR database through the web-based informatics tools, which was utilized in this study, is fee-based. Information about Explorys data access can be found at https://www.ibm.com/watson-health/about/explorys.

## Author Contributions

Q.W. and RX conceived the study, designed the experiments, conducted the analysis and authored the manuscript. P.B.D critically contributed to data interpretation, results, discussion and manuscript preparation. Q.W and R.X had access to the original data.

## Declaration of Competing Interest

Q.W., P.B.D, and R.X have no financial interests to disclose.

## Funding

## References

[1] World Health Organization (WHO). Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports (accessed on October 1st, 2020).

[2] The Centers for Disease Control and Prevention (CDC). Cases in the U.S. https://covid.cdc.gov/covid-data-tracker/#cases_casesinlast7days (accessed on October 1st, 2020).

[3] The Centers for Disease Control and Prevention (CDC). Evidence used to update the list of underlying medical conditions that increase a person's risk of severe illness from COVID-19. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/evidence-table.html (accessed on October 1st, 2020).

[4] Wang QQ, Kaelber DC, Xu R, Volkow ND. COVID-19 risk and outcomes in patients with substance use disorders: analyses from electronic health records in the United States. Mol Psychiatry 2020;14:1–10. doi: 10.1038/s41380-020-00880-7.

[5] Wang QQ, Xu R, Volkow ND. Increased risk of COVID-19 infection and mortality in people with mental disorders: analysis from electronic health records in the United States. World Psychiatry 2020;7. doi: 10.1002/wps.20806.

[6] Wang QQ, Berger NA, Xu R. Analyses of risk, racial disparity, and outcomes of cancers and coronavirus disease 2019 from electronic health 4 records in the US. JAMA Oncol 2020 (in press).

[7] Wang Q.Q., Berger N.A., Xu R., When hematologic malignancies meet COVID-19 in the United States: infections, death and disparities. Blood Rev. https://doi.org/10.1016/j.blre.2020.100775.

[8] The Centers for Disease Control and Prevention (CDC). What to know about liver disease and COVID-19. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/liver-disease.html (accessed on October 1st, 2020).

[9] The Centers for Disease Control and Prevention (CDC). People with certain medical conditions. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html#liver-disease (accessed on October 1st, 2020).

[10] Fix OK, Hameed B, Fontana RJ, et al. Clinical best practice advice for hepatology and liver transplant providers during the COVID-19 pandemic: AASLD expert panel consensus statement [published online ahead of print, 2020 Apr 16]. Hepatology 2020 doi:10.1002/hep.31281. doi: 10.1002/hep.31281.

[11] Mantovani A, Beatrice G, Dalbeni A. Coronavirus disease 2019 and prevalence of chronic liver disease: a meta-analysis. Liver Int. 2020;40(6):1316–20 Epub 2020 Apr 24. doi: 10.1111/liv.14465.

[12] Moon AM, Webb GJ, Aloman C, Armstrong MJ, Cargill T, Dhanasekaran R, et al. High mortality rates for SARS-CoV-2 infection in patients with pre-existing chronic liver disease and cirrhosis: preliminary results from an international registry. J Hepatol. 2020;73(3):705–8 Epub 2020 May 21. doi: 10.1016/j.jhep.2020.05.013.

[13] Ji D, Qin E, Xu J, Zhang D, Cheng G, Wang Y, et al. Non-alcoholic fatty liver diseases in patients with COVID-19: a retrospective study. J Hepatol. 2020 Aug;73(2):451–3 Epub 2020 Apr 8. doi: 10.1016/j.jhep.2020.03.044.

[14] Garrido I, Liberal R, Macedo G. Review article: COVID-19 and liver disease-what we know on 1st May 2020. Aliment Pharmacol Ther 2020;52(2):267–75 Epub 2020 Jun 2. doi: 10.1111/apt.15813.

[15] Fu Y, Zhu R, Bai T, et al. Clinical Features of COVID-19-Infected Patients With Elevated Liver Biochemistries: a Multicenter, Retrospective Study [published online ahead of print, 2020 Jun 30]. Hepatology 2020 10.1002/hep.31446. doi: 10.1002/hep.31446.

[16] Bloom PP, Meyerowitz EA, Reinus Z, et al. Liver biochemistries in hospitalized patients with COVID-19 [published online ahead of print, 2020 May 16]. Hepatology 2020 10.1002/hep.31326. doi: 10.1002/hep.31326.

[17] Singh S, Khan A. Clinical characteristics and outcomes of coronavirus disease 2019 among patients with preexisting liver disease in the United States: a multicenter research network study [published online ahead of print, 2020 May 4]. Gastroenterology 2020 S0016-5085(20)30585-0. doi: 10.1053/j.gastro.2020.04.064.

[18] Louie KS, St Laurent S, Forssen UM, Mundy LM, Pimenta JM. The high comorbidity burden of the hepatitis C virus infected population in the United States. BMC Infect Dis 2012;12:86. doi: 10.1186/1471-2334-12-86.

[19] Sagnelli E, Stroffolini T, Mele A, Imparato M, Sagnelli C, Coppola N, et al. Impact of comorbidities on the severity of chronic hepatitis B at presentation. World J Gastroenterol 2012;18(14):1616.

[20] Önnerhag K, Dreja K, Nilsson PM, Lindgren S. Increased mortality in non-alcoholic fatty liver disease with chronic kidney disease is explained by metabolic comorbidities. Clin Res Hepatol Gastroenterol. 2019;43(5):542–50 Epub 2019 Mar 1. doi: 10.1016/j.clinre.2019.02.004.

[21] Fede G, Privitera G, Tomaselli T, Spadaro L, Purrello F. Cardiovascular dysfunction in patients with liver cirrhosis. Ann Gastroenterol 2015;28(1):31.

[22] Naschitz JE, Slobodin G, Lewis RJ, Zuckerman E, Yeshurun D. Heart diseases affecting the liver and liver diseases affecting the heart. Am. Heart J. 2000;140(1):111–20.

[23] Sipeki N, Antal-Szalmas P, Lakatos PL, Papp M. Immune dysfunction in cirrhosis. World J Gastroenterol 2014;20(10):2564.

[24] Shin EC, Sung PS, Park SH. Immune responses and immunopathology in acute and chronic viral hepatitis. Nat Rev Immunol 2016;16(8):509.

[25] Park SH, Rehermann B. Immune responses to HCV and other hepatitis viruses. Immunity 2014;40(1):13–24.

[26] IBM Explorys EHR solutions. https://www.ibm.com/watson-health/about/explorys (accessed up to October 1st, 2020).

[27] Kaelber DC, Foster W, Gilder J, Love TE, Jain AK. Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data. J Am Med Inform Assoc 2012;19(6):965–72. doi: 10.1136/amiajnl-2011-000782.

[28] Lindberg C. The Unified Medical Language System (UMLS) of the National Library of Medicine. J Am Med Rec Assoc. 1990;61(5):40–2.

[29] SNOMED International. THE SYSTEMATIZED NOMENCLATURE OF MEDICINE – CLINICAL TErms (SNOMED CT). http://www.snomed.org/snomed-ct/why-snomed-ct

[30] IBM Explorys EHR database bibliography categorized by therapeutic area. https://www.ibm.com/downloads/cas/RPKNLL1M) (accessed on August 21, 2020).

[31] Kuritz SJ, Landis JR, Koch GG. A general overview of Mantel-Haenszel methods: applications and recent developments. Annu Rev Public Health 1988;9:123–60. doi: 10.1146/annurev.pu.09.050188.001011.

[32] Yancy CW. COVID-19 and African Americans [published online ahead of print, 2020 Apr 15]. JAMA 2020 doi:10.1001/jama.2020.6548. doi: 10.1001/jama.2020.6548.

[33] The Centers for Disease Control and Prevention (CDC). COVID-19 and older adults. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/older-adults.html (accessed on October 1st, 2020).

[34] Burki T. The indirect impact of COVID-19 on women. Lancet Infect Dis. 2020;20(8):904–5 PMID:32738239. doi: 10.1016/S1473-3099(20)30568-5.

[35] Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. Clin Res Cardiol 2017;106(1):1–9. doi: 10.1007/s00392-016-1025-6.

[36] Coorevits P, Sundgren M, Klein GO, et al. Electronic health records: new opportunities for clinical research. J Intern Med 2013;274(6):547–60. doi: 10.1111/joim.12119.

[37] Ahmad FS, Chan C, Rosenman MB, et al. Validity of cardiovascular data from electronic sources: the multi-ethnic study of atherosclerosis and HealthLNK. Circulation 2017;136(13):1207–16. doi: 10.1161/CIRCULATIONAHA.117.027436.

[38] Safran C, Bloomrosen M, Hammond WE, et al. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. J Am Med Inform Assoc. 2007;14(1):1–9. doi: 10.1197/jamia.M2273.

[39] The Centers for Disease Control and Prevention (CDC). Coronavirus Disease 2019 (COVID-19) https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html (accessed on November 11, 2020)