

## **The role of HIV infection in the clinical spectrum of COVID-19: a population-based cohort analysis based on US National COVID Cohort Collaborative (N3C) Enclave data**

Xueying Yang<sup>1,2,3</sup>\*, Jiajia Zhang<sup>1,2,4</sup>, Siyuan Guo<sup>1,2,4</sup>, Bankole Olatosi<sup>1,2,5</sup>, Sharon Weissman<sup>1,2,6</sup>,

Xiaoming Li<sup>1,2,3</sup>

<sup>1</sup> South Carolina SmartState Center for Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA, 29208

<sup>2</sup> Big Data Health Science Center, University of South Carolina, Columbia, SC, USA, 29208

<sup>3</sup> Department of Health Promotion, Education and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA, 29208

<sup>4</sup> Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA, 29208

<sup>5</sup> Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA, 29208

<sup>6</sup> Department of Internal Medicine, School of Medicine, University of South Carolina, Columbia, SC, USA, 29208

\*Corresponding author:

Xueying Yang, Ph.D.  
SC SmartState Center for Healthcare Quality (CHQ)  
Department of Health Promotion, Education, and Behavior  
Arnold School of Public Health  
University of South Carolina  
Columbia, SC 29208  
Email: [xueyyang@mailbox.sc.edu](mailto:xueyyang@mailbox.sc.edu)

## Abstract

**Background** Evidence of whether people living with HIV (PLWH) are at elevated risk of adverse COVID-19 outcomes is limited. We aimed to investigate this association using the population based National COVID Cohort Collaborative (N3C) data in the US.

**Methods** The harmonized, high-granularity electronic health record data from 54 clinical sites in N3C were used for this study. Logistic and multinomial logistic regression models were employed to estimate the association between HIV infection and hospitalization, mortality and clinical severity of COVID-19, with models being initially adjusted for age and sex, then cumulatively adjusted for race and ethnicity, smoking and obesity, and a broad range of comorbidities. Interaction terms were added to assess modification effect by age, sex, and race.

**Findings** Among a total of 1,436,622 adult COVID-19 cases between January 2020 and May 2021, 13,170 had HIV infection. A total of 26,130 deaths occurred, with 445 among PLWH. PLWH had higher risk of COVID-19 death and hospitalization than non-PLWH after adjusting for age and sex. The associations were attenuated, but remained significant, after adjusting for lifestyle factors and comorbidities (COVID-19 death: [adjusted OR(AOR): 1.38; 95%CI: 1.25-1.54]; hospitalization: [AOR: 1.23; 95%CI: 1.18-1.28]). In terms of COVID-19 disease severity, PLWH were less likely to have mild/moderate illness but more likely to have severe illness when only controlling for demographics, smoking and BMI, although the estimated risk was obviated after controlling for comorbidities. Interaction terms revealed that the elevated risk was higher among older age group, males and black/African American.

**Interpretation** PLWH in the US had increased risk of COVID-19 hospitalization and mortality comparing with non-PLWH. The adverse COVID-19 outcomes might not only be accounted for by HIV but also by other risk factors that are highly prevalent in PLWH.

**Funding** National Center for Advancing Translational Sciences, National Institute of Allergy And Infectious Diseases, US

**Keywords:** COVID-19; HIV/AIDS; mortality; hospitalization; clinical severity

## Introduction

As of May 29 2021, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19), has infected over 169 million people and caused more than 3.5 million deaths worldwide.<sup>1</sup> Since the first confirmed case of COVID-19, the countrywide COVID-19 outbreak has surged quickly, making the United States (US) one of the countries hardest hit by the pandemic.<sup>1</sup> As the pandemic surges in the US, it is important to identify patients at elevated risk of developing severe symptoms to inform clinical management decisions. Older age and presence of comorbidities are recognized to increase the severity of COVID-19.<sup>2</sup> Patients who have malignant disease or solid-organ transplants have overall poorer outcomes of COVID-19,<sup>3</sup> but the evidence is less clear for people with other types of immunocompromising conditions, including people living with HIV (PLWH).<sup>4</sup>

Existing evidence for an association between HIV infection and COVID-19 related outcomes is mixed. Early on in the course of the COVID-19 pandemic, evidence was limited to the case report or case series.<sup>5</sup> According to a systematic review, the COVID-19 prevalence among PLWH was comparable to the general population although there were occasionally reports of atypical, but not more severe, disease course relative to people without HIV (non-PLWH).<sup>5</sup> However, all studies included in this review had very small sample sizes (ranged from 1 to 47) and some lacked a comparison group (i.e., non-PLWH). Later on, emerging data from observational cohort studies showed similar findings,<sup>6-11</sup> while samples of most these studies were all restricted to hospitalized patients. By contrast, several large population-based studies have found conflicting results. Recent large-scale studies conducted in the United Kingdom (UK) and South Africa suggested that PLWH had a higher risk (over double) of COVID-19 mortality than non-PLWH, although different factors were adjusted in different studies.<sup>12,13</sup> A prospective study of patients hospitalized with COVID-19 showed an increased 28-day mortality in PLWH after adjusting for age<sup>14</sup>. One New York study reported a standardized in-hospital mortality ratio of 1.23 for HIV patients.<sup>15</sup> However, one limitation of these population-based studies was that they were not able to adequately adjust for various confounders (e.g., obesity<sup>12</sup>, comorbidities<sup>15</sup>).

Although the studies to date have been informative, it is important to emphasize that salient knowledge gaps remain regarding the role of HIV across the COVID-19 clinical spectrum. Existing literature generally has been limited by small sample sizes,<sup>5</sup> lack of direct comparative data of non-PLWH,<sup>5</sup> inability to adjust for various confounders<sup>12,15</sup> or a biased selection of study population (e.g., restriction to hospitalized patients<sup>7,9,10</sup>). A large, multi-center, representative clinical dataset is needed to fill in these knowledge gaps and to investigate whether HIV is a risk factor for adverse COVID-19 clinical outcomes. Therefore, this study aims to understand the role of HIV infection in affecting the COVID-19 clinical outcomes (e.g., disease severity, hospitalization, mortality) using the National COVID Cohort Collaborative (N3C) data. The representativeness and comprehensiveness of N3C data can overcome the limitations of small sample size, sample selection bias, and lack of comorbidity data.

## **Methods**

### **Data sources**

The N3C Enclave, sponsored by multiple institutes of the US National Institutes of Health,<sup>16</sup> is the largest cohort of US COVID-19 cases and representative controls to date. The N3C is a large, multi-center, and ongoing representative clinical dataset and harmonizes high-granularity electronic health records data for individuals who ever tested COVID-19 since 01/01/2018. N3C data covers data from almost all US geographic regions (particularly in the Southeast, Mid-Atlantic, and Midwest), and representing a racially and ethnically diverse population.

Data ingestion and harmonization in N3C is an iterative process between the Phenotype & Data Acquisition workstream and contributing sites. Contributing sites send updated data payloads approximately weekly. N3C harmonizes site data into Observational Medical Outcomes Partnership (OMOP) version 5.3.1 in partnership with subject matter experts from each common data model (CDM) community. In this process, all data (e.g., laboratory measurements, clinical observations such as vital signs, medications, and clinical conditions) are harmonized and mapped to the OMOP vocabulary. Site data that pass robust data quality assessment pipeline are integrated into the “release” set for use by the community. Details about the data transfer, harmonization, quality, and integration processes have been reported elsewhere.<sup>16</sup> As of the date of this data analysis (May 8, 2021), data from 54 clinical sites was harmonized and integrated into the N3C release set. An Institutional Data Use Agreement was signed between University of South Carolina and National Center for Advancing Translational Sciences (NCATS) N3C Data

Enclave and our study protocol received approval from the institutional review board in University of South Carolina.

### **Cohort definition**

The N3C phenotype<sup>17</sup> is designed to be inclusive of any diagnosis codes, procedure codes, lab tests, or combination thereof that may be indicative of COVID-19 (e.g., CDC coding guidance,<sup>18</sup> PCORnet<sup>19</sup>). N3C includes patients with any encounter after 1/1/2020 who have either 1) one or more of a set of a priori-defined SARS-CoV-2 laboratory tests with a positive result; or 2) one or more of a “strong positive” diagnostic code from the ICD-10 or SNOMED tables; or 3) two or more of the “weak positive” diagnostic codes from the ICD-10 or SNOMED tables. The cohort definition is publicly available on GitHub.<sup>17</sup>

Up until May 8, 2021, a total of 5,830,841 individuals (including both cases and controls) were harmonized and integrated into the N3C release set from 54 clinical sites, with a total of 1,608,854 COVID-19 cases. Of the overall cohort, 1,437,352 adult COVID-19 positive individuals were identified after removing those <18 years of age or with missing value on age. After further removing the individuals with missing data key on demographics (sex, race), 1,436,622 adult COVID-19 positive individuals were included as the study population for the current study, among which 13,170 (0.92%) were co-infected with HIV (Figure 1).

### **Variable definition**

N3C harmonizes data across four clinical data models (ACT Network, PCORnet, OHDSI, and TriNetX) and provides a unified analytical platform in which data are encoded using the OMOP<sup>20</sup> version 5.3.1. The concept sets<sup>21,22</sup> in the OMOP were used to identify each clinical concept (e.g., laboratory measure, conditions, or medication). Data domains extracted by N3C include demographics, encounter details, medications, diagnoses, procedures, vital signs, laboratory results, procedures, and social history. Specific variables included in each domain are listed in each model’s documentation (i.e., Tables). Both concept sets and tables were used to define variables of interest.

#### *People living with HIV*

PLWH was determined through concept sets, which mapped to various domain tables, including any HIV diagnosis codes (ICD-10 codes, SNOMED Codes) in the “*condition occurrence*” table,

HIV lab tests (LOINC codes) in the “*measurement*” table, HIV drug exposure in the “*drug exposure*” table, or combination thereof that may be indicative of HIV diagnosis.

#### *COVID-19 hospitalization, clinical severity and death*

COVID-19 hospitalization in the current study was identified by case insensitive string matching \*inpatient visit\* or \*inpatient critical care facility\* or \* emergency room and inpatient visit \* in the “*Selected Critical Visit*” table. “*Selected Critical Visit*” table was created by N3C work stream to document a COVID-19 related medical encounter. Specifically, N3C defined a single index encounter for each laboratory-confirmed positive patients by selecting encounters that start up to 30 days before or 7 days after the positive test result, or a positive test result during the visit.<sup>23</sup> When multiple encounters met these criteria, it broke ties by preferentially selecting the encounter in which the most severe outcome was observed, then the longest visit, and finally the most recent visit.<sup>23</sup>

Clinical severity was classified using the Clinical Progression Scale (CPS) established by the WHO for COVID-19 clinical research<sup>24</sup>. Based on WHO criterion, N3C placed patients into strata defined by the maximum clinical severity from selected critical visits<sup>23</sup>: “unaffected” (e.g., suspected COVID), “mild”(outpatient, WHO severity 1-3); “mild ED” (outpatient with ED visit, WHO severity 3), “moderate” (hospitalized without invasive ventilation, WHO severity 4-6), “severe” (hospitalized with invasive ventilation or ECMO, WHO severity 7-9); and “mortality/hospice” (hospital mortality or discharge to hospice, WHO severity 10).<sup>23</sup> Due to the small number in certain categories among PLWH, we collapsed some WHO CPS categories and regrouped them into three categories, i.e., unaffected, mild (including mild ED)/moderate, and severe (including mortality/hospice). For the binary death outcome, death was determined through the *death* table.

#### *Life-style factors and comorbidity*

Life-style factors included smoking status and obesity (indicated by Body Mass Index [BMI]). Smoking status was defined by a concept set, whose member comprised of ‘ARIScience-Smoker-JA’, ‘smoker\_NMH’, ‘UVA Former Smoker’, and ‘UVA Current Smoker’ in the ‘*observation*’ and ‘*condition*’ tables. BMI information was retrieved from ‘*patient severity score*’ tables. Encounters beginning on or after 1/1/2018 and before COVID diagnosis are included to provide information about pre-existing health conditions (“lookback data”). The N3C work

stream defined comorbidities based on the ICD codes in the updated Charlson Comorbidity Index.<sup>25</sup> A series of binary variables were used to indicate the presence or absence of each comorbidity, including congestive heart failure, myocardial infarction, peripheral vascular disease, chronic pulmonary disease, dementia, diabetes mellitus, liver disease, chronic kidney disease, rheumatologic disease, cancer, stroke, hemiplegia or paraplegia and peptic ulcer disease.

### **Statistical analysis**

Descriptive statistics were employed to examine the socio-demographics of all the COVID-19 cases by HIV status. The variable distributions between COVID-19 patients with and without HIV infection were summarized and compared using the independent t-test (for continuous variables) or Chi-square test (for categorical variables). For each binary clinical outcome (i.e., hospitalization, death), we used logistic regression analysis to estimate the association between HIV infection and the outcome, initially adjusted for age and sex, then cumulatively adjusted for race and ethnicity, smoking and obesity in sequential models, and eventually adjusted for an array of comorbidities. For COVID-19 severity, the multinomial logistic regressions were applied. Specifically, the HIV infection, demographics, comorbidities, and lifestyle related factors were included in multiple multinomial regressions. Of note, since it requires at least one healthcare encounter information to generate index encounter and define COVID-19 severity, we first conducted multinomial regressions among individuals with at least one healthcare encounter (n=1,396,795; 97.2% of total sample); then we conducted sensitivity analysis to include those without healthcare encounter information in the severity analysis (i.e., using the entire sample). In the sensitivity analysis, we categorized those cases without healthcare encounter information into the first category of CPS (i.e., unaffected).

To investigate whether age, sex, and race could be potential effect modifiers of HIV status, we fitted interaction terms between age (18-49, 50-64, and  $\geq 65$  years), sex, race (white vs black/African American or Asian/other/unknown) and HIV status in all analyses. Besides, we fitted stratified models with HIV infection, demographics, lifestyle and comorbidities on each selected sub-groups (i.e., male, female, 18-49, 56-64,  $\geq 65$ , white, black/African American and Asian/other/unknown). We implemented all analyses using SQL and R and created reproducible pipelines in the Code workbook on N3C Data Enclave.

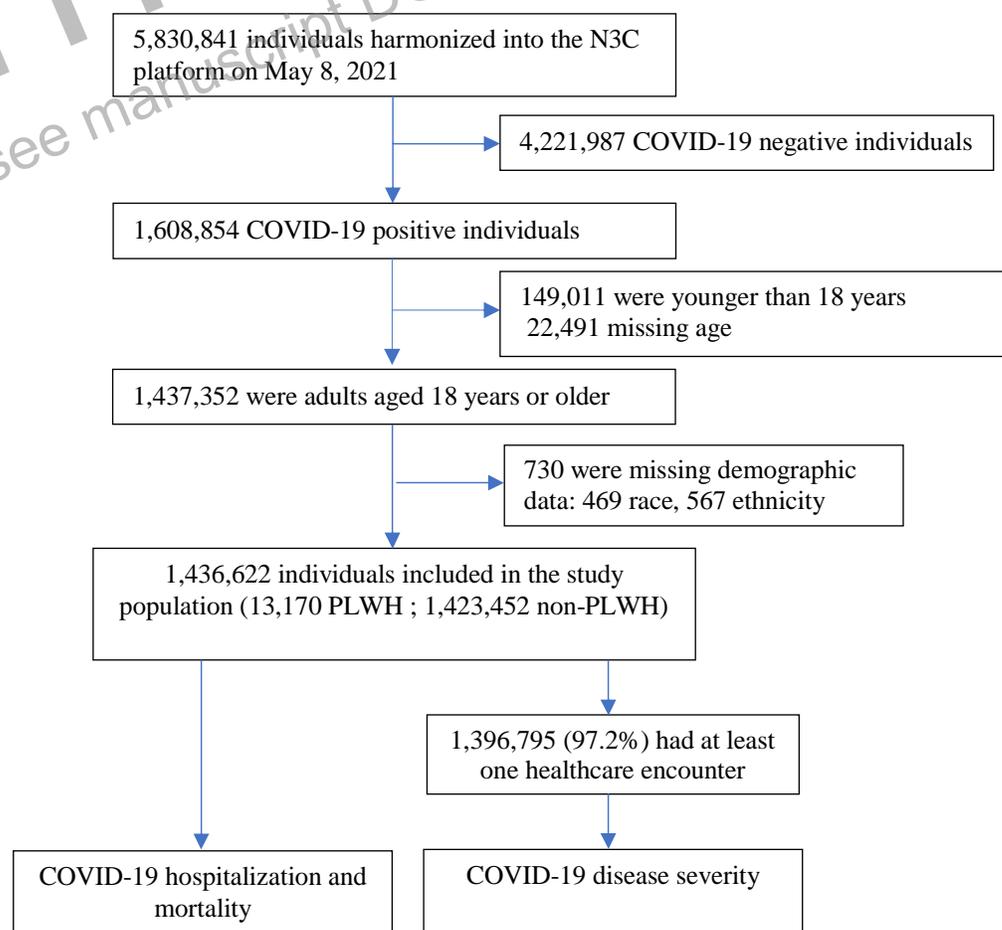
### **Role of the funding source**

The National Center for Advancing Translational Sciences is the primary steward of the N3C data and created the underlying architecture of the N3C Data Enclave, manages the Data Transfer Agreements and Data Use Agreements, houses the Data Access Committee, and supports contracts to vendors to help build various aspects of the N3C Data Enclave. The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## Results

### Background characteristics by HIV status

The cohort profile for the study and exclusion points are displayed below in Figure 1. Compared to non-PLWH, PLWH with COVID-19 diagnosis had similar median age (49 vs 47 years), but there was a considerably lower proportion of people  $\geq 65$  years (14.68% vs 20.59%). PLWH



**Figure 1: Study cohort profile**

compromised of a greater proportion of males (73.20% vs 44.70%) and people of black/African American race (31.07% vs 13.97%); and had higher prevalence of all comorbidities, including diabetes (23.28% vs 16.21%), chronic pulmonary disease (22.58% vs 13.91%), liver disease (16.34% vs 4.96%), chronic kidney disease (13.35% vs 5.97%), cancer (9.15% vs 5.46%) (Table 1).

### **COVID-19 hospitalization**

Among a total of 1,436,622 COVID-19 cases, 262,331 (18.26%) were hospitalized. PLWH proportionally required more COVID-19 related hospitalization than non-PLWH (28.28% vs 18.17%) (Table 1). Crude odd ratio (OR) of COVID-19 hospitalization was higher in PLWH than non-PLWH (OR: 1.78, 95%CI: 1.71, 1.84). The association was attenuated, but remained significant, after adjusting for sequential demographics, lifestyle factors and comorbidities (adjusted odds ratio [AOR]: 1.23, 95%CI: 1.18, 1.28) (Table 2 & Figure 2c & Supplementary Table 2).

### **COVID-19 mortality**

Among the entire cohort of COVID-19 cases, 1.82% (26,130/1,436,622) deaths occurred, with 445 of these deaths being PLWH (Table 1). Crude risk of COVID-19 death was higher in PLWH than non-PLWH (3.38% vs 1.80%; OR: 1.90, 95%CI: 1.73, 2.09). After adjusting for age and sex, the effect size increased (AOR: 2.05, 95%CI: 1.86, 2.26). However, the association attenuated after sequential adjustment for race, ethnicity, lifestyle factors, comorbidities (AOR: 1.38, 95%CI: 1.25, 1.54). (Table 2 & Figure 2d & Supplementary Table 3).

### **COVID-19 disease severity**

Among a total of 1,396,795 COVID-19 cases who had at least one healthcare encounter, 62.33% had mild/moderate illness, and 1.74% had severe COVID-19. Comparing with non-PLWH, PLWH had a higher proportion of severe illness (3.61% vs 1.73%), but a lower proportion of mild/moderate illness (47.15% vs 62.47%) (Table 1). Using unaffected individuals as a reference group in multinomial regression, PLWH had a lower odds of presenting with mild/moderate illness than non-PLWH even after adjusted for all the covariates (adjusted relative risk [ARR]=0.57, 95% CI: 0.55-0.59); in contrast, PLWH had a higher odds of experiencing severe outcomes than non-PLWH after adjusting for sociodemographic and lifestyle factors (ARR=1.25, 95%CI: 1.13-1.38), however, sequential adjustment for an array of comorbidities obviated such

estimated risk (ARR=1.02, 95%CI: 0.92-1.13) (Table 2 & Figure 2a-b & Supplementary Table 1). In the sensitivity analysis, we included individuals without healthcare encounter information and the results were similar to the findings in the models when excluding those without healthcare encounter information.

### **Interaction analysis**

The interaction effect of age and HIV status suggested that the aging process in PLWH exacerbated all the adverse outcomes of COVID-19. PLWH in the older age groups had much higher odds of death (e.g., AOR: 8.43 in PLWH vs 4.26 in non-PLWH in age group of 50-65 years) and hospitalization (e.g., AOR: 3.63 in PLWH vs 2.39 in non-PLWH in age group of  $\geq 65$  years) comparing with non-PLWH in the same age range. As another potential modifier, male sex could also interact with HIV infection in increasing the risk of severe clinical outcomes of COVID-19, yet with a smaller magnitude. Similar results were illustrated in the interaction of race and HIV status, where black/African American race interacted with HIV infection in causing more risk of adverse COVID-19 outcomes. Stratified models revealed that the elevated risk was higher in the similar subgroups (e.g., older age, male sex and black/African American race). (Table 2 & Supplementary Table 4)

### **Discussion**

To the best of our knowledge, this is the largest cohort study to investigate the role of HIV infection in COVID-19 clinical spectrum across the US. Our population-level analysis from N3C data found that PLWH might not be disproportionately vulnerable to SARS-CoV-2 infection but more likely to experience severe symptoms, have higher odds to be hospitalized, and die from COVID-19 although such risk was attenuated or obviated when adjusted for comorbidities. This result illustrated that the adverse COVID-19 outcomes might not only be due to the HIV infection but also some clinical characteristics (e.g., comorbidities) that are highly prevalent in PLWH. For example, PLWH has much higher prevalence of nearly every comorbid condition than non-PLWH within N3C COVID-19 cohort, which is consistent with other studies (including one US study)<sup>6,26,27</sup> and also plausible given the recognized interplay between HIV and comorbidities.<sup>28,29</sup>

Our results showed a smaller but consistent effect (AOR: 1.38; 95%CI: 1.25-1.54) of HIV infection on COVID-19 related mortality with other large population-based cohort studies from

South Africa involving 3.5 million individuals (HR: 2.14; 95%CI: 1.98-3.43)<sup>12</sup> and UK involving 17 million (HR: 2.59; 95%CI: 1.74-3.84)<sup>13</sup>. The smaller effect size in our study might be due to a relatively large number of COVID-19 related deaths among PLWH (n=445) and our ability to accurately control both traditional lifestyle factors and a broad range of comorbidities. For example, although the age profile of the UK study is comparable with ours,<sup>13,14</sup> our HIV sample with COVID-19 diagnosis had more comorbidities, notably a higher prevalence of diabetes (e.g., 23.3% vs 5.5%), chronic pulmonary disease (e.g., 22.6% vs 4.0%), and liver disease (e.g., 16.3% vs 3.4%).<sup>13</sup> In the UK study, the relatively small number of COVID-19 deaths among PLWH (n=25), reflects the younger age distribution of their HIV patients, preventing a definitive conclusions about the role of comorbidities and changes in the role of HIV over time. In addition, the UK sample had an under-representation of HIV cases in the epicenter (e.g., London).<sup>3</sup> In contrast, in the South African study, the HIV-associated risk of COVID-19 death may be overestimated due to the under-ascertainment of some comorbidities and the lack of other residual confounding (e.g., BMI).<sup>12</sup> It is noteworthy that the association in the present study was similar to a New York state-level setting involving 2,988 PLWH with COVID-19 (in-hospital death: sRR: 1.23; 95%CI: 1.07-1.40)<sup>15</sup>, although limitations this study included a lack of control for lifestyle factors and comorbidity data, and a short observation period (March 1-June 7, 2020). Regarding the clinical severity of COVID-19, we found that PLWH were less likely to have mild illness, but more likely to have severe outcomes when only adjusting for demographics and lifestyle factors. The adjustment for comorbidities obviated the estimated risk of severe outcomes among PLWH. This finding suggests that PLWH might show less symptoms at the earlier stage of SARS-CoV-2 infection. Such a protection from the most serious sequelae of COVID-19 might be attributable to the immunocompromised system (avoided the immune(over)reaction) and/or because of the possible anti-SARS-CoV-2 activity of HIV antiretroviral treatments<sup>30</sup>. For hospitalization, our study showed PLWH required more COVID-19 hospitalization (AOR: 1.23; 95%CI: 1.18, 1.28), a level of risk similar to a New York study (sRR: 1.38, 95%CI: 1.29, 1.47)<sup>15</sup> and another US study using TriNETX network data (RR: 1.69; 95%CI: 1.21-2.38), which controlled for BMI and various comorbidities.<sup>6</sup>

In this study, the sex and race disparities in COVID-19 cases were pronounced among PLWH with COVID-19, with 31.1% of PLWH identified as black/African American compared with only 14.0% among non-PLWH, similar scenarios in male (73.2% vs 44.7%). In the stratified

models, older age revealed much higher mortality and hospitalization among PLWH comparing with non-PLWH in the same age range, which was not reported in the UK study. However, we do show similar findings of the larger association between HIV and adverse COVID-19 outcomes among blacks/African Americans and Asians. Male gender played a role in increasing COVID-19 mortality among PLWH, but not for hospitalization. Understanding the reasons for the disproportionately large association between HIV and adverse COVID-19 outcomes in these subgroups will be a priority if effective policies are to be developed to mitigate any increased risks.

Using data from the largest COVID-19 cohort with the heterogeneous population in the US, our study was able to identify PLWH with mild or asymptomatic COVID-19 and examine the different risks of PLWH in SARS-CoV-2 acquisition versus progression to severe disease or death once infected. In this large population-based study, PLWH had an elevated risk of adverse COVID-19 outcomes, which are largely accounted for by the high prevalent factors in HIV (e.g., more black/African Americans, comorbidities) that are parallel with risk factors of COVID-19. The attenuated risk after controlling for comorbidities, which are more prevalent and typically occurring at a younger age among PLWH, indicates that certain underlying medical conditions had a greater influence on COVID-19 outcomes of PLWH. The exacerbated COVID-19 outcomes in black/African American and male PLWH suggested pronounced inequities faced by PLWH in the dual pandemics. PLWH might also need priority considerations given the current wide availability of vaccines against SARS-CoV-2. Given the pandemic's exacerbating effects on health inequities, HIV public health and clinical communities must strengthen services and support for people living with and affected by HIV to protect and accelerate progress made thus far toward the goals of ending the HIV epidemic.

Our study represents the largest population-level analyses to draw on EHR data for HIV and COVID-19, and its robustness and comprehensiveness are major strengths. Several limitations are worth to note. First, although we included around 6 million individuals from the N3C dataset, they are more concentrated in the Southeast, Mid-Atlantic, and Midwest and may not represent the entire COVID-19 population or HIV population in the US. However, the cohort still derives from all US geographic regions. The age, sex, race, ethnicity distributions also indicate a diverse patient cohort that is representative of many segments of the US population. Second, given the scope of the current study, levels of antiretroviral therapy use and HIV progression indicators

(e.g., CD4 count, viral load) were not included in the analysis as these factors would merit separate analyses to investigate if people with virologically controlled HIV who are clinically stable will experience any greater risk for COVID-19 complications than non-PLWH, which is of great clinical significance to COVID-19 research as well.

### **Acknowledgement**

The analyses described in this publication were conducted with data or tools accessed through the NCATS N3C Data Enclave <https://covid.cd2h.org> and supported by NCATS U24 TR002306 and the National Institute of Allergy And Infectious Diseases of the National Institutes of Health under Award Number R01AI127203-4S1. This research was possible because of the patients whose information is included within the data and the organizations (<https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories>) and scientists who have contributed to the on-going development of this community resource <https://doi.org/10.1093/jamia/ocaa196>.<sup>16</sup> The content of this publication and the opinions expressed do not necessarily reflect the views or policies of the National Institutes of Health nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. We also acknowledge support from the N3C Publication Committee in the preparation of this manuscript.

### **Author Contributions**

XY wrote the first draft of the manuscript. JZ set up the statistical test design. SG wrote data preparation code and SQL, R code for data analysis, which was reviewed and verified by JZ. XY prepared tables and figures with input from SG. SW provided clinical input and patient severity predictor considerations and use thereof. BO and XL reviewed and edited the manuscript.

Authorship was determined using ICMJE recommendations.

### **Availability of data and materials**

National Institute of Health's (NIH) National COVID Cohort Collaborative (N3C) has clear procedures for researchers to gain access to the data (1000+ researchers already have access to the data) and as such the patient statistical analysis of this manuscript is transparent and repeatable. <https://ncats.nih.gov/n3c> provides data access request process details.

### **Patient Consent for Publication**

Not Applicable

### **Ethics Approval and Consent to Participate**

National Institute of Health's (NIH) National COVID Cohort Collaborative (N3C) Data Utilization Request Approval committee approved the data utilization request of this project (RP-E72986).

WITHDRAWN  
see manuscript DOI for details

## References

1. Coronavirus Resource Center. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Available at: <https://coronavirus.jhu.edu/map.html>. 2021.
2. Center for Disease Prevention and Control. Underlying Medical Conditions Associated with High Risk for Severe COVID-19: Information for Healthcare Providers. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-care/underlyingconditions.html> Accessed on April 24, 2021. 2021.
3. Waters LJ, Pozniak AL. COVID-19 death in people with HIV: interpret cautiously. *Lancet HIV*. 2021;8(1):e2-e3.
4. Fung M, Babik JM. COVID-19 in Immunocompromised Hosts: What We Know So Far. *Clin Infect Dis*. 2021;72(2):340-350.
5. Baluku JB, Olum R, Agolor C, et al. Prevalence, clinical characteristics and treatment outcomes of HIV and SARS-CoV-2 co-infection: a systematic review and meta-analysis. *medRxiv*. 2020.
6. Hadi YB, Naqvi SFZ, Kupec JT, Sarwari AR. Characteristics and outcomes of COVID-19 in patients with HIV: a multicentre research network study. *AIDS*. 2020;34(13):F3-F8.
7. Karmen-Tuohy S, Carlucci PM, Zervou FN, et al. Outcomes Among HIV-Positive Patients Hospitalized With COVID-19. *J Acquir Immune Defic Syndr*. 2020;85(1):6-10.
8. Miyashita H, Kuno T. Prognosis of coronavirus disease 2019 (COVID-19) in patients with HIV infection in New York City. *HIV Med*. 2021;22(1):e1-e2.
9. Sigel K, Swartz T, Golden E, et al. Coronavirus 2019 and People Living With Human Immunodeficiency Virus: Outcomes for Hospitalized Patients in New York City. *Clinical Infectious Diseases*. 2020;71(11):2933-2938.
10. Stoeckle K, Johnston CD, Jannat-Khah DP, et al. COVID-19 in Hospitalized Adults With HIV. *Open Forum Infect Dis*. 2020;7(8):ofaa327.
11. Mascolini M. COVID-19 Rate No Higher With HIV in Largest US HIV+/HIV- Cohort. *AIDS 2020: 23rd International AIDS Conference Virtual, July 6-10, 2020* Available at: [https://www.natap.org/2020/IAC/IAC\\_105htm](https://www.natap.org/2020/IAC/IAC_105htm). 2020.
12. Boule A, Davies M-A, Hussey H, et al. Risk factors for COVID-19 death in a population cohort study from the Western Cape Province, South Africa. *Clinical Infectious Diseases*. 2020.
13. Bhaskaran K, Rentsch CT, MacKenna B, et al. HIV infection and COVID-19 death: a population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform. *The Lancet HIV*. 2020;8(1):e24-e32.
14. Geretti AM, Stockdale AJ, Kelly SH, et al. Outcomes of Coronavirus Disease 2019 (COVID-19) Related Hospitalization Among People With Human Immunodeficiency Virus (HIV) in the ISARIC World Health Organization (WHO) Clinical Characterization Protocol (UK): A Prospective Observational Study. *Clinical Infectious Diseases*. 2020.
15. Tesoriero JM, Swain CE, Pierce JL, et al. COVID-19 Outcomes Among Persons Living With or Without Diagnosed HIV Infection in New York State. *JAMA Netw Open*. 2021;4(2):e2037069.

16. Melissa H, Christopher C, Kenneth G. The National COVID Cohort Collaborative (N3C): Rationale, Design, Infrastructure, and Deployment. *J Am Med Inform Assoc*. 2020.
17. Github. Phenotype Data Acquisition. Available at: [https://github.com/National-COVID-Cohort-Collaborative/Phenotype\\_Data\\_Acquisition](https://github.com/National-COVID-Cohort-Collaborative/Phenotype_Data_Acquisition). 2020.
18. Center for Disease Control and Prevention. ICD-10-CM Official Coding and Reporting Guidelines April 1, 2020 through September 30, 2020 Available at: <https://www.cdc.gov/nchs/data/icd/COVID-19-guidelines-final.pdf>. 2020.
19. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc*. 2014;21(4):578-582.
20. Voss EA, Makadia R, Matcho A, et al. Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc*. 2015;22(3):553-564.
21. ARIScience. OMOP/N3C Templates and Codes. Available at: [https://www.ariscience.org/y1\\_omop\\_n3c.html](https://www.ariscience.org/y1_omop_n3c.html) Accessed May 27, 2021. 2020.
22. ATLAS. Atlas OHDSI concept sets. Available at: <http://atlas-covid19.ohdsi.org/#/home> Accessed May 27 2021. 2020.
23. Bennett TD, Moffitt RA, Hajagos JG, et al. The National COVID Cohort Collaborative: Clinical Characterization and Early Severity Prediction. medRxiv [Preprint]. 2021 Jan 13:2021.01.12.21249511. doi: 10.1101/2021.01.12.21249511. PMID: 33469592; PMCID: PMC7814838.
24. WHO Working Group on the Clinical Characterisation Management of Covid-infection. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect Dis*. 2020;20(8):e192-e197.
25. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol*. 2011;173(6):676-682.
26. Shalev N, Scherer M, LaSota ED, et al. Clinical Characteristics and Outcomes in People Living With Human Immunodeficiency Virus Hospitalized for Coronavirus Disease 2019. *Clinical Infectious Diseases*. 2020;71(16):2294-2297.
27. Vizcarra P, Pérez-Elías MJ, Quereda C, et al. Description of COVID-19 in HIV-infected individuals: a single-centre, prospective cohort. *The lancet HIV*. 2020;7(8):e554-e564.
28. Duffau P, Ozanne A, Bonnet F, et al. Multimorbidity, age-related comorbidities and mortality: association of activation, senescence and inflammation markers in HIV adults. *AIDS*. 2018;32(12):1651-1660.
29. Shah ASV, Stelzle D, Lee KK, et al. Global Burden of Atherosclerotic Cardiovascular Disease in People Living With HIV: Systematic Review and Meta-Analysis. *Circulation*. 2018;138(11):1100-1112.
30. Cash-Goldwasser S KS, Cobb L, Bochner A, Bradford E and Shahpar C. In-Depth COVID-19 Science Review June 27-July 3, 2020. Resolve to Save Lives. 2020 July 7. Available from <https://prevent.epidemics.org/covid19/science/review/> Accessed on May 13, 2021. 2020.