



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original article

Racial and ethnic disparities in the observed COVID-19 case fatality rate among the U.S. population



L. Philip Schumm, MA^{a,b,*}, Mihai C. Giurcanu, PhD^{a,b}, Kenneth J. Locey, PhD^{b,c}, Jean Czerlinski Ortega, MA^b, Zhenyu Zhang, PhD^{b,d}, Robert L. Grossman, PhD^{b,d,e}

^a Department of Public Health Sciences, The University of Chicago, Chicago, IL

^b Pandemic Response Commons, Chicago, IL

^c Center for Quality, Safety and Value Analytics, Rush University Medical Center, Chicago, IL

^d Center for Translational Data Science, The University of Chicago, Chicago, IL

^e Department of Medicine, The University of Chicago, Chicago, IL

ARTICLE INFO

Article history:

Received 23 February 2022

Revised 9 July 2022

Accepted 30 July 2022

Available online 5 August 2022

Keywords:

Covid-19

Health disparities

Simpson's paradox

ABSTRACT

Purpose: During the initial 12 months of the pandemic, racial and ethnic disparities in COVID-19 death rates received considerable attention but it has been unclear whether disparities in death rates were due to disparities in case fatality rates (CFRs), incidence rates or both. We examined differences in observed COVID-19 CFRs between U.S. White, Black/African American, and Latinx individuals during this period.

Methods: Using data from the COVID Tracking Project and the Centers for Disease Control and Prevention COVID-19 Case Surveillance Public Use dataset, we calculated CFR ratios comparing Black and Latinx to White individuals, both overall and separately by age group. We also used a model of monthly COVID-19 deaths to estimate CFR ratios, adjusting for age, gender, and differences across states and time.

Results: Overall Black and Latinx individuals had lower CFRs than their White counterparts. However, when adjusting for age, Black and Latinx had higher CFRs than White individuals among those younger than 65. CFRs varied substantially across states and time.

Conclusions: Disparities in COVID-19 case fatality among U.S. Black and Latinx individuals under age 65 were evident during the first year of the pandemic. Understanding racial and ethnic differences in COVID-19 CFRs is challenging due to limitations in available data.

© 2022 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Introduction

In the U.S., racial and ethnic minority groups including Black/African American, and Latinx individuals have experienced higher rates of COVID-19 infection and mortality than their White counterparts, as well as higher rates of excess mortality associated with COVID-19[1–3]. Assessing differences in outcomes among those infected is challenging at the national level due to limitations in available data. Studies among subpopulations have often found little or no evidence of racial and ethnic disparities in case fatal-

ity. For example, using data from the Michigan Disease Surveillance System from March 8 through July 5, 2020, Zelner et al. found Black individuals had substantially higher rates of COVID-19 infection and mortality. However, their case fatality rate was only modestly higher than that of White individuals, suggesting that differences in COVID-19 mortality were driven by differences in the rate of infection[4].

Several studies have examined racial and ethnic differences in hospitalization and/or case fatality rates among health care systems, where access to electronic health records permits adjusting for demographic and clinical covariates. Among individuals receiving care in the U.S. Department of Veterans Affairs tested for COVID-19 between February 8 and July 22, 2020, Rentsch et al. found that while Black and Hispanic patients were more likely to test positive than White patients, once infected their 30-day mortality was less and, after adjusting for age, sex, rural/urban residence and comorbidities was equivalent to that of Whites[5].

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

* Corresponding author: Department of Public Health Sciences, The University of Chicago, 5841 South Maryland Ave MC2000, Chicago, IL 60637, Tel.: +773-850-0557.

E-mail address: pschumm@uchicago.edu (L.P. Schumm).

Similarly, Price-Haywood et al. studied COVID-19 positive patients within Ochsner Health in Louisiana and found that although Black patients had roughly twice the odds of hospitalization than White patients, their risk of in-hospital death was similar[6]. Finally, Ogedegbe et al. examined patients within New York University's Langone Health system tested for COVID-19 between March 1 and April 8, 2020[7]. The odds of hospitalization among those positive were similar for Black, Hispanic, and White patients, while among those hospitalized, Black and Hispanic patients were less likely to die despite adjusting for demographics, comorbidities and insurance.

Griffith et al. discuss the potential of collider bias to distort understanding of COVID-19 disease risk and severity[8]. One potential source of collider bias is restricting analyses to a non-representative subset of the population, such as those hospitalized for COVID-19. For example, if Black/African American individuals were more likely to seek and/or receive care later in the acute setting rather than earlier in primary or other non-urgent care settings—perhaps due to lack of access, biases on the part of providers or prior negative experiences with the health system—this may partly account for the higher rates of admission observed in some studies[9,10]. If true, focusing on in-hospital mortality could reduce the apparent difference in case fatality between Black and White individuals due to a larger proportion of Black individuals with COVID-19 being hospitalized.

The purpose of this study was to examine racial and ethnic differences in COVID-19 case fatality rates (CFRs) at the U.S. national level during the first year of the pandemic. We used data from The Atlantic's COVID Tracking Project (CTP), whose Racial Data Tracker was widely regarded as the most complete source of information on race and ethnicity of COVID-19 cases and deaths during this period. We performed a parallel analysis using the Centers for Disease Control and Prevention (CDC) COVID-19 Case Surveillance Public Use data—an independently compiled and regularly updated individual-level data source that contains information on age. Adjusting for age is critical to get an accurate understanding of differences in COVID-19 CFR[11]. We focused on Black/African American, and Latinx individuals and their comparison to Whites, since these categories are reported most widely by a large number of states and permit approximate comparability between the two datasets.

Methods

Data

The COVID tracking project (CTP)

The CTP was a volunteer-run effort started in March 2020 to compile nationwide data on the COVID-19 pandemic and ran until March 2021. During its operation, the CTP was considered an authoritative source of COVID-19 data, and their data have been used in over 1000 academic articles. CTP datasets remain available on their website, and may be used under the Creative Commons CC BY 4.0 license[12].

The COVID Racial Data Tracker, part of the CTP, collected state-level data on cases and deaths separately by race and ethnicity for the purpose of examining the disproportionate impact of the pandemic on minority communities. The dataset contains twice weekly counts of the cumulative number of cases and deaths for each state for which information was available. Separate counts by racial group (White, Black, Latinx, Asian, American Indian/Alaskan Native, Native Hawaiian/Other Pacific Islander, Multiple, or Other) and ethnicity (Hispanic or Non-Hispanic) are also provided. Because information on the joint distribution of race and ethnicity was not provided, we utilized information on racial group only (in many cases, the counts for Latinx were the same as for Hispanic). Information on age is not available.

Our analyses of CTP data used cumulative counts as of February 28, 2021—the last full month for which data were provided. This allowed us to construct a corresponding dataset from the CDC data which are recorded on a monthly basis.

CDC COVID-19 case surveillance public use data

This dataset contains individual-level data on all cases reported to the CDC[13]. The public use dataset contains 19 data elements for each case, including the state and month in which the case was reported, whether the individual died as a result of COVID-19, and individual-level characteristics such as gender (Male or Female), age group (0–17, 18–49, 50–64, and 65+), race (White, Black, Asian, American Indian/Alaskan Native, Native Hawaiian/Other Pacific Islander, Multiple/Other) and ethnicity (Hispanic or Non-Hispanic). We define the three groups for our analysis as follows: White non-Hispanic, Black, and Hispanic non-Black. For comparability with the CTP dataset, we use all cases reported through February 2021.

Statistical analysis

CFR were computed by dividing the total number of deaths by the total number of cases reported as of February 28, 2021. Monthly rates for the CTP data were computed by dividing the total number of new deaths reported in that month by the total number of new cases. While some deaths reported in a month corresponded to cases reported from previous months, sensitivity analyses including cases from the 2–3 previous months yielded similar results. Monthly rates for the CDC data were computed as the proportion of cases reported in that month that resulted in death; this is therefore a leading indicator since some of those deaths occurred in subsequent months. The following states or territories were excluded from all analyses of the CDC data: nine states (AK, DE, HI, MO, NE, SD, TX, VI, WV) reported no deaths, with survival status missing for all or nearly all cases; three states reported either only deaths (WA), nearly all deaths (IL) or half as many deaths as cases (RI), also with survival status missing for a large fraction of cases; and one state (GA) and Puerto Rico were outliers in a log-log plot of deaths versus all cases (including those with missing survival status). These exclusions reduced our analyses of CDC data to 38 states.

Since the distribution of racial and ethnic groups differs across states, state-specific differences in the CFR may bias comparisons between groups. Thus, we calculated CFR ratios comparing Black and Latinx to White individuals separately by state and estimated an overall CFR ratio using a random effects model:

$$\hat{\theta}_j = \theta + u_j + \varepsilon_j$$

where $\hat{\theta}_j$ is the natural log of the CFR ratio for state j , the $u_j \sim N(0, \tau^2)$ represent between-state differences in the log CFR ratio, and the $\varepsilon_j \sim N(0, \hat{\sigma}_j^2)$ represent sampling variability. The between-state variance (τ^2) was estimated using the DerSimonian-Laird method[14], and an estimate of the overall CFR (θ) was obtained as a weighted sum of the $\hat{\theta}_j$; this estimate was then exponentiated to obtain a CFR ratio, together with the endpoints of the corresponding 95% Wald confidence interval. The I^2 measure of heterogeneity (representing the percentage of variability in the $\hat{\theta}_j$ due to between-state differences) is reported[15].

To estimate CFR ratios adjusting for differences in the CFR across time, gender, age group and state, we used the CDC data to compute the number of cases (c_{ijk}) and deaths (d_{ijk}) for each demographic subgroup i (based on the Cartesian product of age group, gender and racial and ethnic group), state j , and month k , and fit the following mixed-effects Poisson model to the data for White, Black, and Latinx individuals age 18 and older[16]:

$$\log(\mu_{ijk}) = \alpha + \log(c_{ijk}) + \mathbf{X}_i\boldsymbol{\beta} + f(k; \boldsymbol{\lambda}) + v_j + w_{jk}$$

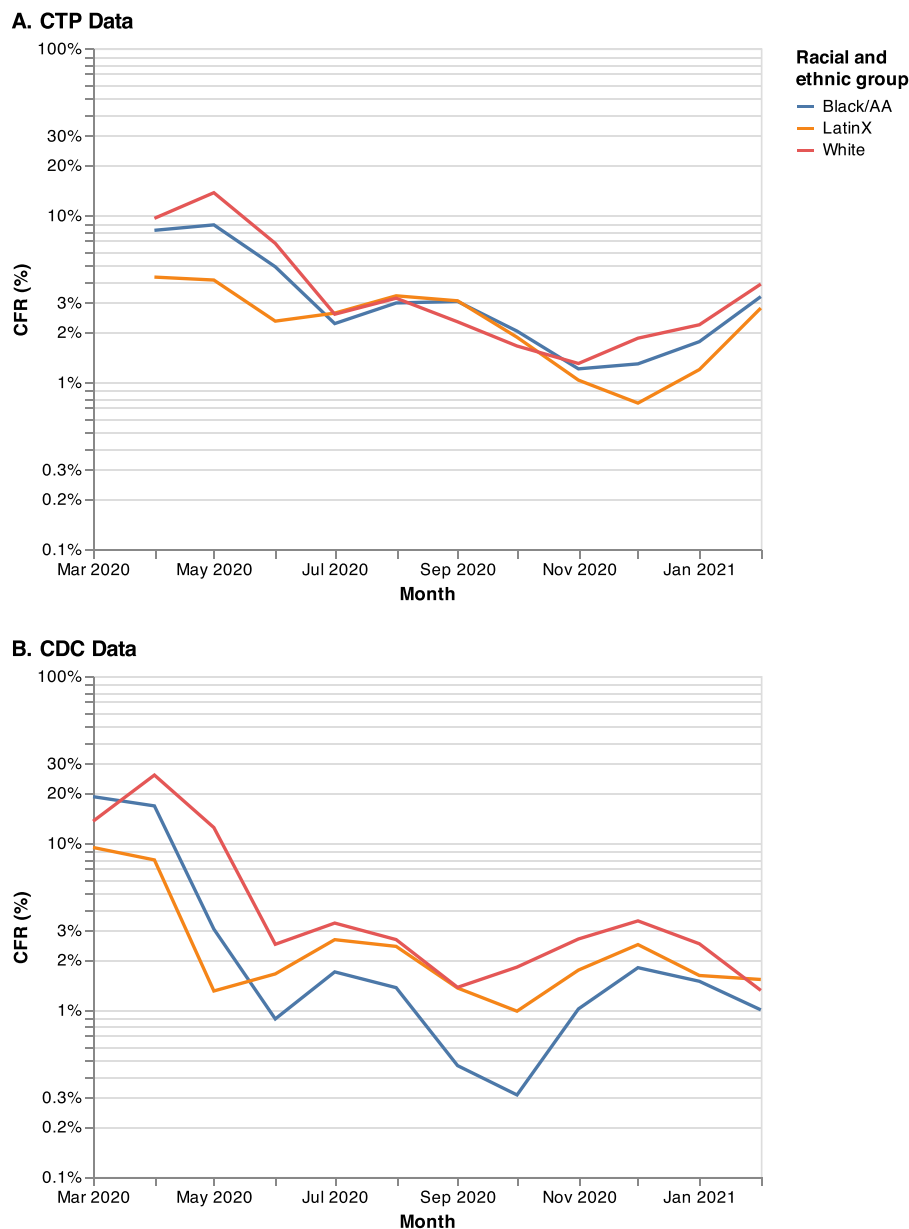


Fig. 1. CFR by month and racial and ethnic group, CTP and CDC datasets.

where μ_{ijk} is the expected number of deaths, \mathbf{X}_i is a vector of discrete covariates describing the i th group including age group, gender, racial and ethnic group and an interaction between age group and racial and ethnic group, and $f(k; \lambda)$ is a restricted cubic spline function of k depending on coefficients λ that captures the change over time (we used five knots based on Harrell’s recommended percentiles) [17]. The term ν_j is a random effect capturing differences between states, and the term w_{jk} is a random effect, nested within state, capturing remaining differences across time within state. Variance estimates were obtained using the clustered form of the robust (sandwich) variance estimator with clustering at the state level [18]. Estimated coefficients (β) were exponentiated to obtain CFR ratios, together with the endpoints of their corresponding 95% Wald confidence intervals. Estimates of the ν_j were obtained using empirical Bayes means and plotted on a map.

All analyses were performed using Stata version 17.0, Python, pandas and geopandas (for data management) and Altair (for plotting)[19]. Data and code are available on GitHub (<https://github.com/pschumm/covid-modeling>).

Results

The CTP dataset includes 28.4 million cases and 512,627 deaths as of February 28, 2021, corresponding to an overall CFR of 1.8% (Table 1). The CDC dataset contains 22.1 million cases over the same period, however survival status is missing for nearly half (10.4 million). Among the other 11.7 million cases, there were 263,398 deaths corresponding to an overall CFR of 2.3%. Removing those states and territories with problematic data (see Methods) left 235,635 deaths with an overall CFR of 2.0%. Both datasets show a similar pattern in the CFR over time, with the CDC curves running ahead of the CTP curves by approximately 1–2 months, as expected (Fig. 1). Based on the CDC data, the CFR decreased from the start of the pandemic through September/October 2020. This decrease was followed by an increase and a subsequent decline, consistent with the estimate of $f(\cdot)$ from the mixed-effects model when fit to the CDC data below (Fig. 2).

The CFR was substantially higher for those 65 and older (13.5%) as compared to those 50–64 (0.8%) and 18–49 (0.1%) (Table 1). No

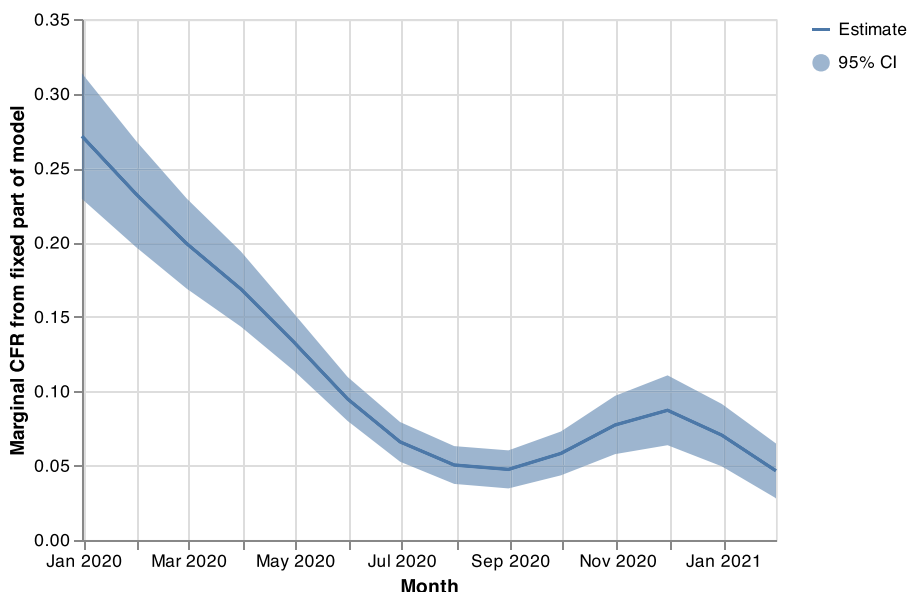


Fig. 2. Estimated overall (i.e., nationwide) time trend in CFR from mixed-effects model fit to CDC data, parameterized using a restricted cubic spline with 5 knots.

Table 1

Case fatality rates (CFR) for COVID tracking project (CTP) and CDC public use datasets through February 2021

	Deaths	Cases	CFR (%)
<i>CTP</i>			
Overall	512,627	28,443,555	1.8
Race and Ethnicity			
White*	268,373	10,309,363	2.6
Black/AA†	61,989	2,359,473	2.6
Latinx‡	71,844	3,666,785	2.0
Other	39,403	2,539,435	1.6
Missing	34,715	9,510,434	0.4
<i>CDC</i>			
Overall	263,398	11,700,999§	2.3
Post-filtering	235,635	11,661,626	2.0
Age group¶			
< 18	0	1,327,892	0.0
18–49	3,728	6,338,709	0.1
50–64	18,689	2,316,705	0.8
65+	213,077	1,580,396	13.5
Missing	141	97,924	0.1
Race and Ethnicity			
White	132,822	3,955,817	3.4
Black/AA	22,735	861,823	2.6
Latinx	26,191	1,158,313	2.3
Other	11,074	854,367	1.3
Missing	42,813	4,831,306	0.9

* Excludes the following states/territories for which counts of cases and/or deaths are not available: AS, MP, ND, NY, PR, and VI.
 † Excludes the following states/territories for which counts of cases and/or deaths are not available: AS, GU, MP, ND, NY, PR, and VI.
 ‡ Excludes 33 states/territories for which counts of cases and/or deaths are not available (23 states remaining).
 § Excludes 10,371,233 cases for which information on survival is not available.
 || Includes data from 38 states that remained after QC filtering.

deaths were reported in the CDC dataset for individuals under 18 years of age. In the CDC dataset, the CFRs for both Black and Latinx individuals were lower than for Whites throughout all months except March 2020 and February 2021 (Fig. 1). In the CTP dataset, CFRs for the three racial and ethnic groups were similar from July 2020 through November 2020, though Black and Latinx individuals still had lower CFRs than Whites before and after this period. Table 2 (Panel A) shows estimated CFR ratios comparing Black and Latinx individuals to Whites, after adjusting for differences in CFR

across states. In the CTP dataset, the CFR ratio comparing Black to White individuals was 0.8 (95% confidence interval, 95% CI [0.7, 0.9]) and comparing Latinx to White individuals was 0.4 (95% CI [0.3, 0.5]); the corresponding CFR ratios for the CDC dataset were even lower at 0.5 (95% CI [0.4, 0.6]) and 0.2 (95% CI [0.2, 0.3]), respectively.

The proportion of cases under age 65 was higher among Black (89%) and Latinx (93%) individuals than among Whites (80%). These differences were even larger for the proportion of cases under age 50 (69% for Black and 78% for Latinx individuals as compared to 58% for Whites). When stratifying by age group, the state-specific CFRs—for states reporting any deaths for a particular age by racial and ethnic group—among those aged 18–49 and 50–64 were higher, on average, among Black and Latinx individuals than among Whites (Fig. 3). The CFR ratios comparing Black to White individuals were 3.9 for ages 18–49, 2.1 for ages 50–64, and 0.9 for ages 65 and older; corresponding CFR ratios comparing Latinx to White individuals were 9.8, 3.0, and 0.9, respectively (Table 2, Panel A).

Results from the mixed-effects model fit to deaths reported in the CDC data are shown in Table 2 (Panel B). Consistent with the CFR ratios reported above, these results demonstrated higher CFRs for the Black and Latinx groups among those under age 65. Specifically, estimated CFR ratios comparing Black to White individuals were 6.0 (95% CI [3.4, 10.8]) for ages 18–49, 2.7 for ages 50–64 (95% CI [1.8, 4.1]), and 1.0 (95% CI [0.8, 1.1]) for ages 65 and older. The corresponding estimated CFR ratios comparing Latinx to White individuals were 4.1 (95% CI [1.9, 8.7]), 2.4 (95% CI [1.5, 3.8]), and 1.2 (95% CI [1.0, 1.4]), respectively. The CFR ratio comparing women to men was 0.8 (95% CI [0.7, 0.8]).

We observed considerable variation between states for both datasets, both in the CFRs and in the CFR ratios comparing racial and ethnic groups (Fig. 3). The value of I^2 was over 95% in all cases. The estimated variance of the v_j was 0.31, corresponding to an increase in the CFR of 75% for a state one standard deviation above the mean. Similarly, the estimated variance of the w_{jk} was 0.26, corresponding to a 67% increase in the CFR for a month one standard deviation above the mean; this variability within state over time is above and beyond that already accounted for by the estimated overall time trend $\hat{f}(\cdot)$ (Fig. 2). Fig. 4 plots estimates of the v_j for all 38 states included in the model.

Table 2

Case fatality rate (CFR) ratios comparing minority racial and ethnic groups to Whites, estimated from state-specific log CFR ratios using random-effects models (CTP and CDC) and based on a mixed-effects Poisson regression model fit to individual-level data (CDC only)*

	Black/AA compared to White			Latinx compared to White		
	No. states	Rate ratio	95% CI	No. states	Rate ratio	95% CI
<i>A. Random-effects models fit to state-specific log CFR ratios</i>						
<i>CTP</i>						
Overall	49	0.8	(0.7, 0.9)	23	0.4	(0.3, 0.5)
<i>CDC</i>						
Overall	26	0.5	(0.4, 0.6)	14	0.2	(0.2, 0.3)
<i>Age group</i>						
18–49	2 (CA & NY)	3.9	(0.3, 44.5)	3 (AZ, CA & NY)	9.8	(4.4, 21.9)
50–64	12	2.1	(1.2, 3.6)	9	3.0	(2.2, 3.9)
65+	26	0.9	(0.8, 1.1)	13	0.9	(0.8, 1.2)
<i>B. Mixed-effects Poisson regression model</i>						
<i>CDC</i>						
<i>Age group</i>						
18–49		6.0	(3.4, 10.8)		4.1	(1.9, 8.7)
50–64		2.7	(1.8, 4.1)		2.4	(1.5, 3.8)
65+		1.0	(0.8, 1.1)		1.2	(1.0, 1.4)

* Mixed effects model also adjusted for gender (estimated CFR ratio comparing women to men was 0.8 (95% CI [0.7, 0.8]) and chronological time (using a restricted cubic spline with 5 knots), and included random effects for state (estimated variance 0.31 (95% CI [0.19, 0.51]) and month within state (estimated variance 0.26 (95% CI [0.19, 0.37])).

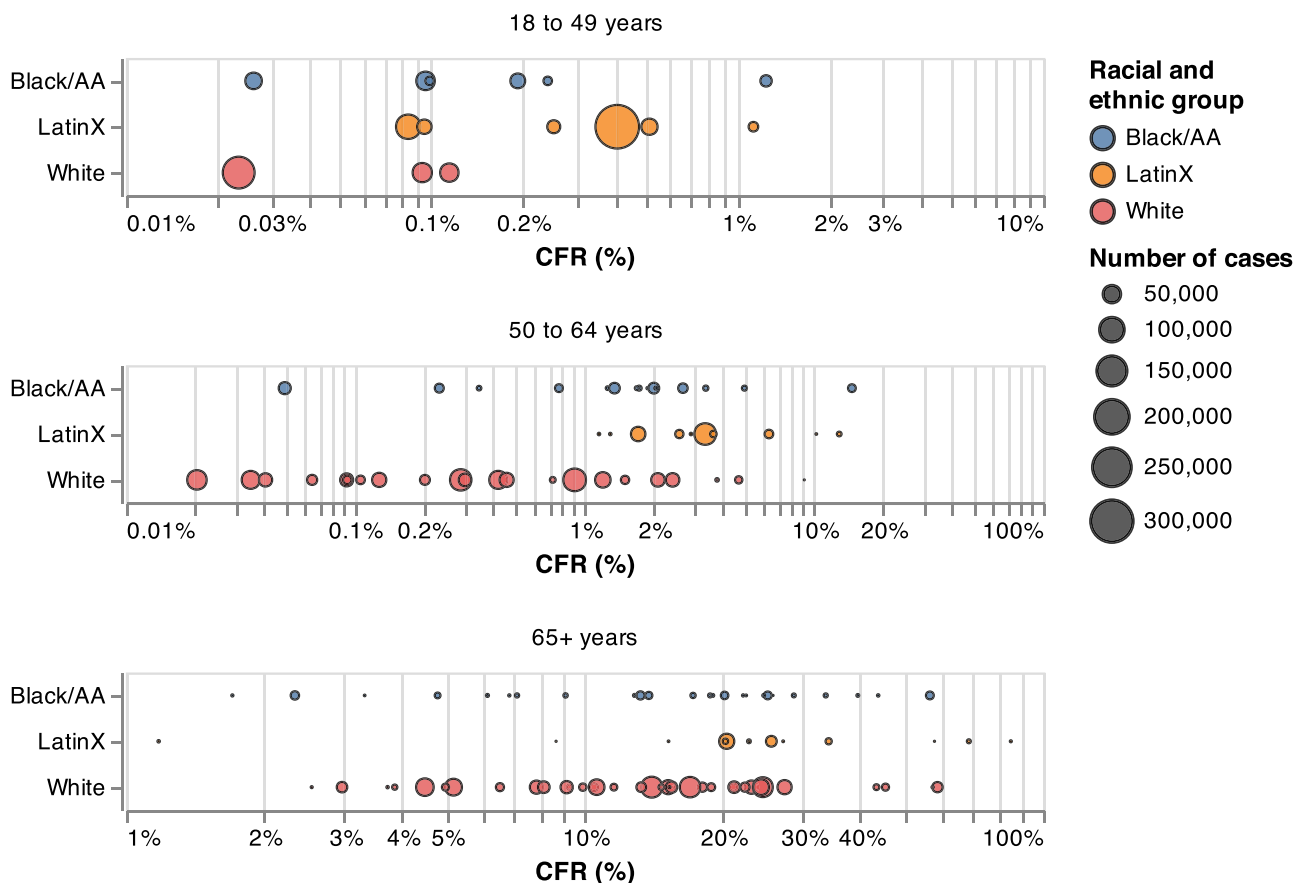


Fig. 3. State-specific CFRs, separately by age group and racial and ethnic group (CDC data). CFRs equal to 0 due to missing data are excluded.

Discussion

While the overall CFR was lower among Black and Latinx individuals than among their White counterparts, their CFRs within age category were higher among those under age 65. This reversal in the direction of the association when stratifying by subgroup is a partial example of what is known as Simpson’s Paradox[20], and

reinforces the point made by Green et al. that COVID-19 CFRs that are not age-specific “may hide more than they reveal.” This observation is in contrast to the CDC’s COVID-19 mortality data from the same period which reveal higher COVID-19 mortality rates among Black and Latinx individuals than among Whites both overall and separately by age group[21].

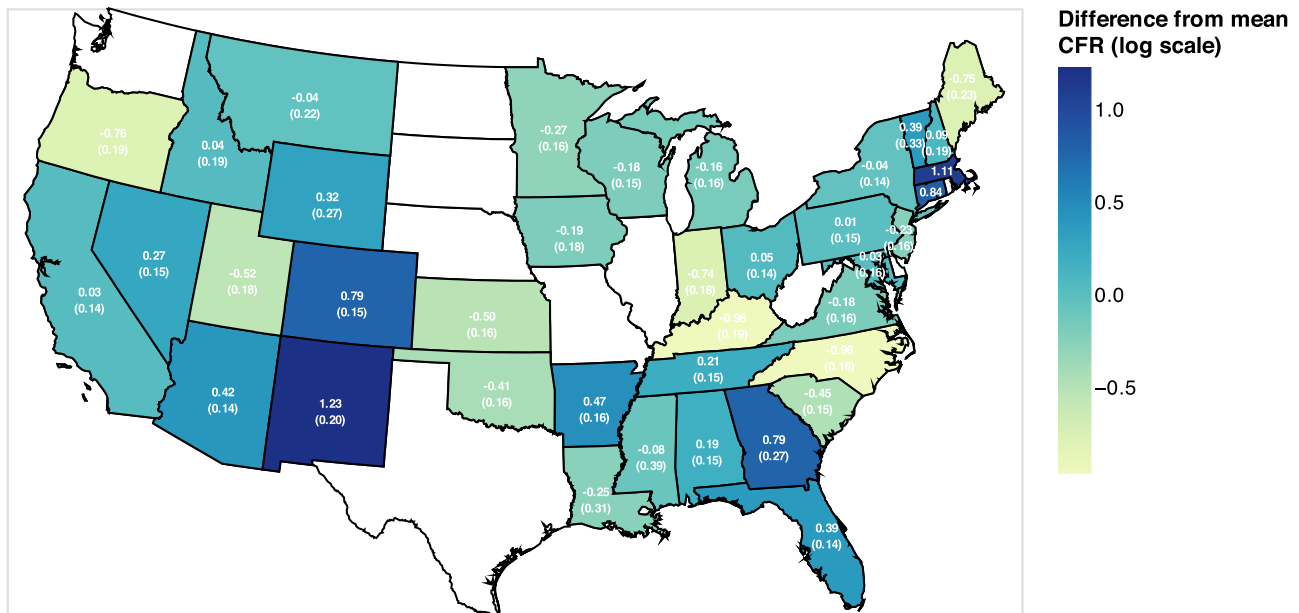


Fig. 4. Estimated differences in CFR across states (CDC data), SEs in parentheses; states not included due to missing data appear as white.

We also found evidence of an interaction between age and racial and ethnic group such that Black and Latinx disparities in COVID-19 CFR increased at each successively younger age group. A similar interaction was reported by Bassett et al.[22], for COVID-19 mortality rates from February 1 through May 20, 2020. This finding deserves further study as it could have substantial implications for health policy—especially for a disease where public attention concerning mortality, at least during the early stages of the pandemic, was primarily focused among the oldest ages.

Our results reflect differences in the *observed* CFR; that is, the number of deaths divided by the number of reported cases. For COVID-19, this differs substantially from the true, underlying CFR since many infected individuals, especially those who were asymptomatic, were never tested. Moreover, some individuals who tested positive may not appear in the datasets used here. Thus, we expect that much of the change over time, as well as of the substantial variation across states, reflects variation in the rates of testing and reporting. Several studies have found higher testing rates among Black and Hispanic individuals than among Whites, including Rentsch et al. (though the VA is a special population with lower healthcare barriers) and a study of roughly 50 million patients in the Epic health record system[5,23–25]. While there is no way to know how many of these individuals appear in the CTP and CDC datasets analyzed here, higher testing rates among Black and Hispanic individuals would be expected *ceteris paribus* to reduce their CFR relative to Whites. Furthermore, by adjusting for differences across time and between states, our results should be robust to confounding due to differences in the racial mix of cases over time or in the distribution of minorities across states.

The CDC dataset contains only 78% of the cases in the CTP dataset, and of these, information on survival is missing for 47% of cases. In addition, among the cases with survival status for the 38 states included in the analysis, race and ethnicity was missing for 41%. Douglas et al. reviewed reporting of race and ethnicity of COVID-19 cases and deaths from April 12, 2020 through November 9, 2020 and concluded that while there were improvements during this period, significant problems with data quality persisted[26]. Such problems might be expected to introduce both bias and additional variability into our results; indeed, potential variability in reporting of race and ethnicity across states is one

reason we conducted a state-level analysis. Despite these limitations, our results are consistent with prior literature showing improvements in COVID-19 survival over the first 6–8 months of the pandemic and a lower risk of death for women relative to men[27–29].

Finally, our study was also intended to highlight the limitations in data available on race and ethnicity of COVID-19 cases and deaths at the national level. Specifically, using information on the joint distribution of racial and ethnic group and age (available in the CDC dataset only) reveals an entirely different understanding than when not using it, and acquiring such data from representative samples of the national population is sorely needed. In addition, Chowkwanyun and Reed argue convincingly that while identifying racial and ethnic disparities in COVID-19 is important, a lack of context as provided by data on SES and place-based risk, together with appropriate consideration of the role of stress due in part to racial discrimination, “can perpetuate harmful myths and misunderstandings that actually undermine the goal of eliminating health inequities”[30].

References

- [1] Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. *JAMA* 2020;323(24):2466–7. doi:10.1001/jama.2020.8598.
- [2] Gold JAW, Rossen LM, Ahmad FB, Sutton P, Li Z, Salvatore PP, et al. Race, ethnicity, and age trends in persons who died from COVID-19 - United States, May–August 2020. *MMWR Morb Mortal Wkly Rep* 2020;69(42):1517–21. doi:10.15585/mmwr.mm6942e1.
- [3] Rossen LM, Ahmad FB, Anderson RN, Branum AM, Du C, Krumholz HM, et al. Disparities in excess mortality associated with COVID-19 - United States, 2020. *MMWR Morb Mortal Wkly Rep* 2021;70(33):1114–19. doi:10.15585/mmwr.mm7033a2.
- [4] Zelnor J, Trangucci R, Narahariseti R, Cao A, Malosh R, Broen K, et al. Racial disparities in coronavirus disease 2019 (COVID-19) mortality are driven by unequal infection risks. *Clin Infect Dis* 2020;72(5):e88–95. doi:10.1093/cid/ciaa1723.
- [5] Rentsch CT, Kidwai-Khan F, Tate JP, Park LS, King JT, Skanderson M, et al. Patterns of COVID-19 testing and mortality by race and ethnicity among United States veterans: a nationwide cohort study. *Zelnor J, editor. PLoS Med* 2020;17(9):e1003379. doi:10.1371/journal.pmed.1003379.
- [6] Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among Black patients and White patients with Covid-19. *N Engl J Med* 2020;382(26):2534–43. doi:10.1056/NEJMs2011686.
- [7] Ogedegbe G, Ravenell J, Adhikari S, Butler M, Cook T, Francois F, et al. Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in New York City. *JAMA Netw Open* 2020;3(12):e2026881. doi:10.1001/jamanetworkopen.2020.26881.

- [8] Griffith CJ, Morris TT, Tudball MJ, Herbert A, Mancano G, Pike L, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020;11(1):5749. doi:10.1038/s41467-020-19478-2.
- [9] Azar KMJ, Shen Z, Romanelli RJ, Lockhart SH, Smits K, Robinson S, et al. Disparities in outcomes among COVID-19 patients in a large health care system in California: study examines disparities in access and outcomes for COVID-19 patients who are members of racial and ethnic minorities and socioeconomically disadvantaged groups. *Health Aff* 2020 Published online May 21. doi:10.1377/hlthaff.2020.00598.
- [10] Pressman A, Lockhart S, Petersen J, Robinson S, Moreno M, Azar KMJ. Measuring health equity for ambulatory care sensitive conditions in a large integrated health care system: the development of an index. *Health Equity* 2019;3(1):92–8. doi:10.1089/heq.2018.0092.
- [11] Green MS, Peer V, Schwartz N, Nitzan D. The confounded crude case-fatality rates (CFR) for COVID-19 hide more than they reveal—a comparison of age-specific and age-adjusted CFRs between seven countries. Di Gennaro F, editor. *PLoS ONE* 2020;15(10):e0241031. doi:10.1371/journal.pone.0241031.
- [12] The COVID Tracking Project (website). <https://covidtracking.com>. Accessed December 24, 2021.
- [13] Centers for Disease Control and Prevention. COVID Data Tracker. Atlanta, GA: US Department of Health and Human Services, CDC; 2021, December 24. <https://covid.cdc.gov/covid-data-tracker>
- [14] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177–88. doi:10.1016/0197-2456(86)90046-2.
- [15] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21(11):1539–58. doi:10.1002/sim.1186.
- [16] Rabe-Hesketh S, Skrondal A. *Multilevel and longitudinal modeling using stata*. 4th ed. College Station, TX: Stata Press; 2022.
- [17] FEJr Harrell. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2nd ed. New York: Springer; 2015.
- [18] Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. *J Human Resources* 2015;50(2):317–72. doi:10.3368/jhr.50.2.317.
- [19] StataCorp. 2021. Stata: Release 17. Statistical Software. College Station, TX: StataCorp LLC.
- [20] Wagner CH. Simpson's paradox in real life. *Am Stat* 1982;36(1):46–8.
- [21] Shiels MS, Haque AT, Haozous EA, Albert PS, Almeida JS, García-Closas M, et al. Racial and ethnic disparities in excess deaths during the COVID-19 pandemic. March to December 2020. *Ann Intern Med* 2021;174(12):1693–9. doi:10.7326/M21-2134.
- [22] Bassett M.T., Chen J.T., Krieger N. The unequal toll of COVID-19 mortality by age in the United States: quantifying racial/ethnic disparities. HCPDS Working Paper Volume 19, Number 3, 2020. Available at: https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1266/2020/06/20_Bassett-Chen-Krieger_COVID-19_plus_age_working-paper_0612_Vol-19_No-3_with-cover.pdf (Accessed: 8 July 2022).
- [23] Rubin-Miller L, Alban C, Artiga S, Sullivan S. COVID-19 racial disparities in testing, infection, hospitalization, and death: analysis of epic patient data. KFF, September 16, 2020. <https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-racial-disparities-testing-infection-hospitalization-death-analysis-epic-patient-data/> [accessed 29.12.21].
- [24] Reitsma MB, Claypool AL, Vargo J, Shete PB, McCorvie R, Wheeler WH, et al. Racial/ethnic disparities in COVID-19 exposure risk, testing, and cases at the subcounty level in California. *Health Aff* 2021;40(6):870–8. doi:10.1377/hlthaff.2021.00098.
- [25] Escobar GJ, Adams AS, Liu VX, Soltesz L, Chen YI, Parodi SM, et al. Racial disparities in COVID-19 testing and outcomes: retrospective cohort study in an integrated health system. *Ann Intern Med* 2021;174(6):786–93. doi:10.7326/M20-6979.
- [26] Douglas MD, Respress E, Gaglioti AH, Li C, Blount MA, Hopkins J, et al. Variation in reporting of the race and ethnicity of COVID-19 cases and deaths across US States: April 12, 2020, and November 9, 2020. *Am J Public Health* 2021;111(6):1141–8. doi:10.2105/AJPH.2021.306167.
- [27] Griffith DM. Men and COVID-19: a biopsychosocial approach to understanding sex differences in mortality and recommendations for practice and policy interventions. *Prev Chronic Dis* 2020;17. doi:10.5888/pcd17.200247.
- [28] Kadri SS, Sun J, Lawandi A, Strich JR, Busch LM, Keller M, et al. Association between caseload surge and COVID-19 survival in 558 U.S. hospitals. March to August 2020. *Ann Intern Med* 2021;174(9):1240–51. doi:10.7326/M21-1213.
- [29] Nguyen NT, Chinn J, De Ferrante M, Kirby KA, Hohmann SF, Amin A. Male gender is a predictor of higher mortality in hospitalized adults with COVID-19. *PLoS ONE* 2021;16(7):e0254066. doi:10.1371/journal.pone.0254066.
- [30] Chowkwanyun M, Reed Jr AL. Racial health disparities and Covid-19—caution and context. *N Engl J Med* 2020;383(3):201–3.