Quantifying the Distribution of Racial Inequalities in COVID-19 Mortality in the United States

Lauren Woyczynski

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2020

Committee:

Emmanuela Gakidou
Laura Dwyer-Lindgren
David Pigott

Program Authorized to Offer Degree:
Global Health

©Copyright 2020
Lauren Woyczynski

University of Washington

Abstract

Quantifying the Distribution of Racial Inequalities in COVID-19 Mortality in the United States

Lauren Woyczynski

Chair of the Supervisory Committee:
Emmanuela Gakidou
Department of Health Metric Sciences

As of August 31, 2020, the COVID-19 pandemic has killed over 183,000 people in the United States (Institute for Health Metrics and Evaluation, 2020). The pandemic has disproportionately affected racial minorities across the United States. It is well established in the United States that, due to structural racism in many forms, racial minorities suffer large health disparities as compared to White populations. As the COVID-19 pandemic has unfolded, tracking these disparities is paramount for continuing to advocate for targeted interventions and increased resources for vulnerable populations. In this study, the distribution of racial inequalities in COVID-19 mortality is quantified, as well as the variation in disparities across age groups and locations. Using publicly available data from the US Centers for Disease Control (CDC), this study finds high and significant relative risks of mortality for racial and ethnic minorities as compared to the non-Hispanic White population in the United States. At the national level, the age-standardized risks total 3.46 (3.41-3.51) for the non-Hispanic Black group, 3.01 (2.97-3.07) for the Hispanic and Latino group, 2.48 (2.31-2.64) for the non-Hispanic American Indian or Alaska Native group, and 1.58 (1.53 – 1.63) for the Non-Hispanic Asian group. Relative risks vary across age groups, with especially high risks for younger age groups, and across locations, with significantly higher than the national relative risks in certain locations for each race/ethnicity group. This analysis makes clear the need for more racially-stratified data and continued monitoring of disparities over time as the pandemic continues.

I. STUDY OBJECTIVE AND SPECIFIC AIMS

This study is an analysis of racial and ethnic disparities in deaths due to the novel coronavirus (COVID-19) in the United States in 2020. The National Center for Health Statistics (NCHS) at the US Centers for Disease Control (CDC) collates and publishes state-level data on COVID-19 deaths that includes stratification by race and ethnicity, age, and location. Using this information, this study aims to understand the relative risk of death due to COVID-19 by racial and ethnic categorization.

The specific aims of this study are the following:

1. To utilize imputation methods to account for suppression of small death counts in CDC data.
2. To identify the variation in relative risk of COVID-19 mortality by race and location across the United States.
3. To assess the age structure related variation across relative risks by race for COVID-19 deaths.

II. BACKGROUND & SIGNIFICANCE

On February 29, 2020, the first death due to COVID-19 was reported in the United States. By the end of March, nearly 5,000 people had died of COVID-19 in the US (Institute for Health Metrics and Evaluation, 2020). Around this time, media reports of disparities among COVID deaths began to circulate. AP News reported that all 8 people who had died of COVID-19 in Milwaukee County were Black, despite being only 38% of the population. They also reported that “the majority of Milwaukee’s confirmed cases were concentrated in African American neighborhoods” (Bauer, 2020). In early April, the New York Times reported that “the coronavirus is infecting and killing black people in the United States at disproportionately high rates” (Eligon et al., 2020). At this point, most cities and states were not reporting case and death data by race, making it impossible to understand the full picture. The Atlantic, in collaboration with Boston University Center for Antiracist Research, published the COVID Racial Data Tracker on April 15th, spurred by a series of essays written by Ibram X. Kendi on the urgent need for collection and reporting of demographic data to fully understand the outbreak and protect vulnerable communities (Barry et al., 2020). On May 8th, 2020, after continuing pressure from activists and organizers, the CDC released its first iteration of the “Deaths involving coronavirus disease 2019 (COVID-19) by race and Hispanic origin group and age, by state” database (Data.CDC.gov, 2020). Dr. Rhea Boyd presented to a US Congress subcommittee for health in mid-June on the inequitable burden of COVID-19 on non-White populations. Along with a summary of evidence, she presented several paths forward, including mandated reporting of COVID-19 inequities by race and ethnicity, universal COVID-19 testing, workplace protections, continuation of federal relief programs, and addressing the needs of incarcerated populations (Boyd, 2020). At the state and local levels, some communities began reporting and attempting to combat disparities early on. For example, New Mexico launched interventions targeted at their

large Native American population in June to combat the disproportionate impact of COVID-19 on tribal communities (New Mexico Human Services Department, 2020). The New York Times later sued the CDC under the Freedom of Information Act, to acquire data on cases by race and ethnicity at the county level (Oppel Jr et al., 2020). Though race disparities in COVID-19 infections and outcomes have been known and felt by communities from early in the pandemic, it has taken considerable organizing to gain public access to timely and trusted demographic data on COVID-19 infections and mortality.

As data has become more available, some literature has been published that quantifies the racial disparities of COVID-19. A group at the Harvard Center for Population and Development Studies found that COVID-19 mortality rates were higher for all racial/ethnic minorities, with especially high rate ratios among younger adults. (Bassett et al., 2020). The Color of Coronavirus project and the COVID Racial Data Tracker, created by APM Research Lab and the Boston University Center for Antiracist Research, respectively, have published weekly updates analyzing the continuing disparities in COVID-19 cases and deaths (APM Research Lab Staff, 2020; Barry et al., 2020). Both of these resources have continually found that Black and Indigenous populations have experienced the heaviest losses due to COVID-19. A recently published paper in the Annals of Epidemiology found that while disproportionately black counties make up 20% of all US counties, they accounted for 52% of COVID-19 diagnoses and 58% of COVID-19 deaths nationally (Millett et al., 2020). The rate of COVID-19 cases was shown to be higher in counties with a higher proportion of Hispanic/Latino or Black residents after adjustment for sociodemographic characteristics, comorbidities, and socioeconomic determinants. (Rodriguez-diaz et al., 2020). Though ecological analysis of county demographic and COVID-19 cases is possible at the sub-state level with current data, racially-disaggregated mortality data has not yet been made available at the county level across all states as of August 2020.

It is well understood at this point that COVID-19 has disproportionally affected Black, Indigenous, and People of Color (BIPOC) in both cases and deaths. However, data has only been made available in certain combinations of critical stratifying variables. County-level data that is stratified by race allows ecological analysis of comorbidities and social and environmental data within a county. However, distribution of these county-level factors likely is heterogeneous and differs significantly by race/ethnicity. To fully understand the forms of racism that affect COVID-19 outcomes in the United States, we must also understand both “the mechanisms by which it may be operating, and other intersecting forms of oppression (such as based on sex, sexual orientation, age, regionality, national, religion, or income)” (Boyd et al., 2020). Many researchers and activists continue to advocate for data that includes information on these variables. This analysis will quantify the existing disparities in COVID-19 outcomes by race and ethnicity using the most granular data available to show both the ubiquity of disparities in the US and the heterogeneity in the magnitude of disparities. This analysis will also suggest key data gaps and present suggestions for data publication that will aid in understanding the mechanisms

that create the disparities quantified here. In this study, the intersections of racism, age, and geography in relation to COVID-19 mortality will be studied in an attempt to quantify and explain facets of the racial disparities of COVID-19 mortality in the United States.

III. METHODS

This study is a population-level analysis of death certificate data. The goal of the analysis is to understand racial health inequalities in COVID-19 mortality in the United States. Age and geographic location will be studied to understand the range of disparities present, and where specific intervention is most urgently needed.

A. DATA SOURCES

The NCHS at CDC publishes weekly updates on COVID-19 deaths data stratified by selected demographic characteristics(Data.CDC.gov, 2020). NCHS uses death certificate data to tabulate provisional death counts for all 50 states and the District of Columbia. Deaths due to COVID-19 are identified using ICD-10 code U07.1. When this code is listed as the underlying cause of death or listed as "probable" or "presumed", the deaths is counted as a COVID-19 death. Laboratory confirmation of COVID-19 is not required. State-based reporting of death data to the NCHS causes lags in the database. The NCHS website lists that "currently, 60% of all US deaths are reported within 10 days of the date of death, but there is significant variation between states." Due to this lag and the manual time needed to code COVID-19 deaths, the database lists a potential lag of 1-8 weeks from real time sources with an average of a 1-2-week lag. The data presents cumulative deaths by state, age, and race and Hispanic origin group, since 02/01/2020. The dataset used in this analysis was downloaded on 08/03/2020. The race/ethnicity data is reported in the following categories:

- Hispanic or Latino
- Non-Hispanic White
- Non-Hispanic Black
- Non-Hispanic American Indian or Alaska Native (AIAN)
- Non-Hispanic Asian
- Non-Hispanic Native Hawaiian or Other Pacific Islander (NHOPI)
- Non-Hispanic More than one race

Death certificate data comes from the National Vital Statistics System (NVSS). Demographic information, including race and ethnicity, is recorded by a funeral director in NVSS data, unlike Census data, where race and ethnicity are self-reported. In a 2016 CDC report, agreement in death certificate and census classification of race and ethnicity is almost 100% for White and Black populations (Arias et al., 2016). However, only 51% of decedents who self-identified as American Indian or Alaska Native (AIAN) were classified as such on their death certificate. Similarly, only 84% of Asian or Pacific Islander (API) decedents and 92% of Hispanic of Latino decedents were identified correctly. These biases in death certificate race classifications, as well

as small population size in some locations, creates difficulties in accurately measuring mortality by race.

Age is reported by 10-year age groups from 5-85, with additional "Under 1", "1-4 years" and "85 years and above" categories. Entries that include 1-9 deaths are suppressed in the dataset due to NCHS confidentiality standards.

State-level population data stratified by race and Hispanic origin was obtained from the 2019 Census Bureau estimates (United States Census Bureau, 2020).

B. ANALYSIS METHODS

1. Imputation of Suppressed Counts

The NCHS data used in this analysis suppresses counts between 1-9 deaths in an age/race/location category due to NCHS confidentiality standards. Figure 1 shows the interval censoring patterns among all age/race/location groups. Missingness is high among young adult age groups and among race/ethnicities that have small populations. Importantly, COVID-19 has lower risk of death for younger age groups, so accumulating the number of deaths to surpass the suppression threshold takes longer. This censoring exists across many CDC data sets, including CDC WONDER (Wide-Ranging Online Data for Epidemiologic Research). For rare outcomes, suppression can be high, and ignoring suppression due to small counts leads to biased inference. Tiwari et al. addresses this suppression by suggesting an algorithm for estimating age-standardized rates in which suppressed age-specific counts are replaced with estimates based on the county's age-specific population size and the state-wide average for that age group (Tiwari et al., 2014). This model is straightforward but presents several limitations. Tiwari et al.'s solution does not allow for any within-state heterogeneity across race/ethnicity groups among imputed data points. It also does not limit the imputed death count to be between 1 and 9, so in states with large between-race heterogeneity, the estimate can be implausible.

Other potential approaches to this problem involve more complex Bayesian models that account for the spatial structures in the data, such as an multivariate conditionally autoregressive (MCAR) model (Quick, 2019). These estimates generally provided highly smoothed results. For the application of COVID-19, smoothing across spatial dimensions may be inappropriate due to the potential importance of state-level policy on COVID-19 related outcomes.

For this dataset, only certain unsuppressed aggregations of the data are available. State-level cumulative mortality data is available from the CDC by age groups across all races. We do not have aggregate information across race/ethnicity groups at the state level, national level, or across all age groups. This lack of information limits the possible imputation methods. In this analysis, the assumption is made that to get an age-standardized relative risk by race at the national level, interval censored data between 1 and 9 is not of high importance, and the midpoint of 5 can be used. However, when looking at disparities at the state and age group level, these imputations are much more important. In this analysis, the national race- and age-specific relative risk will be used to impute missing state level race- and age-specific mortality rates. So,

given a state *s*, $y_{s,a}$ is the known total death count across all race/ethnicity categories in a specific age group. For example, state *s* and age group *a* may have unsuppressed values for races $j = 1, \dots, n$. The remaining suppressed counts can then be imputed as follows:

$$\sum_{i=1}^{n} y_{a,s,r_i} = y_{s,a} - \sum_{i=1}^{n} y_{a,s,r_j} \quad (1)$$

Then, for each paired combination of missing y_{a,s,r_i}s:

$$\frac{y_{a,s,r_1}}{pop_{a,s,r_1}} \times RR_{a,r_1} = \frac{y_{a,s,r_2}}{pop_{a,s,r_2}} \times RR_{a,r_2} \quad (2),$$

where RR_{a,r_i} is the national race- and age-specific crude relative risk. The system of equations formed by equation (1) and equation (2) is solved to impute the missing y_{a,s,r_i}. This method assumes that the national-level relative risk for a specific age group is sufficient for estimating the mortality rate at the state level for an age/race category. This in turn implies that the relationship between races and across locations is more similar than the relationship between races across age groups.

Crude relative risks are calculated using deaths counts and population estimates. Relative risks are calculated using Non-Hispanic White as the reference group. Non-Hispanic White is the largest population group in the United States. Therefore, it is logical for disparities to be measured in comparison to the majority group. Both age-specific and age-standardized relative risks are estimated. Due to small death counts, ages under 35 were aggregated to a single age group.

The gamma method for directly standardized rates was used to calculate the confidence intervals for the age-standardized mortality rates (Fay & Feuer, 1997). No adjustment was made to account for multiple testing.

To calculate the confidence interval of the relative risk, the following formulas were used:

$$se\left(\frac{X}{Y}\right) = \sqrt{\frac{\mu_x^2}{\mu_y^2} \times \left(\frac{se_x^2}{\mu_x^2} + \frac{se_y^2}{\mu_y^2}\right)} \quad (3)$$

$$CI\left(\frac{X}{Y}\right) = \frac{X}{Y} \pm 1.96 \times se\left(\frac{X}{Y}\right) \quad (4)$$

IV. RESULTS

An analysis of the crude relative risks highlighted the importance of age structure to the relationship between race/ethnicity and COVID-19 mortality. Figure 2 shows the risk ratios for the US by age group and race/ethnicity category, with Non-Hispanic White as the reference category.

Though the NCHS data used in this analysis presents seven race/ethnicity categories, only Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Asian, Non-Hispanic American Indian or Alaska Native, and Hispanic or Latino categories will be presented here. Only 1.3% of non-zero entries in the dataset for the Non-Hispanic Native Hawaiian or Other Pacific Islander group and 6.1% of the Non-Hispanic More than one race group are unsuppressed data. Due to high suppression, these categories are omitted from the results shown. When run in the analyses, no results for these two groups were statistically significant.

At the country level, the all-age and age-standardized relative risk of COVID-19 death are presented in Table 1, with Non-Hispanic White as the reference group. The national age structure from the 2019 census estimate is used here as the age standard.

	All-age Relative Risk	Age-standardized Relative Risk
Non-Hispanic Black	1.964 (1.935 – 1.993)	3.461 (3.410 – 3.512)
Non-Hispanic American Indian or Alaska Native	1.364 (1.279 – 1.449)	2.480 (2.318 – 2.643)
Non-Hispanic Asian	0.866 (0.850 – 0.905)	1.577 (1.527 – 1.627)
Hispanic or Latino	1.178 (1.159 – 1.197)	3.018 (2.967 – 3.069)

Table 1. US Relative Risks for COVID-19 mortality by race.

Though the highest proportion of deaths take place in older age groups, Figure 2 shows that the disparities are highest among younger age groups. All age groups across the four race/ethnicity categories presented here show an increased risk of death from COVID-19 for non-White populations. In the United States, the population age structure differs by race. All-age relative risk estimates may be heavily influenced by the underlying age structure within a race/ethnicity group and it is important in this context to provide both age-standardized and all-age relative risk estimates.

In Figure 3, age-standardized relative risks by race/ethnicity category and location are shown. Though most states show increased risk of COVID-19 death for non-White populations, there is large geographic heterogeneity in the magnitude of the relative risk. For the American Indian or Alaska Native group (AIAN), the risks range from nearly 1 to up to 26. The four states with high relative risks for AIAN populations correspond to states with large Native American populations, including Mississippi, New Mexico, Arizona, and Wyoming. The risks are significantly above 1 for Non-Hispanic Black populations across most locations, apart from several low-population states, including West Virginia, Iowa, Washington, New Mexico, Oregon, Nebraska, and Oklahoma. Age-standardized risks significantly greater than 1 for Non-Hispanic Black populations range from 6.57 (6.01 – 7.14) in Wisconsin to 1.04 (1.02 – 1.07) in New York. The risks are not as uniform in their direction across locations for the Hispanic or Latino and Non-

Hispanic Asian categories. In summary, 28 of 51 states had relative risks significantly greater than one (without adjustment for multiple testing) among Hispanic or Latino, 8 of 51 for AIAN, 14 of 51 for Non-Hispanic Asian, and 34 of 51 for Non-Hispanic Black.

Many of the state-specific relative risks reflect the overall national trend. Figure 4 shows differentiation from the national-level relative risk. In the Non-Hispanic Black group, states that show much higher relative risks than the national risk are centered largely in the Midwest, including statistically significant higher risks in Wisconsin, Michigan, Minnesota, and Missouri. For the Non-Hispanic American Indian or Alaska Native group, disparities are largely centered in three states: New Mexico, Wyoming, and Mississippi. These states all have large populations in this category, compared to the rest of the United States. Analyzing these spatial trends can provide further insight into the drivers of these disparities.

V. DISCUSSION

Here, we discuss the race-associated differences in COVID-19 mortality, but do not study the underlying drivers of these differences. We see that at the state level, the magnitude of these disparities is heterogeneous. To make this research actionable, more emphasis must be placed on the drivers of these disparities as well as the causes of substantial subnational variation. Race is generally thought to measure a "combination of social class, culture and genes". However, race only serves as a proxy for these variables. Race is an assigned, contextual variable, a social classification in a country that is race-conscious (Jones, 2001). "Although ethnicity reflects cultural heritage, race measures a societally imposed identity and consequent exposure to the societal constraints associated with that particular identity", therefore, "race is an excellent measure of exposure to racism" (Jones, 2001). Phelan and Link conclude that racial inequalities in health exist primarily because of racism's effect on socio-economic status, though there is also strong evidence that racism, through inequalities in "power, prestige, freedom, neighborhood, and healthcare" also has a fundamental association with health, independent of socio-economic status (Phelan & Link, 2015). Pirtle suggests several mechanisms through which racial capitalism is a fundamental cause of socioeconomic inequalities within COVID-19, including pre-existing comorbidities, risk factors such as health care access and flexible access to resources (Pirtle, 2020). Camara Phyllis Jones suggests three categories of racism that directly affect health: institutional racism, personally mediated (interpersonal) racism, and internalized racism (Jones, 2001).

Institutional racism is defined as "differential access to the goods, services, and opportunities of society by race". Institutional racism could affect COVID-19 in many ways. Access to healthcare and testing sites could cause differential COVID-19 outcomes. Access to resources and services that effect exposure such as essential work status and available transportation methods could also explain racial disparities in outcomes (Gould & Shierholz, 2020). Though information on health insurance coverage and job status is available by race/ethnicity through the American Community Survey, to fully assess the impact of these mechanisms through COVID-19, this information must be made available for COVID-19 patients specifically. Alternatively, COVID-19 outcomes could be reported at a granular spatial resolution, such as the county level, or census

block group level in urban areas, where American Community Survey data can be more reasonably used as a proxy.

Interpersonal racism is defined as “prejudice and discrimination, where prejudice is differential assumptions about the abilities, motives, and intents of others by race, and discrimination is differential actions towards others by race” (Jones, 2001). Differential treatment within the healthcare system is one of the main pathways for this mechanism.

The last mechanism is internalized racism, which is defined as “acceptance by members of the stigmatized races of negative messages about their own abilities and intrinsic worth” (Jones, 2001). This can manifest as adoption of risky health behaviors. Over time, these behaviors can result in comorbidities that make one more susceptible to death due to COVID.

The Resolve to Save Lives published a report titled “Tracking COVID-19 in the United States: From Information Catastrophe to Empowered Communities” which outlines the current landscape of data availability in the US and lays out a framework of 15 essential indicators that states need to report. They suggest that 11 of these, listed in Table 2 below, must be stratified by age, sex, race and ethnicity to “understand disease spread and adequately target interventions to control it” (Resolve to Save Lives & Vital Strategies, 2020). Currently, only 2% of exact essential indicators are reported. 38% of indicators were reported in some way, but without adequate stratifications. 60% of indicators were not reported at all. Reporting these essential indicators with stratifications by demographic variables will be key for targeting interventions towards the correct areas for most marginalized populations.

Indicator	Stratification
New confirmed and probable cases and per capita rates by rate with 7-day moving average	Age, sex, race, ethnicity, and zip code Outbreaks vs community
Percentage of new cases epidemiologically linked to at least one other case, stratified by whether part of known outbreak or not, with threshold	Age, sex, race, ethnicity Outbreaks vs community
New screening (e.g. antigen) and diagnostic (e.g. PCR) testing per capita rates by date, with threshold, with 7-day moving average	Age, sex, race, ethnicity
Percentage of screening (e.g. antigen) and diagnostic (e.g. PCR) tests positive by date, with threshold, with 7-day moving average CLI	Age, sex, race, ethnicity
COVID-19 daily hospitalization per capita rates and 7-day moving average	Age, sex, race, ethnicity
New COVID-19 confirmed and probable deaths and per capita rates with 7-day moving average	Age, sex, race, ethnicity, and zip code Outbreaks vs community
Diagnostic (e.g. PCR) test turnaround time (specimen collection to test report), by week	Age, sex, race, ethnicity

Time from specimen collection to isolation of cases, by week	Age, sex, race, ethnicity
Percentage of cases interviewed for contact elicitation within 48 hours of case specimen collection, including all people with positive tests who reside in the jurisdiction, by week	Age, sex, race, ethnicity
New infections among health care workers not confirmed to have been contracted outside of the workplace, by week	Age, sex, race, ethnicity

Table 2. Resolve to Save Lives Essential Indicators for Effective COVID-19 Response (Resolve to Save Lives & Vital Strategies, 2020)

As shown in the results, the relative risks of death due to COVID vary substantially for a single race/ethnicity across states. The drivers of these disparities are currently unknown and not addressed in this paper. Importantly, federal guidance on COVID-19 mandates and closures have been sparse. Therefore, state-level government responses have been varied on stay-at-home orders, school closures, eviction moratoriums, and mask mandates, for example. There has also been wide variation across states in the scale-up of available testing and contact tracing. Lack of government support at the state level could be exacerbating existing current racial health disparities by not providing safety measures for the most vulnerable populations. State and local governments should be continually assessing disparities. Disparities were reported in Milwaukee early in the pandemic (Bauer, 2020). This analysis shows that these disparities persist at the state level in Wisconsin. Although New Mexico has targeted health interventions towards the Native American populations, the disparities remain stark (New Mexico Human Services Department, 2020). State-level governments should monitor continued disparities across race groups and continue to gather and analyze stratified demographic data to create more focused interventions.

The results show almost universally higher risks for COVID-19 deaths for non-White populations across all states. There are many states where significant differences from the non-White group are not shown. The lack of statistical significance does not necessarily indicate that racism does not impact the COVID-19 mortality in these populations but could signify a lack of statistical power to detect a difference since the outcome is still rare. Structurally, accurate data generation may itself be racialized. For the Non-Hispanic American Indian or Alaska Native group especially, four states show relative risks that are significantly below one. Due to the bias caused by high levels of incorrect matching of death certificate to self-reported race/ethnicity category in this group, these results are uninterpretable. Excess mortality due to COVID-19 is also of interest. By only measuring direct deaths due to COVID-19, indirect deaths through excess mortality due to COVID-19 are not taken into account, potentially masking a greater true disparity.

There are many important related analyses to pursue in the future. The NCHS data has a significant potential lag of 1-8 weeks. Because this lag is not specified in the data, inter-state comparisons may not be comparing data from the same time period. State-level dashboards are generally updated more quickly. Comparison of NCHS data to state-level dashboards can provide a robustness check against potential bias caused by the NCHS data lag. Time series

analysis would be an important addition to the literature. Understanding how these disparities change over time will be key to understanding what interventions are effective for what populations. Currently, only cumulative data is available across all states. As data becomes available on incidence and testing, analysis of case fatality will be important to understand how disparities may differ between exposure and incidence versus mortality.

VI. CONCLUSION

It is clear that the COVID-19 pandemic has disproportionately burdened racial and ethnic minorities in the US. The disparities are stark and exist across almost all age groups and locations. The non-Hispanic Black, American Indian and Alaska Native, and Hispanic and Latino groups all have experienced death due to COVID that far surpasses their representation in the US population, as compared to the White population. Institutional, interpersonal, and internalized racism all play a role in the mechanisms causing these disparities. Iteratively published, stratified, and high-resolution data on COVID-19 related outcomes must be made publicly available to continue to track disparities and further solidify the areas in which targeted interventions can reduce the burden of COVID-19 on racial and ethnic minorities in the Unites States.

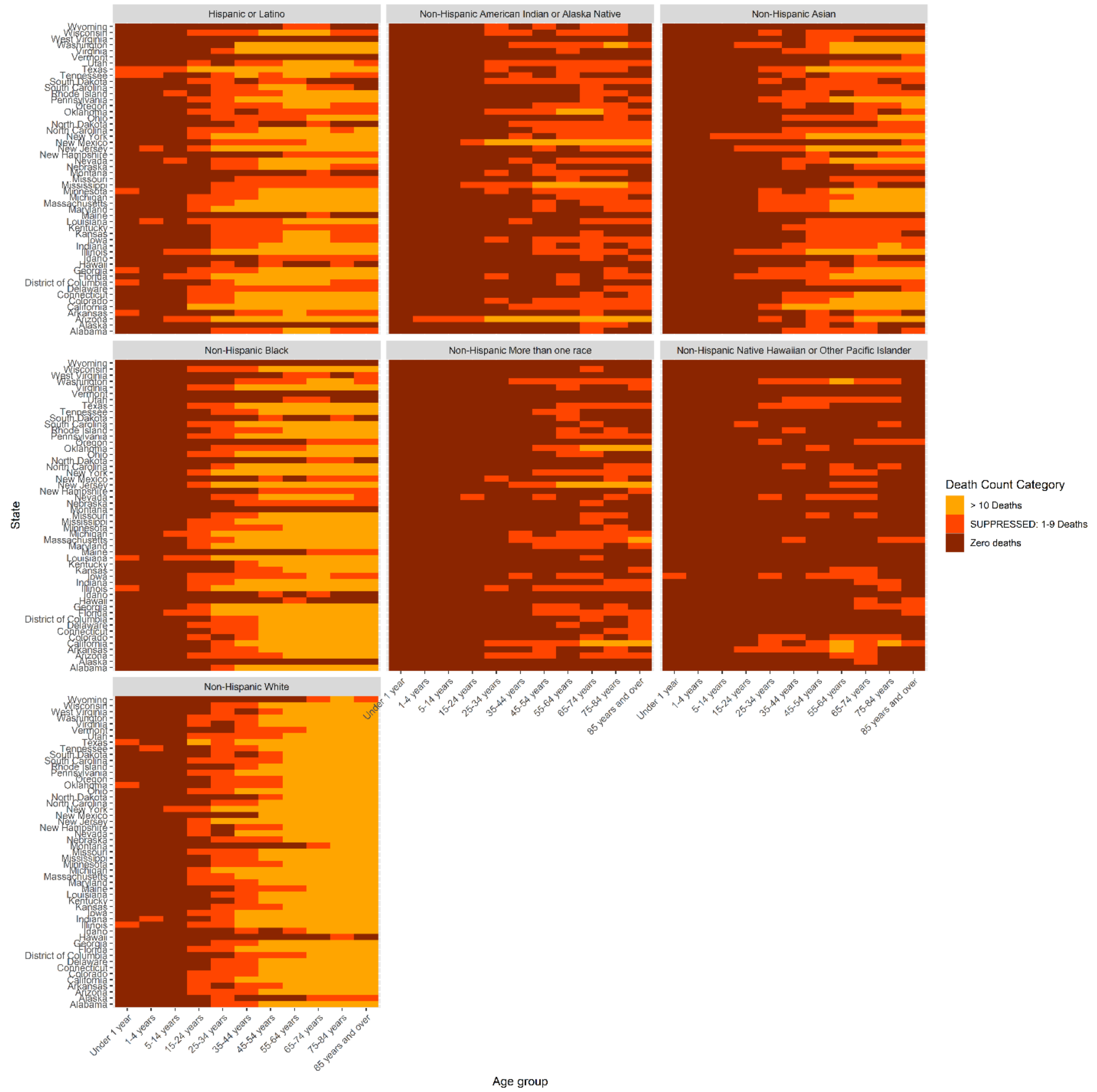

Figure 1. Heat map showing data availability by age group, race/ethnicity, and state. Data that is interval censored and suppressed for confidentiality is shown in orange.

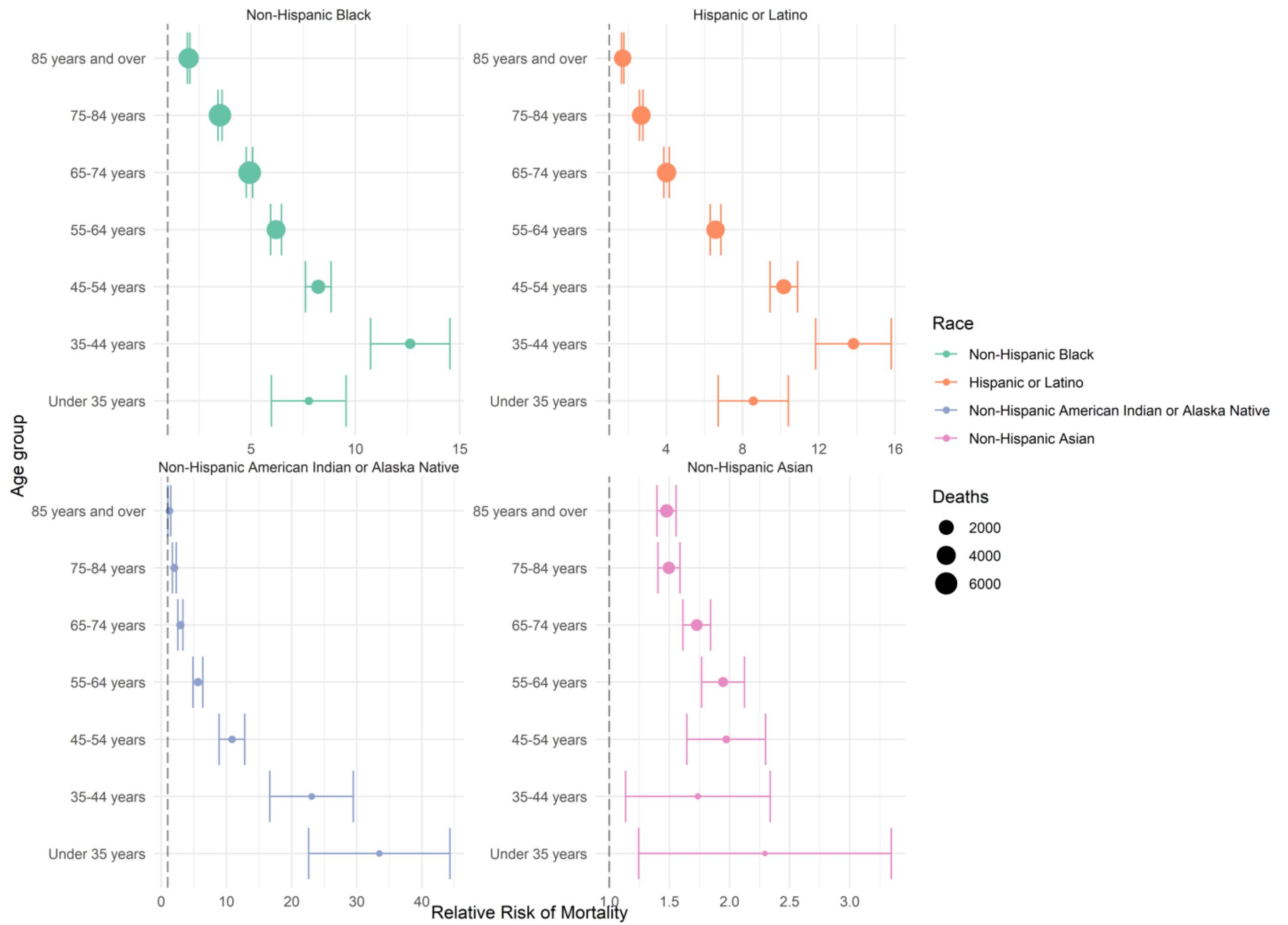

Figure 2. Age-specific relative risks across the US

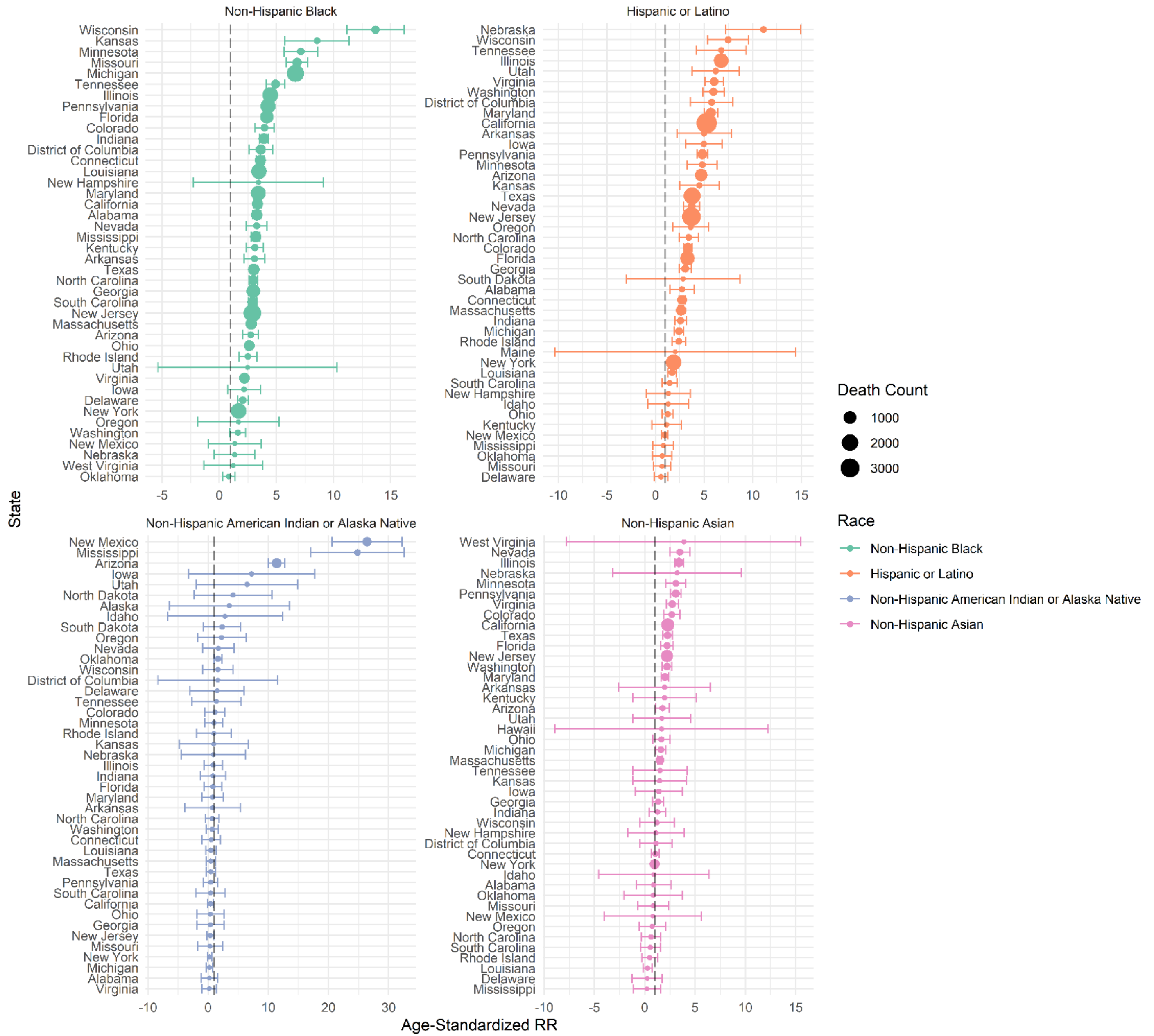

Figure 3. Age-standardized relative risks by state

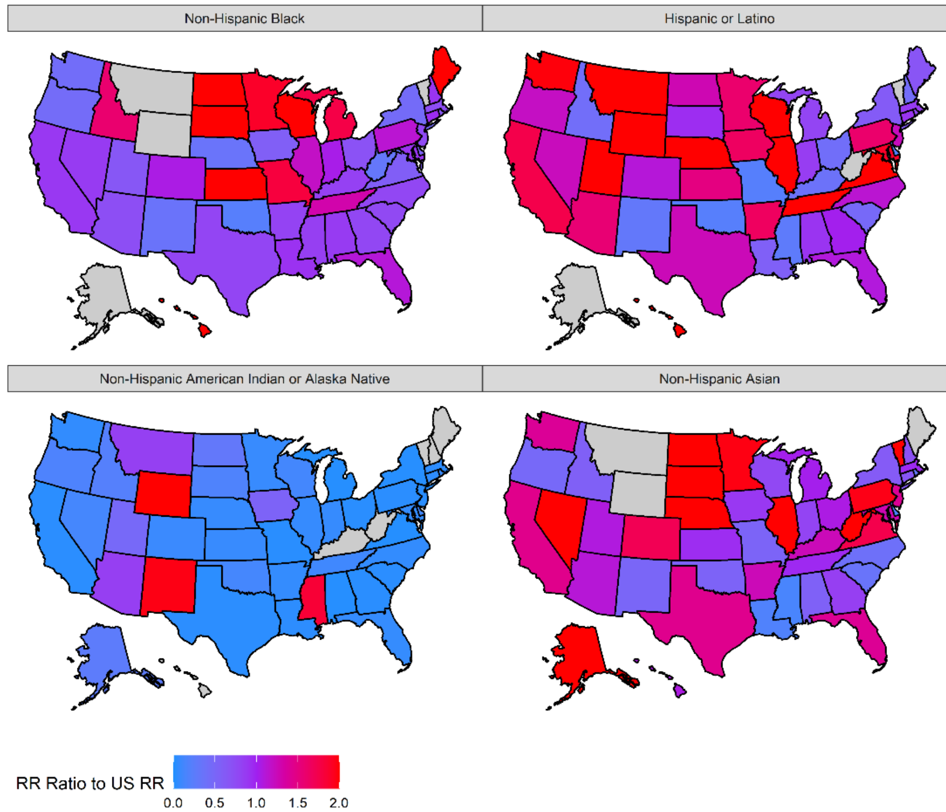

Figure 4. Map of difference to overall US RR to state-specific RR. Each panel is a race/ethnicity category. States in gray have had zero deaths in that race/ethnicity category.

VII. REFERENCES

APM Research Lab Staff. (2020). *The Color of Coronavirus: COVID-19 Deaths by Race and Ethnicity in the U.S.* https://www.apmresearchlab.org/covid/deaths-by-race

Arias, E., Heron, M., & Hakes, J. (2016). The Validity of Race and Hispanic-origin Reporting on Death Certificates in the United States: An Update. *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research*, *172*, 1–21.

Barry, M., Kendi, I. X., Lee, R., Nunez, A. G., Pretrella, C., & Simes, J. (2020). *COVID Racial Data Tracker*. https://covidtracking.com/race/

Bassett, M. T., Chen, J. T., & Krieger, N. (2020). *The unequal toll of COVID-19 mortality by age in the United States: Quantifying racial/ethnic disparities*. *19*(3). https://doi.org/10.20955/r.85.67

Bauer, S. (2020, March 27). Milwaukee's black community hit hard by coronavirus. *AP News*.

Boyd, R. W. (2020). *The injustice of Inequitable Disease: Addressing Racial Health Inequities amid the COVID-19 Pandemic*. https://www.congress.gov/116/meeting/house/110812/witnesses/HHRG-116-IF14-Wstate-BoydR-20200617.pdf

Boyd, R. W., Lindo, E. G., Weekd, L. D., & McLemore, M. R. (2020). *Inequities, On Racism: A New Standard For Publishing On RAcial Health*. Health Affairs Blog. https://doi.org/10.1377/hblog20200630.939347

Data.CDC.gov. (2020). *Deaths involving coronavirus disease (COVID-19) by race and Hispanic origin group and age, by state, NCHS. Updated May 2020. Data provided by the National Center for Health Statistics.* https://data.cdc.gov/NCHS/Deaths-involving-coronavirus-disease-2019-COVID19/ks3g-spdg

Eligon, J., Burch, A. D. S., Searcey, D., & Oppel Jr, R. A. (2020, April 7). Black Americans Face Alarming Rates of Coronavirus Infection in Some States. *New York Times*. https://www.nytimes.com/2020/04/07/us/coronavirus-race.html

Fay, M. P., & Feuer, E. J. (1997). Confidence Intervals for Directly Standardized Rates: A Method Based on the Gamma Distribution. *Statistics in Medicine*, *16*(December 1995), 791–801.

Gould, E., & Shierholz, H. (2020). *Not everybody can work from home: Black and Hispanic workers much less likely to be able to telework*. Working Economics Blog.

Institute for Health Metrics and Evaluation. (2020). *COVID-19 Projections*. https://covid19.healthdata.org/united-states-of-america

Jones, C. P. (2001). Invited commentary: "Race," racism, and the practice of epidemiology. *American Journal of Epidemiology*, *154*(4), 299–304. https://doi.org/10.1093/aje/154.4.299

Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., Honermann, B., Lankiewicz, E., Mena, L., Crowley, J. S., Sherwood, J., & Sullivan, P. S. (2020). Assessing differential impacts of COVID-19 on black communities. *Annals of Epidemiology*, *47*, 37–44. https://doi.org/10.1016/j.annepidem.2020.05.003

New Mexico Human Services Department. (2020). *NM Outreach Campaign to Address COVID-19 in Tribal COmmunities*. https://www.newmexico.gov/2020/06/16/nm-outreach-campaign-to-address-covid-19-in-tribal-communities/

Oppel Jr, R. A., Gebeloff, R., Lai, K. K. R., Wright, W., & Smith, M. (2020, July 5). The Fullest Look Yet at the Racial Inequity of Coronavirus. *New York Times*.

https://www.nytimes.com/interactive/2020/07/05/us/coronavirus-latinos-african-americans-cdc-data.html

Phelan, J. C., & Link, B. G. (2015). *Is Racism a Fundamental Cause of Inequalities in Health ?* https://doi.org/10.1146/annurev-soc-073014-112305

Pirtle, W. N. L. (2020). *Racial Capitalism : A Fundamental Cause of Novel Coronavirus (COVID-19) Pandemic Inequities in the United States*. https://doi.org/10.1177/1090198120922942

Quick, H. (2019). Estimating county-level mortality rates using highly censored data from CDC WONDER. *Preventing Chronic Disease*, *16*(6), 1–9. https://doi.org/10.5888/pcd16.180441

Resolve to Save Lives, & Vital Strategies. (2020). *Tracking COVID-19 in the United States: From Information Catastrophe to Empowered Communities*. https://preventepidemics.org/wp-content/uploads/2020/07/Tracking-COVID-19-in-the-United-States-Report.pdf

Rodriguez-diaz, C. E., Guilamo-ramos, V., Mena, L., Hall, E., Crowley, J. S., Baral, S., Prado, G. J., Marzan-, M., Beyrer, C., Sullivan, P. S., & Millett, G. A. (2020). Risk for COVID-19 infection and death among Latinos in the United States: Examining heterogeneity in transmission dynamics. *Annals of Epidemiology*. https://doi.org/10.1016/j.annepidem.2020.07.007

Tiwari, C., Beyer, K., & Rushton, G. (2014). The impact of data suppression on local mortality rates: The case of cdc wonder. *American Journal of Public Health*, *104*(8), 1386–1388. https://doi.org/10.2105/AJPH.2014.301900

United States Census Bureau. (2020). *State Population by Characteristics: 2010-2019*. https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html#par_textimage_673542126