Trends in age distribution of COVID-19 cases, hospitalizations, and deaths by race in the United States

Madeleine Short Fabic, MHS¹ & Yoonjoung Choi, MPH, DrPH²

INTRODUCTION

COVID-19 cases are quickly growing across the United States with numerous states reporting that the proportion of cases among young people is ballooning.[1] Meanwhile, the pandemic is amplifying existing racial and ethnic health disparities, disproportionately harming Black, Indigenous, and Hispanic people.[2]–[4] Yet another challenge exposed by the pandemic is our limited ability to identify and track demographic trends in COVID-19 cases, hospitalizations, and deaths over time. Because COVID-19 data are typically presented cumulatively and by only one demographic characteristic[5], [6]—e.g. by race, by age, or by sex – more complex demographic patterns have been slow to emerge.

Understanding and communicating complex demographic trends is imperative to recognize population-level vulnerabilities and inform tailored public health responses. Using the latest COVID-19 Case Surveillance Public Use Data by the Centers for Disease Control and Prevention (CDC), we aim to: a) assess one dimension of reporting quality-- data completeness; and b) examine national time-trends in the age pattern of COVID-19 cases, hospitalizations, and deaths overall as well as by race and ethnicity.

¹ United States Agency for International Development

² iSquared. Corresponding author (yj.choi@isquared.global)

Views expressed are the authors' and do not represent either of the authors' organization affiliations.

METHODS

The COVID-19 Case Surveillance Public Use Dataset was first published in May 2020. An updated version is released monthly.[7] It includes all reported individual, de-identified COVID-19 cases, with information on sex, age in nine 10-year groups, and race/ethnicity in seven groups: Hispanic, non-Hispanic American Indian/Alaskan Native (AIAN), non-Hispanic Asian, non-Hispanic Black, non-Hispanic multiple or other races, non-Hispanic Native Hawaiian/Other Pacific Islander, and non-Hispanic White. It also has hospitalization and survival status of each reported case. These data make it possible to monitor trends in COVID-19 cases, hospitalizations, and deaths by multiple demographic characteristics simultaneously.

Using the most recently released data (July 31, 2020), we assess completeness of demographic information by week of reporting to CDC. We then calculate the weekly number of COVID-19 cases, hospitalizations, and deaths by 10-year age group – for the country as well as for each of the seven race/ethnicity groups available in the data. The age-pattern of weekly cases, hospitalizations, and deaths is compared across race and ethnicity groups. Data analysis was performed using R Studio version 3.5.2, and a reproducible code is available.

RESULTS

Data quality: coverage and completeness of demographic information

The most recent version includes cases reported as of 16 July 2020, totaling 2,668,175 cases and 113,194 deaths. According to The COVID Tracking Project, these totals align with those of June 30/June 20, respectively.[8] Age and sex data were missing for few cases (0.1% and 2%, respectively) and even fewer deaths. Conversely, race/ethnicity data were missing for 47% of cases and 14% of deaths. Race/ethnicity reporting has not improved since reporting began (Figure 1).

Age-patterns: national and by race/ethnicity

Overall, younger people account for an increasingly larger proportion of cases with the largest proportional increases among ages 0-29. In late-March, 50% of weekly cases were among people ages <50, rising to 70% in late-June. The age distribution of hospitalizations and deaths has also shifted. In late-March, people ages <50 accounted for 25% of weekly hospitalizations and 20% of weekly deaths. By late-June, they accounted for 32% and 34%, respectively.

Cases are trending younger over time across all race/ethnicity groups, but the change is most marked among Hispanic and non-Hispanic AIAN people. In late-March, 16% of weekly cases among Hispanics were among ages 0-29; by late-June, the proportion was 30% (Figure 2). Among non-Hispanic AIAN, 16% of cases in late-March were among ages 0-29; by late-June, the proportion was 36% (data not shown).

Age distribution of hospitalizations and deaths have remained relatively constant by race/ethnicity over time. Still, among Hispanics the age distribution of those hospitalized is much younger than other racial groups. By late-June, 13% of weekly hospitalizations among Hispanics were among people ages 0-29, the highest proportion among all groups. Additionally, racial and ethnic disparities in the age-distribution of hospitalizations and deaths have widened over the months. Compared with non-Hispanic whites, non-Hispanic Black, non-Hispanic AIAN, and Hispanic people are experiencing increasingly larger proportions of hospitalizations and deaths by the young. For example, in late-March people ages <50 accounted for 3% of weekly deaths among non-Hispanic whites, 9% among non-Hispanic Blacks, and 12% among Hispanics. By late-June, people ages <50 still accounted for 3% of weekly deaths among non-Hispanic whites, but grew to 11% of weekly deaths among non-Hispanic Blacks and 16% among

Hispanics. [Annex 1](#) presents the weekly pattern of cases, hospitalizations, and deaths for all race/ethnicity groups.

DISCUSSION

Race/ethnicity reporting has been persistently deficient over the months since the pandemic began. We know that COVID-19 disproportionately impacts Black, Indigenous, and people of color, but without consistent and complete data, the full scope of racial/ethnic inequities remains unknown. These ongoing data gaps are unacceptable and must be immediately remedied at national, state, and local levels. Specifically, CDC, as well as every State, Tribal government, and territory must require race and ethnicity reporting for all cases. In addition, the public use dataset is released only monthly and does not include geographic variables (e.g. state). This inhibits our understanding of geographic variation in cases, hospitalizations, and deaths, let alone state-level variation in data completeness and quality. For example, because the data did not include a variable for state, we could not use this dataset to pinpoint which states account for the limited race and ethnicity reporting or whether any state has improved its race and ethnicity reporting over time.

Using updated national case-level surveillance data, we present rarely published COVID-19 cases, hospitalizations, and deaths disaggregated by both age and race/ethnicity. Across all racial/ethnic groups, the proportion of cases among younger people has grown. The Hispanic working age population has been especially hard hit. At least four factors could be at play for this shifted age-pattern among cases: a) increasing and expanding test availability; b) younger people facing more exposure than before as they have returned to the workforce with more states advancing their re-opening phases; c) older people avoiding unnecessary risk more effectively than before; and/or d) younger people increasingly

taking unnecessary risk. Reasons for this shift in COVID-19 cases need further investigation. In the meantime, we need to use harm reduction strategies to help people navigate risk.[9]

So far, age pattern shifts in hospitalizations and deaths have been less pronounced. This could be due to multiple factors, including known reporting lags, increased testing capturing milder conditions more prevalent among younger people, the timing of the course of infection, access to formal healthcare among younger people, and/or the protection youth affords from severe complications of the virus.

Finally, in light of the heterogeneity of the US population—including the racial/ethnic health disparities resulting from historic and ongoing racism—we recommend multivariate disaggregation of case, hospitalization, and death data be standard practice and incorporated into publicly available data visualizations, including CDC’s COVID Data Tracker.[5] We reviewed the websites of major groups presenting COVID data visualizations for public consumption, but, unfortunately, we only found univariate demographic disaggregation and visualizations and no information on demographic patterns over time.[5], [8], [10], [11] Cumulative data and univariate disaggregation are both necessary, but insufficient for informing public health actions.

In summary, COVID-19 highlights and amplifies existing racial and ethnic disparities. We first of all urge to implement measures that can improve reporting of race and ethnicity. Our analysis shows unequal and changing age-patterns among cases, hospitalizations, and deaths by race and ethnicity. Age-pattern differences between whites and other races are widening. These disparities are the result of ongoing systemic and institutional racism,[4] which must be addressed immediately, for the duration of the pandemic, and continuously thereafter.

REFERENCES

- [1] K. H. Courage, "Why more young people are getting sick in the latest Covid-19 outbreaks," *Vox*, 18-Jul-2020.
- [2] G. A. Millett *et al.*, "Assessing differential impacts of COVID-19 on black communities," *Ann. Epidemiol.*, vol. 47, pp. 37–44, 2020.
- [3] Centers for Disease Control and Prevention, "Health Equity Considerations & Racial & Ethnic Minority Groups (Updated July 24, 2020)," 2020. [Online]. Available: <https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html>. [Accessed: 04-Aug-2020].
- [4] M. K. Evans, L. Rosenbaum, D. Malina, S. Morrissey, and E. J. Rubin, "Diagnosing and Treating Systemic Racism," *N. Engl. J. Med.*, vol. 383, pp. 274–276, 2000.
- [5] Centers for Disease Control and Prevention, "Demographic Trends of COVID-19 cases and deaths in the US reported to CDC, CDC COVID Data Tracker." [Online]. Available: <https://www.cdc.gov/covid-data-tracker/#demographics>. [Accessed: 04-Aug-2020].
- [6] M. Boulicault, A. C. Danielsen, J. Bruch, A. Tarrant, A. Borsa, and S. Richardson, "Socially Relevant Variables In US State COVID-19 Surveillance Reporting: A Report Card," *Heal. Aff. Blog*, vol. July 14, 2, 2020.
- [7] Centers for Disease Control and Prevention, "COVID-19 Case Surveillance Public Use Data," 2020. [Online]. Available: <https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf>. [Accessed: 01-Aug-2020].
- [8] "The COVID Tracking Project." [Online]. Available: <https://covidtracking.com/>. [Accessed: 04-Aug-2020].
- [9] Centers for Disease Control and Prevention, "Guidance for Businesses & Employers: Plan, Prepare and Respond to Coronavirus Disease 2019," 2020. [Online]. Available:

<https://www.cdc.gov/coronavirus/2019-ncov/community/guidance-business-response.html>.

[Accessed: 04-Aug-2020].

- [10] Harvard GenderSci Lab, "GenderSci Lab COVID Project," 2020. [Online]. Available:

<https://www.genderscilab.org/gender-and-sex-in-covid19/>. [Accessed: 04-Aug-2020].

- [11] Johns Hopkins University Center for Systems Science Engineering, "Coronavirus COVID-19 global cases by the center for systems science and engineering.," 2020. [Online]. Available:

<https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf>

6. [Accessed: 07-Apr-2020].

Figure 1. Racial and ethnic distribution of weekly COVID-19 cases in the United States

Note: only weeks with 20 or more cases are presented.

NH: Non-Hispanic, AIAN: American Indian/Alaskan Native, NHPI: Native Hawaiian/Other Pacific Islander.

Figure 2. Age-distribution of weekly COVID-19 cases, hospitalizations, and deaths by race/ethnicity: Hispanic, non-Hispanic Black, and non-Hispanic white

Note: only weeks with 20 or more observations (cases, hospitalizations, and deaths, accordingly) are presented. Hospitalization and deaths by the week of case report, not necessarily actual week of hospitalization or death.