4-28-2020

Case reports of COVID-19 in the United States by poverty, gender and race: A data review protocol—Study description

Charlotte Brasseux

Thoai Ngo
Population Council

Recommended Citation

Brasseux, Charlotte and Thoai Ngo. 2020. "Case reports of COVID-19 in the United States by poverty, gender and race: A data review protocol—Study description." New York: Population Council.

Title: Case reports of COVID-19 in the United States by poverty, gender and race: A data review protocol[1]

Date: April 26 2020

Authors: Charlotte Brasseux[1], Thoai D Ngo[2,3]

Affiliations:
[1]MPH candidate at Mailman School of Public Health at Columbia University
[2]Poverty, Gender and Youth Program, Population Council, New York, New York, USA
[3]The GIRL Center, Population Council, New York, New York, USA

Contacts:
Thoai D Ngo, PhD, MHS
Program Director
Poverty, Gender and Youth Program
Population Council
 tngo@popcouncil.org.

Charlotte Brasseux, MPH candidate
Mailman School of Public Health at Columbia University
cab2315@cumc.columbia.edu

POPULATION COUNCIL
IDEAS. EVIDENCE. IMPACT.
One Dag Hammarskjold Plaza
New York, NY 10017
www.popcouncil.org

Abstract

The COVID-19 pandemic is evolving fast and affecting rich and poor countries; however, the social determinants of the infections, deaths and recoveries have not been well characterized. While there is increasing discussion of COVID-19 socio-demographics in the media, no systematic compilation of data exists that synthesizes what and how official reporting platforms disaggregate socio-demographic information of COVID-19 cases, hospitalizations and deaths. This review will provide a comprehensive overview of what and how federal and local health officials report COVID-19 cases. Having this information nationally and sub-nationally can help health officials to deploy a more targeted response effort such as testing, treatment, and contact tracing. This information could be useful for future vaccine development. This data review can reveal gaps in our public health data system, which can lead to recommendations on how to improve data standardization, disaggregation, and reporting, particularly for tracking outbreaks. This protocol may be replicated in other countries to understand how socio-demographic factors are being reported and their relationship to COVID-19 cases.

---

[1] This document is evolving due to the nature of the COVID-19 response and will be updated as needed.

Background

The World Health Organization (WHO) Commission on Social Determinants of Health asserts that political, social, and economic forces shape the circumstances in which people live. These circumstances give rise to avoidable health inequities, both within and between countries, which are strongly linked to degrees of social disadvantage[1]. In an article from the *Annual Review of Public Health*, Adler and Rehkopf describe that "physical and social environments, including a person's home, school, work, neighborhood, and community, vary by socio-economic status (SES) and affect the likelihood of individuals' exposure to both health-damaging conditions and health-protecting resources"[2]. Factors including race/ethnicity, gender, income, occupation, education level, and geographic location all have an impact on health outcomes, such as morbidity and mortality rates. The first study in the United States (US) to examine socioeconomic disparities within race/ethnicity using individual-level data from the 1960 matched records of people 25 years of age and older. This study found that compared to whites, age-adjusted, all-cause mortality rates were 34% higher for nonwhite females and 20% higher for males. The same study found that males and females who are the least educated experience 64% and 105% higher mortality rates, respectively, compared to those most educated[2]. Similar trends are apparent in maternal mortality rates. Recent data from the Center for Disease Control and Prevention (CDC) shows that non-Hispanic black women have a rate of 37.1 deaths per 100,000 live births, compared to non-Hispanic white women who have a rate of 14.7 deaths per 100,000 live births[3]. Despite the recognition that addressing these disparities is critical to tackling health inequity and protecting public health, data systems, research and policies that directly tackle these key social determinants continue to lag behind in the US and globally.

During the ongoing COVID-19 pandemic, the data on COVID-19 confirmed cases and deaths in the US are increasingly revealing that people of color are more severely affected by the novel coronavirus. On April 7, the surgeon general first acknowledged the racial disparities of COVID-19 based on case reports from a few states, including Louisiana, Michigan, and Illinois[4]. Analysis of these reports suggested that majority-black counties had three times the infection rate and six times the death rate as majority-white counties[4]. As more states began to release socio-demographic data related to COVID-19, similar trends appeared. In New York, black people make up nine percent of the population but account for 17 percent of the deaths[5]. In addition to the racial disparities, COVID-19 is affecting poor people more severely and is infecting and killing more men than women, is killing older populations compared to young age groups, and is infecting urban populations more than rural ones[6, 7]. On April 17 2020, the CDC released the first report on COVID-19 containing socio-demographic data collected from only 14 states during the month of March. The report stated that only 18% of residents in the sample population were black, yet blacks made up 33% of those hospitalized with COVID-19. In comparison, while 59% of residents in the sample population were white, they made up just 45% of those hospitalized with COVID-19[8]. Public health and medical experts are stressing the fact that people of color, in general, are not more susceptible to the virus, but they are more susceptible to experiencing more severe illness and worse health outcomes as a result of the virus due to the existing health inequities in the US[9].

Historically, epidemics disproportionately affect vulnerable communities, including those living in poor and crowded neighborhoods, women and girls, gender minority individuals, and people

living in poverty. In the article *Health Inequities and Infectious Disease Epidemics*, Quinn and Kumar describe, "poverty, inequality, and social determinants of health create conditions for the transmission of infectious diseases, and existing health disparities or inequalities can further contribute to unequal burdens of morbidity and mortality"[10]. During a pandemic, these inequities become more visible due to the greater impact on vulnerable populations. During the H1N1 pandemic, Quinn and Kumar found that those who were at the highest risk of exposure to H1N1 were those of lower socioeconomic status and were likely to have less access to care if they contracted the disease[11]. Similarly, Dr. Anthony Fauci compared COVID-19 to the HIV pandemic, where HIV affected the gay communities at higher rates compared to the general population, and the COVID-19 pandemic is bringing attention to the racial disparities on health in the US[9]. 2018 data from the CDC continues to show that Black/African American gay, bisexual, and other men were more affected by HIV than any other group in the US[12].

Understanding the socio-demographic trends of COVID-19 could help with deploying preventive and mitigation measures as well as treatment and relief. Having information on who is most affected by this pandemic will help state level governments and local officials to plan a response and offer more support to communities and individuals who need it the most. Further, extensive research has demonstrated that factors including sex, pregnancy, and comorbid conditions can impact the immune response to vaccines[13]. Therefore, understanding socio-demographic factors of those most affected by COVID-19 is critical in the development of an effective vaccine. More generally, Dr. Lisa A. Cooper, Professor at the Johns Hopkins Bloomberg School of Public Health and Johns Hopkins School of Medicine, explains that "health and public health professionals, administrators, employers, policymakers, and even community advocates can use these data to determine how best to use the vast resources we have in this country to improve the lives of our people"[14].

Prevention, mitigation, treatment and relief efforts cannot occur without adequate data and reporting of socio-demographic factors of COVID-19 cases. As a result of the early data suggesting extreme racial disparities in the COVID-19 pandemic, the federal government is being pressured to not only release racial data on cases and deaths, but also release a plan to "blunt the devastation" on people and communities of color[15]. Although, the WHO does not produce official analyses and routine reporting on the socio-demographic factors related to COVID-19 confirmed cases, hospitalizations, deaths and recoveries. The CDC only began reporting this information in 28% of states on April 17, almost three months after the first reported case in the US. In the US, data revealing racial disparities of COVID-19 come from state governments reports; however, these reports are not standardized between states, and the available information is not systematically synthesized at the national level. Together, this highlights the failure of the US public health data system to capture and analyze this information in order to contribute to the test, treat, trace, and isolate initiative, which is needed to reduce and suppress the spread of coronavirus. Capturing and analyzing this information is critical in recording and tackling health inequity in the United States.

Objectives

We are conducting a data review of COVID-19 to examine how cases are reported in the US and how this reporting has changed over time with the following objectives:

1. To assess whether official or reliable case reports of COVID-19 are disaggregated by key socio-demographic characteristics, age, gender/sex, race/ethnicity, education level, geography, and economic status.

2. To assess how COVID-19 affects different population groups by confirmed cases, symptoms, recoveries and fatality.

3. To describe how the patterns of case reports of COVID-19 change over time and from earlier epicenters to recent epicenters since the first confirmed case.

Methods

Inclusion and Exclusion Criteria

This review will include reports of COVID-19 cases in the US from the WHO, the CDC, state health departments, health departments of major US cities (e.g.…..), and data sources from two commonly used tracking sites: Johns Hopkins University (JHU), and the Institute for Health Metrics and Evaluation (IHME). We define "reports" as daily or weekly situation reports, rather than raw data, released by local, state, and national governments and the two major tracking sites listed above. However, if daily or weekly reports are not available, we will look at the raw data provided by each source. We will specifically review reports of cases since the first confirmed case in the US on January 19, 2020[16].

We will not include reports from news articles, published papers (e.g., papers synthesizing and analyzing government reports), and other tracking sites that compile government data (e.g., worldometer or major new outlets tracking efforts). We will exclude reports from county health departments. While this review focuses on the United States, this method can be applied to review reports in other countries.

Search Engine

We will use Google search engine as the starting point. We will also search well-known COVID-19 tracking websites. These include the JHU Coronavirus Resource Center and the IHME. From these websites, we use a snowball sampling method to identify the sources of data that these websites use, such as state health departments and major city health departments.

Outcomes of interest

We are interested in five primary outcomes in our review: cases, symptoms, hospitalizations, recoveries, and case fatalities.

We define cases with the CDC definition: A confirmed case is defined by meeting confirmatory laboratory evidence for COVID-19. A probable case is defined by i) meeting clinical criteria AND epidemiologic evidence with no confirmatory laboratory testing performed for COVID-19; or ii) meeting presumptive laboratory evidence AND either clinical criteria OR epidemiologic evidence; or iii) meeting vital records criteria with no confirmatory laboratory testing performed for COVID19[17].

The CDC includes the following as symptoms of COVID-19: fever, cough, shortness of breath or difficulty breathing, chills, repeated shaking with chills, muscle pain, headache, sore throat, and new loss of taste or smell[18].

Hospitalizations refer to those hospitalized with lab-tested positive COVID-19.

Recoveries refer to those who have recovered from a confirmed case of COVID-19.

We define case fatalities with the CDC definition: A confirmed death is defined by meeting confirmatory laboratory evidence for COVID-19. A probable death is defined by i) meeting clinical criteria AND epidemiologic evidence with no confirmatory laboratory testing performed for COVID-19; or ii) meeting presumptive laboratory evidence AND either clinical criteria OR epidemiologic evidence; or iii) meeting vital records criteria with no confirmatory laboratory testing performed for COVID19[17].

Socio-demographic indicators

We will examine disaggregation of the following socio-demographic factors in each official data report: age, gender/sex, race/ethnicity, education level, geography, and economic status. In this review, geography is defined as the county in which each person lives. This information will allow us to examine how those in poorer counties are affected compared to those in less poor ones.

Data source

Analysis and Synthesis

We will first describe how the different data sources report on their cases, including the frequency of report, what outcomes, and what socio-demographic characteristics are included. We will also examine how different data sources are disaggregated by these socio-demographic factors in their reporting of confirmed cases, symptoms, hospitalizations, fatalities, and recoveries. For those sources that report demographic factors, we will analyze if differences exist between sub-group populations. This analysis will help us understand what information is available, what information is lacking, and how the reporting of cases changes over time. This will give a more comprehensive and clearer analysis of who is most affected by this pandemic, which in turn can help inform future response efforts.

Implications of our review

This review will provide a comprehensive overview of what and how federal and local health officials report COVID-19 cases. Having this information nationally and sub-nationally can help health officials to deploy a more targeted response effort and can help reveal information that could be useful for testing, treatment, and vaccine development efforts. Further, this data review may reveal gaps in our public health data system in which we can make recommendations on how to improve data standardization, disaggregation, and reporting, particularly for tracking outbreaks. This protocol may be replicated to other countries to

understand how socio-demographic factors are being reported and their relationship to COVID-19 cases.

Proposed Tables:

1. Description of data sources included in the review
2. Description of how COVID-19 cases are reported by states, by socio-demographic factors
3. An analysis of disparities of COVID-19 cases nationally and by states
4. An analysis of how reporting of socio-demographic characteristics change over the course of the COVID-19 epidemic

POPULATION COUNCIL
Ideas. Evidence. Impact.

Sources

[1] CSDH (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. Geneva, World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/43943/9789241563703_eng.pdf?sequence=1

[2] Alder, N., Rehkopf, D. (2007). U.S. Disparities in Health: Descriptions, Causes, and Mechanisms. *Annual Review of Public Health*.

[3] Center for Disease Control. (2020). First Data Released on Maternal Mortality in Over a Decade. https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2020/202001_MMR.htm

[4] Thebault, R., Tran, A., Williams, V. (2020). "The coronavirus is infecting and killing black Americans at an alarmingly high rate". *The Washington Post*. https://www.washingtonpost.com/nation/2020/04/07/coronavirus-is-infecting-killing-black-americans-an-alarmingly-high-rate-post-analysis-shows/?arc404=true

[5] Kendi, I. (2020). "Stop Blaming Black People for Dying of the Coronavirus". *The Atlantic.* https://www.theatlantic.com/ideas/archive/2020/04/race-and-blame/609946/

[6] Human Rights Watch. (2020). US: Address Impact of COVID-19 on Poor. https://www.hrw.org/news/2020/03/19/us-address-impact-covid-19-poor#

[7] Global Health 5050 (2020). COVID-19sex-disaggregated data tracker. https://globalhealth5050.org/covid19/

[8] Center for Disease Control. (2020). Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019.

[9] Greenfieldboyce, N. (2020). How Is The Coronavirus Affecting Black Americans. NPR. https://www.npr.org/2020/04/08/829575332/how-is-the-coronavirus-affecting-black-americans

[10] Quinn, s., Kumar, S. Health Inequities and Infectious Disease Epidemics: A Challenge for Global Health Security. Biosecur Bioterror. 2014 Sep 1; 12(5): 263–273. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4170985/

[11] Quinn, s., et al. Racial Disparities in Exposure, Susceptibility, and Access to Health Care in the US H1N1 Influenza Pandemic. Am J Public Health. 2011 February; 101(2): 285–293. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3020202/

[12] Center for Disease Control. (2020). HIV and African American Gay and Bisexual Men. https://www.cdc.gov/hiv/group/msm/bmsm.html

[13] Klein SL, Jedlicka A, Pekosz A. The Xs and Y of immune responses to viral vaccines. Lancet InfectDis. 2010;10(5):338–49.https://www.ncbi.nlm.nih.gov/pubmed/20417416

[14] Johns Hopkins University. (2020). Coronavirus Resource Center. https://coronavirus.jhu.edu/data/racial-data-transparency

[15] Stafford, K., Hoyer, M., Morrison, A. (2020). Racial toll of virus grows even starker as more data emerges. *AP News.* https://apnews.com/8a3430dd37e7c44290c7621f5af96d6b

[16] Holshue, M., et al. First Case of 2019 Novel Coronavirus in the United States. N Engl J Med 2020; 382:929-936 https://www.nejm.org/doi/full/10.1056/NEJMoa2001191

[17] CDC (2020). https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html

[18] CDC (2020). https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html