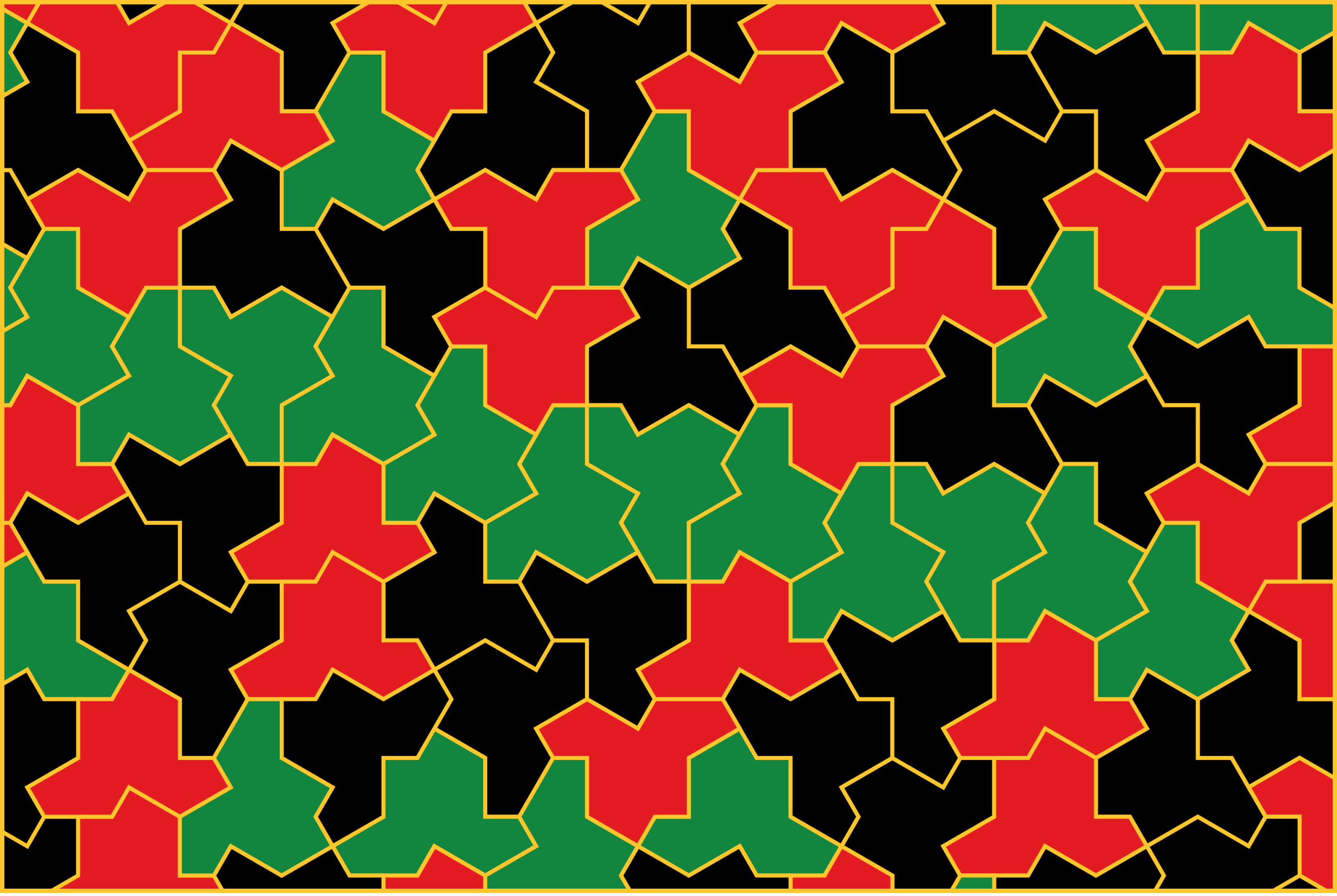

ANTI-RACISM DATA COLLECTION©

Kathryn Youker, Adria Bonillas, Atarah McCoy, Aaniyah Simmons, and the Lawyers' Committee for Civil Rights Under Law

Copyright© November 2023. Cite as: "Kathryn Youker, Adria Bonillas, Atarah McCoy, Aaniyah Simmons, and the Lawyers' Committee for Civil Rights Under Law, *Anti-Racism Data Collection,* November 2023."

POLICY BRIEF

Executive Summary

Reliable racial and ethnic demographic data play a critical role in addressing health disparities. Among other things, reliable data is central to measuring health trends, identifying health disparities between segments of the population, and rooting out inequities in health care. Public health experts have long recognized that disaggregated race and ethnicity data—data that captures diverse subgroups within major race and ethnic categories of people—is a critical component to addressing racial health disparities. However, in previous years there have been no serious efforts at the federal level to collect health data beyond overgeneralized racial and ethnic groups. This has led to significant gaps in the data and, consequently, our understanding of the scope and underlying causes of health disparities among vulnerable communities.

This brief provides a discussion of current data collection methods, their deficiencies, and how those deficiencies led to significant gaps in the health data collected during the COVID-19 pandemic. It then provides potential recommendations for how data collection methods can be improved. The Lawyers' Committee for Civil Rights Under Law conducted an analysis and survey of current data collection practices at the federal and state level to understand the impact of the pandemic on marginalized communities. Our findings from the survey revealed: (1) the existing methods of collecting medical data significantly underrepresented the true impact of the pandemic on these communities, and (2) failed to capture information essential for understanding the root causes of health disparities. To address these problems, this brief recommends updating data collection practices and infrastructure to incorporate social determinants of health and further disaggregation of race and ethnicity classifications to better develop equitable policies that address health disparities and promote civil rights enforcement.

Background

Inadequate data collection efforts by government agencies are nothing new. The federal government has failed to overhaul and expand data collection efforts at various points over the past several decades, which has perpetuated health disparities among marginalized communities. In 1977, the Office of Management and Budget ("OMB") initially developed, in collaboration with other federal agencies, government-wide data collection standards to provide consistent data on race and ethnicity, including for the decennial census, household surveys, and federal administration forms.[1, 2] Development of this protocol stemmed primarily from the federal government's responsibility to enforce civil rights laws.[3] The initial OMB standards consisted of four racial groups: American Indian or Alaskan Native, Asian or Pacific Islander, Black, and White, and two ethnic groups ("Hispanic origin" or "Not of Hispanic origin").[4]

Since its implementation nearly 45 years ago, OMB has revised its data collection methods only one time, resulting in the current standards, which were expanded from four to five racial categories: (1) American Indian or Alaskan Native, (2) Asian, (3) Black or African American, (4) Native Hawaiian or Other Pacific Islander, (5) and White.[5] These standards—implemented in 1997—were expanded to enhance the accuracy of demographic data being collected.[6] However, these standards require federal entities and states to collect only minimum data on race and ethnicity—any additional categorization is discretionary. Additionally, there is no evidence of an enforcement mechanism to ensure these standards are followed.

OMB set the baseline for data collection standards that other agencies should follow. The Department of Health and Human Services ("HHS") and the Center for Disease Control and Prevention ("CDC"), for example, both use the 1997 OMB standards to collect racial and ethnic data within healthcare settings. Notably, this was the standard both agencies used in 2020 when collecting COVID-19 related data.[7, 8]

However, more exacting standards exist. Acknowledging that specific and disaggregated racial and ethnic data is critical for eliminating health disparities, in 2011 HHS promulgated guidelines in accordance with a mandate under Section 4302 of the then newly-enacted Affordable Care Act ("ACA").[9] The guidance, entitled "Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language and Disability Status" added 14 racial and five ethnic subgroups beyond the minimum requirements outlined in the 1997 OMB standards.[10] Yet, HHS limited the application of these categories and explained they were only required for "national population health surveys."[11] Though nothing legally precluded HHS from adopting more detailed standards and applying them to all public health programs and activities, including COVID-19 surveillance, the agency declined the opportunity to do so.[12]

As of this writing, OMB is undergoing a review process of its data collection standards with the goal of ensuring the comparability of race and ethnicity across datasets and to maximize the quality of data by ensuring collection protocols are consistent and based on rigorous evidence.[13] Revisions have yet to be implemented. Moreover, HHS promulgated in 2022 a proposed rule entitled "Nondiscrimination in Health Programs or Activities," which would add data collection standards for entities covered by the ACA, i.e., all health programs, activities, and entities that receive HHS funding.[14] The Lawyers' Committee submitted comments in support of this proposal and provided additional recommendations urging adoption of an enforceable, national standard for data collection that mandates the collection of minimum disaggregated data across health programs and activities.[15] HHS has yet to publish a final rule.

Methodology and Data Collection

At the outset of the COVID-19 pandemic, on April 6, 2020, the Lawyers' Committee, along with over 400 medical professionals, issued a letter to the CDC and HHS urging the federal government to release comprehensive race and ethnic demographic data concerning COVID-19 diagnoses, tests, and outcomes.[16] The agencies eventually released a report revealing that the HHS and CDC had significant gaps in collecting race and ethnic demographic data for a considerable portion of people seeking treatment for COVID-19.[17] In order to understand the reasons behind these substantial gaps, the Lawyers' Committee conducted an analysis of existing federal standards and regulations governing COVID-19 data collections. This analysis crossexamined current data collection standards and categories with the HHS's 1965 Minority Health Task Force report titled "Report of the Secretary's Task Force on Black and Minority Health," the OMB standards, and standards promulgated pursuant to the ACA.

The Lawyers' Committee and our pro bono partners also surveyed each state's COVID-19 website to evaluate the different data collection standards and processes across the country.[18] Each state was examined to see if they collected and published information regarding the number of cases, hospitalizations, ICU admissions, testing, deaths, and vaccinations specifically categorized by race and ethnicity.

Findings from Our Work

Our research revealed that the gaps in data were a consequence of poorly implemented policies and standards. Specifically, the data collection tools used during the pandemic relied on the outdated 1997 OMB demographic standards, which were known to be inadequate for addressing health disparities among communities of color and other systemically marginalized groups.[19] The CDC and HHS's use of the 1997 OMB standards resulted in glaring holes in racial and ethnic data. Notably, racial and ethnic data was missing for nearly one-third of more than 250 million people who had received at least one dose of the COVID-19 vaccine, and only 66% of this data for Covid-19 cases was available nearly two years into the pandemic.[20]

Our research further identified problems with available data at the state level that were multi-fold. First, many states did not publish data according to federal standardized categories and most states failed to publish disaggregated data for each racial and ethnic group.[21] We were unable to identify any instances in which OMB attempted to enforce compliance with the baseline standards. Our survey showed that only 9 states and the District of Columbia collected and published racial and ethnic demographic data for COVID-19 tests.[22] The pandemic's impact on diverse subsets of people thus went entirely unreported.

Second, no two states reported the same race or ethnicity data.[23] The lack of uniformity made it difficult for public health professionals to compare and analyze data. For instance, only a handful of states published data for income status, housing type, and employment status.[24] No federal COVID-19 data surveillance tools requested sexual orientation and gender identity information, with only two states—California and Rhode Island—proactively tracking this information.[25] This created deeper data gaps and inconsistencies, making it even more difficult to fully capture the pandemic's toll on impacted communities. The overly broad OMB standards for data collection allowed states to collect data however they saw fit. As a result, states continued using insufficient data collection infrastructures which varied widely in metrics and processes.

The consequences of the Covid-19 pandemic have largely mirrored the devastating effects of systemic health inequities that have plagued communities of color and other marginalized groups in the United States for centuries. The pandemic thus presented an unparalleled chance for federal agencies to implement initiatives that address deficiencies in demographic data. Despite this opportunity, there has been minimal progress in advancing the collection of disaggregated race and ethnicity data.

Recommendations

In light of our findings, we believe the federal government should take steps to collect more complete and valuable data that reveals health outcomes of individuals at the intersections of multiple identities and socio-economic factors. To achieve this, we recommend the following steps be taken:

OMB should implement revised standards to improve the quality and usefulness of race and ethnicity data. Better national standards for data collection practices are critical to efforts to narrow health disparities, yet the 1997 OMB standards oversimplify diverse experiences and fail to recognize the unique circumstances of each community. Federal data collection systems further neglect to include social determinants of health that can combine to direct health outcomes, including primary language, disability status, housing status, pregnancy status, employment status and type, insurance status, and sexual orientation and gender identity, all of which play a significant role in contributing to existing health disparities. OMB should complete the review of its data collection standards, revise its 1997 standards, and adopt robust provisions that ensure, at a minimum, the collection of disaggregated racial and ethnic data. We further recommend OMB take steps to ensure states comply with data collection requirements.

HHS should adopt a final rule that includes a mandatory data collection provision. HHS has acknowledged that civil rights data collection plays a vital role in civil rights compliance. We agree and recommend that HHS issue a rule that mandates the collection of minimum disaggregated data across health programs and activities and provides concrete tools for enforcement. We further recommend that any data collection tool HHS implements should, at minimum, collect and publish the following social determinants of health: 1) sexual orientation and gender identity, 2) employment type and income, and 3) housing status.

HHS, the CDC, and other federal agencies should collaborate to collect disaggregated data for future public health emergencies. In the context of the pandemic, we previously recommended that HHS and the CDC collaborate to develop and provide health surveillance forms with disaggregated data categories. These formats should guide labs, health care providers, and hospitals to collect care and outcome data in the following categories: 1) hospitalization rates; 2) ICU admissions; 3) comorbidities; 4) disability status and type; 5) insurance status; and 6) pregnancy status disaggregated by the 2011 HHS standards including 14 races and five ethnicities. We recommend such measures be taken in anticipation of future public health crises.

Conclusion

The COVID-19 pandemic has exposed and intensified the impact of systemic health disparities on communities of color and marginalized groups, underscoring the need for immediate action and change. Unfortunately, the voices of these communities and healthcare advocates have been ignored, perpetuating a system that fails to provide equal service to all members of society. Our research emphasizes the importance of collecting equitable data that considers multiple identities and socio-economic factors when assessing health outcomes. Without such data, barriers to quality healthcare and discriminatory practices in the healthcare sector may persist. We encourage policymakers, healthcare professionals, public health specialists, and community advocates to use our recommendations as a resource to expand healthcare data. By doing so, we believe we can take a crucial step towards addressing disparities highlighted by the pandemic and achieving a more equitable healthcare system in the future.

Acknowledgements

Kathryn Youker, Adria Bonillas, Atarah McCoy, and Aaniyah Simmons are staff at the [Lawyers' Committee for Civil Rights Under Law](), where they work on the [Economic Justice Project](). They are also members of the Anti-Racism Consortium.

Support for this policy brief was provided by the Robert Wood Johns Foundation. The views expressed here do not reflect the views of the Foundation.

References

1. OMB Directive 15: Race and Ethnic Standards For Federal Statistics And Administrative Reporting. CDC WONDER, Centers for Disease Control and Prevention, https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html (accessed 9 Aug. 2023).

2. There are two exceptions to this requirement: if "the collection involves a sample of such size that the data on the smaller categories would be unreliable, or when the collection effort focuses on a specific racial or ethnic group," Centers for Disease Control and Prevention, "OMB Directive 15."

3. Ibid.

4. Ibid.

5. *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*, Office of Management and Budget, 30 Oct. 1997, https://obamawhitehouse.archives.gov/omb/fedreg_1997standards (accessed 9 Aug. 2023).

6. Ibid.

7. "Human Infection with 2019 Novel Coronavirus Case Report Form." Centers for Disease Control and Prevention, US Department of Health and Human Services; "Human Infection with Coronavirus Disease 2019 (Covid-19) Surveillance Worksheet." Centers for Disease Control and Prevention, US Department of Health

and Human Services, www.cdc.gov/ncird/surveillance/downloads/COVID-19-Surveillance-Worksheet-508.pdf; “VaxAdmin_CVRS, Appendix C: CDC Comprehensive Vaccine Data Requirements, Technical Standards & Reporting: COVID-19 Vaccination.” Centers for Disease Control and Prevention, US Department of Health and Human Services, www.cdc.gov/vaccines/covid-19/reporting/requirements/index.html (last accessed Aug 9, 2023).

8. Centers for Disease Control and Prevention, “COVID-19 Vaccination Provider Requirements and Support;” “COVID-19 Lab Data Reporting Implementation Specifications.” *US Department of Health and Human Services*, 2020, www.hhs.gov/sites/default/files/hhsguidance-implementation.pdf; “COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115.” *U.S. Department of Health and Human Services*, 8 Jan. 2021, www.hhs.gov/sites/default/files/covid-19-laboratorydata-reportingguidance.pdf; “Interim Final Rule (IFC), CMS-3401-IFC; Requirements and Enforcement Process for Reporting of COVID-19 Data Elements for Hospitals and Critical Access Hospitals.” *Centers for Medicare and Medicaid Services*, US Department of Health and Human Services, 6 Oct. 2020, www.cms.gov/files/document/qso-21-03-hospitalscahs.pdf-0; “COVID-19 Guidance for Hospital Reporting and FAQs For Hospitals, Hospital Laboratory, and Acute Care Facility Data Reporting.” *US Department of Health and Human Services*, www.hhs.gov/sites/default/files/covid-19-faqs-hospitals-hospital-laboratoryacutecare-facility-data-reporting.pdf.

9. *HHS Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status*, 10 Oct. 2011, https://aspe.hhs.gov/reports/hhs-implementation-guidance-datacollection-standards-race-ethnicity-sex-primary-language-disability-0 (accessed 9 Aug. 2023).

10. Ibid.

11. “HHS Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status.” *Office of the Assistant Secretary for Planning and Evaluation*, U.S. Department of Health and Human Services, 31 Oct. 2011, https://aspe.hhs.gov/basic-report/hhsimplementation-guidance-data-collection-standards-race-ethnicity-sex-primary-language-and-disabilitystatus.

12. Patient Protection and Affordable Care Act of 2010. Pub. L. No. 111-148, 124 STAT. 578–82; Public health surveillance is ”[a]n ongoing, systematic collection, analysis and interpretation of health-related data essential to the planning, implementation, and evaluation of public health practice[.]” “Introduction to Public Health Surveillance.” Centers for Disease Control and Prevention, 15 Nov. 2018, https://www.cdc.gov/training/publichealth101/surveillance.html. “[S]urveys are often done only once, to determine the distribution of risk factors in a population at a point in time.” Bonita, Ruth, et al. “From Surveys to Surveillance.” *Scientific Electronic Library Online (SciELO) Public Health*, Organizacion Panamericana De La Salud, 1 Oct. 2001, https://scielosp.org/article/rpsp/2001.v10n4/223-225/.

13. Initial Proposals For Updating OMB’s Race and Ethnicity Statistical Standards, 88 Fed. Reg. 5375 (Jan. 27, 2023).

14. Nondiscrimination in Health Programs and Activities, 87 Fed. Reg. 47824 (Aug. 4, 2022).

15. Lawyers’ Committee for Civil Rights Under Law, Comment Letter on Proposed Rule on Section 1557 of the

Affordable Care Act (Oct. 27, 2022), https://www.regulations.gov/comment/HHS-OS-2022-0012-67734.

16. "Letter from the Lawyers' Committee for Civil Rights Under Law." Received by Alex Azar, Secretary of the U.S. Department of Health and Human Services, *Lawyers' Committee for Civil Rights Under Law*, 6 Apr. 2020, https://lawyerscommittee.org/wp-content/uploads/2020/04/DHHS-Letter-COVID-19.pdf.

17. "Demographic Trends of COVID-19 Cases and Deaths in the US Reported to CDC." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, https://covid.cdc.gov/covid-datatracker/#demographics (accessed 9 Aug. 2023).

18. Sharma, "50-State Survey"; "State COVID-19 Data by Race." *Johns Hopkins Coronavirus Resource Center*, Johns Hopkins University & Medicine, https://coronavirus.jhu.edu/data/racial-data-transparency (last updated 12 Mar. 2021); "How States Collect, Report, and Act on COVID-19 Race and Ethnicity Data." *National Academy for State Health Policy*, www.nashp.org/how-states-report-covid-19-data-by-raceandethnicity/#tab-id-1.

19. *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*, Office of Management and Budget, 30 Oct. 1997, https://obamawhitehouse.archives.gov/omb/fedreg_1997standards (accessed 9 Aug. 2023).

20. Centers for Disease Control and Prevention, "Demographic Characteristics of People Receiving COVID-19 Vaccinations in the United States."

21. Sharma, "50-State Survey."

22. Ibid.

23. The Lawyers' Committee reviewed each state's public health authority's COVID-19 dashboard, and the data each published to conduct this survey.

24. Sharma, "50-State Survey"; "California's Commitment to Health Equity." *State of California*, 19 Jan. 2022, https://covid19.ca.gov/equity/; "Situation Update for COVID-19." *Minnesota Department of Health*, 17 Mar. 2021, www.health.state.mn.us/diseases/coronavirus/situation.html#map2.

25. Ibid.