Gaps in Available Data Exacerbate Health Disparities and Create Barriers to Change

INTRODUCTION

COVID-19 has exacerbated longstanding disparities and further exposed the strong role of social determinants of health (e.g., poverty, discrimination, access to resources and services, employment, housing, education, and health care) in health outcomes.[i] COVID-19 cases among Native Hawaiian/Pacific Islanders and American Indian/Alaskan Natives are 2.5 and 2 times as high, respectively, as compared to whites. The rate of COVID-19 related deaths among African Americans is twice as high as whites.[ii]

The American Medical Association (AMA) reports that higher incidence of COVID-19 in African American communities can be attributed to, among many, three key factors, including structural-level inequalities and social determinants of health, higher rates of pre-existing chronic conditions — as a result of lower access to health care and population health factors, and increased rate of employment in the essential non-health care labor sector (e.g., public transportation workers).[iii] The findings are not surprising as they highlight centuries of lived experience that major advancements in health care over the last century too often exploit or leave behind communities of color[iv] — this includes advancements in the collection of health data that drives our understanding of health systems.

COVID-19 also illuminated critical gaps in information on race, ethnicity, and health that left holes in our early understanding of disparities in the spread and impact of the pandemic. As of December 2020, the Centers for Disease and Control (CDC) reports race/ethnicity data for just 51% of COVID-19 cases and 78% of COVID-19 deaths.[v] As evidenced by CDC COVID-19 demographic reporting, there remain large gaps in the consistency and uniformity of COVID-19 racial/ethnicity data, which continue to hinder our ability to recognize and act on racial/ethnic COVID-19 disparities in communities of color.[vi] For example, in lieu of reporting actual crude numbers of COVID-19 deaths by race/ethnicity, some states report percentages instead —

resulting in estimations. In addition, gaps remain with respect to information on hospitalizations related to COVID-19 cases.[vii]

Fixing the system means understanding and confronting all that feeds into it. To begin overcoming health disparities, researchers, policymakers, and providers need actionable data that demonstrate how and why people of color experience worse health outcomes than their white counterparts. Without substantial, accurate, disaggregated data, that understanding cannot be realized and we will continue to see higher rates of negative health outcomes for people of color.

This paper is the first in a series that will identify some of the challenges in collecting the types of data that are essential to improving health disparities in the United States; solutions to those challenges and barriers to their implementation; and actionable steps that policymakers can take to drive change. We begin our series by examining three major challenges that perpetuate data gaps: inadequate data collection, technological barriers, and provider and patient concerns.

Figure 1. Process to Seek Actionable Policy Recommendations on Disaggregated Race/Ethnicity Data

- Challenges to national and state data collection of race/ethnicity and correlates of health disparities
- Solutions to advance data collection by race/ethnicity and correlates of health disparities
- Actionable policy recommendations to improve disaggregated race/ethnicity data collection
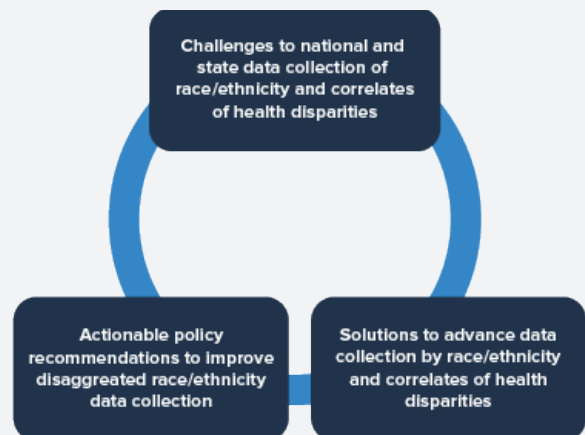
BACKGROUND: HEALTH DISPARITIES SIGNAL GAPS IN HEALTH DATA

Health disparities in the United States are well-documented.[viii] Since publication of the landmark 1985 Heckler Report - the first national report demonstrating health disparities across racial and ethnic groups[ix] - significant research and policy efforts have been devoted to reporting racial and ethnic health disparities. Groundbreaking research across disciplines has transformed the way that we understand the correlation of race, ethnicity, and health. Firstly, sociology literature has revealed the social construction of race through discriminatory laws and practices.[x] Through this lens, we begin to uncover that racial/ethnic disparities are correlated with social determinants (e.g., geography, socioeconomic status, wealth, and physical/built environment) and structural inequities.[xi] Public health research has heightened understanding of the effects of racism on racial health inequities.[xii]

Statistics on maternal mortality serve as alarming evidence of both the impact of health disparities and the need for better data collection. In the United States, American Indian/Alaskan Native and Black women are 2-3 times more likely to suffer a pregnancy-related death than white counterparts.[xiii] Among women with a college degree, Black women are 5 times more likely to die as a result of maternal causes than white mothers. However, because of flawed data collection, researchers did not realize the depth and breadth of the crisis until very recently.

Nearly two decades ago, researchers realized that lack of standardized information collection was resulting in inaccurate counting of maternal deaths. As a solution, in 2003, the National Center for Health Statistics recommended that states include a maternal death checkbox on death certificate forms. It took nearly fifteen years for all states to adopt a standard checkbox, so it was not until 2018 that the United States had data on maternal mortality from all 50 states to better inform understanding of this terrible and avoidable crisis.

Likewise, mortality rates due to COVID-19 have been making somber headlines for several months and deserve the attention they are garnering. Early in the pandemic, reports began to surface that largely Black communities are disproportionately affected by COVID-19.[xiv] Reports developed across the country -- including Dallas,[xv] the Bronx,[xvi] and Detroit[xvii] – that demonstrated inequities in COVID-19 related deaths and hospitalizations across communities with low access to care and widening income inequality. In the case of COVID-19, we understand that the observed racial disparities are the result of decades of systemic racism and social determinants of health, not race itself.[xviii] This pandemic highlights the urgency in improving collection and reporting of race/ethnicity in relation to health outcomes to be able to evaluate and track progress on eliminating health inequities, which in the United States has been strongly associated with race through laws and policies.

Disaggregated data within the dimension of race/ethnicity alone are not sufficient to reduce health inequities since racial health inequities are so closely correlated with social determinants of health. The impacts of race/ethnicity are interconnected with other social dynamics – we need to simultaneously assess these areas to ascertain root causes of racial/ethnic disparities for action. Thus, disaggregated race/ethnicity data must be complemented by data on social determinants of health to truly advance health equity.

The following discussion examines three major challenges to proper data collection that must be overcome to improve data collection and reporting, including improvements to the types of data currently collected, technological barriers to collecting data, and apprehension from both providers and patients when collecting or providing data.

DISCUSSION

Challenge 1: The types of data collected are not mandated, not standardized, and not granular

2

Currently, most aspects of race and ethnicity data collection and reporting are not fully standardized, and this perpetuates the gaps in information that negatively affect people of color. In fact, health data can be collected and reported by numerous entities at the federal, state, local, and provider levels with different requirements or systems, leading to fragmented and disjointed capture of health data. See Appendix.

Standardization

The importance of standardization has been recognized on a national scale at least as far back as 2002, when the Institute of Medicine released a report with this finding:

> "Standardized data collection is critically important in the effort to understand and eliminate racial and ethnic disparities in health care. Data on patient race, ethnicity, and primary language would allow for disentangling the factors that are associated with health care disparities, help plans monitor performance, ensure accountability to enrolled members and payers, improve patient choice, allow for evaluation and intervention programs, and help identify discriminatory practices."

Section 4302 of the Affordable Care Act established standards for the collection and reporting of race, ethnicity, language, sex, and disability data in all publicly funded national health surveys. The federal government has issued guidance to improve standardization of those data, but the standards fail to reach a great number of other reporting entities at the federal, state, and local levels including administrative, billing, and medical records that are key health data collection sources.[xix]

Mandates

While standardization is crucial, the federal government has not enacted the types of mandates needed to require that the data are collected in the first place. A 2018 analysis of 125 CDC-sponsored surveillance and health monitoring data systems active from 2010-2013 found that 20% of those systems did not collect any such data, and a small minority of systems (21%) contained detailed race information; authors defined data on race and ethnicity as "detailed if the system collected more detailed data on race (e.g., Vietnamese) or Hispanic/Latino ethnicity (e.g., Puerto Rican)" than the Office of Management and Budget (OMB) standard five racial and two ethnic

categories discussed below. The systems analyzed included nationally notifiable disease systems, annual health surveys, disease registries, vital records, and hospital discharge data systems.[xx]

A number of states have laws or regulations related to the collection of race and ethnicity data needed to help address disparities. The degree to which those mandates are enforced varies, and enforcement could be key to improvement. According to one expert, enforcing state mandates is difficult because collecting race and ethnicity data is a sensitive topic.[xxi] While states have the ability to reject hospital data when key fields are missing, this is not likely to occur with race and ethnicity data. [xxii]

A survey of hospitals conducted when twenty-two states had a mandate to collect race/ethnicity data revealed that only about 50% of respondents were aware of the mandate.[xxiii] In addition, 5% of hospitals located in states with a mandate to collect race/ethnicity data did not do so. According to the researchers who conducted the survey, "this gap between knowledge and practice, and the apparent lack of enforcement of mandates, demonstrates the need for policies at the state and/or federal level regarding the collection of complete and accurate race/ethnicity data to improve quality of care and reduce disparities."[xxiv]

Granularity

At present, even in the minority of cases when government entities collect data crucial to improving disparities, the data lack the granularity needed for meaningful change. Entities that collect certain demographic data based on federal guidelines for reporting must follow standards set forth by OMB. Specifically, OMB provides minimum standards for collecting data on race across five broad categories: American Indian or Alaska Native (AIAN), Asian, Native Hawaiian or other Pacific Islander (NHOPI), Black, or white.

Two categories for ethnicity are also part of the OMB standards - Hispanic or Latino, and Not Hispanic or Latino. Race and ethnicity have historically been collected as separate questions, despite recent research encouraging otherwise. In fact, OMB standards require the 2-question format for self-identification purposes. If self-identification is not possible, then it allows for observers to use the one question format and encourages them to answer both

race and ethnicity questions in the same question.[xxv]

These broad categories have at least two major drawbacks. First, the categories do not provide enough detail to identify and understand disparities within smaller minority communities or populations. For example, OMB defines "Hispanic" for purposes of data collection as "a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race." It is a broad term that could encompass people from scores of countries, each with unique cultural aspects. A 2015 survey conducted by the Pew Research Center found that 50% of Hispanics most often describe themselves by their family's country of origin.[xxvi] The Asian American community in the United States is similarly heterogenous in background and culture. Due to such heterogeneity, interventions to advance health equity among Hispanic and Asian American populations should be culturally sensitive and relevant to the intended communities. Second, when broad categories fail to align with how individuals self-identify, people tend to overuse the "Some Other Race" category, which represented the third largest group in the 2000 and 2010 US census. [xxvii, xxviii]

When state, local, or private entities collect certain data, they are not required to use the OMB standards, which creates an additional challenge to the lack of granularity. Data collected by those entities may not be comparable enough for researchers and policymakers to understand how health systems are failing certain populations. Gaps in data collection during the COVID-19 pandemic highlight this challenge. A recent article found that at the time of publicatio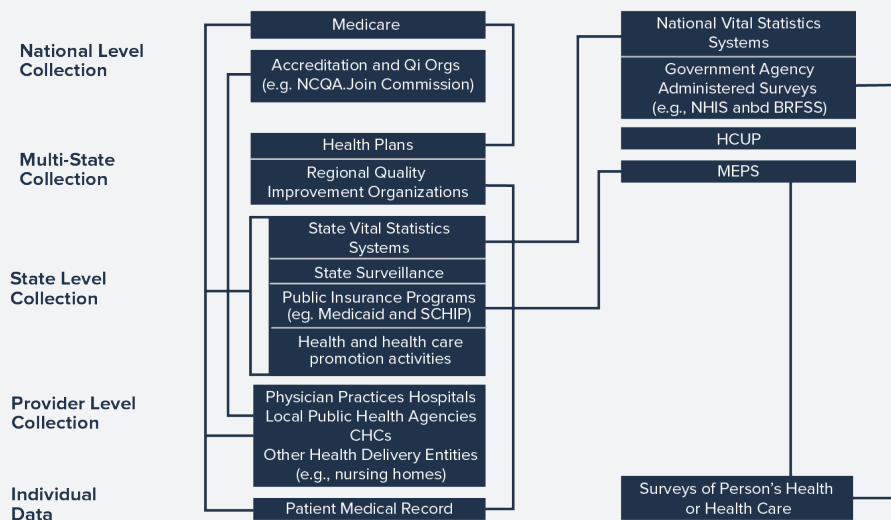n, 39 of 47 states with hospital discharge data systems were collecting race and ethnicity data; however, reporting entities had created their own systems for classification and coding.[xxx] Given this data backdrop, COVID-19 researchers have been faced with navigating a fractured landscape, resulting in crucial time being lost for realizing the incredible disparity.

Challenge 2: Data collection systems are outdated and do not allow for interoperability

Health data are collected and reported through a network of public and private entities that are fed in large part by administrative, billing, and medical records used by hospitals and physicians.

Because patient records play a crucial role in health surveillance and monitoring, as well as understanding health outcomes, experts believe provider registration systems could serve as a reliable source of data useful for researching and alleviating health disparities. One study has found that "[i]nformation on race/ethnicity was most often collected upon admission... or at first visit/new registration. Data were collected less often through health care provider notes or the medical record, and rarely at discharge." Despite this finding, most provider registration systems still have not been required to collect such data.[xxxi] A major reason for the slow uptake of these registration systems is a combination of time and money. Health systems tend to focus resources on care. While understanding health disparities could drive important improvements in care and quality, the need for technological upgrades in registration processes likely do not seem as urgent as providing immediate care to patients. In addition to resource constraints, interoperability of

Figure 2. Agency for Healthcare Research and Quality's snapshot of data flow in a complex healthcare system[xxix]

National Level Collection	Medicare Accreditation and Qi Orgs (e.g. NCQA.Join Commission)	National Vital Statistics Systems Government Agency Administered Surveys (e.g., NHIS anbd BRFSS)
Multi-State Collection	Health Plans Regional Quality Improvement Organizations	HCUP MEPS
State Level Collection	State Vital Statistics Systems State Surveillance Public Insurance Programs (eg. Medicaid and SCHIP) Health and health care promotion activities	
Provider Level Collection	Physician Practices Hospitals Local Public Health Agencies CHCs Other Health Delivery Entities (e.g., nursing homes)	
Individual Data	Patient Medical Record	Surveys of Person's Health or Health Care

systems is a technological barrier that perpetuates gaps in data collection. While some providers or public health entities may have advanced data systems, many in the United States are well behind. Some providers still rely on fax machines or regular email to submit data, which are often missing crucial elements.

During the COVID-19 crisis, the lack of a standard, interoperable system has become a source of frustration for providers, epidemiologists, and other health experts. A scenario in Washington State, where National Guard members were called up to manually enter data that were not received electronically, illustrates the challenge public health officials and researchers face. [xxxii] In a time when contact tracing and other community information is imperative, the ability to see a full picture of health data is crucial but nearly impossible to achieve because of fractured technological systems. A recent report in the New York Times summarizes the situation in the United States: "Data often come to public health authorities using only the information that laboratories need to track the record, not the details that public health officials need to understand the disease."[xxxiii] Details are just as important to understanding disparities as they are to understanding the diseases themselves.

Challenge 3: Providers and patients are apprehensive about race and ethnicity data collection

Challenges for data collection are not limited to regulatory and technical issues. Providers and patients both report discomfort with collecting and providing information that is needed to address disparities in the United States health system.

Self-reported race and ethnicity data is considered the most accurate by many researchers. When another party reports data on race or ethnicity, high instances of inaccurate information have been reported. One study of the accuracy of observer-reported data found that nearly 30% of African Americans, 40% of Hispanics, 70% of Asians, and nearly all Native Americans participating in the study were classified differently than their self-reported information.[xxxiv] The study also provided that "self-reported Asians had the most administrative classifications of 'unknown' and self-reported American Indians the most occurrences of being falsely classified in the administrative data." Those inaccuracies can lead to gaps in data and underestimation of health disparities. For example, the 2006 National Healthcare Disparities Report

"could not produce statistically reliable estimates for most quality and access measures for the NHOPI population, about half of the quality measures and one-quarter of the access measures for the AIAN population, about one-third of the quality measures for Asians, and most quality measures for people of more than one race."[xxxv]

Providers express many concerns, including "privacy, the legality of collecting the information, possible resistance from patients and staff, difficulty recording the data, and uncertainty about whether the data would be useful."[xxxvi] For example, according to at least one expert, providers must sometimes deviate from various standards of care according to individual patient needs. Reluctance to collect certain types of data may stem fear of being labeled racist or negatively evaluated by administrators and payors when deviating from those standards.[xxxvii]

Additionally, providers may not believe racism or disparate outcomes are a problem in their practices, despite evidence that providers may contribute to health disparities unknowingly as a result of implicit or unconscious bias. Recent research shows that even when controlling for income, age, disease severity, and health insurance coverage, Black patients were less likely to receive appropriate care for certain cardiac and kidney conditions and less likely to receive the best treatment for stroke, AIDS, and cancer, though many effective treatment options currently exist.[xxxviii] These findings point to an uncomfortable reality and the lived experience of many people of color—some people in the US are more likely to die because of implicit bias held by providers about their race or ethnicity, not just because they lack access to health care or financial resources. If providers view data collection as relevant only to their own practices, instead of the larger system and the systemic racism that drives disparities, they may be less inclined to see the importance of collecting data.[xxxix]

Concerns from patients are in many ways tied to the legacy of racism and exploitation in the U.S. health systems. History, even very recent history, is replete with examples of medical experimentation on people of color, from the Tuskegee Syphilis Experiment, to the harvesting of cells from Henrietta Lacks, to the history of forced sterilization of Black and Brown women that continues presently, and numerous historical examples causing medical mistrust that resonate in minority communities to this day. Patients understandably cite mistrust and concerns that data will be used for

discriminatory practices or adversely affect care when
surveyed about hesitation in providing certain data.[xl]

CONCLUSION

In the United States, racial disparities in health
outcomes are well-documented and researched, as
is the lack of data needed to address the inequities
in U.S. health systems. Accurate, actionable data
is essential to improving health outcomes for all
Americans and crucial to remedying the disparities
faced by people of color. Missing information related
to race, ethnicity, and language result in our nation's
inability to knock down health barriers, and a lack of
urgency to implement change reinforces that health
disparities in the United States are deeply rooted in
systemic racism and devaluing of Black and Brown
lives.

In this paper, we have explained the importance of
data collection on race/ethnicity and some of the
challenges to improving data collection:

- The types of data currently collected lack the
  granularity and standardization needed to draw a
  complete picture of the health challenges that result
  from systemic racism.

- The systems used to collect data are outdated,
  and communities lack the resources needed
  to implement more advanced, interoperable
  technology.

- Race is a fraught issue, and providers are either
  not aware of the importance of race and ethnicity
  data for the health system as a whole or are wary
  of collecting the information out of fear it could be
  used to question their practices. Patients may also
  be unwilling to share such personal information
  because of the history of exploitation that Black and
  Brown communities have endured.

Perhaps the biggest question in this discussion is,
"why are we still in this position?" As evidenced by
the research cited in this paper, and the mountains
of other studies, surveys, and calls to action, the
importance of data and the large gaps that exist in U.S.
health systems are well-known and well-documented.
What more can be done? Why has it not been done?
In our next paper we will explore some of the solutions
that have been proposed, and also the barriers to
implementation.

References

i   Statement of the American Medical Association. Re: The Disproportionate Impact of COVID-19 on Communities of Color. Submitted to the U.S. House of Representatives Committee on Ways and Means. 8 June 2020. Available at: https://searchlf.ama-assn.org/undefined/document-Download?uri=%2Funstructured%2Fbinary%2Fletter%-2FLETTERS%2F2020-6-8-AMA-Statement-For-the-Record-on-Racial-Disparity-Hearing.pdf

ii   The COVID tracking project, https://covidtracking.com/race/infection-and-mortality-data, accessed December 2, 2020

iii Mark Zarefsky. Why African American Communities are Being Hit Hard by COVID-19. American Medical Association. 13 May 2020. Available at: https://www.ama-assn.org/delivering-care/population-care/why-african-american-communities-are-being-hit-hard-covid-19

iv   Dayna Bowen Matthew, Edward Rodrigue, and Richard V. Reeves, Time for justice: Tackling race inequalities in health and housing, October 19, 2016, https://www.brookings.edu/research/time-for-justice-tackling-race-inequalities-in-health-and-housing/

v   Demographic Trends of COVID-19 cases and deaths in the US reported to the CDC. Centers for Disease Control and Prevention. December 2, 2020. Available at: https://covid.cdc.gov/covid-data-tracker/#demographics. Accessed December 2, 2020.

vi   Brooks OT, Burton LA. Better collection, reporting of race and ethnicity data are essential for fighting COVID-19. STAT. December 2, 2020. Available at: https://www.statnews.com/2020/12/02/comprehensive-reporting-race-ethnicity-covid-19-data/

vii   FAQ Data and Surveillance. CDC. Nov 20, 2020. Available at: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/faq-surveillance.html

viii Thomas A. LaVeist and Lydia A. Issac. (2013). Race, ethnicity, and health: a public health reader, 2nd edition. Jossey-Bass.

ix   Report of the Secretary's Task Force on Black and Minority Health, Task Force on Black & Minority Health, and Task Force on Black and Minority Health. United States. Department of Health and Human Services. Task Force on Black and Minority Health. Washington, D.C.: U.S. Dept. of Health and Human Services, [1985-1986], http://resource.nlm.nih.gov/8602912

x   Dorothy Roberts (2011). Fatal Invention. The New Press.

xi   Samantha Artiga. Health Disparities are a Symptom of Broader Health Inequities. Kaiser Family Foundation. 1 June 2020. Available at: https://www.kff.org/policy-watch/health-disparities-symptom-broader-social-economic-inequities/

xii Camara Phyllis Jones. Levels of Racism: A Theoretical Framework and a Gardener's Tale. Am J Public Health. 2000; 90:1212 – 1215.

xiii Infographic: Racial/Ethnic Disparities in Pregnancy-Related Deaths — United States, 2007–2016. Centers for Disease and Control. Updated February 2020. Available at: https://www.cdc.gov/reproductivehealth/maternal-mortality/disparities-pregnancy-related-deaths/infographic.html

xiv Millett GA, Jones A T, Benkeser D, et al. Assessing differential impacts of COVID-19 on black communities. Annals of Epidemiology. 2020; 47: 37-44. Available at: https://doi.org/10.1016/j.annepidem.2020.05.003

xv   Holly K. Hacker and LaVendrick Smith. COVID-19 has hit black neighborhoods in Dallas hard. Here's why. The Dallas Morning News. May 2020. https://www.dallasnews.com/news/public-health/2020/05/28/covid-19-has-hit-black-neighborhoods-in-dallas-hard-heres-why/

xvi COVID-19 outbreak exposes generations-old racial and economic divide in New York City. ABC News. 20 May 2020. Available at: https://abcnews.go.com/US/covid-19-outbreak-exposes-generations-racial-economic-divide/story?id=70796081

xvii Shamus KJ, Tanner K, and Warikoo N. Wayne County, Detroit account for 47% of state's coronavirus cases. Detroit Free Press. 3 April 2020. Available at: https://www.freep.com/story/news/health/2020/04/03/coronavirus-covid-19-cases-wayne-county-detroit-michigan/5116620002/

xviii Julia Craven. It's not too late to save black lives. SLATE. 21 May 2020. Available at: https://slate.com/news-and-politics/2020/05/covid-19-black-communities-health-disparity.html

xix HHS Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status. Office of ASHPE, US Department of Health and Human Services. 31 October 2011.

xx   Rodriguez-Lainz, Alfonso et al. "Collection of Data on Race, Ethnicity, Language, and Nativity by US Public Health Surveillance and Monitoring Systems: Gaps and Opportunities." Public health reports (Washington, D.C. : 1974) vol. 133,1 (2018): 45-54. doi:10.1177/0033354917745503

xxi Expert interview conducted and summarized by Leavitt partners.

xxii Id.

xxiii Hasnain-Wynia, Romana, Debra Pierce, and Mary A Pittman. 2004. "Who, When, and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals." New York: Commonwealth Fund, 42.

xxiv Id.

xxv OMB (Office of Management and Budget). 1977. Statistical policy directive No. 1 , race and ethnic standards for federal statistics and administrative reporting. http://wonder.cdc.gov/wonder/help/populations/bridged-race/Directive15.html

xxvi Ulmer, Cheryl, Bernadette McFadden, and David R Nerenz. 2009a. "Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement, Defining Categorization Needs for Race and Ethnicity Data." National Academies Press. http://www.ahrq.gov/research/findings/final-reports/iom-racereport/reldata3.html.

xxvii Hasnain-Wynia, Romana, Kevin Van Dyke, Mara Youdelman, Christian Krautkramer, Susan L. Ivey, Robert Gilchick, Erin Kaleba, and Matthew K. Wynia. 2010. "Barriers to Collecting Patient Race, Ethnicity, and Primary Language Data in Physician Practices: An Exploratory Study." Journal of the National Medical Association 102 (9): 769–75. https://doi.org/10.1016/S0027-9684(15)30673-8.

xxviii Howland, Renata E, and Tsu-Yu Tsao. 2020. "Evaluating Race and Ethnicity Reported in Hospital Discharge Data and Its Impact on the Assessment of Health Disparities:" Medical Care 58 (3): 280–84. https://doi.org/10.1097/MLR.0000000000001259.

xxix Howland, Renata E, and Tsu-Yu Tsao. 2020. "Evaluating Race and Ethnicity Reported in Hospital Discharge Data and Its Impact on the Assessment of Health Disparities:" Medical Care 58 (3): 280–84. https://doi.org/10.1097/MLR.0000000000001259.

xxx Hasnain-Wynia, Romana, Kevin Van Dyke, Mara Youdelman, Christian Krautkramer, Susan L. Ivey, Robert Gilchick, Erin Kaleba, and Matthew K. Wynia. 2010. "Barriers to Collecting Patient Race, Ethnicity, and Primary Language Data in Physician Practices: An Exploratory Study." Journal of the National Medical Association 102 (9): 769–75. https://doi.org/10.1016/S0027-9684(15)30673-8.

xxxi Hasnain-Wynia, Romana, Debra Pierce, and Mary A Pittman. 2004. Who, When, and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals. New York: Commonwealth Fund, 42. https://www.commonwealthfund.org/sites/default/files/documents/___media_files_publications_fund_report_2004_may_who__when__and_how__the_current_state_of_race__ethnicity__and_primary_language_data_collection_in_ho_hasnain_wynia_whowhenhow_726_pdf.pdf.

xxxii Kliff, Sarah, and Margot Snager-Katz. The New York Times. July 13, 2020. https://www.nytimes.com/2020/07/13/upshot/coronavirus-response-fax-machines.html

xxxiii Id.

xxxiv Boehmer, Ulrike, Nancy R Kressin, Dan R Berlowitz, Cindy L. Christiansen, Lewis E. Kazis, and Judith A. Jones. 2002. "Self-Reported vs Administrative Race/Ethnicity Data and Study Results." American Journal of Public Health 92 (9): 1471–72. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1447261/.

xxxv Bilheimer, Linda T, and Jane E Sisk. 2008. "Collecting Adequate Data on Racial and Ethnic Disparities in Health: The Challenges Continue." Health Affairs (Project Hope) 27 (2): 383–91. https://doi.org/10.1377/hlthaff.27.2.383.l

xxxvi Hasnain, et al. 2010.

xxxvii Expert interview conducted and summarized by Leavitt Partners.

xxxviii Bridges, Khiara M., Implicit Bias and Racial Disparities in Health Care. American Bar Association. Available from https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/the-state-of-healthcare-in-the-united-states/racial-disparities-in-health-care/

xxxix Expert interview conducted and summarized by Leavit Partners.

xl Rubin, Victor, Danielle Ngo, Angel Ross, Dalila Butler, and Nisha Balaram. 2018. "Counting a Diverse Nation: Disaggregating Data on Race and Ethnicity to Advance a Culture of Health." PolicyLink. https://www.policylink.org/sites/default/files/Counting_a_Diverse_Nation_08_15_18.pdf.

Appendix A. Data Collection Sources, Mandates, Standards, and Granularity

	Mandates	Standardization	Granularity
Federal surveys	Section 4302(a) of the ACA required HHS to establish a data collection standard for national population health surveys that is aimed at understanding health disparities	HHS developed a disaggregated standard based on OMB standard for national population health surveys that rely on self-reporting	HHS standard is more granular than OMB minimum standard, however, it still leaves out major groups such as Middle Eastern/ North African
	Other surveys are subject to various requirements which may or may not require collection of race/ethnicity data	Provider-based surveys rely on provider records which do not follow these standards	Sample sizes are sometimes not large enough to support granular publishing of data
		Not applicable to surveys that are not considered population health surveys	Other Federal surveys have option to use minimum OMB standard
Medicare	Medicare Enrollment Database (EDB) includes race and ethnicity data	The primary data source is administrative records of the Social Security Administration (SSA) with some limited self-reported data	Granularity and quality vary considerably by racial/ ethnic group with the data being more accurate for white and Black populations
		Data collection standards have changed over time and so the race/ethnicity categories vary by beneficiary depending on when the data were collected	
Medicaid and CHIP	Required by Section 4302(b) of the ACA to collect race/ ethnicity data based on the HHS standard developed for Federal population health surveys	Ongoing data quality concerns as evidenced by analysis of the Medicaid Statistical Information System (MSIS) data including large amounts of missing data	HHS standard more granular than OMB minimum standard, however, it still leaves out major groups such as Middle Eastern/ North Africa
State surveys	Race/ethnicity data collection on these surveys is at the discretion of the state/ survey sponsor	Race/ethnicity data collection on these surveys is at the discretion of the state/survey sponsor	Race/ethnicity data collection on these surveys is at the discretion of the state/ survey sponsor
		Lack of standardization among states can make it challenging to aggregate	States can use a level of granularity that allows them to collect data on the populations in their specific state

Vital Records	Race/Ethnicity is a required part of the US Standard Certificate of Death. The US Standard Certificate of Live Birth recommends inclusion of parental race/ethnicity but does not require it. Newborn race/ethnicity is not recorded on the US Standard Certificate of Live Birth.	Ongoing data quality concerns particularly for death records since it is recorded by a third party	The granularity of State collection of vital records is primarily dependent on state compliance and unique needs
		States, Territories, DC, and NYC compromise 57 distinct jurisdictions and can adapt the US standard to meet their local needs	US standard certificates follow OMB Standards and go beyond the required ethnicity/race categories listed in the OMB
Private entities (e.g., hospitals, provider groups, health plans, etc.)	No direct national mandates	Lack of standardization among private entities makes it challenging to aggregate across entities	Lack of granularity or small sample sizes for very granular categories can make it difficult to discern disparities among smaller populations
	Unrelated statutes result in an incidental mandate -- for example, hospitals must comply with Title VI of the Civil Rights Act that prohibits discrimination on the basis of race in order to be reimbursed by Medicare/Medicaid. This results in a requirement to collect race/ethnicity data for the purpose of documenting the absence of discriminatory behaviors.	Data collection processes are at the discretion of the private entities	Varies between entities