# Data gaps for race and ethnicity are holding back antiracism efforts, new report says

June 22 2022, by Andrew Thurston



"We must standardize racial data collection across the United States," according to Ibram X. Kendi, director of the BU Center for Antiracist Research. Credit: Janice Checchio for Boston University Photography

During the first year of the pandemic, Black people accounted for a

disproportionate [15 percent of COVID-19 deaths](#) in the United States. That number is bleak, but it's also probably incomplete: it includes only deaths where an official took the time to log the deceased's race and ethnicity. In many cases, that information was not recorded, meaning the full unequal impact of the pandemic may never be known.

With few national rules on reporting race and ethnicity, such gaps in our knowledge aren't unusual. They don't just hinder what we know about pandemic-era disparities, but also hold back our understanding of how policies on housing, education, and employment affect different communities and people in different ways. Now, a new report from the BU Center for Antiracist Research aims to highlight the nation's race and ethnicity data shortcomings—and offer solutions for fixing them.

*Toward Evidence-Based Antiracist Policymaking: Problems and Proposals for Better Racial Data Collection and Reporting* details the center's own experiences of trying to collect demographic data and gives a roadmap—for the [federal government](#) and state and local officials—for pulling in more robust information.

"Not only has the COVID-19 pandemic exposed that racism is a [public health crisis](#), but the pandemic has also revealed the gaps in reporting by race and ethnicity at the county, state, and federal levels," says Ibram X. Kendi, the center's founding director and BU's Andrew W. Mellon Professor in the Humanities. "We must standardize racial data collection across the United States, as this groundbreaking report shows. This standardization is essential to seeing and eliminating racial disparities."

## A national standard for racial and ethnic demographic data

What we do know about the racial distribution of pandemic

deaths—including that 15 percent figure—didn't come from a federal database or get computed by an algorithm. It was the result of a painstaking manual data collection effort by a team of researchers, journalists, and volunteers at the COVID Racial Data Tracker—a joint project of the Center for Antiracist Research and *The Atlantic*. The tracker ran from April 2020 to March 2021 and attempted to measure the pandemic's impact by race and ethnicity. In the absence of an official database, it became a go-to resource for officials and journalists.

But when the tracker's team first began sourcing COVID-19 data from state officials and local public health authorities, they kept running into the same major problems when it came to race and ethnicity: missing and inconsistent data. Most states had some data linking cases and race, but not all: New York, for example, didn't collect any race data. And although Texas technically did, it only had information for 3 percent of cases. As researchers got further into the numbers, they found even more race-related data gaps: patchy information for hospitalizations and testing, inconsistent reporting methods, rounding errors, variability in racial and ethnic categories, infrequent reporting.

The new report channels those experiences—alongside those from the center's ongoing work to build a racial data tracker for other areas, including houselessness, criminal arrests, and police violence—into 11 recommendations for correcting the nation's data failings. The first is a call for federal leadership on centralizing and standardizing "racial and ethnic demographic data collection and reporting across critical issue areas." Those areas include health, housing, employment, education, environment, and more.

"We think the gold standard is a nationwide system that can ensure standardization," says Neda A. Khoshkhoo, a report coauthor and associate director of policy at the Center for Antiracist Research. "If you want to have an effective policy response in the context of COVID, for

example, and you don't see which groups are being most harmed, how can you craft a solution that takes into account where more resources should be concentrated?"

## A fuller picture of racism

The report recommends the government use financial incentives to nudge states and localities into timely and accurate data collection, create an external oversight board to make sure data is used responsibly and to monitor noncompliance, and review and update current race and ethnicity data standards—something that may be edging closer after the country's chief statistician, Karin Orvis, announced a [formal review into the Office of Management and Budget's standards](#) for collecting and presenting federal data.

Although Khoshkhoo acknowledges that's all going to take time, she says there are changes that even small institutions—from towns to school districts to hospitals, which all have myriad forms and questionnaires—can start making.

"Everyone has a role in this," she says. "There are policy recommendations—for state or local entities, whoever is collecting the data—that can be implemented separately."

Those suggestions include making data public and easily accessible, allowing people to see national statistics *and* drill down to the local numbers, and encrypting and anonymizing data. The report also shares ideas for getting more intersectional data, so race and ethnicity can also be paired with characteristics and variables like sex, gender identity, sexual orientation, disability, income, linguistic usage, and socioeconomic status.

When it comes to the benefits of getting more granular with the data,

Khoshkhoo gives the example of Hawaii. The COVID Racial Data Tracker team found the state broke its Asian, Native Hawaiian, and Pacific Islander statistics into subgroups. That allowed it to spot nuances that would otherwise have been missed: like the fact that Filipino people—who make up 16 percent of the state's population—accounted for 20 percent of COVID cases, while Japanese people, 15 percent of the state's population, made up just 7 percent of its cases.

"The data really allows folks to see trends, to see inequity across group experiences overall," says Khoshkhoo. "When you have data problems, including inaccurate, incomplete, and infrequently updated data, that imprecision really clouds the picture of where the racism is concentrated—that makes it harder to see and then it's harder to respond."

## Rooting out injustice and dismantling racism

Although the report acknowledges the potential for the misuse of data, including "to fuel racist stereotypes and falsely justify racist policies," it says collecting it is vital to a better understanding of racism: "The antiracist solution to racial inequity is not to avoid tracking racial inequities for fear of their misuse," according to the report, "but to actively educate the public about inequities and the policies that contribute to them."

"Racism is everywhere—we know that even in the absence of data," says Khoshkhoo. "The purpose in collecting this data is really to see exactly how and where racism is manifesting."

According to Geoff Holtzman, a postdoctoral quantitative research associate at the Center for Antiracist Research, the current data gaps don't make valid and valuable research on racism impossible—just much harder.

"They mean that in many instances, those who use racial data have to do exponentially more work—finding data, cleaning data, validating data, merging data—to get the data into a reliable state," says Holtzman, who also contributed to the report.

The report concludes that a strong race and ethnicity data infrastructure is essential to "the process of analyzing, contextualizing, and dismantling racism."

"Without racial data there is an erasure of [racial disparities](#)," says Jasmine Gonzales Rose, a report coauthor, Center for Antiracist Research deputy director of research and policy, and a BU School of Law professor. "We simply need racial data to root out racial injustice. This report is a call for action."

  **More information:** Full report: [www.bu.edu/antiracism-center/f … ist-Policymaking.pdf](#)

Provided by Boston University