

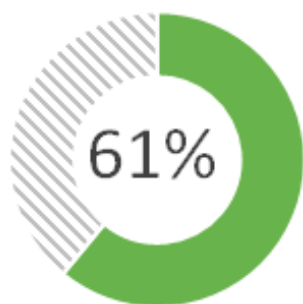
Improving Equity with Data: Vaccine Reporting and Data Visualizations



Collecting accurate and complete race and ethnicity data on routine and COVID-19 vaccinations is critical to addressing vaccine disparities and allocating resources effectively.

Using "unknown" or "other" race and ethnicity categories

During the COVID-19 pandemic, the CDC has mandated that all COVID-19 vaccinations administered be reported in each state's Immunization Information System (IIS), which is a confidential, population-based, computerized database that records all immunization doses administered by participating providers to persons residing in each state. However, large percentages of race and ethnicity data are reported as "unknown" or "other".



As of October 4, 2021, CDC reported that race/ethnicity was known for 61% of people who had received at least one dose of the COVID-19 vaccine. ([Kaiser Family Foundation](#))

New multiracial categories for data reporting and visualization

The CDC, as of Aug 31, changed their reporting to align better with census data. Now, people who reported more than one race were recategorized into the single race category selected in addition to other in the data and in visualizations.

Consolidating race and ethnicity data collected from different sources: The New York Times methods

- The NYT removes unknown data from its visualizations, even when unknown is a large race category.
- Rather than reporting race and ethnicity by dose, states should report by person. Reporting by dose means an individual could be counted two or more times. States that reported race and ethnicity by dose were excluded from the NYT analysis.
- If ethnicity is not reported, it is removed from the ethnicity analysis dataset, and if over 50% of ethnicity is unknown, the state is not included in ethnicity analysis.
- Datasets with large percentages of "other" race categories were excluded from analysis.

Improving Equity with Data: Improving Quality and Completeness in Immunization Data



Accounting for unknown race and ethnicity data: The CDC methods

- The CDC currently uses a method where race and ethnicity data are mutually exclusive. That means, when ethnicity data is missing, non-Hispanic ethnicity is assumed along with the race selected.
- An alternative method is not having race and ethnicity as mutually exclusive categories. When this is the case, it is more difficult to categorize unknown data
- These two methods lead to differences in race and ethnicity analyses, and more studies are needed to determine best practices.

Recategorizing unknown race and ethnicity data: Colorado

Department of Health and Environment methods

- Unknown data could be paired with data from state health information exchange databases to find race and ethnicity data
- To estimate the race and ethnicity breakdown of unknown data, a geographic analysis could be used to understand the general demographics of race and ethnicity in an area based on census tract or zip code
- Bayesian Improved Surname Geocoding could be used to predict race and ethnicity based on geography and surname

Recommendations for improving data quality

From the New York City Department of Health and Mental Hygiene

- Ask patients to self-identify race and ethnicity. Never assume a patient's race or ethnicity.
- Select "Prefer not to answer" if a patient chooses not to disclose race and ethnicity, rather than skipping the question.
- If patients do not understand the terms Hispanic or Latino, share the federal government's definition: The United States Office of Management and Budget defines "Hispanic or Latino" as a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin regardless of race.
- Collect each patient's race, ethnicity, and ZIP code in the electronic health record intake system. Routinely report this data to public health reporting systems and on electronic laboratory reports.
- Develop staff trainings, scripts for asking for race and ethnicity, and talking points and answers to frequently asked questions to help staff accurately collect race and ethnicity data