CDC Found Ways To Use Data To Understand and Address COVID-19 Health Disparities, Despite Challenges With Existing Data
CDC Found Ways To Use Data To Understand and Address COVID-19 Health Disparities, Despite Challenges With Existing Data

As the Nation’s lead public health agency, CDC works to increase national health security and to respond to public health threats as they arise. In doing this, CDC receives public health data from State, Territorial, and local entities (jurisdictions) as well as from providers on a mostly voluntary basis. CDC collects patient-level data on several COVID-19 indicators, including testing, cases, hospitalizations, and deaths. Some of these data contain information on patients’ race and ethnicity, as well as indicators of their socioeconomic status.

Why OIG Did This Review
The COVID-19 pandemic has devastated communities across the United States, and socially vulnerable populations have been disproportionately affected. From the beginning, reports indicated that people of color and people from economically disadvantaged communities were at increased risk of becoming sick from COVID-19, of needing intensive care due to COVID-19, and of dying from COVID-19 compared to members of predominantly White and/or affluent communities. At the same time, reports revealed that demographic data on COVID-19 were incomplete, which could make understanding and addressing disparities more difficult.

How OIG Did This Review
In March and April 2021, we conducted structured interviews with CDC staff responsible for collecting and analyzing COVID-19 data; collaborating with State, local, and Territorial entities and with Tribal Epidemiology Centers (TECs); and developing initiatives to address disparities. We also conducted structured interviews with staff from six jurisdictions and two TECs representing a variety of State, local, Territorial, and Tribal entities. We asked CDC staff how they used data to identify and address disparities and about support they provided to jurisdictions and TECs throughout the pandemic. We asked jurisdictions and TECs about their collaboration with CDC and about challenges they faced related to the collection, reporting, and receipt of COVID-19 data and analysis.

What OIG Found
CDC’s racial, ethnic, and socioeconomic data for COVID-19 testing, cases, hospitalizations, and deaths have limitations and provide an incomplete picture of COVID-19 disparities. Racial and ethnic data associated with COVID-19 reporting are sometimes missing, inconsistent, or inaccurate, while socioeconomic data in this reporting are neither clearly defined nor consistently collected.

CDC has taken steps to supplement and improve these data. CDC analyzed disparities using additional data sources, such as emergency department data on COVID-like illness and qualitative data from ad hoc surveys on attitudes and beliefs. CDC also developed a methodology to identify disproportionately impacted communities of color using Census data. Additionally, CDC has worked with the entities that report COVID-19 data to improve the quality of these data at their source.

CDC reported using both these supplemental data and data on COVID-19 testing, cases, hospitalizations, and deaths to address disparities via technical assistance to partners, targeted interventions, and significant funding investments. For example, CDC has helped partners conduct focus group sessions to learn about challenges faced by disproportionately impacted communities and has used data to determine equitable locations for testing sites. CDC also reported that it has elevated health equity throughout its response efforts. For instance, it created a Chief Health Equity Officer unit and published a health equity strategy to guide its COVID-19 response.

The TECs and Territory we interviewed also noted data limitations that could inhibit their ability to identify and address COVID-19 disparities. The TECs we interviewed reported difficulty accessing public health data from CDC and
States, while the Territory we interviewed reported that it lacked the technical infrastructure to collect and report COVID-19 data.

**What OIG Recommends and Agency Response**
We recommend that CDC: (1) expand efforts both to improve racial and ethnic data associated with COVID-19 and to supplement them with additional data sources and (2) ensure that TECs have timely access to all public health data to which they are entitled. CDC concurred with both of these recommendations.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BACKGROUND</strong></td>
</tr>
<tr>
<td>Methodology</td>
</tr>
<tr>
<td><strong>FINDINGS</strong></td>
</tr>
<tr>
<td>Racial, ethnic, and socioeconomic data associated with COVID-19 testing, cases, hospitalizations, and deaths have limitations and do not provide a complete picture of disparities</td>
</tr>
<tr>
<td>Racial and ethnic data associated with COVID-19 are sometimes incomplete, inconsistent, or inaccurate</td>
</tr>
<tr>
<td>Socioeconomic data associated with COVID-19 are not clearly defined or consistently collected</td>
</tr>
<tr>
<td>Data limitations result in an incomplete picture of disparities in COVID-19 testing, cases, hospitalizations, and deaths</td>
</tr>
<tr>
<td>CDC has taken steps to supplement and improve data on COVID-19 testing, cases, hospitalizations, and deaths</td>
</tr>
<tr>
<td>CDC used existing data sources and collected additional data related to COVID-19 disparities to supplement COVID-19 data</td>
</tr>
<tr>
<td>CDC has worked to support labs and jurisdictions in addressing data collection and reporting challenges unique to COVID-19</td>
</tr>
<tr>
<td>CDC reported using multiple data sources and analyses to inform efforts to address COVID-19 disparities</td>
</tr>
<tr>
<td>Some Tribal partners and at least one Territory have experienced challenges addressing COVID-19 disparities due to problems accessing or managing data</td>
</tr>
<tr>
<td><strong>CONCLUSION AND RECOMMENDATIONS</strong></td>
</tr>
<tr>
<td>CDC should expand its efforts both to improve racial and ethnic data associated with COVID-19 and to supplement them with additional data sources</td>
</tr>
<tr>
<td>CDC should ensure that Tribal Epidemiology Centers have timely access to all public health data to which they are entitled</td>
</tr>
<tr>
<td><strong>AGENCY COMMENTS AND OIG RESPONSE</strong></td>
</tr>
<tr>
<td><strong>APPENDIX</strong></td>
</tr>
<tr>
<td><strong>ACKNOWLEDGMENTS AND CONTACT</strong></td>
</tr>
<tr>
<td><strong>ABOUT THE OFFICE OF INSPECTOR GENERAL</strong></td>
</tr>
<tr>
<td><strong>ENDNOTES</strong></td>
</tr>
</tbody>
</table>
BACKGROUND

Objectives

1. To identify limitations in the racial, ethnic, and socioeconomic data associated with COVID-19 testing, cases, hospitalizations, and deaths used by the Centers for Disease Control and Prevention (CDC).
2. To examine how CDC uses these racial, ethnic, and socioeconomic data in its response to COVID-19 disparities.
3. To describe CDC’s challenges, successes, and lessons learned in using these data and supporting jurisdictions and Tribal Epidemiology Centers in response to the COVID-19 pandemic.

The COVID-19 pandemic has had a historic impact on the Nation, with Black or African American, American Indian or Alaska Native (AI/AN), Asian, Native Hawaiian or Other Pacific Islander, and Hispanic or Latino individuals and members of economically disadvantaged communities often hard-hit. Consistent with wider health trends, available data indicate that members of these communities have been more likely to experience poor outcomes from COVID-19 than members of predominantly White and/or affluent communities. Specifically, throughout the pandemic, reports indicated that Hispanic or Latino, Black, AI/AN, and Asian individuals had higher rates of infection, hospitalizations, and deaths compared to their White counterparts, and residents of lower-income neighborhoods are more likely not only to contract COVID-19 but also to die from it. Furthermore, data also indicate that members of other groups, such as people who are experiencing homelessness, people with disabilities, and people who are incarcerated, face increased risk of poor outcomes from COVID-19.

Early in the pandemic, concerns arose that CDC was not collecting, or not reporting, the racial and ethnic information for COVID-19 testing and cases. CDC is the Nation’s lead public health agency, responsible for providing health information and responding to health emergencies. We conducted this evaluation to identify limitations in CDC’s racial, ethnic, and socioeconomic data for COVID-19 and to examine how CDC used these data to understand and address COVID-19 disparities.
Public Health Reporting

CDC relies on various State, Territorial, and local entities, hereinafter referred to as jurisdictions, to collect and report data for public health purposes. Providers such as laboratories, hospitals, and outpatient health care facilities submit most of these data, including data on COVID-19 testing, cases, hospitalizations, and deaths. In general, the authority to mandate public health reporting rests with States and Territories, not the Federal Government. Jurisdictions reporting to CDC typically send anonymized data (i.e., data that cannot be traced to a specific person) to CDC. In most instances, reporting to CDC is voluntary.

Changes during the COVID-19 pandemic introduced mandatory reporting of some public health data to Federal agencies. Under the Coronavirus Aid, Relief, and Economic Security (CARES) Act in March 2020, Congress required every lab that performs or analyzes a test intended to detect or diagnose a possible case of COVID-19 to report test results to the Secretary of the Department of Health and Human Services (HHS). The statute authorized the Secretary to prescribe the form and manner, and timing and frequency, of such reporting. HHS’s subsequent guidance established August 1, 2020, as the implementation date for this reporting, provided instructions to all laboratories on the specific required data elements to be reported daily to CDC, and outlined avenues for reporting, including through existing public health reporting mechanisms. Beginning in May 2020, the Centers for Medicare & Medicaid Services (CMS) also began requiring aggregate weekly reporting of COVID-19 data from nursing homes.

CDC’s Data Modernization Initiative

CDC’s efforts to modernize public health data and reporting accelerated after the COVID-19 pandemic exposed gaps in the public health data system. CDC’s data modernization initiative includes several activities, such as creating interoperable a

---

a Providers submit public health data to public health departments. Some public health departments report these data to CDC directly, while others report to CDC via the public health department of a higher jurisdiction. For example, some local public health departments report directly to CDC while others report via the public health departments of their States.

b In May 2020, CMS began requiring skilled nursing facilities and nursing facilities to report, on a weekly basis, confirmed and suspected COVID-19 infections, total deaths, and COVID-19 deaths among residents and staff, per 85 Fed Reg 27550.
systems in public health and health care to make data sharing easier, upgrading and modernizing technology infrastructure, and strengthening the public health workforce. The data modernization initiative also includes increasing the use of both electronic case reporting and electronic lab reporting.\(^c\)\(^,\)\(^14\) One goal of this initiative is to make important data, including racial and ethnic data, more complete.

Congress has provided $1.1 billion in support of data modernization activities since shortly before the pandemic began. Prior to the pandemic, CDC received dedicated modernization funding, for the first time, through its fiscal year (FY) 2020 appropriations.\(^,\)\(^15\) When the pandemic hit, Congress provided additional funds in the CARES Act to support rapid modernization.\(^,\)\(^16\) In addition, CDC’s FY 2021 appropriations and the American Rescue Plan Act provided additional funding for data modernization activities.\(^,\)\(^17\)

CDC expanded several activities from its data modernization initiative during the pandemic.\(^,\)\(^18\) For example, CDC rapidly increased the use of electronic lab reporting from zero jurisdictions being able to report directly to CDC to 56 jurisdictions (representing all of the lab-testing volume in the country) by April 2021.\(^,\)\(^19\) In addition, CDC also greatly increased the use of COVID-19 electronic case reporting on two fronts: (1) all States, the District of Columbia, and 12 large, local jurisdictions are now capable of receiving electronic case reports, up from a handful in 2019, and (2) as of May 2022, more than 12,400 health care facilities can send case reports to jurisdictions electronically.\(^,\)\(^20\)

**COVID-19 Data on Testing, Cases, Hospitalizations, and Deaths**

Entities that report COVID-19 data to CDC transmit these data through several new and existing reporting systems. The data transmitted to CDC are individual- or patient-level data but are anonymized, meaning they cannot be traced to a specific person. CDC maintains several systems that collect COVID-19 data beyond testing, cases, hospitalizations, and deaths, but they are not the focus of this evaluation.\(^d\)

**Testing.** CDC receives information about COVID-19 tests directly from the labs and providers that conduct the tests and indirectly through public health departments as part of the CARES Act reporting requirement. CDC compiles data on positive and

\(^c\) Electronic case reporting is the automated, real-time exchange of case report information between electronic health records and public health departments. This enables health departments to obtain more complete data from patient health records, including racial and ethnic information. Electronic lab reporting is the transmission of digital laboratory reports from labs to public health departments, health care systems, and CDC. Please see CDC, *What is eCR?*, at [https://www.cdc.gov/ecr/what-is-ecr.html](https://www.cdc.gov/ecr/what-is-ecr.html) and CDC, *What is ELR?* at [https://www.cdc.gov/elr/about.html](https://www.cdc.gov/elr/about.html).

\(^d\) This evaluation did not focus on data for vaccinations or CDC’s efforts to address disparities in vaccination, despite existing concerns, as that campaign was in the early stages during our data collection period. OIG is conducting an evaluation assessing the effectiveness of State vaccination systems (OEI-05-22-00010). See OIG, *Use of States’ Immunization Information Systems To Monitor COVID-19 Vaccinations* (OEI-05-22-00010), forthcoming (see OIG Work Plan at [https://oig.hhs.gov/reports-and-publications/workplan/summary/wp-summary-0000571.asp](https://oig.hhs.gov/reports-and-publications/workplan/summary/wp-summary-0000571.asp)).
negative tests and disseminates these data along with a 7-day positivity rate on its public COVID Data Tracker.\textsuperscript{a,21} HHS's guidance details the required data to be reported, including specifics about the test itself, the result, where the test was performed, and the patient's race, ethnicity, and ZIP Code.\textsuperscript{22}

**Cases.** CDC receives information about COVID-19 cases, along with other notifiable conditions, through a disease surveillance system. Local health departments work with health care providers, hospitals, and labs to collect data on COVID-19. Providers and local public health departments report data to their State and Territorial health departments, which then report to CDC. CDC and the Council of State and Territorial Epidemiologists coordinate to develop standard surveillance case definitions for reporting each notifiable condition, along with the demographic information (i.e., race and ethnicity), location, and outcome data that should be reported.\textsuperscript{f,23}

**Hospitalizations.** CDC's Coronavirus Disease 2019-Associated Hospitalization Surveillance Network (COVID-NET) collects data about hospitalizations associated with laboratory-confirmed COVID-19 cases. Hospitals in 99 counties in the participating 14 States report detailed data for each case including the age, sex, location, race, ethnicity, clinical history, and clinical outcomes for each hospitalized COVID-19 patient.\textsuperscript{24}

CDC also has access to hospital data from other sources. The National Syndromic Surveillance Program contains data on emergency department visits that may be related to COVID-19. The Premier Healthcare Database contains hospital discharge data representing about 20 percent of U.S. hospital admissions, including encounters related to COVID-19. Additionally, CDC has access to some aggregated data on COVID-19 hospitalizations through HHS's Unified Hospital Dataset within HHS Protect.\textsuperscript{g}

**Deaths.** CDC receives information about deaths via death certificates filed with State and Territorial vital statistics offices. Death certificates are dually completed by funeral directors and medical certifiers (e.g., physicians, medical examiners, and coroners). In April 2020, CDC issued guidance indicating that COVID-19 should be specified on death certificates if it "played a role in the death."\textsuperscript{25} CDC's recommended

\textsuperscript{a} Testing data track individual tests and results, not patients. While testing data are not person-based, they do contain personal information about the people being tested, such as race and ethnicity.


\textsuperscript{g} HHS’s Unified Hospital Dataset is maintained on the HHS Protect data platform. It contains data on COVID-19 hospitalizations, such as admissions, inpatient utilization, and ICU utilization. For more information on the Unified Hospital Dataset, please see [https://covid.cdc.gov/covid-data-tracker/#abouthospitaldata](https://covid.cdc.gov/covid-data-tracker/#abouthospitaldata). For more information on the HHS Protect platform, please see the HHS Protect Public Data Hub at [https://protect-public.hhs.gov/](https://protect-public.hhs.gov/).
standard death certificate includes the decedent’s cause and place of death in addition to race, ethnicity, ZIP Code, education, occupation, and industry.26

Racial, Ethnic, and Socioeconomic Data

CDC requests that jurisdictions submit demographic data (including racial, ethnic, and socioeconomic data) but has indicated that it does not have the authority to require this information for COVID-19 cases, hospitalizations, and deaths.27 As authorized by the CARES Act and laid out in HHS guidance, labs are required to report race, ethnicity, and ZIP Code for COVID-19 test results, which ultimately are reported to CDC.28

Racial and Ethnic Data. CDC’s reporting instructions generally use the standard minimum Federal Office of Management and Budget (OMB) categories for race and ethnicity.29 For race, the categories are: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White; and for ethnicity: Hispanic or Latino and Not Hispanic or Latino.29 However, jurisdictions may choose not to use these categories. Rather, jurisdictions may choose to define race and ethnicity differently. In addition, jurisdictions often develop their own reporting tools and data systems that may use different racial and ethnic categories when the data are collected.

Exhibit 1 provides examples of how information on race and ethnicity may flow to select CDC systems tracking COVID-19 testing, cases, hospitalizations, and deaths.

---

h A standard death certificate includes more racial and ethnic categories than the minimum OMB categories. Please see the U.S. Standard Certificate of Death at https://www.cdc.gov/nchs/data/dvs/DEATH11-03final-acc.pdf for more information.
Exhibit 1: CDC collects racial and ethnic data for COVID-19 testing, cases, hospitalizations, and deaths from providers and public health departments.

Testing

Providers
Collect race and ethnicity data from patients

Labs
Confirm or disconfirm COVID-19

Public health departments
May conduct followup for missing race and ethnicity data

Dataset of positive and negative test results

Cases

Providers
Collect race and ethnicity data from patients

Labs
Confirm COVID-19

Public health departments
May conduct followup for missing race and ethnicity data

National Notifiable Diseases Surveillance System (NNDSS)

Hospitalizations

Hospitals
Collect race and ethnicity data from patients

Surveillance officers
In the 14 States participating in infectious disease surveillance programs prior to COVID-19

COVID-NET

Deaths

Funeral directors
Record race and ethnicity data with assistance from friends or relatives of deceased

State vital records offices

National Vital Statistics System (NVSS)

Note: This exhibit represents racial and ethnic data on COVID-19 hospitalizations that CDC obtains through COVID-NET reporting. CDC’s additional sources of COVID-19 hospitalizations data, such as the National Syndromic Surveillance Program and the Premier Healthcare Database, may also contain racial and ethnic data.


Socioeconomic Data. While socioeconomic data may be defined in various ways depending on the entity collecting or analyzing the data, CDC collects a range of variables for COVID-19 data that could be used as socioeconomic indicators. For instance, CDC’s suggested COVID-19 case report form requests information about residence at illness onset, occupation, and industry.30 As another example, the recommended standard death certificate includes information about education, occupation, and industry.31
Federal Response to COVID-19 Disparities

CDC, Congress, and the Administration have taken numerous steps to address health disparities resulting from the COVID-19 pandemic. Several key steps are described below.¹

**CDC.** On January 21, 2020, CDC activated its Emergency Operations Center in response to the COVID-19 public health emergency.³² ³³ The Emergency Operations Center, which is activated for the duration of a public health emergency, uses an incident management system to coordinate CDC’s response to an emergency.³⁴ As part of its response to COVID-19, CDC worked to collect COVID-19 data, facilitate more rapid identification of COVID-19 hotspots, and consolidate available data. In addition, it deployed personnel to address staffing needs in State, Tribal, local, and Territorial health departments and monitored potential outbreaks and hotspots to mitigate the spread of COVID-19.

To address the importance of health equity as part of its COVID-19 response, CDC released the **COVID-19 Response Health Equity Strategy** in August 2020. The strategy provides “an evidence-based, comprehensive, and coordinated framework for reducing COVID-19 disparities.”³⁵ CDC’s strategy prioritizes expanding the evidence base, which includes having timely, complete, and representative data for race and ethnicity.³⁶

**Congress.** Since early in the pandemic, Congress has taken several steps to help address COVID-19 disparities. For instance, given concerns about the CDC’s lack of racial and ethnic data related to COVID-19, Congress passed legislation that requires the Secretary, in coordination with other appropriate departments and agencies, to disclose racial, ethnic, and other demographic data related to COVID-19.³⁷ In addition, the American Rescue Plan Act included $250 million in funding to develop and support a community-based workforce to increase COVID-19 vaccinations in underserved communities.³⁸ The American Rescue Plan Act also included funding to support State and local health departments’ COVID-19 response capabilities, which may enable them to improve access to testing, especially among communities of color.³⁹

**Administration.** President Biden signed two Executive Orders in January 2021 related to COVID-19 and health equity. One establishes a COVID-19 Health Equity Task Force to develop recommendations for mitigating the health inequities caused or exacerbated by the pandemic and preventing future inequities.⁴⁰ Another requires

---

¹ This section is not intended to serve as a comprehensive review of all Federal actions to respond to COVID-19 disparities.

² A September 2020 GAO report found that CDC’s strategy lacked critical details to achieve the priority of collecting complete data by race and ethnicity. GAO recommended that CDC determine whether having authority to require States and jurisdictions to report racial and ethnic information is necessary and, if so, to seek such authority from Congress. It also recommended that CDC involve key stakeholders, such as public health departments, to improve the completeness and consistency of demographic data. Please see [https://www.gao.gov/products/gao-20-701](https://www.gao.gov/products/gao-20-701) for more information.
increased coordination among Federal agencies regarding COVID-19 data collection, production, and analysis of data, including key equity indicators, as well as sharing such data with State, local, Tribal, and Territorial authorities.\textsuperscript{41}

Related Office of Inspector General Work

This report is part of the Office of Inspector General’s (OIG’s) strategic efforts to oversee HHS’s response to the COVID-19 pandemic, address health equity, and ensure that HHS programs have complete, accurate, and timely data. Additional OIG work in support of these priorities includes, for example, assessments of the completeness and accuracy of racial and ethnic data for Medicare and Medicaid beneficiaries.\textsuperscript{42} OIG’s goals include both ensuring that HHS is protecting the health and welfare of all Americans and promoting the effectiveness of HHS programs.

Methodology

Scope

This report focuses on the limitations of CDC’s racial, ethnic, and socioeconomic data for COVID-19 testing, cases, hospitalizations, and deaths, and how CDC used these data to understand and address COVID-19 disparities. This report does not focus on CDC’s use of racial, ethnic, and socioeconomic data for COVID-19 vaccinations, as at the time of our planning and data collection the vaccination campaign was just beginning. Our primary data collection occurred in spring 2021, with updates received from CDC through January 2022.

Data Collection

In March and April 2021, we conducted structured interviews with CDC staff responsible for collecting and analyzing COVID-19 data, collaborating with jurisdictions and Tribal Epidemiology Centers (TECs), and developing initiatives to address disparities. We also conducted structured interviews covering the same topics with staff from six jurisdictions and two TECs representing a variety of State, Territorial, local, and Tribal entities. The six jurisdictions selected for interviews were Mississippi, Nevada, North Dakota, Puerto Rico, Texas, and the District of Columbia. The two TECs selected for interviews were the Navajo Epidemiology Center and the Urban Indian Health Institute.\textsuperscript{k} We considered several factors when selecting jurisdictions and TECs for interviews, including COVID-19 burden and the degree to which jurisdictions reported racial and ethnic data. The purpose of our interviews with the selected jurisdictions and TECs was to provide useful, qualitative insights and context regarding experiences with collecting, reporting, or using racial, ethnic, and

\textsuperscript{k} The Urban Indian Health Institute is the TEC that serves Urban Indian Health programs across the country. Please see https://www.uihi.org/ for more information.
socioeconomic information associated with COVID-19 testing, cases, hospitalizations, and deaths.

We asked CDC staff how they have used racial, ethnic, and socioeconomic data related to COVID-19 testing, cases, hospitalizations, and deaths to identify and address disparities; any challenges they have faced using the data; and how they have overcome those challenges. We also asked about lessons learned during the pandemic. Furthermore, we asked how CDC uses these data to support jurisdictions and TECs, and how CDC uses these data when working with Federal agencies. Lastly, we asked CDC about its approach to health equity and how the Chief Health Equity Officer unit works throughout CDC to address COVID-19 disparities. In some cases, we received written responses to questions in lieu of an interview.

We asked staff from the jurisdictions and TECs about the racial, ethnic, and socioeconomic data they share with or receive from CDC, challenges they have encountered with the data, and how they have collaborated with CDC to use these data to address COVID-19 in their jurisdictions or among Tribes.

In some instances, we requested supporting documentation or submitted followup questions. This allowed us to clarify interview responses and review specific documents or publications discussed during the interviews.

Data Analysis

We analyzed our interview data and supporting documents to determine how CDC used racial, ethnic, or socioeconomic data to address COVID-19 disparities, and to understand barriers and challenges to using data. We looked for common themes across CDC, the jurisdictions, and TECs.

Limitations

Our analysis relied on interviews and documents provided by CDC staff and staff from the jurisdictions and TECs. We did not directly analyze CDC’s racial, ethnic, and socioeconomic data for COVID-19 testing, cases, hospitalizations, or deaths. We reviewed publicly available materials to supplement the information provided in interviews and documents. However, we did not independently verify the accuracy of the information we received. In addition, the experiences and perspectives of the six jurisdictions and two TECs that we interviewed do not necessarily reflect the experiences and perspectives of all jurisdictions and TECs.

Standards

We conducted this study in accordance with the Quality Standards for Inspection and Evaluation issued by the Council of the Inspectors General on Integrity and Efficiency.
Racial, ethnic, and socioeconomic data associated with COVID-19 testing, cases, hospitalizations, and deaths have limitations and do not provide a complete picture of disparities

Racial and ethnic data associated with COVID-19 are sometimes incomplete, inconsistent, or inaccurate

While the quality of racial and ethnic data associated with COVID-19 has improved since the beginning of the pandemic, completeness and consistency continue to vary. Race and ethnicity are still missing for many COVID-19 tests, cases, and hospitalizations. In addition, jurisdictions and health care providers define racial and ethnic categories differently, which means these data may not have the same meaning across jurisdictions. Furthermore, racial and ethnic data associated with COVID-19 testing, cases, hospitalizations, and deaths that are not self-reported may be inaccurate.

Incomplete. Although race and ethnicity for testing and cases have become more complete since reporting began, both testing and case data continue to lack racial and ethnic information for a significant portion of data. Specifically, as of January 2022, CDC reported that 67 percent of COVID-19 testing data and 34 percent of COVID-19 case data were missing race and ethnicity.43 By contrast, race and ethnicity for deaths were nearly 100 percent complete.

Challenges at various stages in the data reporting process contribute to CDC’s incomplete racial and ethnic data on COVID-19 tests and cases. As shown in Exhibit 2, COVID-19 data may flow to CDC through several stakeholders, each of which may experience challenges that impact data completeness.1

---

1 Stakeholders include the patient, the provider, health care entities, public health departments, and CDC.
Exhibit 2: Data generally flow through multiple stakeholders before reaching CDC.
Each stakeholder faces obstacles that contribute to the incompleteness of racial and ethnic data on COVID-19.

For example, patients may be uncomfortable sharing racial or ethnic information with health care providers or public health entities, and providers may be reluctant to request this information. When these incomplete data reach public health departments, the departments may lack the staff or resources needed to conduct followup. Also, health care providers and other entities that report data to public health departments sometimes report data via fax, mail, or email, which strains public health department resources as staff must manually enter these data. Even if racial and ethnic data are successfully collected, one State and CDC reported that some State laws prohibit them from sharing these data with CDC.

Furthermore, available data may not be representative of the U.S. population. For example, as of July 2021 jurisdictions varied widely in the completeness of racial and ethnic data they reported for COVID-19 testing. As a result, racial and ethnic data for COVID-19 testing may be skewed toward the racial and ethnic composition of jurisdictions reporting more complete data rather than representative of the demographics of the United States. Additionally, because COVID-NET is a population-based surveillance system, the racial and ethnic data CDC reports through it cover only about 10 percent of the U.S. population, representing data from the 14 States participating in it. While these hospitalizations data have nearly complete racial and ethnic information (99 percent), and the population covered is demographically similar to the U.S. population, COVID-NET data are not generalizable to the country.

Inconsistent or inaccurate. Racial and ethnic data associated with COVID-19 are, in some cases, inconsistent or inaccurate with respect to the populations they describe. Although standard minimum categories are defined for Federal use, officials from CDC and two of the jurisdictions we interviewed reported that variation exists in the categories used at the jurisdictional level to collect racial and ethnic data. Generally, jurisdictions have discretion in how they collect demographic data for public health surveillance. As a result, there is variation in the categories used to collect racial and...
ethnic data, and the same categories across jurisdictions and reporting entities may refer to different groups.

Some of the inconsistency of racial and ethnic data for COVID-19 arises from the way various jurisdictions aggregate or disaggregate racial and ethnic categories. For example, some State health departments combine categories for race when reporting cases to CDC such as placing AI/AN people in the same category as Native Hawaiians or Other Pacific Islanders. Additionally, jurisdictions vary in whether they treat Hispanic or Latino as a racial category or as an ethnic category. As a result, two people of the same race and ethnicity residing in different jurisdictions could appear in public health reporting to belong to different demographic groups. This impedes the ability of CDC or outside researchers to compare disparities across jurisdictions or to identify national or regional trends. See Exhibit 3 for an example.

Exhibit 3: Inconsistent racial and ethnic categories across jurisdictions could obscure the impact of COVID-19 on certain racial or ethnic groups.

State A's case reporting form lists American Indian or Alaska Native as a stand-alone racial category. If this person lived in State A, they would be reported as American Indian or Alaska Native both in the State and to CDC.

State B's case reporting form lists a combined Other racial category that includes both American Indian or Alaska Native and Native Hawaiian or Other Pacific Islander. If this person lived in State B, they would be reported as “Other” in the State, obscuring their race as American Indian.


Finally, racial and ethnic information for COVID-19 deaths may be inaccurate. Racial and ethnic data for COVID-19 deaths are nearly 100 percent complete, but previous analyses have indicated that death certificates for as many as 40 percent of AI/AN people misclassify them as White or other races. As a result, COVID-19 death data may understate the burden of COVID-19 on AI/AN people. More accurate death data would help CDC understand the true impact of COVID-19 on vulnerable populations.

---

\[^m\] CDC reported that the issue of racial and ethnic misclassification likely extends to any data type where race and ethnicity are reported by proxy.
**Socioeconomic data associated with COVID-19 are not clearly defined or consistently collected**

CDC does not set a standard for reporting socioeconomic data elements across COVID-19 testing, cases, hospitalizations, and deaths. Socioeconomic data can include a wide range of elements such as ZIP Codes, health insurance status, and occupations, and reporting entities are often the decisionmakers regarding which measures, if any, to collect. Two jurisdictions we interviewed reported that guidance and standards would be needed to improve completeness and standardization of socioeconomic data at the individual level.

In lieu of patient-level COVID-19 socioeconomic data, both CDC and jurisdictions reported using community-level data such as data from the Census Bureau or CDC’s Social Vulnerability Index (SVI). These community-level data enable CDC and jurisdictions to identify and understand community vulnerabilities and target interventions to address them. However, while CDC reported these data are useful for steering community-level interventions, they do not allow for direct analysis of individual socioeconomic disparities in COVID-19 testing, cases, hospitalizations, or deaths.

**Data limitations result in an incomplete picture of disparities in COVID-19 testing, cases, hospitalizations, and deaths**

Although CDC used racial, ethnic, and socioeconomic data to report on COVID-19 disparities, its analyses had gaps and limitations. For example, an August 2020 CDC report on COVID-19 among AI/AN people included only 23 States because the remaining States’ case data lacked sufficient racial and ethnic information. CDC has not published findings about disparities in COVID-19 cases for AI/AN people in the more than half of States that lacked sufficient data. Furthermore, in a 2020 report that CDC conducted on COVID-19 racial and ethnic disparities in “hotspot” counties, CDC had to limit its analysis to counties where at least 50 percent of case data included race. This meant that CDC could include in its analysis only 79 of the 205 counties it identified as hotspots—just 38.5 percent. Racial and ethnic data may have improved since these analyses were conducted.

---

*SVI is a tool that helps emergency response planners and public health officials identify, map, and plan support for communities that will most likely need support before, during, and after a public health emergency. It ranks each census tract on 15 American Community Survey variables to determine each tract’s relative social vulnerability. Please see [https://svi.cdc.gov/Documents/FactSheet/SVIFactSheet.pdf](https://svi.cdc.gov/Documents/FactSheet/SVIFactSheet.pdf) for more information.

Additionally, CDC may lack sufficiently granular data to understand certain disparities, even with complete racial and ethnic data. For example, some jurisdictions in our small sample reported that information on individuals’ Tribal affiliation would be useful for understanding and addressing COVID-19 in Tribes. However, jurisdictions vary as to whether they collect this information. If the data are not collected in the jurisdictions, CDC cannot receive them. Without data on Tribal affiliation, it may be more difficult for CDC, States, and Tribal partners to identify or respond to Tribal-specific outbreaks or inequities.

**CDC has taken steps to supplement and improve data on COVID-19 testing, cases, hospitalizations, and deaths**

CDC reported that it has attempted to overcome limitations in demographic data for COVID-19 testing, cases, hospitalizations, and deaths in several ways. First, CDC supplements these data by using information from multiple existing data sources and by conducting additional data collection via ad hoc surveys and listening sessions. CDC reported that this combination of supplemental data and the demographic data on COVID-19 testing, cases, hospitalizations, and deaths has been sufficient to demonstrate the existence of racial, ethnic, and socioeconomic disparities and to begin to understand their nature and magnitude. Second, CDC took steps to support the entities that report COVID-19 data to CDC to improve the quality of COVID-19 data at their source.

“While more robust race and ethnicity data would provide a more comprehensive view of the pandemic, the data reported to CDC and shared with Federal partners has been sufficient to document major disparities.”

— CDC official

**CDC used existing data sources and collected additional data related to COVID-19 disparities to supplement COVID-19 data**

Recognizing the limitations of racial, ethnic, and socioeconomic data associated with COVID-19 testing, cases, hospitalizations, and deaths, CDC has leveraged its broad array of existing public health data sources to bolster its understanding of disparities. CDC combines data on COVID-19 testing, cases, hospitalizations, and deaths with

---

\[\text{Research note: The standard death certificate includes an option to report Tribal affiliation for AI/AN individuals. Please see the } U.S. \text{ Standard Certificate of Death at } \text{https://www.cdc.gov/nchs/data/dvs/DEATH11-03final-acc.pdf} \text{ for more information.}\]
these other data sources, such as emergency department and Census data, to better understand COVID-19 disparities. For example:

- CDC uses clinical data from research and health plan networks and emergency department data on COVID-like illness.\(^q\) CDC reported that its emergency department data, which come directly from electronic health records, can offer more complete racial and ethnic data because multiple records are automatically scanned for information on patients’ race and ethnicity.
- CDC also developed a methodology to identify disproportionately impacted communities of color at the county level using Census data on race and ethnicity.\(^49\)
- CDC has used a geospatial tool—the SVI—to identify geographic areas that may be at higher risk for severe outcomes due to social factors.

CDC also gathers additional qualitative data that provide information it could not get even from complete quantitative COVID-19 data, such as insights into beliefs and attitudes. CDC uses surveys and listening sessions to gather these data on the impact of COVID-19 on diverse populations. For example, CDC conducted a survey on mental health during the pandemic for which racial and ethnic data were more than 99 percent complete.\(^50\) Furthermore, CDC conducted 18 listening sessions with diverse audiences to learn more about the beliefs and needs of disproportionately impacted populations. These surveys and listening sessions captured information on topics such as vaccine hesitancy, mask adherence, and beliefs about COVID-19, none of which are included in patient-level COVID-19 data.

**CDC has worked to support labs and jurisdictions in addressing data collection and reporting challenges unique to COVID-19**

The COVID-19 pandemic presented unique challenges that impacted how data were collected and reported. Although these challenges were not specific to racial, ethnic, or socioeconomic data, they complicated general COVID-19 data collection activities and therefore may have contributed to data quality concerns. CDC has taken steps to support labs and jurisdictions in collecting and reporting better data for COVID-19, including improved racial, ethnic, and socioeconomic data.

The scale of the COVID-19 pandemic created challenges for labs that were reporting testing data. To increase testing capacity, many labs that had not previously done routine testing opened new testing sites. CDC and jurisdictions reported that many of these new testing sites were unfamiliar with the reporting process, and many were not set up to report data electronically or to collect information on race, ethnicity, or socioeconomic status. Additionally, in contrast to most public health reporting, labs that do COVID-19 testing must report COVID-19 results for both positive and negative results.\(^51\) This meant that some labs had to develop new electronic reporting

\(^q\) CDC reported there is no broad Federal requirement for States to report emergency department data to CDC. However, some funding opportunities may require States to report these data.
systems to satisfy the requirement to report negatives. These factors have strained the resources of labs already stretched thin.

CDC took steps to support labs in reporting these data. Following the release of HHS guidance on the CARES Act—which required labs to report race, ethnicity, and ZIP Code—CDC provided labs with additional information and technical assistance for establishing electronic reporting. For example, CDC clarified which labs must report testing data, how labs could report data using standard terminology, and answered frequently asked questions. However, CDC also acknowledged that labs cannot control providers’ data collection practices or patients’ willingness to provide demographic data, stating that “laboratories should make every reasonable effort to provide” the required data elements.

Given broader concerns about inconsistency, CDC worked to standardize COVID-19 case reports—including racial and ethnic data—by creating a reporting worksheet for and providing related technical assistance to jurisdictions. The worksheet uses the minimum standard Federal racial and ethnic categories, but may not be sufficient to standardize racial and ethnic data for cases if jurisdictions’ systems are not set up to capture data at the same level of specificity as the Federal categories. For example, some jurisdictions collect racial data across multiple Federal race categories (such as Asian and Native Hawaiian or Other Pacific Islander) only in a single, umbrella category.

**CDC reported using multiple data sources and analyses to inform efforts to address COVID-19 disparities**

CDC reported that it uses the variety of data and analyses described above to inform its efforts to address racial, ethnic, and socioeconomic disparities in COVID-19 through technical assistance to partners, targeted interventions, and funding.

CDC also reported that it elevated health equity as a priority in its pandemic response, which has influenced its use of data to address disparities. CDC included a health equity unit in its COVID-19 incident response structure. For the first time, CDC placed a senior leader—the Chief Health Equity Officer—on its incident management leadership team. The health equity unit has centered data in its activities, explicitly calling for improved racial, ethnic, and socioeconomic data on COVID-19 and espousing “data-driven approaches” to advancing health equity.

**Technical assistance.** Jurisdictions and Tribal entities may request technical assistance from CDC to analyze challenges and identify data-driven solutions. As of December 2021, CDC reported that it had responded to more than 300 such requests related to health disparities. For example, CDC has responded to requests from Tribes and Indian Health Service units for support in epidemiology, data management, and health communication. In one case, CDC deployed staff to the Navajo Nation to help the Tribe assess water access and construct water-access points. This supported COVID-19 prevention efforts by increasing Tribal residents’ access to water for...
handwashing. CDC reported that it also routinely monitors COVID-19 data and has conducted outreach to populations that appear to need additional support. For example, CDC proactively reached out to Tribal communities when surveillance data indicated high COVID-19 burden.

Exhibit 4: See below for examples of technical assistance CDC provided to partners.

Partnered with the National Center for Farmworker Health to build capacity in organizations serving agricultural workers and develop COVID-19 mitigation strategies for farmworker communities

Deployed a tailored tool to help Tribes integrate COVID-19 data and visualize the impact of COVID-19 on Tribal communities

Held focus group discussions with members of Marshallese and Hispanic or Latino communities in Arkansas to identify COVID-19-related challenges and needs


Designing and implementing targeted interventions. CDC reported that it has partnered with Federal agencies to use racial, ethnic, and socioeconomic data to design and implement targeted interventions to address disparities. CDC reported that it has used these data to determine which interventions and what guidance are needed, and for whom. For instance, CDC used data to determine locations for community-based testing sites, to ensure that its messaging and ad campaigns were targeted to the communities most at risk, and to inform the distribution of masks purchased by the Government.

COVID-19-related funding opportunities. CDC reported that it addresses disparities in COVID-19 by providing funding opportunities including cooperative agreements paired with technical assistance with an emphasis on health equity. In some cases, CDC has offered funding specifically to address COVID-19 disparities among disproportionately affected communities, as in the $2.25 billion grant opportunity CDC posted through its Center for State, Tribal, Local, and Territorial Support in March 2021.55 This grant aims to address COVID-19-related health disparities among people living in rural areas as well as racial and ethnic minority groups. CDC has also offered $300 million to support community health workers in addressing COVID-19-related health disparities in high-risk populations, including communities of color.56 In other cases, CDC awarded funds to the CDC Foundation, which gave jurisdictions resources to support health equity.57 CDC also reported that it considers health equity

5 The CDC Foundation is an independent nonprofit created by Congress to marshal philanthropic and private resources for CDC. For more information, see https://www.cdcfoundation.org/.
issues, unintended impacts on minority populations, and alignment with the Health Equity Strategy when reviewing proposals for funding.

Some of these health equity funding opportunities also promote data quality improvement. For example, the March 2021 opportunity includes strategies for jurisdictions to increase and improve data collection and reporting for populations disproportionately impacted by COVID-19. It encourages recipients to improve data collection by race and ethnicity, to educate providers on the importance of racial and ethnic data, and to disseminate health equity data and other materials that are culturally and linguistically tailored, among other activities.57

Some Tribal partners and at least one Territory have experienced challenges addressing COVID-19 disparities due to problems accessing or managing data

Despite their status as public health authorities, the two TECs we interviewed reported difficulty accessing COVID-19 data on AI/AN people from both CDC and States. These challenges, which were also reported in media accounts and testimony from a nonprofit Tribal organization in the summer of 2020, stemmed from staff confusion regarding the public health authority of the TECs.58, 59, 60 Specifically, the partners we interviewed reported that CDC staff and State health departments denied several of their data requests throughout the pandemic, citing privacy concerns and questioning Tribal partners’ authority to access these data. CDC published guidance in 2015 affirming the TECs had access to HHS data under the Indian Health Care Improvement Act,61 and CDC reported that it took steps to provide these partners with access to datasets on HHS Protect and that the issue had been resolved. However, the TECs we interviewed reported that CDC staff and State health departments did not consistently know or apply the 2015 guidance, and that they continued to face challenges due to poor-quality racial data and logistical hurdles. For example, both TECs reported they had to negotiate across multiple Federal, State, and Tribal entities to access the COVID-19 data they needed.

Additionally, the Territory we interviewed experienced challenges addressing COVID-19 because it lacked the technical infrastructure needed to collect, analyze, and report COVID-19 data. CDC provided this Territory, whose population is mostly people of color, with additional support to limit its COVID-19 burden. Specifically, CDC deployed staff to assist the Territory in quickly establishing IT systems for conducting public health surveillance and managing COVID-19 data.
CONCLUSION AND RECOMMENDATIONS

To combat health disparities most effectively both during the COVID-19 pandemic and more broadly, CDC needs adequate data on health outcomes for diverse populations. The pandemic has highlighted and exacerbated existing racial, ethnic, and socioeconomic health disparities. At the same time, the pandemic has strained the Nation’s public health system and exposed gaps in its infrastructure, including limited demographic information in key public health data. While CDC has made efforts to improve its racial and ethnic data for COVID-19 testing, cases, hospitalizations, and deaths, some data remain incomplete, inaccurate, or inconsistent.

CDC has opportunities to improve the way it works with jurisdictions, labs, and hospitals to collect and use data to address COVID-19 disparities. Despite data limitations, CDC has been able to improve its understanding of COVID-19 disparities and take steps to address them. CDC faces choices about the best balance of approaches for ensuring adequate resources to support an equitable response. Furthermore, TECs and certain jurisdictions, such as Territories, may need more support from CDC to address COVID-19 due to problems accessing data or lack of resources.

We recommend that CDC:

Expand its efforts both to improve racial and ethnic data associated with COVID-19 and to supplement them with additional data sources

CDC should expand its two-pronged approach of: (1) improving racial and ethnic data for COVID-19 testing, cases, hospitalizations, and deaths; and (2) using additional data sources to enhance its insights into racial, ethnic, and socioeconomic disparities. Both efforts are valuable, and CDC should determine the appropriate balance of activities and resource investments to support each.

As part of its efforts to improve the quality of racial and ethnic data associated with COVID-19, CDC should pursue several additional steps. CDC should assess the breadth of jurisdictions’ need for support in developing adequate data collection and management systems and use its findings to inform the type and level of support it provides (e.g., technical assistance, staff deployments, funding). CDC could also use this information to inform its ongoing data modernization initiative. Additionally, CDC could work with jurisdictions and labs to encourage or incentivize the adoption of the minimum Federal standards for racial and ethnic categories. Finally, CDC could
consider taking actions to encourage the collection of Tribal affiliation, such as
developing guidance on how it could be standardized and used, and best practices
from jurisdictions already collecting it. CDC could also consider whether and how
those data could be submitted to CDC.

Ensure that Tribal Epidemiology Centers have timely access to all public health data to which they are entitled

CDC should work with TECs to understand which HHS data or systems they lack access to and what barriers they have faced, and work to resolve them. CDC should also ensure that all CDC staff working with data and with Tribal partners are aware of TECs’ status as public health authorities and CDC’s legal obligations toward them to provide access to HHS data. GAO has made similar recommendations for HHS and CDC to take actions to improve TECs’ access to HHS data. CDC could also clarify TECs’ authority for States, if needed.
CDC concurred with our first recommendation to expand its efforts to improve and supplement racial and ethnic data associated with COVID-19. In its response, CDC described several actions it is taking to improve and supplement these data, including planning a $3 billion funding opportunity for public health agencies to address infrastructure and workforce needs and collaborating with the Council of State and Territorial Epidemiologists to survey jurisdictions about their challenges with collecting and sharing racial and ethnic data. We hope to see CDC continue its efforts to improve racial and ethnic data going forward, beyond the expiration of the COVID-19 public health emergency.

CDC concurred with our second recommendation to ensure that TECs have timely access to all public health data to which they are entitled. In its response, CDC stated that it planned to develop written guidance for TECs on requesting public health data.

In our draft report, we also recommended that CDC assess the costs and benefits of standardizing and collecting socioeconomic data at the individual level, compared to the costs and benefits of using socioeconomic data at the community level. CDC disagreed with our recommendation but affirmed the value of socioeconomic data for understanding and addressing COVID-19 disparities. In its response, CDC provided support for its conclusion that collecting community-level socioeconomic data through surveys was more economically efficient and feasible than attempting to collect individual-level socioeconomic data and linking it to case surveillance data. CDC also described how its Data Modernization Initiative would help improve community-level socioeconomic data. The considerations and assessments that CDC included in its response met the intent of our recommendation and, as such, we did not include it as a recommendation in our final report. We appreciate CDC’s commitment to improving socioeconomic data at the community level and to using these data to advance health equity in its COVID-19 response.

For the full text of CDC’s comments, see the Appendix.
TO: Gregory E. Demske  
Acting Principal Deputy Inspector General

FROM: Centers for Disease Control and Prevention (CDC)

DATE: May 6, 2022


Attached is the Centers for Disease Control and Prevention (CDC) response to the Office of Inspector General (OIG). This response details CDC’s planned actions regarding recommendations contained in the OIG report, “CDC Found Ways To Use Data To Understand and Address COVID-19 Health Disparities, Despite Challenges With Existing Data,” OEI-05-20-00540.

CDC appreciates the opportunity to review and comment on this report prior to its final release.

Sincerely,

[Signature]

Rochelle P. Walensky, MD, MPH  
Director, CDC, and  
Administrator, ATSDR
OIG Recommendation (1)
OIG recommends CDC expand efforts to improve racial and ethnic data associated with COVID-19 and to supplement them with additional data sources.

CDC Response:
CDC concurs with the OIG recommendation. CDC is committed to continuing to expand its efforts to improve racial and ethnic data associated with coronavirus disease 2019 (COVID-19) and supplement them with additional data sources. CDC has several ongoing activities to improve race and ethnicity data, as well as additional data sources. Examples include:

- A currently forecasted funding opportunity through CDC’s Center for Surveillance, Epidemiology, and Laboratory Services, CDC-RFA-OE22-2203, *Strengthening U.S. Public Health Infrastructure, Workforce, and Data Systems*, with an estimated funding amount of $3 billion. This funding is a first of its kind, approached as non-categorical and cross-cutting, and intended to help meet critical infrastructure and workforce needs in the short term; it should also make possible strategic investments that will have lasting effects on public health agencies across the United States. To that end, it will support strategically strengthening public health capacity and systems related to the workforce, foundational capabilities, data modernization, physical infrastructure, and support from national public health partners.

- CDC’s National Center for Health Statistics (NCHS) is committed to improving data on race and ethnicity collected from death certificates. In Fiscal Year 2022, NCHS began targeted outreach and training for funeral directors, who are often responsible for proxy reporting of race and ethnicity of the decedent on the death certificate. This training reinforces existing NCHS guidance for funeral directors to consult with the decedent’s family or close friends when determining race and ethnicity. NCHS has also focused on the misclassification of race and ethnicity for the American Indian and Alaska Native (AI/AN) population. To help correct misclassification of race and ethnicity on U.S. death certificates and to compute more reliable estimates, NCHS uses linked death certificate and self-reported census data. NCHS issued a recent report that provided an adjusted mortality profile for the AI/AN population for 2019. This report used linked census and National Vital Statistics System (NVSS) mortality data to assess the level of misclassification by select characteristics, generate correction factors, and apply these corrections to observed NVSS mortality data. This work produced more reliable mortality statistics for the AI/AN population. NCHS will continue to expand efforts to correct misclassification on death certificates.

- CDC continues efforts to promote the adoption of Electronic Case Reporting (eCR). For example, as of January 1, 2022, eCR is required by the Centers for Medicare & Medicaid Services’ Promoting Interoperability Program for eligible hospitals and critical access hospitals and the Merit-Based Incentive Payment System eligible clinicians.

- CDC collaborated with the Council for State and Territorial Epidemiologists (CSTE) on a survey of jurisdictions to learn their experiences with barriers to collecting and sharing race and ethnicity data. Released in April 2022, *Addressing Gaps in Public Health Reporting of Race and Ethnicity Data for COVID-19 – Findings & Recommendations Among 45 State & Local Health Departments* findings suggest that multiple factors, including incomplete healthcare provider and laboratory reporting of race and ethnicity data to public health; the lack of a modern, interoperable public health data infrastructure; and the need for additional staff at public health agencies hampered the ability to identify and respond to the enormous disparities in the impact of COVID-19 on racial and ethnic minority groups. CDC will work with CSTE to assess next steps based on the recommendations from the Assessment.
**OIG Recommendation (2)**

OIG recommends CDC assess the costs and benefits of standardizing and collecting socioeconomic data at the individual level.

**CDC Response:**

CDC non-concurs with the OIG recommendation to assess the costs and benefits of standardizing and collecting socioeconomic data at the individual level for COVID-19 testing, cases, hospitalizations, and deaths. Throughout the COVID-19 response, CDC has been considering ways to use data to understand the impact of socioeconomic status and other important factors on COVID-19 outcomes. At this time, working to improve completeness of existing data fields, using existing data sources for analyses, and leveraging the data modernization initiative will lead to greater improvements for use of socioeconomic and related data than conducting the described assessment.

At the same time, CDC strongly agrees that there is value in using socioeconomic data to identify, understand, and address health disparities. CDC understands the costs and benefits of using these data, and the agency’s focus is on assessing how to effectively collect and use these data as well as leveraging existing data, such as the American Community Survey and other population-level data sources.

It is far more economically efficient and feasible to collect community-level socioeconomic data through surveys than to collect individual-level data and link it to case surveillance data. The latter would require state and local public health staff to contact individuals to ask questions about their socioeconomic status and access non-public health databases that they likely do not have access to in order to link socioeconomic data to case reports. These efforts would require coordination with several other agencies; CDC could not collect these data independently.

Examples of efforts to use socioeconomic data include a wide range of work by NCHS. NCHS has a robust data linkage program that links their own population and provider surveys, at the individual person level, to other sources of health-related data including Medicare and Medicaid health insurance and federal housing assistance programs to better understand the role of socioeconomic status and social support programs in individual health outcomes. These linkages create new sources of data that combine survey reported information on personal and socioeconomic characteristics, health conditions and health care access and utilization with more detailed information on program participation found in administrative data files. These linked data resources are used by researchers and other analysts to inform evidence-based policymaking.

NCHS’s National Health Interview Survey (NHIS) collects information on a variety of income sources as well as estimates of total combined family income. The NHIS asks adult participants about their work status in the week before the interview, main reason for not working, hours worked, work benefits for those working, and days of sick leave taken. The NHIS also asks about educational attainment, measured as the highest level of school or highest degree completed. On a rotating basis, the NHIS asks about the sample adult’s occupation, industry, and work activities through a series of questions allowing for verbatim text responses. Those responses are then coded by statistical clerks at the U.S. Census Bureau into standardized industry and occupation codes.

NCHS’s National Health and Nutrition Examination Survey (NHANES) collects the following information on socioeconomic data from survey participants

- Income: annual household income, annual family income; ratio of family income to poverty
- Education: highest grade or level of school completed, or highest degree received
- Employment status.
NCHS population survey data from NHIS and NHANES can be linked at the person level to death certificates, administrative data, and other statistical data to understand the health of the American population.

CDC agrees that socioeconomic data are important in health equity analyses; however, CDC does not agree with the specific recommendations provided. CDC does not plan to focus on the costs and the authority to collect standardized socioeconomic data at the individual level for COVID-19 testing, cases, and hospitalizations. Health equity remains central to CDC’s COVID-19 response and we will continue to identify existing data sources for community-level analyses as well as linkages, such as those described above.

Too often, the data needed to support public health actions are neither structured nor standardized and are not collected in a consistent manner. As part of CDC’s Data Modernization Initiative, CDC is working with the Office of the National Coordinator for Health Information Technology and state, tribal, local, and territorial partners via U.S. Core Data for Interoperability (USCDI) and USCDI+ to address these complex and critical issues. This work will benefit socioeconomic status data in addition to demographic and other important data.

- Adopting USCDI will help make public health data more compatible with healthcare providers, payers, and other partners.
- Development of a nationwide forum will enable practitioners to identify mission-critical data elements and send a consistent signal about what matters and what is most important to public health. The data identified through USCDI+ will help provide a common baseline to support public health data sharing more broadly.
- Strengthening existing standards management and harmonization activities across public health will aid in identifying and overcoming barriers to efficient data sharing.
- Technical support will ensure USCDI and USCDI+ are adopted in ways that help improve the timeliness, relevance, and reliability of data shared to support public health purposes.

**OIG Recommendation (3)**

OIG recommends CDC ensures that tribal partners have timely access to all public health data to which they are entitled.

**CDC Response:**

CDC concurs with OIG’s recommendation. However, CDC requests OIG revise the language used in this recommendation to focus specifically on Tribal Epidemiology Centers (TECs) and not on all tribal partners. As described in Objective 3, OIG focused on TECs, not all tribal partners. CDC requests this recommendation coincide with the OIG objective and is revised to read: “CDC ensures that Tribal Epidemiology Centers (TECs) have timely access to all public health data enabling them to fulfill their public health mission.”

This OIG recommendation aligns with the recommendations that the Government Accountability Office (GAO) issued to CDC, Indian Health Service, and Health and Human Services (HHS) in its March 2022 report, “Tribal Epidemiology Centers: HHS Actions Needed to Enhance Data Access, GAO-22-104698.” CDC is committed to ensuring that TECs have access to the epidemiological data they need to ensure they can fulfill their public health missions.

During the COVID-19 pandemic, CDC has worked with TECs to ensure that they have access to timely COVID-19 case surveillance and vaccine administration data. Looking beyond the COVID-19 pandemic, CDC understands that there is a need to ensure a systematic process to share public health data with TECs. CDC agrees that there is a need to make data sharing easier, less burdensome, and clearer for
TECs. To implement this recommendation, CDC will utilize the approach for the GAO report and develop written guidance for TECs on how to request data.
Acknowledgments

Nicole Hrycyk served as the team leader for this study, and Abigail Wydra served as the lead analyst. Others in the Office of Evaluation and Inspections who conducted the study include Camille Harper, Jonathan Jones, and Rebekah Schwartz. Office of Evaluation and Inspections staff who provided support include Rob Gibbons and Mike Novello.

This report was prepared under the direction of Laura Kordish, Regional Inspector General for Evaluation and Inspections in the Chicago regional office; Adam Freeman, Deputy Regional Inspector General; and Hilary Slover, Assistant Regional Inspector General.

Contact

To obtain additional information concerning this report, contact the Office of Public Affairs at Public.Affairs@oig.hhs.gov. OIG reports and other information can be found on the OIG website at oig.hhs.gov.

Office of Inspector General
U.S. Department of Health and Human Services
330 Independence Avenue, SW
Washington, DC 20201
The mission of the Office of Inspector General (OIG), as mandated by Public Law 95-452, as amended, is to protect the integrity of the Department of Health and Human Services (HHS) programs, as well as the health and welfare of beneficiaries served by those programs. This statutory mission is carried out through a nationwide network of audits, investigations, and inspections conducted by the following operating components:

**The Office of Audit Services (OAS)** provides auditing services for HHS, either by conducting audits with its own audit resources or by overseeing audit work done by others. Audits examine the performance of HHS programs and/or its grantees and contractors in carrying out their respective responsibilities and are intended to provide independent assessments of HHS programs and operations. These audits help reduce waste, abuse, and mismanagement and promote economy and efficiency throughout HHS.

**The Office of Evaluation and Inspections (OEI)** conducts national evaluations to provide HHS, Congress, and the public with timely, useful, and reliable information on significant issues. These evaluations focus on preventing fraud, waste, or abuse and promoting economy, efficiency, and effectiveness of departmental programs. To promote impact, OEI reports also present practical recommendations for improving program operations.

**The Office of Investigations (OI)** conducts criminal, civil, and administrative investigations of fraud and misconduct related to HHS programs, operations, and beneficiaries. With investigators working in all 50 States and the District of Columbia, OI utilizes its resources by actively coordinating with the Department of Justice and other Federal, State, and local law enforcement authorities. The investigative efforts of OI often lead to criminal convictions, administrative sanctions, and/or civil monetary penalties.

**The Office of Counsel to the Inspector General (OCIG)** provides general legal services to OIG, rendering advice and opinions on HHS programs and operations and providing all legal support for OIG’s internal operations. OCIG represents OIG in all civil and administrative fraud and abuse cases involving HHS programs, including False Claims Act, program exclusion, and civil monetary penalty cases. In connection with these cases, OCIG also negotiates and monitors corporate integrity agreements. OCIG renders advisory opinions, issues compliance program guidance, publishes fraud alerts, and provides other guidance to the health care industry concerning the anti-kickback statute and other OIG enforcement authorities.


22 HHS, *COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115*, March 8, 2022. This guidance was originally issued on June 4, 2020, and updated on March 8, 2022. (See Endnote 11 for URL information.)


28 CARES Act § 18115(a); HHS, *COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115*, March 8, 2022. This guidance was originally issued on June 4, 2020, and updated on March 8, 2022. (See Endnote 11 for URL information.)


52 HHS, COVID-19 Pandemic Response, Laboratory Data Reporting: CARES Act Section 18115, March 8, 2022. This guidance was originally issued on June 4, 2020, and updated on March 8, 2022. (See Endnote 11 for URL information.)

53 CDC, How to Report COVID-19 Laboratory Data, updated on January 26, 2021. (See Endnote 50 for URL information.)

54 Ibid.


