



Assessment of the CDC's New COVID-19 Data Reporting

Version 1.0, May 18, 2020



Table of Contents

Executive summary	5
About The COVID Tracking Project	7
Report contributors	7
The two datasets: state and federal	8
Data comparison	10
Methodology notes	11
Case counts and deaths	13
Testing data	13
Known systemic discrepancies between the datasets	19
Lack of historical data	19
Specimens vs. people	19
Inconsistent reporting by commercial laboratories	19
Positive rate ranges vs. absolute numbers of positive tests	20
Known sources of the federal government's COVID-19 data	20
Closing the gaps	24
We need the CDC's public health data leadership	25
Conclusion	26
Appendix A: State analysis	27
Appendix B: State orders on reporting COVID-19 data	48
Appendix C: Disease surveillance and reporting protocol	52
Appendix D: Web accessibility audit	54



Executive summary

During the week of May 9, 2020, the Centers for Disease Control and Prevention launched a [new COVID-19 data dashboard](#), including new national and state-level case counts, death counts, and testing data. The COVID Tracking Project at *The Atlantic*, which has been compiling and publishing COVID-19 data from state public health authorities since March 7, 2020, has completed an initial analysis of this data.

Highlights from the analysis:

- The new **case and death counts** from the CDC show a high degree of concordance with official state-reported data. If these numbers continue to be regularly reported and aligned, The COVID Tracking Project will begin using the CDC's case and death counts in our public reporting and API.
- The new **testing data** from the CDC, however, differs from official testing data reported by state health departments. **In 29 states, the raw numbers fall within 10% of each other, while in 13 states, the data diverges by 25% or more.** Adjusting for different reporting methodologies does not fully explain these differences.
- Small variations in these datasets are to be expected, but large gaps are cause for concern. For many states, the CDC publishes **higher testing numbers than the states themselves report**, which raises questions about the structure and integrity of both state and federal data reporting.
- Another point of contrast between the CDC's new reporting and the official state data compiled by The COVID Tracking Project is that **the CDC has not released historical, state-level testing data for the first three months of the outbreak.** Until we can reconcile the CDC's new data with the state-reported data that makes up our historical dataset, the new data is of limited use to disease modelers and other COVID-19 data users.



As part of our accountability mission, The COVID Tracking Project team will do everything we can to **understand and close the gaps between the state and federal data**. That work begins with this assessment, and will continue as we integrate the CDC data into our dataset and data API.

This report is licensed [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/). Please attribute it to “The COVID Tracking Project at *The Atlantic*.” You can contact us anytime at <https://covidtracking.com/contact>.



About The COVID Tracking Project

The COVID Tracking Project at *The Atlantic* is a volunteer-driven organization dedicated to collecting and publishing the data required to understand the COVID-19 outbreak in the United States. Since early March, we have grown from a tiny team with a spreadsheet to a project with hundreds of volunteer data-gatherers, epidemiologists, infectious disease scientists, reporters, data scientists, visualization experts, and other dedicated contributors.

Every day, our team compiles data on COVID-19 testing and patient outcomes from all 50 states, 5 territories, and the District of Columbia. Our dataset is currently used by Johns Hopkins University, multiple disease modeling and public policy research groups, and newsrooms around the world. It has also been cited by the White House. As of mid-May, our data API, which allows sites and apps to import our dataset automatically, receives nearly two million requests per day.

Report contributors

Jesse Anderson
Sonya Bahar
Jeanette Beebe
Maya Emmons-Bell
Hannah Birch
Artis Curiskis
Alice Fairbank Goldfarb
Amanda French
Emily R. F. Gottlieb
Jeremia Kimelman
Erin Kissane
Sam Klein
Julia Kodysh
Olivier Lacan
Betsy Ladyzhets
Zach Lipton
Alexis Madrigal

Michal Mart
Robinson Meyer
Kevin Miller
Quang P. Nguyen
Kara Oehler
Michael A. Parks
Jordan Gass-Pooré
Prajakta Ranade
Jessica Malaty Rivera
Nicole S. Rivera
Kara Schechtman
Ryan Scholl
Isabel Sepúlveda
Heather Succio
Allen Tan
Dylan Thurston



The two datasets: state and federal

At The COVID Tracking Project, we compile official COVID-19 data **from US states and territories** to arrive at national summary figures. This process aligns with [early guidance](#) from the CDC's Director of the National Center for Immunization and Respiratory Diseases in a March 3 briefing:

States are reporting results quickly, and in the event of a discrepancy between CDC and state case counts, the state case counts should always be considered more up to date.

In practical terms, official state and territory data sources have been **the only comprehensive, public sources of data** on case counts, deaths, and testing throughout the first three months of the COVID-19 outbreak in the United States. During this period, more than **10 million tests were performed, more than 1.4 million cases were discovered** in the US, and **more than 80,000 Americans died**.

The COVID Tracking Project was formed on March 7 to compile a daily snapshot of COVID-19 data from all states, until the CDC began publicly posting that data. When the project began, our founding team worked under three linked assumptions:

1. the CDC would soon begin publishing state-by-state and national summary statistics;
2. the CDC dataset would be higher quality and more complete than our volunteer-gathered data; and
3. the CDC's dataset would harmonize with the official reports published by states and territories, because COVID-19 data would flow from labs and hospitals through the states to the CDC.



We wrote in our original project FAQ that once the CDC began providing testing data, we would “keep our data going for a while to make sure the data matched up, and then we’d call it quits.”

Ten weeks after our project was formed, the CDC began publishing national and state-level case counts, death counts, and some testing data. This let us test our assumptions about their dataset. Our initial analysis suggests that the federal and state datasets exhibit substantial discrepancies, raising concerns about the testing data available from both the CDC and the states.

In an ideal world, there would be a single coherent dataset about the COVID-19 outbreak in the United States, coordinated and maintained by the CDC in partnership with state health departments. But currently we have two major datasets—one from the states, one from the CDC—that differ in several important ways. Until the CDC and state-published datasets are reconciled, this will be a source of duplicated effort and uncertainty.

In light of this analysis, The COVID Tracking Project is taking four actions:

1. Publishing an analysis comparing national and state-level data from state public health authorities with the newly available data from the CDC.
2. Documenting our understanding of the structural and methodological reasons for the divergences between the two datasets.
3. Proposing potential next steps for reconciling the two datasets and ensuring full transparency for data sources.
4. Continuing our daily data collection from state and territorial public health authorities until such time as we can responsibly replace our daily compilations with the CDC’s data.
 - a. For national and state-level case counts and death counts, we are preparing to switch to direct ingestion of the CDC’s



data within the next ten days, if the data is frequently updated and aligned with official state counts during that period. The COVID Tracking Project will privately maintain our own count from state public health authorities as a backup data source until we are confident that it will not be needed.

- b. For national and state-level testing data, we have set an internal benchmark for adopting the federal data: when the CDC and state testing datasets can be aligned or corrected to within 10% of each other across all states and territories, we will switch to direct data ingestion of the CDC data for the relevant metrics. As with case and death counts, we will privately maintain our own count from state public health authorities as a backup data source until we are confident that it will not be needed.

Data comparison

The new **national totals** from the CDC for cases, deaths, and test counts are well aligned with the **national totals** that the COVID Tracking Project has compiled from the states and the District of Columbia.¹ The **national test positivity** rate from the CDC is also well aligned with the one produced by averaging positivity rates reported by the states.

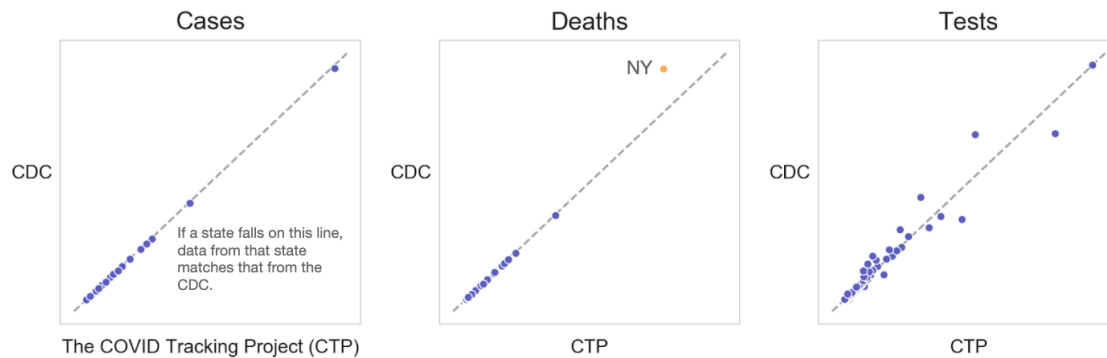
At the state level, **case and death counts** are also well aligned between the two datasets with minor (and well-understood) exceptions, but **testing data** remains more difficult to reconcile, as can be seen in Figure 1.

¹ We have excluded data from the five US territories we track in our state-by-state and national analysis, because the territorial data published by the territories themselves remains too low-quality and incomplete to allow for appropriate comparison.



Comparing State-Reported Cases, Deaths, and Tests with the CDC Data

The values reported by each state's public health agency and the CDC match up well for case counts and the number of deaths, but the number of tests performed is more divergent.



Note: All units are in absolute numbers.

Sources: The COVID Tracking Project and the CDC COVID Data Tracker.



Figure 1

We have provided **detailed state-by-state comparisons of state and CDC data in Appendix A**. In conducting our comparison, we also performed an informal **web accessibility review**, available in **Appendix D**.

This report highlights the most significant disparities we found between the two datasets.

Methodology notes

All data and calculations for this report are posted for public review at <https://github.com/COVID19Tracking/cdc-comparison>.

We standardized our analysis on the CDC data file current at 6pm ET on May 16. That file contains two different timestamps for the three categories of data—case counts, death counts, and test counts—that we assessed:

- "Case and Death data updated as of May 16 2020 5:45PM"
- "Testing data updated as of May 13 2020 12:00AM"



Therefore, to achieve a sound comparison, we had to use different timepoints in The COVID Tracking Project dataset we compared to the CDC data:

- For case and death data, since the CDC timestamp was May 16, we compared it to our latest published data at that point, which was also May 16.
- For testing data, since the CDC timestamp was May 13, we chose to compare it to our published May 14 data, which is closest to reflecting the May 13 reports processed by states. Some state data for this date is an exact match with CDC data timestamped May 13, which suggested to our team that we were comparing the best pair of dates.

Our resulting [primary data file](#)² therefore contains **CDC data from the CDC May 16 file, CTP case/death data from May 16, and CTP testing data from May 14**. The COVID Testing Project testing data from May 14 that we use for this analysis *differs slightly from published CTP May 14 testing data* for two reasons:

- The CDC reports test counts for specimens tested. Therefore, if a state provides **a total count by specimen** and **a total count by people**, we prefer the **count by specimen** for this comparison. [On the public CTP website](#), however, we have displayed the testing count **by people** whenever possible. We have published a [data file](#) to record which states have indicated that they are counting by specimen or by people. Four states provided both counts on May 14. For 17 states, we do not know which count they are using.
- Another reason the totals in the analysis may differ from the totals released on our website is that, on our website, we *include probable cases in our positives number*, but these should be excluded for total testing counts.

² https://github.com/COVID19Tracking/cdc-comparison/blob/master/merged_data_for_analysis.csv



Case counts and deaths

The death and case counts provided by the states and by the CDC **match up well at the national and state level**, as can be seen in the left and middle panels of Figure 1.

The exception is New York State, where the CDC's death count of 27,755 is 5,277 higher than The COVID Tracking Project's count of 22,478. Since April 14, a substantial number of deaths (currently roughly 5,000) that fit the CDC's definition of a probable COVID-19 death were [reported in New York City](#) but *not* publicly reported as such by New York State. The CDC appears to include those deaths in their New York State total, while the COVID Tracking Project currently does not, as we currently only compile official data from state public health authorities.

Testing data

Testing data provided by the states and by the CDC **matches up well at the national level**, but **diverges to a greater degree at the state level**, as can be seen from the third panel in Figure 1.

Through May 16, the states and District of Columbia reported a total of 10.5 million tests, and the CDC reported 10.8 million.³ Both datasets indicate that the United States has substantially ramped up testing capacity since early March. As of this writing, **the CDC's national test positivity rate is 13.2%** and the national test positivity rate we calculate from **state reports** for the equivalent date is **13.9%**.

The similarities at the national level, however, mask substantial discordance at the state and territory level:

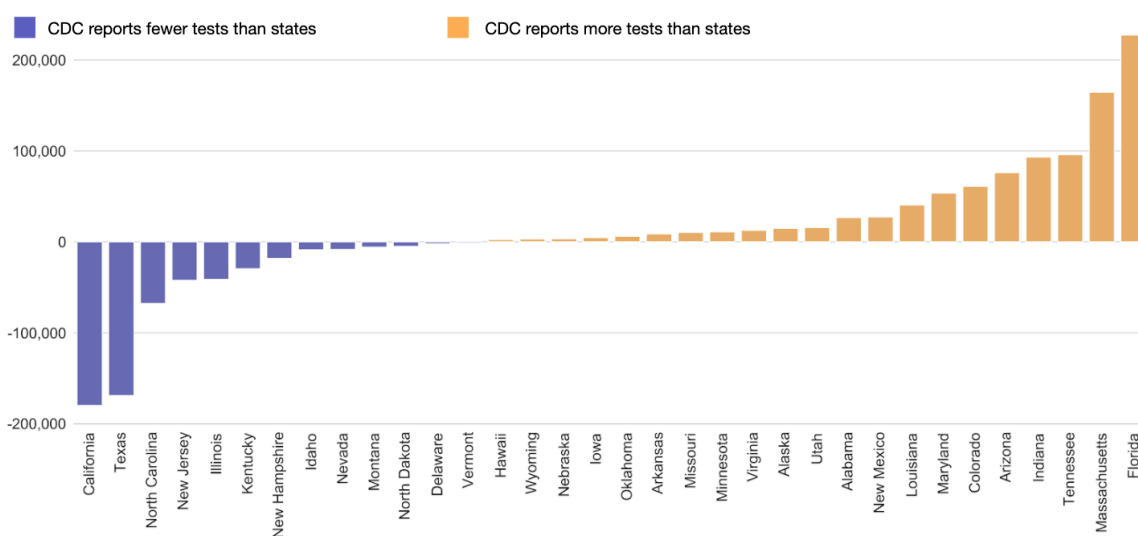
³ We privately capture specimens tested where that figure is available, but do not yet publish those figures. We are using those private numbers to calculate this total figure for specimens tested. State-by-state notes on differences can be found in Appendix B.



- **28 states and the District of Columbia's test numbers fall within 10% of the total test number reported by the state**, and only a few match precisely;
- **22 states fall outside that range**—and some of the discrepancies are very large on a percentage or absolute basis;
- **13 of the total test numbers published by the CDC diverge from state reporting by more than 25%.**
- Reporting from the territories shows **even wider discrepancies.**

Comparing State-Reported Test Count Data with the CDC Data

There are large discrepancies in the test counts reported by the CDC and the data from each state's public health agency.



Notes: All units are in absolute numbers. States are only shown if the testing count differs by 5% or more.

Sources: The COVID Tracking Project and the CDC COVID Data Tracker.



Figure 2

Major testing data discrepancies by absolute numbers

Relative to data reported from the states, in absolute numbers, the ten largest discrepancies in either direction are in Florida, California, Texas, Massachusetts, Tennessee, Indiana, Arizona, North Carolina, Colorado, and Maryland, as can be seen in Figure 2. For example, the CDC has reported 179,955 fewer tests than what the state of California is reporting,



while the CDC has reported 227,456 more tests than what the state of Florida is reporting. The precise numerical values of these discrepancies are given in Table 1.

Differences in Test Counts From States and the CDC			
State	State Reports	CDC Reports	Difference
Florida	691653	919109	227456
California	1104651	924696	-179955
Texas	623284	454133	-169151
Massachusetts	410032	574645	164613
Tennessee	302317	398173	95856
Indiana	160239	253619	93380
Arizona	134338	210388	76050
North Carolina	219268	151449	-67819
Colorado	112505	173626	61121
Maryland	178454	232086	53632

Table 1

Major testing data discrepancies by percentages

The differences between the two test counts are also notable in percentage terms. Relative to data reported from the states, by percentage difference, the ten largest discrepancies in either direction are in Indiana, Arizona, Colorado, New Hampshire, Alaska, Massachusetts, Florida, Tennessee, North Carolina, and Maryland. These are listed in Table 2.

Differences in Test Counts From States and the CDC by Percentage			
State	State Reports	CDC Reports	Percent Difference
Indiana	160239	253619	58%
Arizona	134338	210388	57%



Colorado	112505	173626	54%
New Hampshire	37739	19450	-48%
Alaska	31762	46589	47%
Massachusetts	410032	574645	40%
Florida	691653	919109	33%
Tennessee	302317	398173	32%
North Carolina	219268	151449	-31%
Maryland	178454	232086	30%

Table 2

These divergences appear in both directions. Sometimes the state test count is higher than the CDC's; other times, the CDC's count is higher. We had expected to find instances in which the CDC had lower numbers than states, as all but two state governments have publicly issued orders or other official guidance requiring that all COVID-19 test results be reported to their state public health departments (please see **Appendix B: State reporting orders**). Participation in the United States Department of Health and Human Services' national data collection systems, on the other hand, appears to be voluntary. We have documented publicly available facts about those systems below.

We had not expected to see the CDC report a substantially *higher* test count than reported by any of the states. Given that reporting to state public health departments is generally mandatory, this is an unsettling discovery.

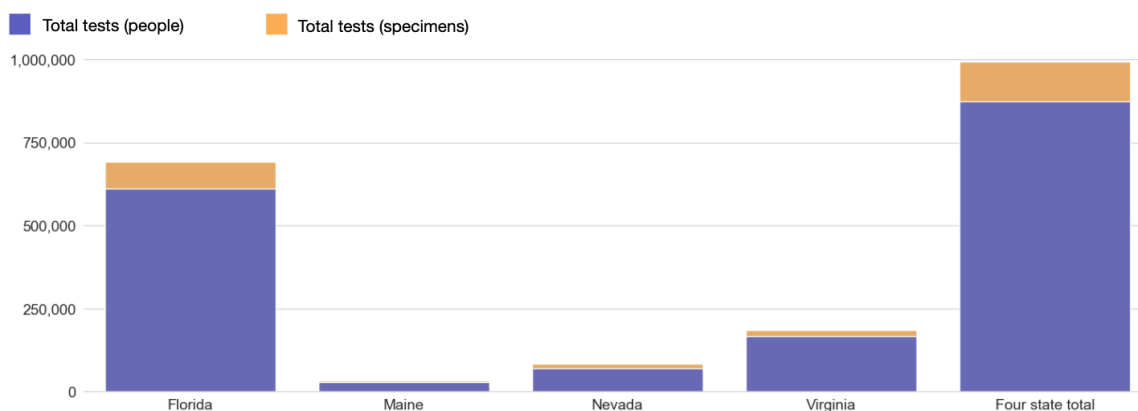
In some cases, this difference could be explained either by data-cleaning work at the federal level that re-sorted tests into different states, or by defensible methodological differences. Most states publicly report “people tested,” rather than “total tests.” Because some people receive multiple tests over time, the number of total tests is likely to be higher than the number of people tested. Four states—Florida, Maine, Nevada,



Virginia—helpfully reported both metrics (Figure 3), and in those states, there are **between 12% and 20% more total tests than people tested**.

State-Reported Testing Counts for People and Specimens

Many states publicly report total "people" tested rather than total number of "specimens" tested. In the four states that reported testing counts for both people and specimens, the two counts are similar.



Note: All units are in absolute numbers.

Source: The COVID Tracking Project.



Figure 3

Using these ratios, it is possible to create a range of possible specimen counts for each state that reports "people tested."⁴ For example, in Indiana, which we believe from direct outreach to be reporting **people tested, rather than specimens**, the CDC reports a higher test number. If we theorize that this is because the state is reporting in people and the CDC is reporting in specimens, we can increase the state's count by 12% to 20%, and this narrows the gap between Indiana reporting and the CDC's. But there are other states, such as North Carolina—which is unclear about whether its test number represents people or specimens—where the state reports *more* tests than the CDC. For these states, doing the same adjustment widens the gap between state and

⁴ To do so, we simply multiplied 1.12*the number of people tested to get a lower specimen count number and 1.20*the number of people tested to get a higher specimen count number.



CDC numbers. As a whole, when we make these approximate specimen count adjustments, the data discrepancies do not suddenly resolve.

The two biggest outlier states in absolute terms—California and Florida—both report specimen numbers, so the problem with these statistics cannot only be attributed to the discrepancy between specimens and people tested.

In the most extreme case, Florida reports about 700,000 total specimens tested, but the CDC reports more than 900,000 tests for the state. Florida has issued clear emergency directives to report all tests to the state⁵. Though we do not have enough evidence to assess what's happening, the federal count suggests three possibilities:

1. Some laboratories are reporting only to the federal government;
2. Some of the federal government's counted tests are duplicates; or
3. Some labs reported negative tests to the wrong state, and the federal agencies sorted out the misclassification later. (This last explanation seems likelier for states with small state-federal discrepancies and less likely for states that report hundreds of thousands more or fewer tests for their state than the CDC does.)

We have also tried applying several simple temporal adjustments to account for misaligned timestamps, but none have made the data match up—which stands to reason, because the divergences run in both directions.

It is vital that all parties work to resolve the discrepancies between what states report and what the CDC reports so that the United States can provide a single source of known facts about this outbreak.

⁵ Emergency order at http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/_documents/emergency-rule-reporting-results.pdf with technical guidance at http://www.floridahealth.gov/diseases-and-conditions/disease-reporting-and-management/_documents/hospital-reporting-covid-lab-results.pdf



Known systemic discrepancies between the datasets

Lack of historical data

The COVID Tracking Project's archive of state data extends back to early March. The CDC has not yet released any historical numbers, and the newly released dataset isn't sufficiently similar to the one The COVID Tracking Project has collected from states to allow us to connect our historical time series with the official CDC data. This poses problems for people trying to use the data to model the outbreak. It also prevents us from doing deeper analysis on when divergences between federal and state numbers began or narrowed.

Specimens vs. people

Most states report the number of **people** they've tested because people are the unit for many other public health functions. Others report the number of **specimens** that have been tested (also often referred to as **total tests**). The COVID Tracking Project dataset therefore contains mixed units. In all cases, **the CDC indicates that it exclusively reports the number of specimens** tested; the testing data portion of the COVID-19 Data Tracker includes "Specimens Tested and Reported by US Laboratories: Commercial and Reference, Public Health, and Hospital". As noted above, four states helpfully report both metrics, and in those states, there are 1.12-1.20 tests per person.

Inconsistent reporting by commercial laboratories

To our knowledge, all laboratories that process COVID-19 tests in the United States have reported positive test results to both state authorities and the CDC since the beginning of the outbreak. Through time, however, not all commercial laboratories have reported all negative test results. When it became clear that testing capacity was an important number to know, most governors issued executive orders requiring that all test results be reported to those states. It may be that some states are still missing some tests from those commercial laboratories. For example, Indiana's "total tests" statistic comes with a health department warning:



“Number of tests is provisional and reflects only those reported to ISDH. Numbers should not be characterized as a comprehensive total.”

Positive rate ranges vs. absolute numbers of positive tests

States have consistently reported an **absolute number of positive tests and/or cases**. These numbers have been recorded by our data entry team. The CDC now provides an **absolute case count**, which appears to match state data well. But, within their testing data, the CDC data groups states into broad, cumulative positive-rate ranges: 0-5%, 6-11%, 11-20%, and 21-30%. It is unclear how to translate these numbers into daily positive rates for further analysis by the Johns Hopkins tracker and the many research projects that rely on our data.

Known sources of the federal government’s COVID-19 data

To understand why we see significant discrepancies between state and CDC test counts, members of The COVID Tracking Project who are on staff at *The Atlantic* requested clarification from an HHS spokesperson. We received a boilerplate description of the CDC tracker, including the statement that “The data presented are aggregate data reported to CDC from state health departments and territorial jurisdictions.”

HHS declined to comment on the record. What follows is, therefore, a list of the milestones in the history of federal collection of COVID-19 data in the United States.

- Throughout February, the CDC [reported the total number of people tested](#) in the United States. **The agency stopped on February 29.**
- **In March, the White House Coronavirus Task Force** established a relationship⁶ with [a consortium of five large reference laboratories](#)—LabCorp, BioReference Laboratories, Quest Diagnostics, Mayo Clinic Laboratories, and ARUP Laboratories—to report their numbers directly to the federal government.

⁶ Some details of this relationship can be inferred from the March 29 and April 10 letters.



- **On March 29, Vice President Mike Pence, in his role as head of the Coronavirus Task Force, [sent a letter](#)** to hospital administrators requesting that they submit a spreadsheet of their testing data to an email inbox.
- **Beginning in April 2020, the HHS contracted data-analysis firm Palantir Technologies** to build out a system known as “HHS Protect” to integrate more than 150 data sources. In total, [HHS has signed three contracts with Palantir with a combined value of \\$26.9 million](#).
- **On April 10, Alex Azar, Secretary of the HHS, [sent another letter](#)** to hospital administrators with a key provision about testing data. HHS Protect had developed sufficiently so that hospitals could submit their information in new ways, instead of by emailing a spreadsheet. Hospitals were encouraged to:
 - Upload the data directly to the HHS Protect platform, the data pipeline built by Palantir Technologies;
 - Submit the data to a state public health authority, which would in turn submit it to a FEMA regional administrator⁷; or
 - Submit the data directly from a hospital’s software to HHS/CDC.
- On the [May 11 Clinical Laboratory COVID-19 Response Weekly Call](#) held by the CDC’s Division of Laboratory Systems, Jason Hall, of the CDC Division of Preparedness and Emerging Infections and also serving in the **CDC Emergency Operations Center Data Analytics Task Force**, described a change in the way laboratories should report data to federal authorities. Hall referred to the two letters listed above and stated that testing data from US hospital laboratories was (at the time of the call) meant to be reported directly to HHS via the HHS Protect System.

Hall then announced that “in an effort to ensure that state and local health departments have the data they need for local decision-making and the streamlined reporting requirements on all

⁷ This option requires that the state receive and share authorization from FEMA to allow for state reporting to meet federal reporting directives. The HHS letter explains the option as follows: “Provide directly to their State if the state is reporting complete information daily to the FEMA Regional Administrator and their state has shared a written notification from FEMA confirming the reporting requirements are being met.”



the US hospital laboratories, CDC is going to begin handling reporting into the HHS Protect system.” Hospital laboratories, he said, should submit their test reports to “state and large local health departments, which will, in turn, send de-identified reports to CDC on your behalf, and we’ll be able to report those to the Department of Health and Human Services, HHS Protect system. So this will obviate the need for US hospital laboratories to report directly into HHS.”

On the same call, Hall stated that **“as of Friday May 8...CDC began reporting laboratory testing data publicly, based on what the states are sending to us, through our CDC COVID-19 data tracker website.** [...] right now these data are aggregated at the state level. We’re receiving them from states right now, and a few territorial jurisdictions as well. We’re receiving them on the county level. So in future updates, we’re going to be showing county-level maps, and having those made available publicly as well.”

- Also on the **May 11 Clinical Laboratory COVID-19 Response Weekly Call**, Jasmine Chaitram, Associate Director for Laboratory Preparedness in the Division of Laboratory Systems, read a question from a call participant noting that some commercial labs were already reporting directly to the CDC.

The data reporting relationships between state authorities and the CDC, HHS, FEMA, and Palantir are still not entirely clear, nor is the provenance of the data on the CDC COVID Tracker. Based on the letters, background information, and call transcripts described above, we would expect that the CDC’s COVID-19 Data Tracker is being generated exclusively from data reported by state public health departments. If this were the case, however, **we would not expect to see the CDC publish test counts that are more than 200,000 higher in Florida than Florida’s own official count or more than 150,000 lower in California and Texas than those states’ official counts.**

This degree of discordance between state and federal databases suggests substantial differences in reporting and publishing methods. The



COVID Tracking Project does not know where in the reporting chain those differences emerge.



Closing the gaps

We realize that there are organizational limitations and political complexities that remain invisible to us, the private citizens, working outside official channels to collect and publish this data. But **as the sole public source of compiled national and state-data for the first three months of the pandemic**, we feel a deep responsibility to our data users, including government agencies, public health research projects, worldwide media organizations, and the people of the United States of America.

It is with this perspective that we issue this report and urge the CDC and their governmental and private partners to resolve the inconsistencies present in the current release of public data. We know dedicated people at all levels of government are working to improve the data quality within their agencies—and harmonize across different jurisdictions. Towards that end, from the outside, looking across the state and federal data, several next steps seem clear:

- **Investigate and close the gaps in testing data** we have identified in this report, with particular attention to the disparate use of the [CDC/CSTE standardized case definition and criteria](#);
- Offer **transparent, detailed, up-to-date sourcing information** for all COVID-19 data published by the federal government;
- Issue clear guidelines for the **separate reporting of viral and antibody tests** by state public health departments;
- Offer all public COVID-19 data in a **fully accessible online format** according to the provisions of Title III of the Americans with Disabilities Act (see **Appendix D** for our brief accessibility review of the data tracker);
- **Release the additional COVID-19 data**, including hospitalization rates, patient outcomes, and detailed demographic information, that we believe HHS to be collecting;



- **Provide a clear roadmap for next steps in disease surveillance and reporting** with regard to testing, case counts, death counts, and known future COVID-19 data metrics concerning therapeutics and vaccines.

We need the CDC's public health data leadership

Since its formation in 1946, the CDC has been the nation's cornerstone for disease prevention and health promotion and efforts. As a federal agency within the Department of Health and Human Services (HHS), its primary role is to protect the United States from threats that endanger the public health. To accomplish this, the CDC conducts clinical research and provides critical data to policymakers.

US public health professionals look to the CDC for scientific leadership, expertise, and guidance on a macro level. For decades, the CDC has coordinated efforts across states and standardized epidemiological data and methods, giving us a nation-wide snapshot of new diseases as they form. In the case of COVID-19, it took more than 15 weeks from the first reported case in the US for the CDC to release their [COVID-19 Data Tracker](#). In the absence of coordinated protocols at the national level, the decentralized datasets produced by US states and territories are now fraught with discrepancies in how case counts, completed tests, and death tolls are reported.⁸

The launch of the CDC's new COVID Data Tracker is a major step—ideally, disease modelers, researchers, and public health authorities would be working from the same data. The general public, too, should be able to trust that there is *one set of numbers on which they can rely*. No dataset is perfect, but there is value in unified data—and the CDC is

⁸ In addition to defensible methodological inconsistencies between states that make it difficult to compare numbers across states or achieve accurate national totals, more serious problems have emerged. At least one state health authority has publicly [blamed a lack of guidance from the federal government](#) for their choice to mix viral testing and antibody testing counts—a deceptive practice that The COVID Testing Project has independently confirmed is occurring in at least four other states.



uniquely positioned to unify and reconcile the discordant datasets from the states and territories.

Conclusion

The public needs reliable, consistent data about the outbreak—and the best possible compiler and provider of that data is the CDC. **By providing unified case and death counts, national summary test numbers, and the beginnings of state testing data, the CDC has taken a huge positive step in public reporting.**

The current discrepancies between what states report and what the CDC reports—and the lack of historical data in the federal dataset—mean that The COVID Tracking Project cannot yet end our data compilation process and replace it with the CDC's numbers. More importantly, the people of the United States are left with two divergent databases of COVID-19 testing data from official government sources.

We believe it to be of vital importance to address these divergences and restore the partnership of state and federal public health authorities to provide a single, consistent, maximally useful database that every US research team, governmental agency, newsroom, and member of the public can rely on.



Appendix A: State analysis

This appendix provides concise analysis of each state's individual situation. The comments primarily address variance in the number of tests reported by the state and the CDC.⁹ There is one large known difference between the datasets that we address for each state. The CDC claims to report the number of “specimens tested.” Some states also report specimens tested, but others report “people tested” or are unclear about the units of their testing data. These differences generate substantial uncertainty in how to compare the numbers that states report with the new ones from the CDC.

Four states report both people tested and specimens tested, so we used these states to determine the average ratio of specimens tested to people tested.

State	Specimens Tested	People Tested	Ratio of Specimens to People Tested
Florida	691,653	609,574	1.13
Maine	33,035	28,357	1.16
Nevada	82,993	69,484	1.19
Virginia	185,551	165,486	1.12
Total	993,232	872,901	Average: 1.14¹⁰

We then applied that ratio to states that appear to report only people tested in order to predict how many specimens that state would report if it were reporting specimens. In some cases, this calculation narrows the gap, bringing the data within 10% of the CDC number. In others, the calculation widens the gap, pushing what appear to be numbers in alignment outside of the 10% margin.

The CDC labels its data in multiple places as “specimens tested.” However, it is worth calling attention to New York and the District of Columbia, which report identical numbers as the CDC while labeling their data as “people tested.” The numbers match, but the units do not.

⁹ We have not included the territories in this report because their divergences were very large and seemed to stem from other issues than those covered in this report.

¹⁰ This ratio results from summing the specimens and the people tested and dividing by four. Adding the ratios and dividing *those* by four gives a ratio of 1.15.



The basic assessment here is that simply correcting (or trying to correct) for a difference in reporting units does not resolve the discrepancies between the datasets. However, it may be an important factor in explaining the gap for an individual state, as for example Indiana, which shows the widest discrepancy in the dataset.

All data on which these analyses are based is available in our GitHub repository at <https://github.com/COVID19Tracking/cdc-comparison>.

Alabama

Alabama	Cases	Deaths	Tests
State	11,523	485	141,985
CDC	11,642	484	168,687
% Difference	1%	0%	19%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data differs by more than 10% from the CDC number, if we estimate the number of specimens tested, the state’s number would fall within 10% of the CDC number. It is unclear if there is a major discrepancy of statistical significance.

Alaska

Alaska	Cases	Deaths	Tests
State	392	10	31,762
CDC	388	10	46,589
% Difference	1%	0%	47%

The CDC and Alaska both report “specimens tested.” However, after comparing the total testing data reported by the CDC and this state, the differences in the data exceed 25%. It is reasonable to believe there are major discrepancies of statistical significance.



Arizona

Arizona	Cases	Deaths	Tests
State	13,631	679	134,338
CDC	13,169	651	210,388
% Difference	4%	4%	57%

The CDC and Arizona both report “specimens tested.” However, after comparing the total testing data reported by the CDC and this state, the differences in the data exceed 25%. It is reasonable to believe there are major discrepancies of statistical significance.

Arkansas

Arkansas	Cases	Deaths	Tests
State	4,578	98	75,818
CDC	4,463	98	84,496
% Difference	3%	0%	11%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data differs by more than 10% from the CDC number, if we calculate a reasonable number of specimens tested, the state’s number would fall within 10% of the CDC number. It is unclear if there is a major discrepancy of statistical significance.

California

California	Cases	Deaths	Tests
State	76,793	3,204	1,104,651
CDC	74,936	3,108	924,696
% Difference	2%	3%	16%

The CDC and California both report “specimens tested.” However, after comparing the total testing data reported by the CDC and this state, the



differences in the data exceed 10%. It is reasonable to believe there are major discrepancies of statistical significance.

California shows the largest absolute discrepancy between what the state reports and the CDC. It could be that not all testing sites in the state are reporting to the Federal system. “We cannot speculate on the reason for the difference in number of reported tests performed,” a California Department of Health spokesperson told us. “CDPH updates the California testing data daily in our News Releases. You may want to contact the CDC for an answer to your question.”¹¹

Colorado

Colorado	Cases	Deaths	Tests
State	21,232	1,150	112,505
CDC	21,131	1,150	173,626
% Difference	0%	0%	54%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate “specimens tested” figure, neither state number would fall within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.

Connecticut

Connecticut	Cases	Deaths	Tests
State	36,703	3,339	149,562
CDC	36,085	3,285	151,175
% Difference	2%	2%	1%

The CDC reports “specimens tested.” We believe Connecticut is reporting specimens tested, though the state has not confirmed this. If that is the case, the state’s numbers line up precisely with the CDC’s. It is reasonable to believe there are probably no major discrepancies of statistical significance.

¹¹ The CDC has not responded to our inquiries for clarifications about this state or any other.



Delaware

Delaware	Cases	Deaths	Tests
State	7,547	286	36,857
CDC	7,373	271	34,793
% Difference	2%	6%	6%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data falls within 10% of the CDC’s number, if we calculate an approximate “specimens tested” number, the state’s total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

District of Columbia

District of Columbia	Cases	Deaths	Tests
State	7,042	375	32,999
CDC	6,871	368	32,999
% Difference	2%	2%	0%

The CDC and District of Columbia testing numbers match precisely, but the CDC says they report specimens tested and the District of Columbia says it reports “people tested overall.” Despite the units mismatch, it seems unlikely there is testing *data* discrepancy.

Florida

Florida	Cases	Deaths	Tests
State	44,811	2,040	691,653
CDC	42,940	1,917	919,109
% Difference	4%	6%	33%

The CDC reports “specimens tested.” This state reports both people and specimens tested. After comparing the total testing data reported by the CDC



and this state, the differences in the data exceed 25%. It is reasonable to believe there are major discrepancies of statistical significance.

Importantly, Florida requires that all tests be reported to the state, so it is difficult to explain the CDC reporting *more* tests than the state itself.

Georgia

Georgia	Cases	Deaths	Tests
State	37,147	1,592	285,881
CDC	36,680	1,557	282,988
% Difference	1%	2%	1%

The CDC and Georgia both report “specimens tested.” After comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Hawaii

Hawaii	Cases	Deaths	Tests
State	638	17	38,881
CDC	587	17	41,561
% Difference	9%	0%	7%

The CDC and Hawaii both report “specimens tested.” After comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Idaho

Idaho	Cases	Deaths	Tests
State	2,389	73	33,556
CDC	2,389	73	24,627
% Difference	0%	0%	27%



The CDC and Idaho both report “specimens tested.” However, after comparing the total testing data reported by the CDC and this state, the differences in the data exceed 25%. It is reasonable to believe there are major discrepancies of statistical significance.

Illinois

Illinois	Cases	Deaths	Tests
State	92,457	4,129	512,037
CDC	90,369	4,058	470,698
% Difference	2%	2%	8%

The CDC reports “specimens tested.” We believe Connecticut is reporting specimens tested, too, though the state has not confirmed this. If that is the case, the state’s numbers line fall within 10% of the CDC’s. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Indiana

Indiana	Cases	Deaths	Tests
State	27,280	1,741	160,239
CDC	26,655	1,691	253,619
% Difference	2%	3%	58%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate “specimens tested” figure, neither state number would fall within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.

Iowa

Iowa	Cases	Deaths	Tests
State	14,328	346	89,294
CDC	14,049	336	93,959
% Difference	2%	3%	5%



The CDC reports “specimens tested.” This state reports “people tested.” Based on five states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. This is true both evaluating the raw (“people tested”) numbers, and after calculating a rough “specimens tested” adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Kansas

Kansas	Cases	Deaths	Tests
State	7,886	172	57,544
CDC	7,886	172	60,337
% Difference	0%	0%	5%

The CDC reports “specimens tested.” This state reports “people tested.” Based on five states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing the testing data reported by the CDC and this state, the difference falls within 10%. This is true both evaluating the raw (“people tested”) numbers, and after calculating a rough “specimens tested” adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Kentucky

Kentucky	Cases	Deaths	Tests
State	7,444	332	117,395
CDC	7,444	332	87,753
% Difference	0%	0%	25%

The CDC reports “specimens tested.” Kentucky reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate “specimens tested” figure, neither state number is within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.



Louisiana

Louisiana	Cases	Deaths	Tests
State	34,117	2,479	247,588
CDC	33,903	2,448	288,133
% Difference	1%	1%	16%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data differs by more than 10% from the CDC number, if we calculate a reasonable number of specimens tested, the state’s number would fall within 10% of the CDC number. It is unclear if there is a major discrepancy of statistical significance.

Maine

Maine	Cases	Deaths	Tests
State	1,648	70	33,035
CDC	1,648	70	33,038
% Difference	0%	0%	0%

The CDC and Maine both report “specimens tested.” After comparing the total testing data reported by the CDC and this state, the numbers are a nearly perfect match. It is reasonable to believe there are no major discrepancies of statistical significance, though this state’s reporting contains considerable complexities beyond the scope of this analysis.

Maryland

Maryland	Cases	Deaths	Tests
State	37,968	1,957	178,454
CDC	37,968	1,957	232,086
% Difference	0%	0%	30%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for



each person tested, 1.12 to 1.19 tests are completed. While the adjusted approximate "specimens tested" figure brings the state and CDC numbers closer together, they still differ by more than 10%. It is reasonable to believe there are major discrepancies of statistical significance.

Massachusetts

Massachusetts	Cases	Deaths	Tests
State	84,933	5,705	410,032
CDC	83,421	5,592	574,645
% Difference	2%	2%	40%

The CDC reports "specimens tested." This state reports "people tested." Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While the adjusted approximate "specimens tested" figure brings the state and CDC numbers closer together, they still differ by more than 10%. It is reasonable to believe there are major discrepancies of statistical significance.

Michigan

Michigan	Cases	Deaths	Tests
State	50,504	4,880	345,403
CDC	50,079	4,825	361,485
% Difference	1%	1%	5%

The CDC and Michigan both report "specimens tested." After comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Minnesota

Minnesota	Cases	Deaths	Tests
State	14,969	709	128,752
CDC	14,240	692	139,893
% Difference	5%	2%	9%



The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. This is true both evaluating the raw (“people tested”) numbers, and after calculating a rough “specimens tested” adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Mississippi

Mississippi	Cases	Deaths	Tests
State	11,123	510	105,326
CDC	11,123	511	105,326
% Difference	0%	0%	0%

The CDC and Mississippi both report “specimens tested.” After comparing the total testing data reported by the CDC and this state, the numbers align precisely. It is reasonable to believe there are no major discrepancies of statistical significance.

Missouri

Missouri	Cases	Deaths	Tests
State	10,675	589	126,935
CDC	10,456	576	137,274
% Difference	2%	2%	8%

The CDC reports “specimens tested.” This state reports “people tested.” Based on five states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. This is true both evaluating the raw (“people tested”) numbers, and after calculating a rough “specimens tested” adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.



Montana

Montana	Cases	Deaths	Tests
State	468	16	24,549
CDC	466	16	18,701
% Difference	0%	0%	24%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate “specimens tested” figure, neither state number would fall within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.

Nebraska

Nebraska	Cases	Deaths	Tests
State	9,772	119	53,427
CDC	9,772	119	56,879
% Difference	0%	0%	6%

The CDC reports “specimens tested.” This state reports “people tested.” Based on five states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. This is true both evaluating the raw (“people tested”) numbers, and after calculating a rough “specimens tested” adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Nevada

Nevada	Cases	Deaths	Tests
State	6,662	345	82,993
CDC	6,629	354	74,579
% Difference	0%	3%	10%



The CDC and Nevada both report "specimens tested." After comparing the total testing data reported by the CDC and this state, the difference in the data falls right at 10%. It is unclear if there are major reporting discrepancies.

New Hampshire

New Hampshire	Cases	Deaths	Tests
State	3,464	159	37,739
CDC	3,464	159	19,450
% Difference	0%	0%	48%

New Hampshire (NH) is reporting the total number of "people" tested as opposed to the number of "specimens" tested as published by the CDC. New Hampshire has also previously been reporting combined PCR and antibody tests in its [testing numbers](#), promising that the state would separate out these numbers as soon as possible. As of 5/15, the total number of individuals with antibodies tested is 3,913. However, numbers reported by CDC (5/14) for NH so far is 18,289 (48%) less than what the state reports. Unfortunately, this discrepancy cannot be explained either by the differences in unit of tests reported (people or specimens) or by the prior inclusion of antibody tests.

New Jersey

New Jersey	Cases	Deaths	Tests
State	145,089	10,249	451,696
CDC	143,905	10,138	409,320
% Difference	1%	1%	9%

The CDC and New Jersey both report "specimens tested." After comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

New Mexico

New Mexico	Cases	Deaths	Tests
State	5,662	253	115,011
CDC	5,662	253	142,431



% Difference	0%	0%	24%
--------------	----	----	-----

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data differs by more than 10% from the CDC number, if we calculate a reasonable number of specimens tested, the state’s number would fall within 10% of the CDC number. It is unclear if there is a major discrepancy of statistical significance.

New York

New York	Cases	Deaths	Tests
State	348,232	22,478	1,298,757
CDC	343,304	27,755	1,298,757
% Difference	1%	19%	0%

The The CDC and New York testing numbers match precisely, but the CDC says they report specimens tested and New York says it reports “total persons tested.” Despite the units mismatch, it seems unlikely there is a testing *data* discrepancy.

The death number discrepancy here is a result of the state of New York not counting approximately 5000 “probable” COVID-19 deaths that [have been reported by the city of New York](#).

North Carolina

North Carolina	Cases	Deaths	Tests
State	17,982	652	219,268
CDC	17,129	641	151,449
% Difference	5%	2%	31%

The CDC reports “specimens tested.” This state reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate “specimens tested” figure, neither state number would fall within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.



North Dakota

North Dakota	Cases	Deaths	Tests
State	1,848	42	50,311
CDC	1,761	42	45,251
% Difference	5%	0%	10%

The CDC reports “specimens tested.” North Dakota reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. If we use this ratio to create an approximate “specimens tested” figure for North Dakota, the gap between the state and CDC numbers grows. It is reasonable to believe there are major discrepancies of statistical significance.

Ohio

Ohio	Cases	Deaths	Tests
State	27,474	1,610	231,795
CDC	26,954	1,581	237,120
% Difference	2%	2%	2%

The CDC reports “specimens tested.” We believe this state’s “total tested” number is a report of “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. If we use this ratio to create an approximate “specimens tested” figure for Ohio, the gap between the state and CDC numbers grows to 10%. It is unclear if there are major discrepancies of statistical significance.

Oklahoma

Oklahoma	Cases	Deaths	Tests
State	5,237	288	112,647
CDC	4,971	288	118,751
% Difference	5%	0%	5%



The CDC and Oklahoma both report "specimens tested." After comparing the total testing data reported by the CDC and Oklahoma, we find that the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Oregon

Oregon	Cases	Deaths	Tests
State	3,612	137	86,679
CDC	3,541	137	84,053
% Difference	2%	0%	3%

The CDC reports "specimens tested." Oregon reports "people tested." Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While the raw numbers from Oregon and the CDC are similar, if we use this ratio to create an approximate "specimens tested" figure, the gap between the numbers grows to more than 10%. It is unclear if there are major discrepancies of statistical significance.

Pennsylvania

Pennsylvania	Cases	Deaths	Tests
State	61,611	4,403	311,195
CDC	60,622	4,342	301,916
% Difference	2%	1%	3%

The CDC reports "specimens tested." Pennsylvania reports "people tested." Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While the raw numbers from Pennsylvania and the CDC are similar, if we use this ratio to create an approximate "specimens tested" figure for Pennsylvania, the gap between the numbers grows to more than 10%. It is unclear if there are major discrepancies of statistical significance.

Rhode Island

Rhode Island	Cases	Deaths	Tests
State	12,434	489	101,601



CDC	12,219	479	98,403
% Difference	2%	2%	3%

The CDC reports “specimens tested.” It’s not clear if Rhode Island is reporting “specimens tested” or “people tested.” If they are reporting “specimens tested,” then the numbers match well. Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While the raw numbers from Rhode Island and the CDC are similar, if we use this ratio to create an approximate “specimens tested” figure, the gap between the numbers grows to more than 10%. It is unclear if there are major discrepancies of statistical significance.

South Carolina

South Carolina	Cases	Deaths	Tests
State	8,407	380	102,535
CDC	8,407	380	98,474
% Difference	0%	0%	4%

The CDC reports “specimens tested.” South Carolina appears to be reporting “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data falls within 10% of the CDC’s number, if we attempted to calculate a possible “specimens tested” number, the state’s total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

South Dakota

South Dakota	Cases	Deaths	Tests
State	3,959	44	26,473
CDC	3,887	44	27,465
% Difference	2%	0%	4%

The CDC reports “specimens tested.” South Dakota reports “people tested.” Based on five states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Still, after comparing



the total testing data reported by the CDC and South Dakota, we find that the difference in the data falls within 10%. This is true both evaluating the raw ("people tested") numbers, and after calculating a rough "specimens tested" adjusted number. It is reasonable to believe there are probably no major discrepancies of statistical significance.

Tennessee

Tennessee	Cases	Deaths	Tests
State	17,288	295	302,317
CDC	17,052	290	398,173
% Difference	1%	2%	32%

The CDC reports "specimens tested." This state reports "people tested." Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. Either in the raw form or adjusted to create an approximate "specimens tested" figure, neither state number would fall within 10% of what the CDC is reporting. It is reasonable to believe there are major discrepancies of statistical significance.

Texas

Texas	Cases	Deaths	Tests
State	46,999	1,305	623,284
CDC	45,198	1,272	454,133
% Difference	4%	3%	27%

The CDC reports "Specimens tested." Texas is reporting a mixed unit testing number, it seems, mostly composed of "specimens." Texas has also previously been reporting combined PCR and antibody tests together, which could inflate its test total, though it is not known by how many tests. Unfortunately, the discrepancy cannot be explained either by the differences in unit of tests reported (people or specimens) or by the prior inclusion of antibody tests.

Utah

Utah	Cases	Deaths	Tests
State	7,068	78	160,119



CDC	7,012	78	175,808
% Difference	1%	0%	10%

The CDC and Utah both report "specimens tested." After comparing the total testing data reported by the CDC and this state, we find that the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Vermont

Vermont	Cases	Deaths	Tests
State	934	53	22,505
CDC	933	53	21,018
% Difference	0%	0%	7%

The CDC reports "specimens tested." Vermont reports "people tested." Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state's raw data falls within 10% of the CDC's number, if we attempted to calculate an approximate "specimens tested" number, the state's total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

Virginia

Virginia	Cases	Deaths	Tests
State	29,683	1,002	185,551
CDC	29,683	1,002	198,217
% Difference	0%	0%	7%

The CDC and Virginia both report "specimens tested." After comparing the total testing data reported by the CDC and this state, the difference in the data falls within 10%. It is reasonable to believe there are no major discrepancies of statistical significance.

Washington

Washington	Cases	Deaths	Tests
------------	-------	--------	-------



State	17,951	992	261,080
CDC	17,951	992	255,104
% Difference	0%	0%	2%

The CDC reports “specimens tested.” Washington appears to be reporting “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data falls within 10% of the CDC’s number, if we attempted to calculate a possible “specimens tested” number, the state’s total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

West Virginia

West Virginia	Cases	Deaths	Tests
State	1,457	64	68,713
CDC	1,447	64	65,283
% Difference	1%	0%	5%

The CDC reports “specimens tested.” West Virginia reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data falls within 10% of the CDC’s number, if we attempted to calculate an approximate “specimens tested” number, the state’s total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

Wisconsin

Wisconsin	Cases	Deaths	Tests
State	12,187	453	133,873
CDC	11,685	445	128,430
% Difference	4%	2%	4%

The CDC reports “specimens tested.” Wisconsin reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data falls within 10% of the CDC’s number, if we attempted to calculate an



approximate “specimens tested” number, the state’s total would vary by more than 10% from the CDC number. It is unclear if there is a major discrepancy of statistical significance.

Wyoming

Wyoming	Cases	Deaths	Tests
State	716	7	15,417
CDC	716	7	18,553
% Difference	0%	0%	20%

The CDC reports “specimens tested.” Wyoming reports “people tested.” Based on four states that report both numbers, our analysis suggests that for each person tested, 1.12 to 1.19 tests are completed. While this state’s raw data differs by more than 10% from the CDC number, if we calculate a hypothetical number of specimens tested, the state’s number would fall within 10% of the CDC number. It is unclear if there is a major discrepancy of statistical significance.



Appendix B: State orders and guidance on reporting COVID-19 data

State	Order from Governor or Health Department	Included in “Reportable Disease” List (if gov./health dept. order not available)
AK	Alaska Section of Epidemiology (SOE) Guidance for Coronavirus Disease 2019 (COVID-19) Testing in Alaska	
AL	Certification of Emergency Rules Filed with Legislative Services Agency	
AR	Standardized Case Definition and Notification for Coronavirus Disease 2019 (COVID-19)	Arkansas Department of Health (ADH) Mandatory Reportable Diseases List and Instructions
AZ	State of Arizona Executive Order 2020-13 Enhanced Surveillance Advisory COVID-19	
CA	Changes Letter for Providers	Title 17, California Code of Regulations (CCR) §2500, §2593, §2641.5- 2643.20, and §2800-2812 Reportable Diseases and Conditions
CO	Public Health Order 20-17 Hospital Data Reporting for COVID-19	
CT	Change to the List of Reportable Laboratory Findings	
DC	Health Notice for District of Columbia Health Care Providers Updated Priorities and Procedures for COVID-19 Testing and Guidelines for Clearance	
DE	Evaluating and Testing Persons for Coronavirus Disease 2019	
FL	Notice of Emergency Rule	
GA	COVID-19: Guidance for Healthcare Professionals	Disease Reporting



HI		Hawaii Laboratory Reporting Requirements
IA	Designation of positive and negative SARS-CoV-2 laboratory results as immediately electronically reportable	
ID	Health Advisory for Idaho Healthcare Providers: Guidance for Novel Coronavirus Disease (COVID-19) Specimen Collection and Testing	
IL	We found no orders or guidance from the state of Illinois.	
IN	State Health Commissioner's Order Requirements for Reporting COVID-19 Information to the State Department of Health	
KS	Guidance for CLIA-approved high and moderate complexity laboratories to begin COVID-19-2 Testing	
KY	Reportable Disease Requirements	
LA	COVID-19 Lab Reporting and COVID-19 Reportable Conditions	
MA	Order of the Commissioner of Public Health	
MD	Amended Directive and Order Regarding Various Healthcare Matters	
ME		Temporary Updates to the Notifiable Diseases and Conditions List - SARS CoV-2 and COVID-19 Deaths
MI	Michigan State and Local Public Health COVID-19 Standard Operating Procedures	
MN		Notification Letter to Minnesota Health Care Providers, Hospitals, and Clinical Laboratories Regarding Reporting of COVID-19/SAR
MO	Update: Reporting COVID-19 Cases	
MS	Mississippi Public Health Laboratory SARS-CoV-2 (virus that causes COVID-19) Test Information (MPHL)	



MT	37.114.204 : REPORTS AND REPORT DEADLINES - Administrative Rules of the State of Montana	
NC	May 15, 2020 (replaces version dated April 20, 2020) To: All North Carolina Clinicians and Laboratories From: Zack Moore, M	
ND		HEALTH ADVISORY
NE	Updated Guidance on Evaluating and Testing Persons for Coronavirus Disease 2019 (COVID-19) NEBRASKA DEPARTMENT OF HEALTH AND HUMAN SERVICES GUIDANCE DOCUMENT	
NH	Coronavirus Disease 2019 (COVID-19) Outbreak, Update # 5	
NJ	EXECUTIVE ORDER NO. 111	
NM		New Mexico Department of Health 2020 Weekly Infectious Disease Report Week 19 May 03, 2020 - May 09, 2020 (included in Notifiable Infections Diseases Weekly Report, line 1 of table)
NV	Updated Testing Criteria for COVID-19 Division of Public and Behavioral Health	
NY	DATE: April 30, 2020 TO: Clinical Laboratories, Limited Service Labs, and Other Entities Supporting SAR- CoV-2 Testing FROM: Ne	
OH	Director's Order to Laboratories in Ohio to Report Results of COVID19 Tests	
OK	Sixth Amended Executive Order 2020-13	March 30, 2020 Dear Laboratory Colleagues, 63 OS § 1-503 requires the Oklahoma State Department of Health ("OSDH") promul
OR	Temporary Administrative Order 333-018-0900	OREGON PUBLIC HEALTH DIVISION REPORTING FOR LABORATORIES
PA	Order of the Secretary of the Pennsylvania Department of Health Requiring Hospitals to	GUIDANCE ON HOSPITALS' RESPONSES TO COVID-19 UPDATED May 12, 2020 The Department of



	Make Daily Reports of Specified Data Regarding Supplies and Equipment	Health (Department) has received questio
RI	Executive Order - FOURTEENTH SUPPLEMENTAL EMERGENCY DECLARATION - TESTING, CRITICAL SUPPLIES AND HOSPITAL CAPACITY REPORTING	List of Reportable Diseases: Department of Health
SC	Updated reporting criteria for 2019 novel coronavirus disease (COVID-19)	
SD		Reportable Diseases - SD Dept. of Health
TN	COVID-19 Case Reporting Guidance	
TX	Governor Abbott Issues Executive Order To Strengthen Reporting Capabilities	Information for Laboratories
UT	Executive Order Declaring a State of Emergency Due to Infectious Disease COVID-19 Novel Coronavirus	
VA		Virginia Reportable Disease List
VT		Infectious Disease Reporting
WA	WASHINGTON STATE REGISTER	List of Notifiable Conditions
WI	New requirements for reporting cases and patients under investigation for COVID-19	Disease Reporting Wisconsin Department of Health Services
WV	Order Regarding Covid19 As A Reportable Disease or Condition	
WY		Wyoming Department of Health Reportable Diseases and Conditions



Appendix C: Disease surveillance and reporting protocol (pre-COVID-19)

The National Notifiable Diseases Surveillance System (NNDSS) helps public health authorities monitor, control, and prevent about 120 diseases [includes infectious diseases, foodborne outbreaks and noninfectious conditions like lead poisoning]. Approximately 3,000 public health departments gather and use data on these diseases to protect their local communities. Through NNDSS, CDC receives and uses these data to “keep people healthy and defend America from health threats.” These data inform the CDC Morbidity and Mortality Weekly Reports (MMWR).

Pre COVID-19, the 2020 NNDSS Event Code List included 145 entities released on [December 10, 2019](#). The Event Code List was updated to 147 entities on [May 4, 2020](#) to include Event Code **11065 for “Coronavirus Disease 2019 (COVID-19)”**. Note that the spreadsheet includes a column that distinguishes which events are considered “nationally notifiable.”

Jurisdictional laws and regulations mandate reporting of cases of specified infectious and noninfectious conditions to health departments. The health departments work with healthcare providers, laboratories, hospitals, and other partners to obtain the information needed to monitor, control, and prevent the occurrence and spread of these health conditions. The CDC Division of Health Informatics and Surveillance (DHIS) supports NNDSS by receiving, securing, processing, and providing nationally notifiable infectious diseases data to disease-specific CDC programs.

Integrated surveillance information systems in public health departments are primary sources of data for NNDSS. These systems are based on the National Electronic Disease Surveillance System (NEDSS) architectural standards. By encouraging the use of and helping to support standards-based public health surveillance systems, NEDSS helps public health agencies accept electronic data exchanges from healthcare systems and enables health departments to create and send standards-based case



notifications to CDC for NNDSS. Currently, jurisdictions can send case notifications by using different standards; NMI is working to provide a single, standardized message format to transmit data to CDC.

NEDSS Base System (NBS), a CDC-developed information system, helps jurisdictions manage reportable disease data and send notifiable diseases data to CDC. To date, 22 health departments (19 states; Washington, DC; Guam; and U.S. Virgin Islands) use NBS to manage public health investigations and transfer general communicable disease surveillance data to CDC.

CDC is currently modernizing the infrastructure supporting NNDSS, referring to it as the NNDSS Modernization Initiative (NMI). It is a multi-year initiative to increase the robustness of the technological infrastructure to make it more user-friendly, standardized, with helpful exchange mechanisms.



Appendix D: Web accessibility audit

This is an accessibility audit of the [CDC COVID Data Tracker](#) (herein referred to as “the website”) as of May 17, 2020. This report follows the guidelines outlined in the US Access Board’s [Section 508 standards](#), specifically:

E205.4 Accessibility Standard. Electronic content shall conform to Level A and Level AA Success Criteria and Conformance Requirements in WCAG 2.0 (incorporated by reference, see 702.10.1).

All requirements outlined in the report are Level A and AA WCAG success criteria. This audit should not be considered an exhaustive measure of the website’s accessibility, but instead a highlight of major accessibility problems that may prevent people with a wide variety of abilities from accessing critical information on the site.

Overview

The website uses a combination of tools to build different map interfaces, which results in inconsistent browsing experiences for people who rely on keyboard navigation, screen readers, and other assistive devices.. When navigating between sections, it is impossible to know that the main portion of the page has changed.

Some of the maps are presented with accessible table versions of data, which is a helpful feature for a wide variety of users. However, the school closures and social impact maps do not. This renders the information on these pages inaccessible to people with screen readers or zoomed-in screens.

The U.S. Cases page uses nonstandard buttons that prevent people from switching between “Total Cases,” “Rates,” etc. Users without a mouse or pointer will not be able to use these buttons.

Results

[Success criterion 1.1.1](#) - Non-text content

Inaccessible data in Social Impact

Quick page selector: #map under “Social impact”



The social impact map provides no accessible version of the table for people with a screen reader or other assistive technologies. A table should be available with the same data.

Inaccessible data in School Closures

Quick page selector: `#shapeMap` under “School closures”

The school closure map provides no accessible version of the table for people with a screen reader or other assistive technologies. A table should be available with the same data.

[Success criterion 1.4.3](#) - Contrast

Low-contrast links

Quick page selector: `.btnMapDataToggle`

The buttons on U.S. cases that read “Total Cases,” “Cases in Last 7 Days,” etc., are light grey (`#bbbbbb`), and are on a white background. The color contrast for these two colors is 1.91, which is much lower than the minimum standard of 4.5 set in the success criterion.

[Success criterion 1.4.4](#) - Resize text

The page sets the viewport to not be scalable for mobile users. This can negatively affect people with low vision on touch devices.

```
<meta name="viewport"
content="initial-scale=1,maximum-scale=1,user-scalable=no">
```

[Success criterion 2.4.3](#) - Focus order

TabIndex used to add label

Quick page selector: `#widget_1`, `#widget_2`, for example

WCAG failure: [F44](#)

The top-line totals like “Total cases in the USA” are wrapped in an element with a title and tabindex. This makes the screen reader read the title of the element, then all the text, then the title again on exit, while the content in the element is already readable.



Success criterion 2.4.6 - Headings and heading order

Heading order in page header

Quick page selector: `.header-text-wrapper`

The page header has an inappropriate header order (h1, followed by an h4). The h4 is merely a descriptive part of the main page title and should be an appropriate element.

Non-header elements used as header

Quick page selector: `.cv-bold`, `#mainContent_Title` for example

The section labels like “Total Cases by Jurisdiction” should be headers, to provide appropriate landmarks for people with screen readers to jump to different sections of the page.

Success criterion 3.2.5 - Change on request

User is notified on a change of context

Quick page selector: buttons under `#navButtons`

WCAG Failure: [F61](#)

The buttons that switch between maps in the main navigation do not alert the user of new context, update the page title, or update the URL of the page. There is no notification to the user that the page content has changed when the button is activated.

Success criterion 4.1.2 - Name, Role, Value

Buttons must have discernible text

Quick page selector: `#btnUStableExport`, `#hamburgerMenu`

The buttons for downloading the CSV data and expanding the map have no discernable text. People with screen readers will not be able to tell what the button is for. Add an `aria-label` attribute or non-visible, readable content within the button.

On mobile devices, the menu button is also missing text and is not readable.



Custom elements used as control

Quick page selector: `butttton`

WCAG Failure: [F15](#)

The buttons labeled “Total Cases”, “Cases in Last 7 Days”, etc. act as buttons, but use the tag name “butttton” (with an extra “t”). This makes them inaccessible to people with keyboards. People with screen readers can access them, but will not know they are buttons.

Map and chart toggles are not buttons

Quick page selector: `#map-toggle-container`, for example

The buttons to switch between the chart and map view are div elements, and are not reachable with a keyboard alone.