

About HHS Protect

At the beginning of the pandemic, it became clear that HHS needed a central way to make data collected by various operating divisions, including CDC, CMS, HRSA, and others, visible to first responders at federal, state, and local levels and we needed to collect this data as fast as possible. To fulfill this need, HHS built HHS Protect, a secure data ecosystem powered by eight commercial technologies for sharing, parsing, housing, and accessing COVID-19 data and driven by four principles: transparency, sharing, privacy, and security.

Read on to learn more.

What is HHS Protect?

The HHS Protect ecosystem is a secure platform for authentication, amalgamation, and sharing of healthcare information, so that the U.S. government can harness the full power of data for the COVID-19 response. U.S. healthcare data has often been fractured and inaccessible. With Protect, more than 200 disparate data sources are brought together into one ecosystem that integrates data across federal, state, and local governments and the healthcare industry. It provides a holistic view of the U.S. healthcare system so decision makers informed by Protect have near-real-time information to guide action and save lives with a data-driven COVID-19 response.

What date did the HHS Protect platform begin operating?

HHS Protect became operational on April 10, 2020.

Is the data in HHS Protect public?

View the public data dashboards here: <https://protect-public.hhs.gov/>

What data does it aggregate?

Through HHS Protect, we have access to hospital-specific data, like inpatient bed utilization, ICU bed utilization, percentage of inpatient beds occupied by COVID-19 patients, and number of COVID-19 cases. We also have insight into the supply chains of large healthcare distributors. By integrating this data together into one system, we can help federal, state and local leaders make strategic decisions and maximize resources.

We have more than 200 datasets integrated in the system, including:

- Multiple COVID-19 case count sources to ensure comprehensive visibility
- Hospital capacity, utilization, inventory and supply from states and territories
- Supply chain data from government and industry
- Diagnostic labs testing data
- Census population and demographic statistics
- Testing locations
- State policy actions

- COVID-19 and influenza-like-illness emergency department data
- Private-sector partner contributions

How is data collected?

We have multiple pathways for collecting data from the many data sources we rely on. These reporting pathways include submitting directly to HHS Protect, to state health departments and state hospital associations, or through TeleTracking.

For hospitals specifically, we have two pathways for reporting data. Hospitals may report directly to HHS through TeleTracking or to state health departments or state hospital associations, which then share data with HHS.

The NHSN will continue to collect information on another critical area for COVID-19—data from nursing homes and long-term care facilities.

Why did the CDC's National Healthcare Safety Network (NHSN) step down from data collection for COVID-19 from hospitals, leaving TeleTracking and State Health Departments/State Hospital Associations as the only collection mechanisms for data?

Success in fighting this pandemic requires flexible and real-time access to data. To accomplish this need, on July 15, HHS changed reporting procedures for hospitals to streamline data into the COVID-19 response. With TeleTracking, HHS is able to create new data fields and collect data from the more than 6,000 hospitals in the country in only 1-3 days. This very same process required several weeks to accomplish using the National Healthcare Safety Network. With a novel virus like this coronavirus, scientists are learning new things and asking new questions daily. It's important the federal response be able to adapt in real time to these changing needs for the safety and health of the country.

The NHSN will continue to collect information on another critical area for COVID-19—data from nursing homes and long-term care facilities. NHSN will also continue its efforts around data collection for the fight against antibiotic resistance. This data collection change has zero effect on CDC's access to data, and HHS continues to rely heavily on the CDC's experts in analyzing the data.

Is the data collected from hospitals different now that it is collected through TeleTracking?

The data has not changed, but the display in the dashboard will look slightly different.

NHSN dashboards previously only included information reported from about half of America's hospitals. Its reports included statistical estimates and extrapolation to account for missing data from hospitals that did not report. The new dashboards will include information from over 80% of America's hospitals.

HHS accepts partial data submissions given the variability and maturity of data reporting from small and large hospitals. Given this, these dashboards include all data reported to HHS

regardless of whether all data fields were completed by the hospital. We believe this complete transparency will create greater understanding of COVID-19 across the country and lead to improved data collection over time.

Will this data be manipulated by political appointees within the government?

No. HHS Protect was built and designed by the Department's career staff with the input of scientists and doctors. This is a system designed to be above politics and manipulation. The public health professionals who work at HHS are solemnly devoted to the health and wellness of all of America and committed to empowering Americans with the same knowledge guiding the decision makers in the COVID-19 response.

How will you ensure the data is never manipulated?

As one of the underlying principles of HHS Protect, HHS is committed to transparency with the American public about what we know, when we know it.

HHS will monitor and correct any attempts at data manipulation. External data submissions are provided under specific data use agreements. Prior to being allowed to submit to HHS Protect, all external data sources are identified and registered. External data are validated daily.

All data received by HHS Protect is immediately captured, recorded and time stamped before it is accessible by HHS Protect or others that may subsequently have access to the data. When users log in to HHS Protect, every data element and every data set has a record of lineage that is built on a hashing technology. This allows HHS to track when the data was curated, when it was parsed, and when it was accessed.

For any data that's shared outside the Protect ecosystem, HHS uses a hashing technology to ensure the integrity of the underlying data set. Meaning once the data is shared outside of Protect, if a person were to try to change the underlying data that was shared, there would be a record of the change that was made.

All of the data is time-stamped.

Does CDC have access to this data?

Yes, hundreds of CDC staff have access to the system. CDC has been delegated the ability to directly authorize CDC personnel for access to HHS Protect. Appropriate CDC personnel have sole authority to grant access to CDC personnel for HHS Protect.

How are you protecting American's health data?

First, we protect access to the HHS Protect system. Access to the HHS Protect system is only granted to authorized government employees and contractors, who are granted access as necessary by mission need. We authenticate and authorize every user to ensure only mission

essential activity is occurring within HHS Protect.

Second, data provided to HHS Protect does not contain direct identifiers, meaning that there is no personally identifiable information (PII) in the system.

HHS has made the security and protection of the data involved a top priority. Least-privilege and National Institute of Standards and Technology (NIST) cybersecurity frameworks have been applied to support confidentiality, integrity, and availability. These are higher standards than typically applied to protecting healthcare data in many other parts of the American healthcare system. Controls and platforms are tested for vulnerabilities, which are mitigated quickly, and mechanisms are in place to prevent exfiltration of data.

Does HHS Protect contain data from electronic health records (EHR) and if so, how do you protect this data?

HHS Protect does have EHR data; it does not have personally identifiable information (PII). EHR data is specific only to the county level. Examples of EHR data include: hospitalization utilization, COVID-19 case counts, ventilator usage, and positive test percentages.

HHS has made the security and protection of the data involved a top priority. Least-privilege and National Institute of Standards and Technology (NIST) cybersecurity frameworks have been applied to support confidentiality, integrity, and availability. These are higher standards than are applied to protecting healthcare data in many other parts of the American healthcare system. Controls and platforms are tested for vulnerabilities, which are mitigated quickly, and mechanisms are in place to prevent exfiltration of data.

Where can I find the multiple contracts that support HHS Protect?

Contracts for HHS Protect are made available on government contracting websites just like any other contracts. Data use agreements are not contracts and are not typically made public. HHS has partnerships with many entities including, federal, state, tribal, local, and territorial governments; colleges and universities; and private sector companies. It is only through the huge number and diversity of our data partners that HHS has been able to accomplish what we have so far and maintain the ability to continue providing vital information products in the future.

Is HHS Protect capturing demographic data?

HHS has directed hospitals, testing companies and others to submit demographic data and HHS Protect contains the data that has been submitted so far. We continue to work with hospitals and testing companies to expand this so we can better understand how COVID-19 affects demographic subgroups and disproportionately impacts minority populations.