



COVID-19 Case Surveillance Public Use Data

Case Surveillance

View Data

Visualize

Export

API



Currently, CDC provides the public with COVID-19 case surveillance data in two ways: an 11 data element public use dataset of the line-listed dataset of all COVID-19 cases shared with CDC and a 31 data element [restricted access dataset](#) of the line-listed dataset of all COVID-19 cases shared with CDC.

The following are true for both the public use dataset and the restricted access dataset:

- Data elements can be found on the COVID-19 case report form located at www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf.
- Data are considered provisional by CDC and are subject to change until the data are reconciled and verified with the state and territorial data providers.
- Data are suppressed to protect individual privacy.
- The datasets include all cases with an initial report date of case to CDC at least 14 days prior to the creation of the previously updated datasets. This 14-day lag will allow case reporting to be stabilized and ensure that time-dependent outcome data are accurately captured.

- Datasets are updated monthly.
- Both datasets are created using CDC's operational Policy on Public Health Research and Nonresearch Data Management and Access and include protections designed to protect individual privacy.
- For more information about data collection and reporting, please see <https://wwwn.cdc.gov/nndss/data-collection.html>
- For more information about the COVID-19 case surveillance data, please see <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/faq-surveillance.html>

Overview

The COVID-19 case surveillance system database includes individual-level data reported to U.S. states and autonomous reporting entities, including New York City and the District of Columbia (D.C.), as well as U.S. territories and states. On April 5, 2020, COVID-19 was added to the Nationally Notifiable Condition List and classified as “immediately notifiable, urgent (within 24 hours)” by a Council of State and Territorial Epidemiologists (CSTE) Interim Position Statement (Interim-20-ID-01). CSTE updated the position statement on August 5, 2020 to clarify the interpretation of antigen detection tests and serologic test results within the case classification. The statement also recommended that all states and territories enact laws to make COVID-19 reportable in their jurisdiction, and that jurisdictions conducting surveillance should submit case notifications to CDC. COVID-19 case surveillance data are collected by jurisdictions and shared voluntarily with CDC.

For more information:

<https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/08/05/>

The deidentified data in the public use dataset include demographic characteristics, exposure history, disease severity indicators and outcomes, clinical data, laboratory diagnostic test results, and comorbidities. All data elements can be found on the COVID-19 case report form located at www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf.

The Case Surveillance Task Force and Surveillance Review and Response Group (SRRG) within CDC's COVID-19 Response provide stewardship for datasets that support the public's access to COVID-19 data while protecting individual privacy.

COVID-19 Case Reports

COVID-19 case reports have been routinely submitted using standardized case reporting forms. On April 5, 2020, CSTE released an Interim Position Statement with national surveillance case definitions for COVID-19 included. Current versions of these case definitions are available here:

<https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/>.

All cases reported on or after were requested to be shared by public health departments to CDC using the standardized case definitions for lab-confirmed or probable cases. On May 5, 2020, the standardized case reporting form was revised. Implementation of case reporting using this new form is ongoing among U.S. states and territories.

Data are Considered Provisional

- The COVID-19 case surveillance data are dynamic; case reports can be modified at any time by the jurisdictions sharing COVID-19 data with CDC.
- CDC may update prior cases shared with CDC based on any updated information from jurisdictions.
- National case surveillance data are constantly changing. For

instance, as new information is gathered about previously reported cases, health departments provide updated data to CDC. As more information and data become available, analyses might find changes in surveillance data and trends during a previously reported time window. Data may also be shared late with CDC due to the volume of COVID-19 cases.

- Annual finalized data: To create the final NNDSS data used in the annual tables, CDC works carefully with the reporting jurisdictions to reconcile the data received during the year until each state or territorial epidemiologist confirms that the data from their area are correct.

Data Limitations

To learn more about the limitations in using case surveillance data, visit [FAQ: COVID-19 Data and Surveillance](#).

Data Quality Assurance Procedures

CDC's Case Surveillance Section routinely performs data quality assurance procedures (i.e., ongoing corrections and logic checks to address data errors). To date, the following data cleaning steps have been implemented:

- Questions that have been left unanswered (blank) on the case report form are re-classified to a Missing value, if applicable to the question. For example, in the question "Was the individual hospitalized?," where the possible answer choices include "Yes," "No," or "Unknown," the missing value is re-coded to Missing if the respondent did not answer the question.

- Logic checks are performed for date data. If an illogical date has been provided, CDC reviews the data with the reporting jurisdiction. For example, if a symptom onset date that is in the future is reported to CDC, this value is set to null until the reporting jurisdiction updates this information appropriately.

- The initial report date of the case to CDC is intended to be completed by the reporting jurisdiction when data are submitted. If blank, this variable is completed using the date the data file was first submitted to CDC.

- Additional data quality processing to recode free text data are ongoing. Data on symptoms, race and ethnicity, and healthcare worker status have been prioritized.

Data Suppression for the 11 data element dataset

To prevent release of data that could be used to identify people, data cells are suppressed for low frequency (<5) records and indirect identifiers (date of first positive specimen). Suppression includes uncommon combinations of demographic characteristics (sex, age group, race/ethnicity). Suppressed values are re-coded to the NA answer option.

Restricted Access Dataset

A restricted access, detailed version of the line-listed dataset of all COVID-19 cases shared with CDC is available for public use. The restricted access dataset includes 31 data elements, including state of residence and county of residence, with data suppression to ensure protection of individuals' privacy. This dataset may be most suitable for researchers and others with specific analytic questions.

For more information on the restricted access dataset including the data elements, registration information, and data use restrictions, please see [LINK TO THE DOCUMENTS](#) For questions, please contact ASK SRRG (eocevent394@cdc.gov).

Additional COVID-19 Data

COVID-19 data are available to the public as summary or aggregate count files, including total counts of cases and deaths by state and by county. These and other data on COVID-

19 are available from multiple public locations:

<https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html>

<https://www.cdc.gov/covid-data-tracker/index.html>

<https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html>

<https://www.cdc.gov/coronavirus/2019-ncov/php/open-america/surveillance-data-analytics.html>

Less

Updated December 4, 2020

Data Provided by CDC Case
Surveillance Task Force

Featured Content Using this Data

Demographic Trends of
COVID-19 cases and death...



External
Content

CDC is working with states to pro

About this Dataset

Mute Dataset

Updated

**December 4,
2020**

Data Last
Updated

December 4,
2020

Metadata Last
Updated

December 4,
2020

Date Created

May 15, 2020

Views

442K

Downloads

23.8K

Data Provided
by

CDC Case
Surveillance
Task Force

Dataset
Owner

Data.CDC.gov
Administrator

Contact Dataset Owner

Common Core

Publisher	CDC
Contact Name	Surveillance Review and Response Group
Contact Email	eocevent394@cdc.gov
Suggested Citation	Centers for Disease Control and Prevention, COVID-19 Response. COVID-19 Case Surveillance Public Data Access, Summary, and Limitations (version date: December 04, 2020).
Bureau Code	009:20
Program Code	009:020

Data Quality

Geographic Coverage	US
Temporal Applicability	2020-01-01/2020-11-19
Update Frequency	Monthly

[Show More](#)

What's in this Dataset?

Rows

8.4M



Columns

11

Each row is a

Deidentified Patient

Columns in this Dataset

Column Name	Description	Type
cdc_report_dt	Initial case report date to CDC	Date & Time  

Data Type Floating Timestamp	API Field Name cdc_report_dt			
pos_spec_dt	Date of first positive specimen collection	Date & Time	📅	▼
Data Type Floating Timestamp	API Field Name pos_spec_dt			
onset_dt	Symptom onset date, if symptomatic	Date & Time	📅	▼
Data Type Floating Timestamp	API Field Name onset_dt			
current_status	Case Status: Laboratory-confirmed case; Probable case	Plain Text	T	▼
Data Type Text	API Field Name current_status			
sex	Sex: Male; Female; Unknown; Other	Plain Text	T	▼
Data Type Text	API Field Name sex			
age_group	Age Group: 0 - 9 Years; 10 - 19 Years; 20 - 39 Years; 40 - 49 Years; 50 - 59 Years; 60 - 69 Years; 70 - 79 Years; 80 + Years	Plain Text	T	▼
Data Type Text	API Field Name age_group			
Race and ethnicity (combined)	Race and ethnicity (combined): Hispanic/Latino; American Indian / Alaska Native, Non-Hispanic; Asian, Non-Hispanic; Black, Non-Hispanic; Native Hawaiian / Other Pacific Islander, Non-Hispanic; White, Non-Hispanic; Multiple/Other, Non-Hispanic	Plain Text	T	▼

Data Type
Text

API Field Name
race_ethnicity_combined

[Show All \(11\)](#)

Table Preview

View Data

Create Visualization

cdc_r... :	pos_s... :	onse... :	curre... :	sex :	age... :	Race ... :	hosp... :	icu_yn
2020/11/10	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	No	Unknow
2020/11/14	2020/11/10	2020/11/10	Laborator...	Male	10 - 19 Ye...	Black, No...	No	No
2020/11/19	2020/11/10	2020/11/09	Laborator...	Male	10 - 19 Ye...	Black, No...	No	No
2020/11/14	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/13	2020/11/10	2020/11/10	Laborator...	Male	10 - 19 Ye...	Black, No...	No	No
2020/11/17	2020/11/10	2020/11/08	Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/14	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	No	Missing
2020/11/10	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/10	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Unknown	Unknow
2020/11/17	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/10	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/09	2020/11/10	2020/11/09	Laborator...	Male	10 - 19 Ye...	Black, No...	No	No
2020/11/10	2020/11/10		Laborator...	Male	10 - 19 Ye...	Black, No...	Missing	Missing
2020/11/15	2020/11/10	2020/11/08	Laborator...	Male	10 - 19 Ye...	Black, No...	No	No

< Previous [Next](#) >

Showing Deidentified Patients 1 to 14 out of 8,405,079

Related Content Using this Data



Public

COVID-19 Case Surveillance Public Use Data Profile



December
4, 2020

16K Views

[About CDC](#)

[Jobs](#)

[Funding](#)

[Policies](#)

[Privacy](#)

[FOIA](#)

[No Fear Act](#)

[OIG](#)

[Podcasts/RSS](#)

[Contact Us](#)



Powered By  Socrata