CDC CaseTrace

A healthcare outbreak investigation tool.

User Manual V1

Introduction

CDC's Division of Healthcare Quality Promotion (DHQP) advances patient safety by building on decades of expertise in preventing healthcare-associated infections (HAIs) and improving healthcare quality. CDC CaseTrace builds on DHQP's experience and expertise, translating that knowledge into tools that support safer patient care and support HAI outbreak investigations. CDC CaseTrace helps outbreak investigators and healthcare facilities identify potential HAI transmission events within healthcare settings, enabling more timely detection of transmission patterns and informing effective infection prevention and control measures.

This dashboard is designed to assist outbreak investigators and healthcare facilities in understanding patient movement in and among healthcare facilities. It identifies times and locations of potential patient exposures to the outbreak pathogen.

In each session, the user uploads their de-identified CDC CaseTrace-supported linelist using the provided template. Once uploaded, the user can visualize patient stays over time and by location with additional information of when the patient tested positive for the outbreak pathogen, received a treatment, or experienced an outcome of interest. A summary and detailed description of overlapping patient stays are presented in table format as well.

Linelists uploaded to this tool should not contain any identifiable information (e.g., patient or facility names) to ensure confidentiality. All data, visualizations, and tables are deleted from the server after each session is completed. This manual will walk through the step-by-step process of using this dashboard, providing visual snapshots and important information for the user.

Table of Contents

Linelist Setup

Columns

This tool requires a linelist which has **the same columns as those in the <u>Downloadable linelist template</u>** and that is in .xlsx or .csv format to ensure the tool performs properly. The required columns include: **patient** (a de-identified patient ID), **admission** (date of admission), **discharge** (date of discharge), **facility**, **unit**, **location**, **test**, and **test_date**. Optional columns include **WGS**, **event** and **event_date**. Please see the section <u>Linelist Dictionary</u> or the Linelist Dictionary tab on the dashboard for information about what each column is. Any additional columns contained in the linelist which are not listed above will not be used by the dashboard.

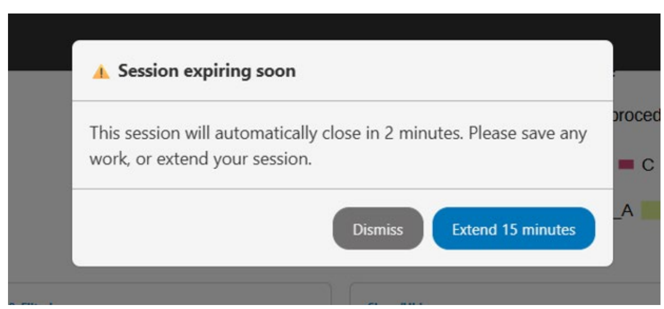
Rows

This tool requires linelists that are in the “long” format, meaning that **each row represents each stay, event, or test occurrence for a patient**. An event or test that occurred during a stay can be entered in the same row of the linelist under the appropriate columns (e.g., events under Event and Event_date, testing under Test and Test_date). However, if the patient received multiple tests in a single stay or multiple events occurred during a stay, there will need to be a separate row in the linelist for each of the subsequent tests or events with the same information entered for the patient, facility, location, and admission and discharge dates (see figure below for an example). For the overlap calculations, the tool will deduplicate all stays before calculating the number of days overlapping between patients.

patient	facility	unit	location	admission	discharge	test	test_date	wgs	event	event_date
104	ACH 1	H	007	2/2/2026	2/9/2026				scope	2/4/2026
104	ACH 1	H	007	2/2/2026	2/9/2026				procedure 1	2/9/2026
104	ACH 2	Q	020	2/9/2026	2/17/2026					
104	ACH 1	H	014	2/17/2026	2/20/2026	screening	2/17/2026		procedure 2	2/17/2026
104	ACH 1	H	014	2/17/2026	2/20/2026	PCR	2/20/2026	abc2		

Session timeouts

Each session will time out after 5 minutes of inactivity or 30 minutes regardless of whether there is activity. There will be a two-minute warning before the 30-minute timeout which will allow the user to extend their session by 15 minutes. Once a session times out, the linelist is deleted from the temporary memory and all the settings are reset. Refreshing the browser will start a new session and reset the 30-minute timer.



Starting the Session

Opening the Dashboard

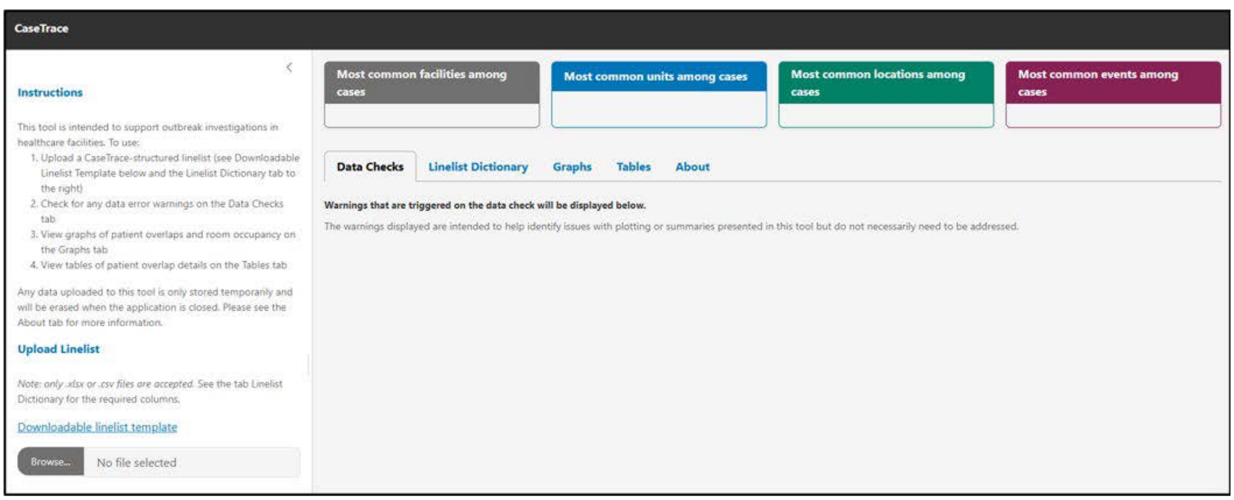
The dashboard operates using any commonly used web browser (e.g., Microsoft Edge, Google Chrome). To access the interactive dashboard, use any browser and navigate to the URL: https://cdcposit.cdc.gov/CaseTrace/

If the URL fails or the dashboard disconnects from the server unexpectedly, please send an email with the steps taken and any screenshots to CDC-CaseTrace@cdc.gov.

Layout of the Dashboard

The CDC CaseTrace tool shows step-by-step instructions and where to upload the linelist on the left-side panel. At the top of the page, there are 4 informational boxes which are detailed in the section Informational boxes. Below the boxes are 5 tabs which include:

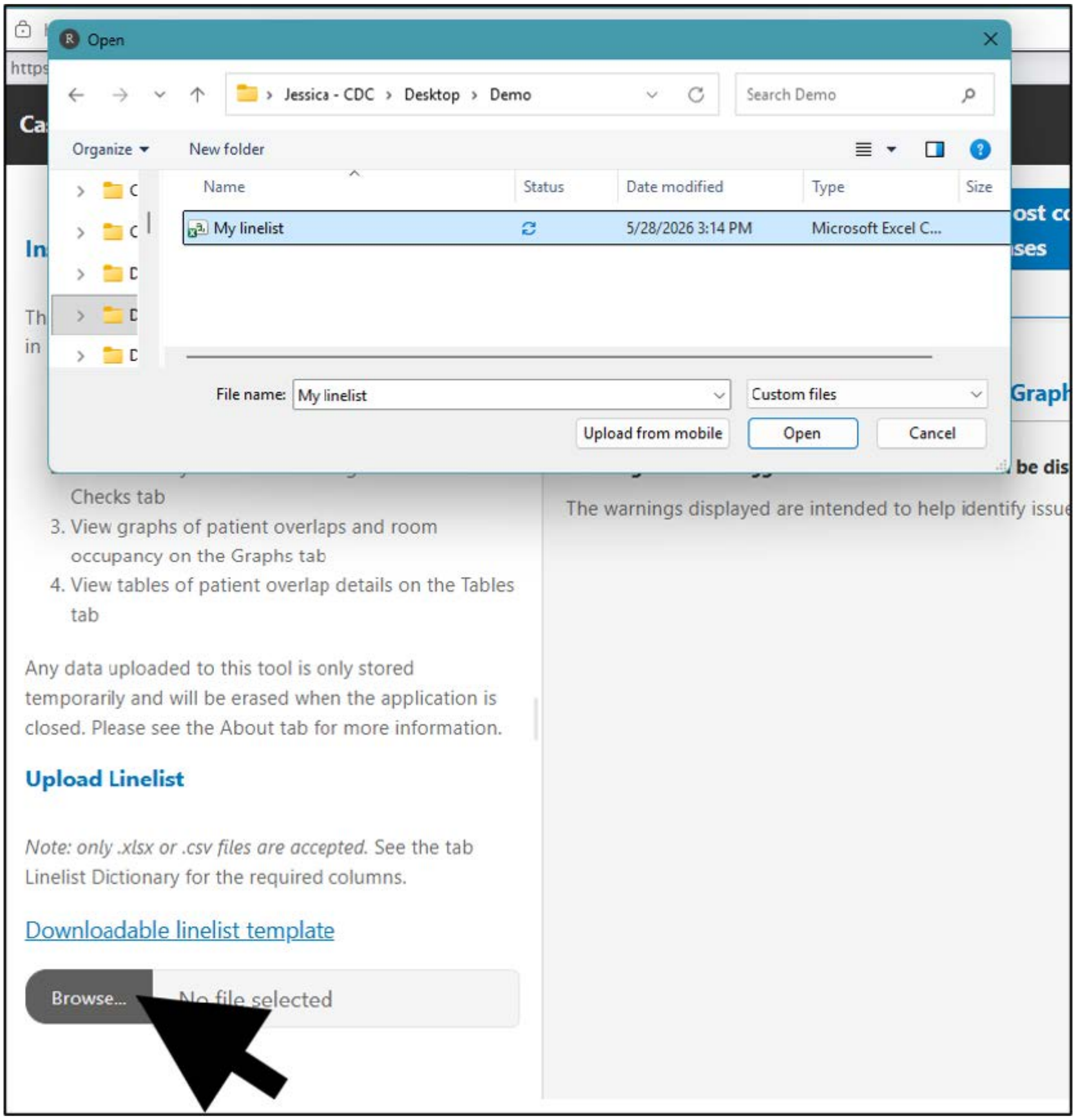
- Data Checks
- Linelist dictionary
- Graphs
- Tables
- About



Uploading a Linelist

To begin a session, the user uploads their CDC CaseTrace-structured linelist (see section Linelist Setup). The uploaded file must have **the same columns as those in the Downloadable linelist template to ensure the tool performs properly**. The user can either reformat an existing linelist to have the same column names (the column order does not matter and the column names are not case sensitive) or can download the template and populate it with their own data. Please see the Linelist Dictionary tab or section in this document for more information.

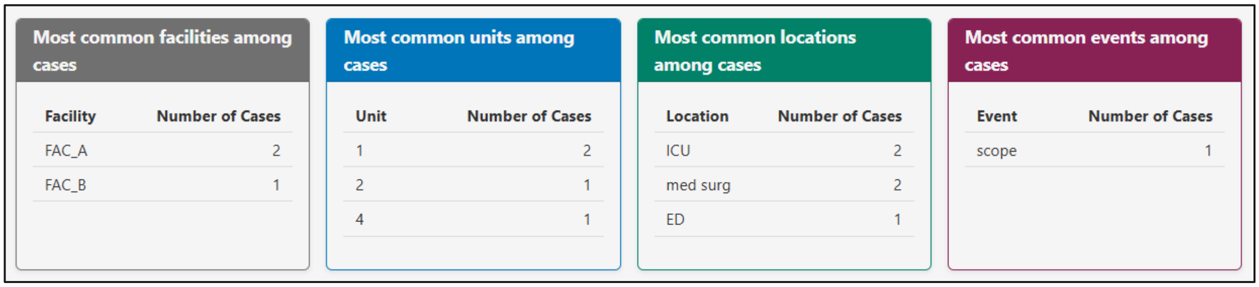
The user can then locate the "**Browse…**" button below the instructions, which once clicked, will open a file explorer window. The user can then navigate to the linelist which they would like to upload to the tool. Only excel (xlsx) and comma separated value (csv) files are supported for upload.



Navigating the Dashboard

Informational boxes

At the top of the dashboard, four boxes display high-level summaries intended to guide the user toward locations or events that may be of most interest.



From left to right the boxes display:

- The facilities most common among the patients who had any positive test, in dark grey
- The units most common among the patients who had any positive test, in blue
- The locations most common among the patients who had any positive test, in green
- The most common events among the patients who had any positive test, in maroon

Tab 1: Data Checks

The dashboard opens on the Data Checks tab. Once the user has uploaded a linelist, the tab will display the linelist in a table and will flag any potential issues. These warnings are intended to help identify issues with plotting or summaries presented in this tool, but not all need to be addressed. There are six possible warnings, which are listed in Table 1 below.

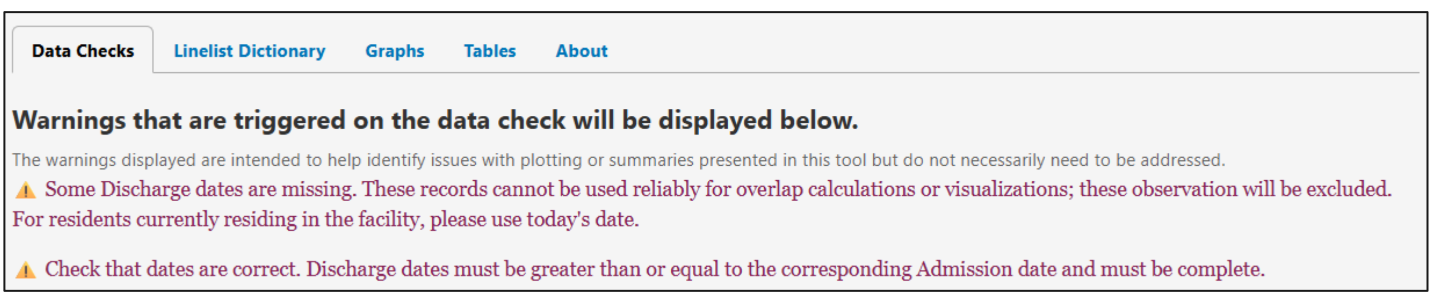


Table 1: Data Checks Warnings and Their Meanings

Warning	Meaning
Your data is missing the following required column(s):	The uploaded linelist is missing one or more of the expected columns (see Linelist Dictionary tab or section). The dashboard will not populate the graphs or tables without the required columns.
***[#]* records are missing a Facility name/Location information**	This is informational and alerts the user that these records may not be displayed in the graphs and tables depending on the selected options.
Some patient identifiers are missing.	All records must include a patient identifier to be included in the graphs and tables.
Some admission dates are missing.	Records missing admission dates cannot be reliably used for calculating patient overlaps or for visualization and will be excluded.
Some discharge dates are missing.	Records missing discharge dates cannot be reliably used for calculating patient overlaps or for visualization and will be excluded. **For patients currently staying in the facility, the user can insert today's date into the linelist before uploading.**
Check that dates are correct.	Records which have a discharge date that occurs before the admission date will be flagged and will be excluded.

Tab 2: Linelist Dictionary

This informational tab provides definitions for each column in the linelist template. The user may download a linelist template by clicking the "**Download Template Linelist**" button.

Uploaded linelists are expected to have columns named patient (a de-identified patient ID)**, admission** (date of admission)**, discharge** (date of discharge)**, facility** (identifier)**, unit, location, test, and test_date.**

For data to appear in graphs, **patient**, **admission**, and **discharge** fields must not be missing and the discharge date must be greater than or equal to the admission date. **Facility**, **unit** and **location** columns can be left blank.

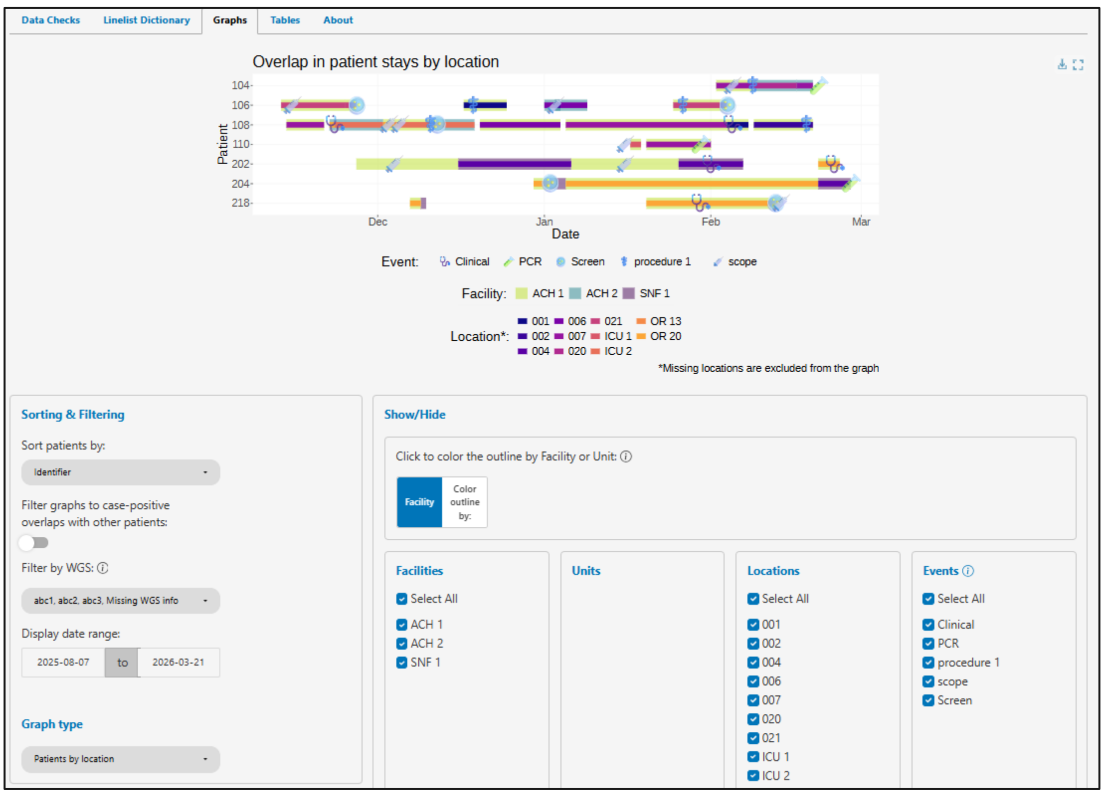
Optional columns include **WGS**, **event**, and **event_date**. **WGS** can contain whole genome sequencing information or pathogen identifiers for filtering graphs and tables. **Event** contains procedures, outcomes, or shared medical instruments of interest for the outbreak investigation. **Event_date** is the date on which the **event** occurred.

Download Template Linelist

Patient	Facility	Unit	Location	Admission	Test	Test_date	WGS	Event	Event_date	Discharge
Anonymized identifier for each patient	(Case-sensitive) Facility where the patient was seen	(Case-sensitive) Macro location within the facility (e.g., unit); can be left blank	(Case-sensitive) Location within the facility (e.g., room, ward); can be left blank	Date of admission to the location	Pathogen testing done on the patient at this location and facility	Date of pathogen testing	(Optional column) Whole genome sequencing data or pathogen detected	(Optional column) Procedure, use of a shared medical instrument, etc. on the patient at this location and facility	(Optional column) Date of the procedure, use of the shared medical instrument, etc.	Date of discharge from the location

Tab 3: Graphs

The Graphs tab displays interactive visualizations of patient stays over time. The graph may take a few moments to load depending on the size of the data.



Sorting & Filtering

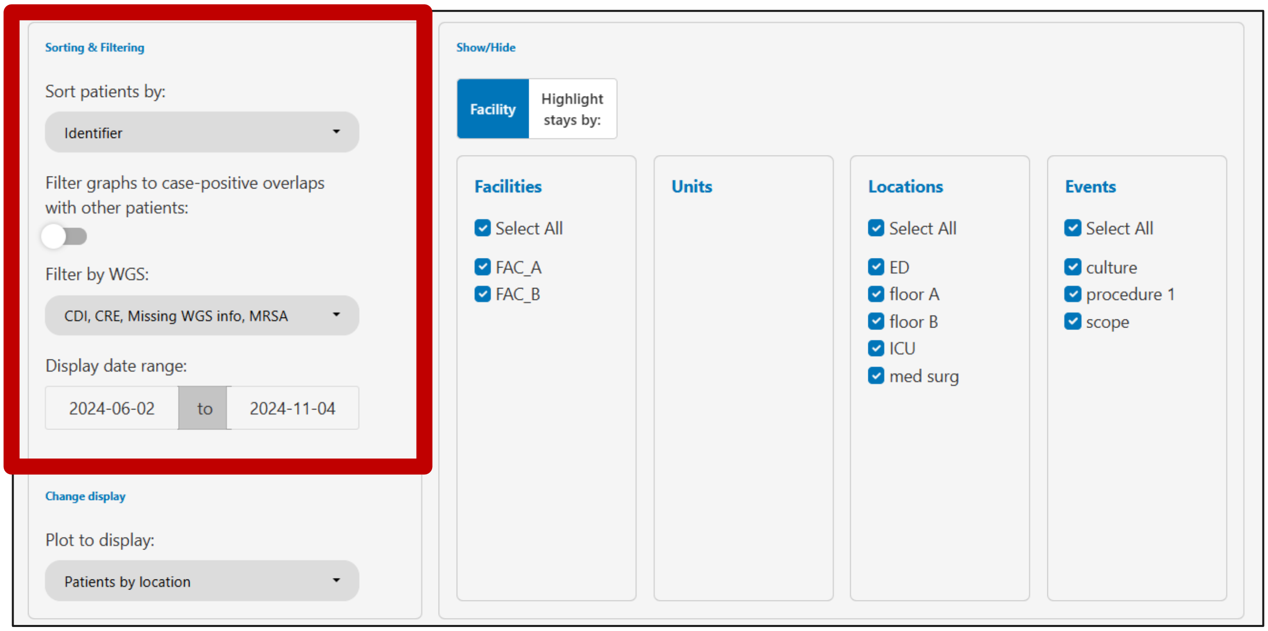
The left-hand box on the Graphs tab contains sorting and filtering controls:

- **Sort patients by:** Patients on the y-axis (i.e., the left, vertical axis) can be sorted by their de-identified patient ID, First culture (date of first positive test), or Location.

- **Filter graphs to case-positive overlaps with other patients:** This toggle filters the graph to only show patients who either had a positive test or who overlapped in time and location with a patient who tested positive at some point.

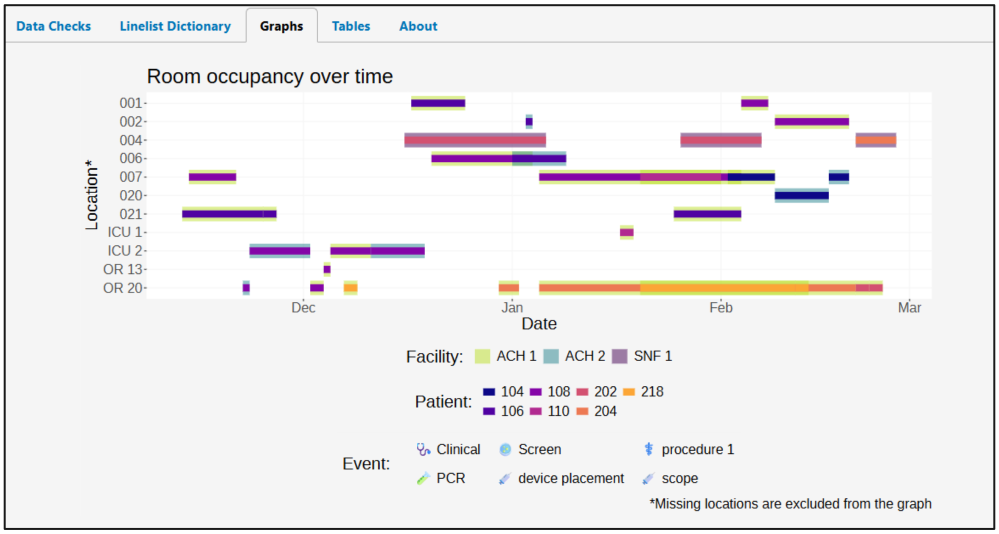
- **Filter by WGS:** Filters the graph to show only patients matching the selected whole genome sequencing (WGS) result(s). Multiple values can be selected. Patients with missing WGS information are listed as "Missing WGS info."

- **Display date range:** A date range selector selects the time period to display in the window. Patient stays that extend beyond the selected range are clipped to the range boundaries but are not removed from the data.



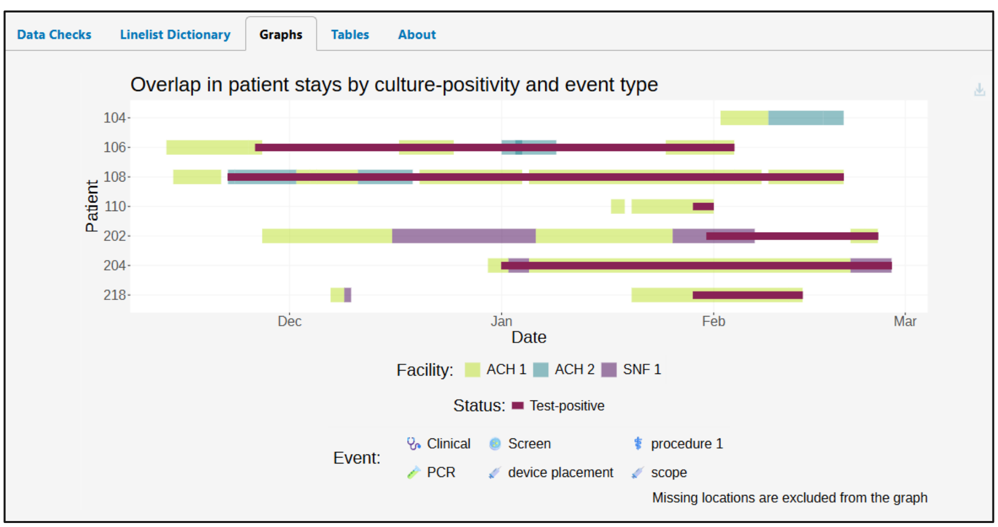
Change Display

Three graph types are available and can be selected using the "Change display" dropdown under the Sorting & Filtering section:



- **Patients by location:** Each patient's stay is shown as a horizontal bar, colored by facility or unit (thick bar) and location (thinner bar overlaid). Events and test dates (from the Events and Test columns in the linelist) are marked with contextual icons along the patient's timeline.

- **Patients by positivity:** The same patient stay bars are presented horizontally but are overlaid by a thinner bar to indicate the time at which the patient first tested positive and all time points after.



Room occupancy: Room or location occupancy over time is displayed, allowing the user to see which patients occupied a given location at the same time or in sequence.

Show/Hide

The right-hand box on the Graphs tab allows the user to control which elements are visible in the graph:

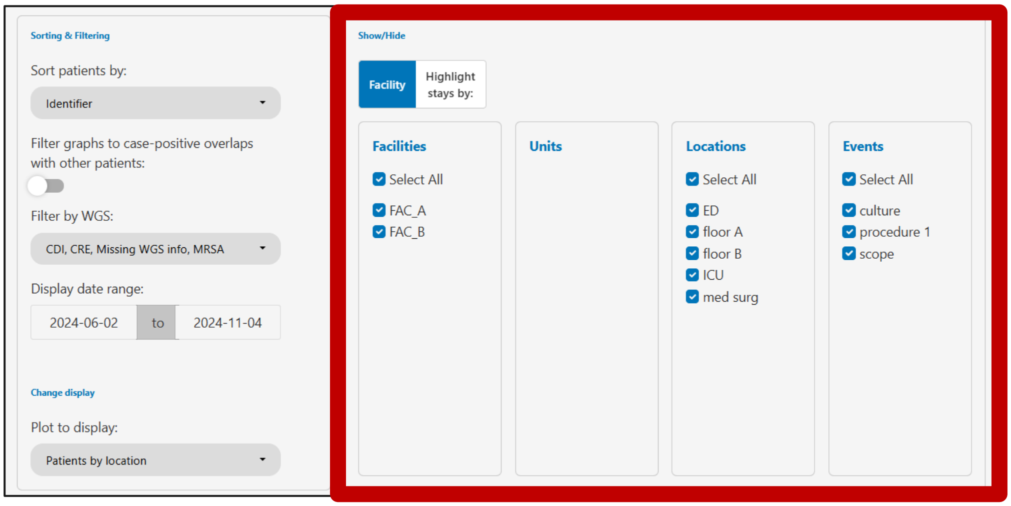
- **Highlight stays by — Facility / Unit toggle:** Using this toggle switches the thick background bar between coloring by Facility and coloring by Unit. When set to Facility, a list of facilities appears as checkboxes; when set to Unit, a list of units appears instead. Unchecking a facility or unit will hide those bars from the graph.

- **Linked location filtering** — when a facility or unit is checked or unchecked, the Locations list automatically updates to show only locations within the selected facility or unit.

- **Locations:** Checkboxes for each location in the data are available. Unchecking a location will hide the corresponding thin location bar from the graph.

- **Events:** Checkboxes for each event or test type in the data are available. Unchecking an event type will hide those event icons from the graph.

- **Select All checkboxes** — each group (Facilities, Units, Locations, Events) has a Select All toggle that includes all items when selected and removes all items when unselected.



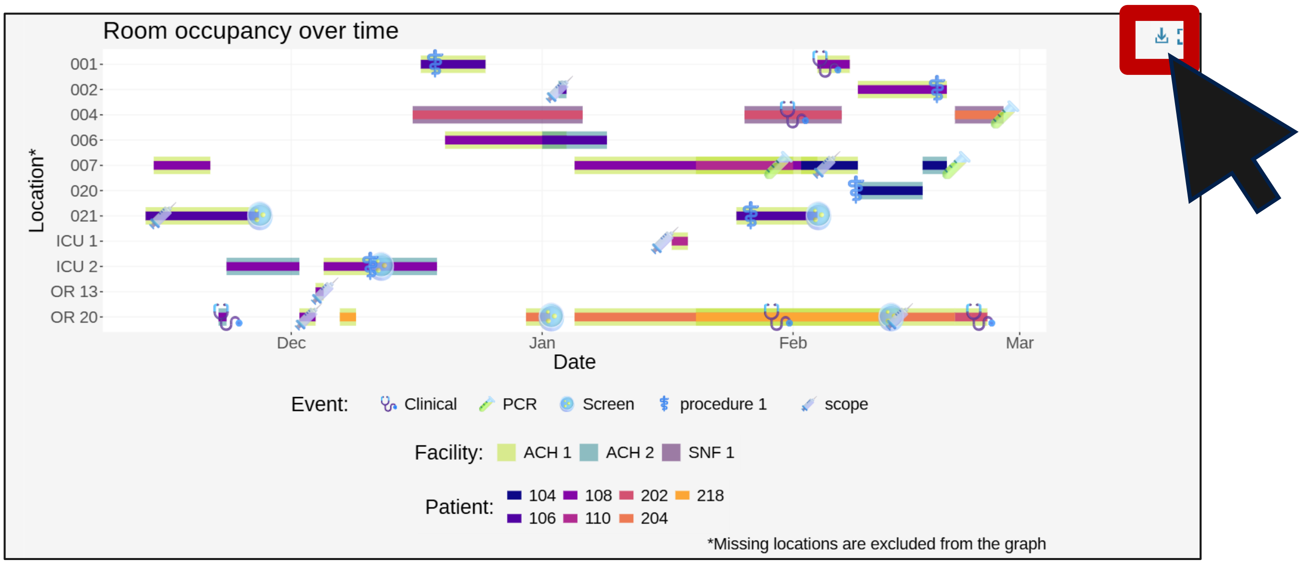
Interactive Features

The graphs are interactive and support the following actions:

- **Hover:** Hovering over any bar or event icon will display a tooltip showing the Facility, Unit, Location, Patient, and date range for that element.

Exporting the Graph

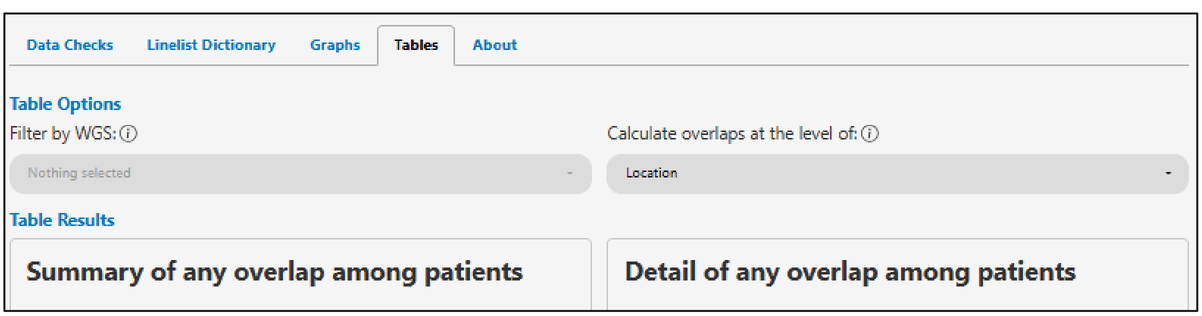
The graph can be exported as a PNG image using the download icon in the top-right toolbar of the graph. The image will download to the browser's default downloads folder.



Tab 4: Tables

The Tables tab presents quantitative summaries of patient overlap. Two filters at the top of the tab apply to all tables:

- **Filter by WGS column:** This filter limits the tables to patients matching the selected WGS result(s). Multiple values can be selected simultaneously by clicking each option in the dropdown. This will apply to all tables in the Tables tab.
- **Show overlaps in:** This filter controls the spatial level at which overlaps are calculated and summarized. Options are Facility, Unit, or Location. Selecting Location will show the most granular overlap information. This will apply to all tables in the Tables tab.



How overlaps are calculated

An “overlap” is defined as the time when a patient is present in the same facility/unit/location (depending on the option selected) as another patient inclusive of the date of admission and discharge. The tables on this tab will present the number of days of overlap between patients at varying levels of detail. Patients admitted on the same day as another patient is discharged will be counted as one day of overlap.

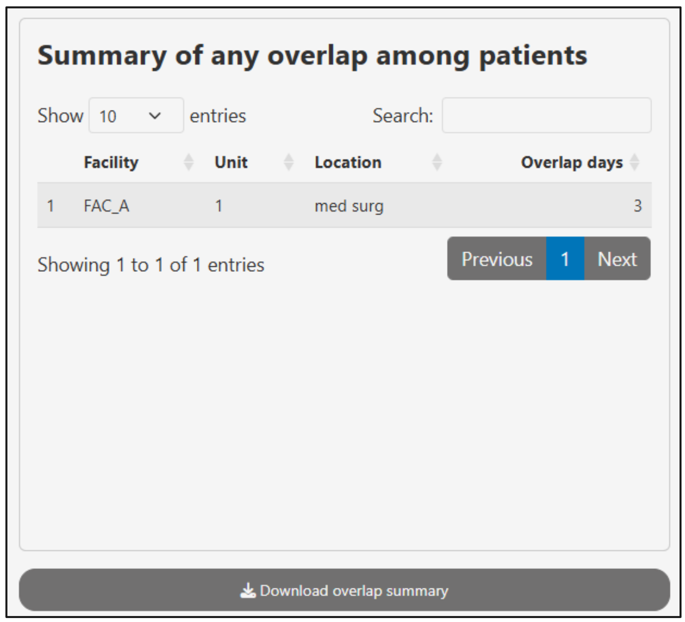
Examples:

1. Patient A is admitted to location X on May 1 and discharged on May 5. Patient B is admitted to location X on May 3 and is discharged June 1. The location overlap between the patients is 3 days.

2. Patient A is admitted to location X on May 1 and discharged on May 5, *then readmitted May 10 and discharged on May 15*. Patient B is admitted to location X on May 3 and is discharged June 1. The location overlap between the patients is 9 days (3 + 6).

Summary of any overlap among patients

The left table displays a summary of all patient-pair overlaps aggregated to the selected geographic level. Each row represents a unique facility-unit-location combination (depending on the "Show overlaps in" selection) and shows the total number of overlapping patient-days across all patient pairs in that location inclusive of admission and discharge date.



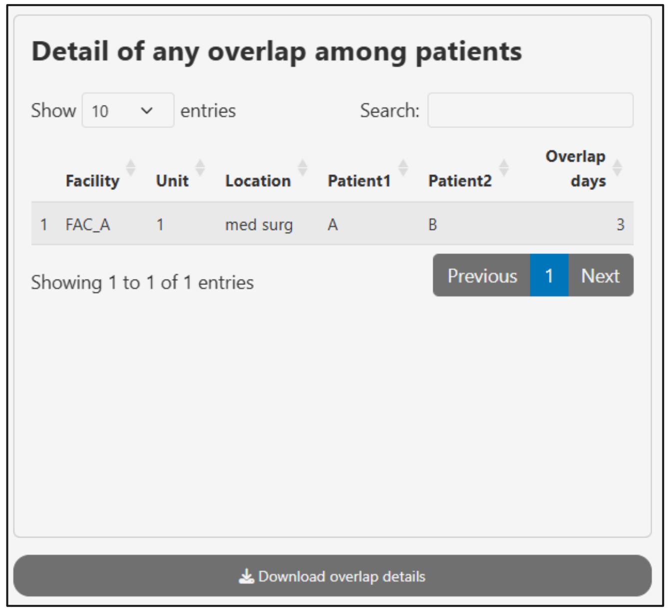
The table can be sorted by any column by clicking the column header, and can be searched using the Search box. The number of entries shown per page can be adjusted using the "Show entries" dropdown.

The table can be downloaded as a CSV file using the **"Download overlap summary"** button at the bottom of the tab. The file will download to the browser's default downloads folder.

Detail of any overlap among patients

The right table displays each individual patient pair that overlapped, along with the number of overlapping days. Each row represents one unique patient pair within a location. This table provides the most detailed view of which specific patients were in the same place at the same time.

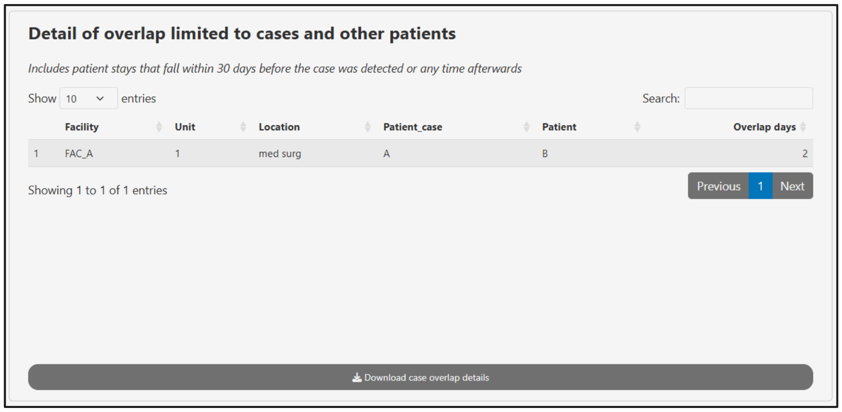
The table can be downloaded as a CSV file using the **"Download overlap details"** button at the bottom of the tab. The file will download to the browser's default Downloads folder.



Detail of Overlap Limited to Cases and Other Patients

A third table below the two above shows overlaps specifically between case-positive patients and any other patient. This table is limited to patient stays that fall within 30 days before the patient-case tested positive or any time afterwards, making it particularly useful for identifying potential exposure events in close proximity to a confirmed case.

The table can be downloaded as a CSV file using the **"Download case overlap details"** button. The file will download to the browser's default Downloads folder.



Tab 5: About

The About tab provides background information about the tool including its purpose, key definitions, disclaimers regarding data security, and known limitations. Users are encouraged to review this tab before interpreting results.

Key definitions displayed on this tab:

- **Overlap:** The number of days each patient pair overlaps within each facility, unit, or location, until both patients are positive or discharged are displayed. Overlapping days are counted even if neither patient tests positive.
- **Case-positive overlap:** This displays the number of days each patient pair overlaps within each facility and location given that Patient 1 is case-positive, until Patient 2 is case-positive or either patient is discharged.

Key limitations displayed on this tab:

- This tool does not account for transmission events that could have occurred before patients were tested
- Overlap is not proof of transmission taking place between two patients
- Linelist accuracy determines the output quality; missing or imprecise admission/discharge dates or facility/unit/location information can produce misleading results
- All outputs of this tool should be reviewed by the user for accuracy

Disclaimers

CDC CaseTrace is an analytic and visualization tool intended to support healthcare outbreak investigation activities by helping users identify patient stays that overlap by time and location. CDC CaseTrace outputs are intended for situational awareness and hypothesis generation and CDC recommends they be interpreted in combination with epidemiologic, laboratory, infection prevention, and facility-specific information.

The tool does not determine whether transmission occurred, identify a source patient, establish directionality of transmission, or replace professional public health judgment. Patient overlap shown in graphs or tables indicates potential temporal and spatial co-presence only.

This user manual describes how to prepare data for CDC CaseTrace and how to navigate and interpret the dashboard. It does not establish CDC policy, clinical guidance, regulatory requirements, or legal obligations. Users are responsible for ensuring that data uploaded to the tool are authorized for use and comply with applicable privacy, security, institutional, and jurisdictional requirements.

Linelists uploaded to CDC CaseTrace should not contain direct identifiers or other information that could identify individual patients unless use of such data has been explicitly authorized through appropriate privacy, security, and governance processes. Users should also consider whether combinations of dates, facilities, units, locations, and rare events could create re-identification risk.

AI Use Statement

Claude was used to assist in some troubleshooting of backend dashboard code. Claude was also used to help draft the first draft of this user manual.

All AI-assisted content was reviewed and edited by the authors/developers, who are responsible for the final dashboard, code, analyses, interpretations, and user manual. AI tools were not used as authors or as independent authorities for scientific, policy, legal, privacy, or security determinations.