

How to Use NSSP–ESSENCE CCI and DDI Data Quality Filters

Why Data Quality Metrics Matter

The dynamic nature of high-volume, timely syndromic surveillance (SyS) data can lead to fluctuation in data quality metrics and influence one's interpretation of trend data. To adjust for this, NSSP has a process for integrating chief complaint informative (CCI) and discharge diagnosis informative (DDI) data quality metrics. The key is to use CCIs and DDIs that are *informative*. A CC or DD is considered informative if it is complete (not blank) and is not on the predefined list of terms that NSSP finds *noninformative*. "Noninformative" refers to a chief complaint with a non-null value that adds negligible value to SyS.¹ An example of a noninformative chief complaint on the list is "REFERRAL." For a list of Noninformative Chief Complaints (NICC), see the [NSSP Data Dictionary](#).

This job aid will help you assess completeness and informativeness of chief complaints (CC) and discharge diagnoses (DD) for a site² and multiyear period. It's a two-step process: first, you need to calculate the CCI and DDI percentages for the period of interest; then, by following data quality checks, you can determine which CCI and DDI filters to use when analyzing trends for a particular syndrome or CCDD category. The second step is to use this information to decide which DDI and CCI data quality filters apply.

1. Assess CC and DD Informative Percentages for Period of Interest

Steps:

1. Using Chrome or Microsoft Edge, navigate to NSSP–ESSENCE via this link: https://essence2.syndromicsurveillance.org/nssp_essence/servlet/HomePageServlet.
2. Select "Query Portal" from the home page.
3. Select "Facility Location (Full Details)" under Datasource."
4. Select your site name from the "Site" query field. Select "Discharge Diagnosis Informative" under "Data Quality Filters" and choose "Yes" (see Figure 1).
5. (Optional) To limit to emergency department (ED) encounters, select "Has Been Emergency" and choose "Yes."
6. Change "Time Resolution" from "Daily" to "Weekly."
7. Select "Discharge Diagnosis Informative" for "As Percent Query."
8. Modify "Start Date" and "End Date" to correspond to the multiyear period you want to assess.
9. Choose the "Time Series" option at the bottom of the Query Wizard to visualize the weekly DDI percentages over the selected date range.

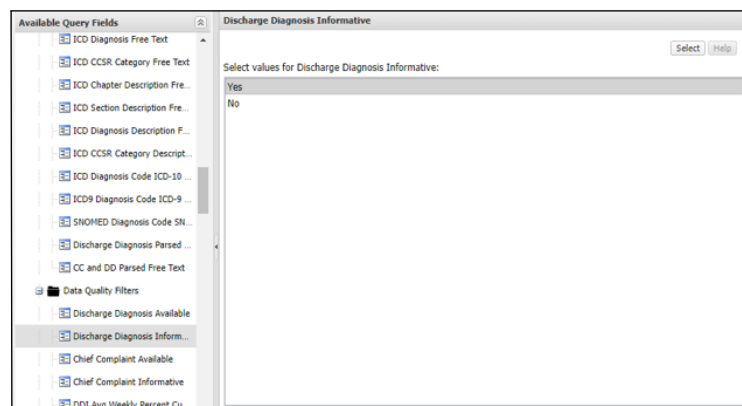


Figure 1: Select "Yes" for "Discharge Diagnosis Informative" under "Data Quality Filters."

For example, if an epidemiologist in a public health department wants to examine the weekly percentage of ED encounters with an informative discharge diagnosis between week 1 of 2020 and week 7 of 2023, then the Query Wizard portal will display as shown in Figure 2. The user would need to select a site after making these selections.

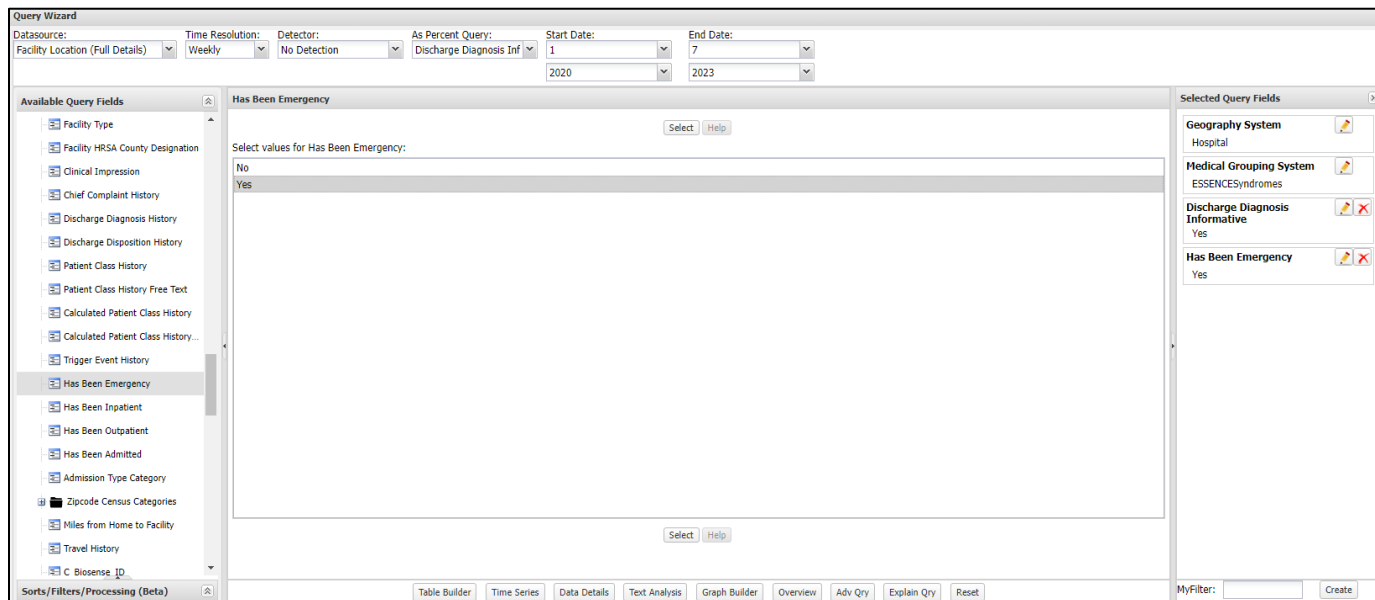


Figure 2: ESSENCE Query Wizard with selected query fields for examining weekly DDI percentages from January 2020 through February 2023

Figure 3 shows a weekly DDI percentage time series for national ED encounters. The decrease in percent DDI at the end of the time series is expected and reflects discharge diagnoses filling in over the most recent week. Discharge diagnoses are not as timely as near real-time chief complaints and can take 3 to 4 days to fill in. (Keep in mind that ICD coding decisions can be influenced by medical billing, which is the original intent of ICD diagnosis codes.)

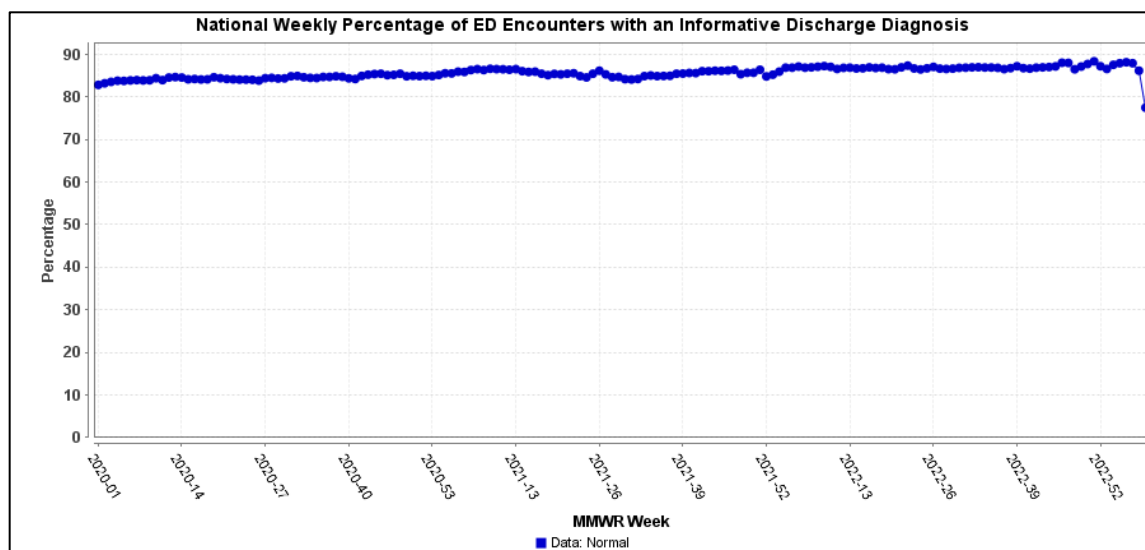


Figure 3: Weekly DDI percentages for national NSSP–ESSENCE ED encounters during January 2020 to February 2023

This assessment will identify if substantial fluctuations in percent DDI or percent CCI have occurred within the site over a multiyear period. If large fluctuations have occurred (i.e., large increases in percent DDI or percent CCI), then any analyses of syndromic trends for a particular syndrome, CCDD category, or query could be implicitly biased by changes in data quality over the analytic period of interest rather than by increases in ED encounters identified by the definition due to public health events.

Example

A site administrator wants to examine weekly counts or percentages of ED encounters identified by a CCDD category or query that is largely (or completely) driven by diagnosis codes between 2018 and 2022. Figure 4 displays a simulated time series for a scenario in which weekly DDI percentages have an overall increase with percent DDI about 50% at the start of the assessed period (January 2018) and percent DDI about 93% at the end of the assessed period (December 2022).

The site administrator chose to apply DDI data quality filters. Had the site administrator NOT applied DDI data quality filters when analyzing time series for a diagnostic query or definition over the same multiyear period, then potential increases in counts or percentages of encounters identified by the query or definition could reflect overall increases in completeness of discharge diagnoses.

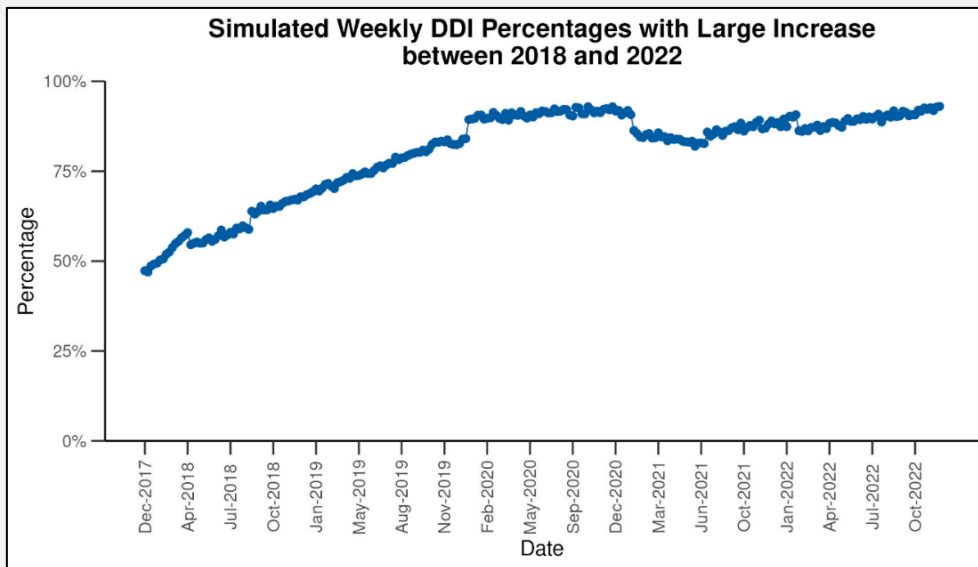


Figure 4: Simulated scenario in which a site has a large, overall increase in the percentage of ED encounters with an informative discharge diagnosis during a multiyear period spanning January 2018 through December 2022

The next section shows CCI and DDI data quality filters that can be used to address this problem. Data quality filters can limit data pulls for a query or definition to facilities with consistently high completeness of CCs or DDs over the period being analyzed.

Note This process can be repeated to examine weekly CCI percentages. Simply substitute "Chief Complaint Informative" for "Discharge Diagnosis Informative."

2. Select DDI and CCI Data Quality Filters

If you've followed the guidance so far, you're ready to make an informed selection. The Chief Complaint Informative (CCI) and Discharge Diagnosis Informative (DDI) data quality filters described previously are for assessing completeness and informativeness of CCs and DDs for a site over time. Traditionally, CCI and DDI average weekly percent filters are used with an operator (i.e., greater than or greater than or equal to) to limit data to facilities with *consistently high* completeness/informativeness. As of February 2023, Nssp-ESSENCE³ offered six variations of these filters for CCI and DDI that correspond to the years to date that will be analyzed for syndromic trends. (Figure 5).

You can determine the CCI and DDI average weekly percent data quality filters for site facilities by calculating the weekly percentage of ED encounters with an informative chief complaint or discharge diagnosis. Weekly percentages are averaged across all weeks for the time range associated with the selected filter. The date range assessed is the first *MMWR* week of *n* years ago to date minus a 2-*MMWR*-week lag from today. Hovering over the filter name in the ESSENCE user interface will display a tip summarizing how the data quality metric is calculated.

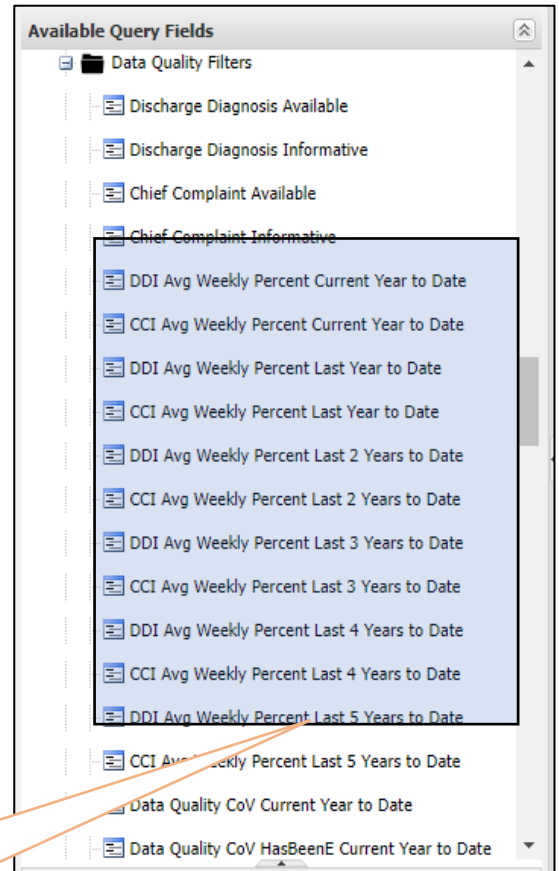
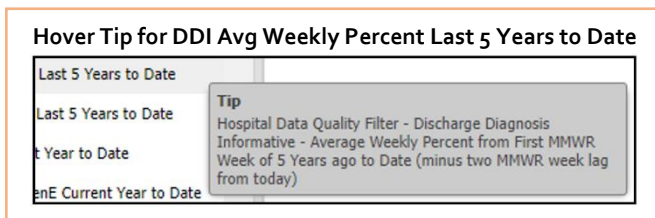


Figure 5. CCI and DDI filters are available to analyze different periods of interest.

Most often, the DDI average weekly percent data quality filters are used with an operator of "Greater Than or Equal" and a threshold of 75% to limit analyses to facilities with a weekly average percent DDI of 75% or higher. The CCI average weekly percent data quality filters are not as widely used. Most sites have high CCI percentages—80% or higher—with little variation over time and almost no impact on interpretation of syndromic trends.

The preceding steps will help you determine whether to use the CCI average weekly percentage data quality filters. Depending on the consistency of completeness of discharge diagnoses for your site, you may want to use percentage cutoffs other than the widely used cutoff of 75%. The [Data Quality Filter Matrix Rnssp Template](#) in RStudio Workbench provides a detailed summary of the number of facilities from your site that would be included for various DDI percentage cutoffs individually and in conjunction with cutoffs for coefficient of variation (CoV) filters.

Build the Query

Follow these steps to build a query for a diagnostic CCDD category and limit facilities to those with an average weekly DDI percent of 75% or higher:

Steps:

1. Navigate to the Query Wizard in NSSP–ESSENCE and select “Facility Location (Full Details)” under “Datasource.”
2. Choose your site name under “Site.”
3. Choose the CCDD category of interest under “CC and DD Category.”
4. (Optional): To limit to ED encounters, select “Has Been Emergency” and choose “Yes.”
5. Choose the “DDI Avg Weekly Percent” filter corresponding to the period.
6. Select “Greater Than or Equal” from the operator drop-down list (Figure 6).
7. Populate the percentage cutoff in the text box under the selected operator.
8. Select “Weekly” under “Time Resolution.”
9. Choose values for “Start Date” and “End Date” that are consistent with the amount of time corresponding to the DDI average weekly percent filter you chose.
10. Depending on the analysis you are planning, select either “Time Series” or “Table Builder” at the bottom of the Query Wizard.

The screenshot shows a window titled "DDI Avg Weekly Percent Last 3 Years to Date". At the top right, there are "Select" and "Help" buttons. Below this, there is a dropdown menu labeled "Select an operator for D" with a downward arrow. The dropdown menu is open, displaying the following options: "Equal", "Does Not Equal", "Less Than", "Less Than or Equal", "Greater Than", "Greater Than or Equal" (which is highlighted with a grey background), "Between", and "Is Null". To the right of the dropdown menu, there is a text input field. Below the text input field, there are another "Select" and "Help" buttons.

Figure 6: Operator selections

Example

An analyst with a public health department wants to determine the weekly count or percentage of ED encounters with diagnosed RSV using the CDC Respiratory Syncytial Virus DD v1 CCDD category between January 2020 and February 2023 (present date). The analyst would select “DDI Avg Weekly Percent Last 3 Years to Date” data quality filter. The Query Wizard portal would appear as it does in Figure 7.

The screenshot shows the ESSENCE Query Wizard interface. At the top, there are fields for 'Data source' (Facility Location (Full Details)), 'Time Resolution' (Weekly), 'Detector' (No Detection), 'As Percent Query' (No Percentage Query), 'Start Date' (1/2020), and 'End Date' (7/2023). Below this, the 'Available Query Fields' list includes 'DDI Avg Weekly Percent Last 3 Years to Date' which is selected. The central configuration area shows a filter set to 'Greater Than or Equal' with a value of '75'. The 'Selected Query Fields' panel on the right includes 'Geography System' (Hospital), 'Medical Grouping System' (ESSENCE Syndromes), 'DDI Avg Weekly Percent Last 3 Years to Date' (Greater Than or Equal 75), 'CC and DD Category' (CDC Respiratory Syncytial Virus DD v1), and 'Has Been Emergency' (Yes). At the bottom, there are buttons for 'Table Builder', 'Time Series', 'Data Details', 'Text Analysis', 'Graph Builder', 'Overview', 'Adv Qry', 'Explain Qry', and 'Reset', along with a 'MyFilter' field and a 'Create' button.

Figure 7: ESSENCE Query Wizard with selected query fields and DDI average weekly percent filter selections for examining weekly ED encounters with diagnosed RSV from January 2020 through February 2023. The user would need to select a site after making these selections.

Resources

[NSSP Technical Resources and Standards](#): Go to *NSSP Data Dictionary*.

Notes

1. Centers for Disease Control and Prevention. [What's a Noninformative Chief Complaint \(NICC\)? | CDC](#) [Internet]. Atlanta (GA): CDC; [cited 2023 Mar 27]. National Syndromic Surveillance Program. Available from: www.cdc.gov/nssp.dqc/articles/whats-a-nicc.html
2. The National Syndromic Surveillance Program (NSSP) groups facilities under administrative authorities called sites. A *site* is a public health department (local, state, county, or federal) that serves as the administrative hub for any number of facilities that will transmit data to the BioSense Platform. Although the term “site” is often used interchangeably with public health “jurisdiction,” the terms are not the same. “Sites” originated as logical groupings of public health departments that operated in the same geographic areas—but not necessarily under the same public health authority. Within a site, facility data are pooled to get a better understanding of community health.
3. ESSENCE (Electronic Surveillance System for the Early Notification of Community-based Epidemics) is a disease surveillance system developed by the Johns Hopkins University Applied Physics Laboratory. ESSENCE is a surveillance tool used by CDC’s National Syndromic Surveillance Program to collect, process, monitor, and analyze emergency department and urgent care visit data.

This job aid was produced in collaboration with the Wisconsin Department of Health Services by the Centers for Disease Control and Prevention (CDC); Office of Public Health Data, Surveillance, and Technology; Division of Health Informatics and Surveillance. The job aid reflects best practices of CDC’s National Syndromic Surveillance Program (NSSP).