PERFORMING DATA QUALITY CHECKS FOR GROUP DATA IN NHSN: *C. DIFFICILE* INFECTION (CDI) LABID EVENT REPORTING
This document was prepared by the CDC LTCF NHSN support team to provide information for group users on how to review the data submitted by the nursing homes in their groups. Data quality checks are important to ensure that the information entered into HNSN is accurate and complete. This verification will prevent misrepresentation of disease trends. This document discusses the steps and reports available in NHSN that can help in verifying data accuracy for facilities participating in C. difficile infection (CDI) LabID event reporting.
Contents

WHEN ARE DATA CONSIDERED COMPLETE IN NHSN? ................................................................. 3

STEPS FOR PERFORMING QUALITY CHECKS OF NHSN DATA .................................................. 3
  Generating Datasets .................................................................................................................. 3
  Analysis .................................................................................................................................... 4

HOW TO VERIFY FACILITY LEVEL INFORMATION? .............................................................. 5

HAVE THE FACILITIES CORRECTLY COMPLETED THE MONTHLY REPORTING PLAN? ............. 6

HAVE THE FACILITIES CORRECTLY ENTERED CDI EVENTS IN NHSN? ................................. 7

HAVE THE FACILITIES CORRECTLY ENTERED DENOMINATOR DATA INTO NHSN? ................. 9

ANALYSIS TIPS: .......................................................................................................................... 10
  HOW TO CALCULATE AN ESTIMATE FOR THE NUMBER OF RESIDENT DAYS USING NUMBER OF BEDS IN A FACILITY? .................................................................................................................. 10
  HOW TO ACCESS FACILITY NAME AND ORGID? .................................................................. 12

DEFINITIONS ............................................................................................................................... 13
WHEN ARE DATA CONSIDERED COMPLETE IN NHSN?

Data are considered complete in NHSN once facilities do **ALL** of the following:

- Complete the Annual Facility Survey
- Complete a monthly reporting plan to indicate which module(s) and event(s) a facility plans to report for each month in which facility plans to perform surveillance and submit data to NHSN
- Submit event data to NHSN using the NHSN Multidrug-Resistant Organism & *Clostridium difficile* Infection (MDRO/CDI) Module protocols
- Indicate “no events” on monthly summary data page if no CDI events were found for the month
- Submit monthly summary data to NHSN
- Resolve all outstanding alerts

STEPS FOR PERFORMING QUALITY CHECKS OF NHSN DATA

The NHSN application has several analysis reports that can be accessed to validate completeness and accuracy of data being submitted to NHSN, including: (1) Facility Survey Data Line List; (2) Monthly Reporting Plan Line List; (3) Event Level Data Line List; and (4) Summary Level Data Line List.

Generating Datasets

Before accessing analysis reports in NHSN, datasets must be generated to ensure the most recent data are included in the reports. Generating datasets will take a snapshot of the data that have been submitted.

*Remember: Datasets must be generated prior to analyzing the data or the data will not be current.*

To generate datasets:

On your NHSN landing page, select “Analysis” then “Generate Data Sets”
You will see a page titled “Generate Data Sets.” Next, check the box highlighted in the green circle below, before selecting “Generate New.” Generating datasets may take a few minutes; however, you can still navigate within the NHSN application while the dataset is generating.

Analysis

Once your dataset has been successfully generated, select “Analysis” then “Reports.”

*Remember: Make sure that you have generated a new dataset before completing this step or you will risk analyzing an older dataset without the most recent updates.*

To analyze NHSN data:

On your NHSN landing page, select “Analysis” then “Reports”
* Please note: The reports shown below are generated from test facilities and are hypothetical.

HOW TO VERIFY FACILITY LEVEL INFORMATION?

**Facility Survey Data Line List.** To verify facility level information, a report can be generated that summarizes data from the Annual Survey. The report is a line list of facility survey data and contains information such as CCN, facility affiliation, number of beds, and average daily census.

To create a line list of facility survey data, navigate to the “Advanced folders”: Open the “Facility-level data” folder
1. Select “Line listing – Facility survey data”
2. Select “run report”

*Remember: Make sure that you have generated a new dataset before completing this step or you will risk analyzing an older dataset without the most recent updates.*

---

Data Source: Annual survey

```
<table>
<thead>
<tr>
<th>Facility Org ID</th>
<th>Survey Year</th>
<th>CMS Certification Number</th>
<th>Certification</th>
<th>Affiliation</th>
<th>Average Daily Census</th>
<th>Number of Beds</th>
<th>Primary C. df Test Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>39455</td>
<td>2014</td>
<td>324569</td>
<td>DUAL</td>
<td>IFS</td>
<td>50</td>
<td>34</td>
<td>EIA</td>
</tr>
<tr>
<td>39455</td>
<td>2015</td>
<td>324569</td>
<td>MCARE</td>
<td>IFS</td>
<td>77</td>
<td>92</td>
<td>NAAT</td>
</tr>
<tr>
<td>39455</td>
<td>2016</td>
<td>324569</td>
<td>MCARE</td>
<td>IFS</td>
<td>100</td>
<td>200</td>
<td>ToxOul</td>
</tr>
<tr>
<td>41141</td>
<td>2014</td>
<td></td>
<td>DUAL</td>
<td>IFS</td>
<td>88</td>
<td>100</td>
<td>EIA</td>
</tr>
<tr>
<td>41141</td>
<td>2016</td>
<td></td>
<td>DUAL</td>
<td>IFS</td>
<td>95</td>
<td>100</td>
<td>NAAT</td>
</tr>
</tbody>
</table>
```
Common Errors in Annual survey data include:
- Incorrect or missing CCN number
- Incorrect census data (example: average daily census greater than number of beds)
- Incorrect CDI test method -- verify that “other” or “toxigenic culture” has not been selected as primary testing method

HAVE THE FACILITIES CORRECTLY COMPLETED THE MONTHLY REPORTING PLAN?

Monthly Reporting Plan line list. This report summarizes the monthly plans reported by each facility in your group. The monthly reporting plan line list allows group-users to review the organisms that the facilities plan to report monthly. Additionally, this report allows group-users to verify that a completed monthly reporting plan has been submitted for each month.

To create a line list of monthly reporting plans, navigate to the “Advanced folders”:
1. Open the “Plan data” folder
2. Select “Line listing – Monthly Reporting Plan”
3. Select “run report”

Remember: Make sure that you have generated a new dataset before completing this step or you will risk analyzing an older dataset without the most recent updates.
Common Errors in Annual survey data include:

- CDI reporting plan missing for one or more months
- Incorrect organism selected for surveillance (example: an organism such as VRE or creEcoli selected for one month but missing for other months)

HAVE THE FACILITIES CORRECTLY ENTERED CDI EVENTS IN NHSN?

Line Listing – All events. The line listing of an event allows group-users to view all of the events reported by each facility each month.

To create a line list of all events, navigate to the “Advanced folders”:
1. Open the “Event-Level Data” folder
2. Select “Line listing – All Events”
3. Select “run report”

Remember: Make sure that you have generated a new dataset before completing this step or you will risk analyzing an older dataset without the most recent updates.

Report: Line List for all Events
Common Errors in Event level data include:

- Missing reporting month -- If a facility does not enter an event for a given month, that month will not appear in the line list of events. Missing months should be further investigated to verify whether there were no events to report for the month or if a facility has incomplete reporting for the month.

Assess accuracy of '0' events reported 6 months or more

- If no event reported, follow-up with NH and ask:
  In the past month......
  - a) Did you submit any stool specimens for C. diff testing to lab?
  - b) Did you start any residents on treatment for CDI?
HAVE THE FACILITIES CORRECTLY ENTERED DENOMINATOR DATA INTO NHSN?

**All Summary Data line list.** The line list for all summary data allows group users to view the denominator data entered by each facility in a group by month including the number of resident days and number of admissions on CDI treatment. Even if a facility has no events for a given month, the monthly summary data form is still required and provides further information on the events entered by the facility. Therefore, this table can be used along with the line listing of events to verify that all events have been entered correctly in NHSN.

To create a line list of all summary data, navigate to the “Advanced folders”:
1. Open the “Summary-Level Data” folder
2. Select “Line listing – All Summary Data”
3. Select “run report”

**Remember:** Make sure that you have generated a new dataset before completing this step or you will risk analyzing an older dataset without the most recent updates.

**Report: Line List for all Summary Data**

Data Source: Monthly Summary Data Form
Common Errors in Monthly Summary data include:

- Number of resident days =0
- Incorrect number of residence days (see section analysis tip)
- Missing events:
  - “No Events” column:
    - NoEvents = N – Events entered into NHSN for the month
    - NoEvents = “Y” – No events to report for the month
    - No Events= “ “ – Incomplete or missing event for the month
- CDI event incorrectly entered as part of the “ Number of Admissions on C. diff Treatment,”
  - Look for possible inconsistencies for example: number of admissions on C. diff treatment that is greater than the number of beds (see Facility Survey Data Line List).

ANALYSIS TIPS:

How to calculate an estimate for the number of resident days using number of beds in a facility?

The number of beds from the “Facility Survey Data Line List” is helpful in calculating an estimate of the number of resident days reported in the summary data line list.

Variables used:

- **Number of beds**: Obtained from facility survey line list
- **Occupancy**: How many beds are occupied in the facility? You can choose between 50% - 100% in order to estimate the resident days when the facility is between at half capacity to full capacity.
  - 100% = 1
  - 50% = 0.50
- **Number of days**: 30 days (1 month)

**Estimated number of resident days** = number of beds * occupancy* number of days
1. Look up the number of beds reported for the 2016 survey year. In this example, we will look up the number of beds for OrgID 39455 for the 2016 survey year.

   **National Healthcare Safety Network**  
   **Line Listing of Facility Survey Data**  
   As of: October 26, 2017 at 1:32 PM  
   Data Range: All LTC Survey

<table>
<thead>
<tr>
<th>Facility Org ID</th>
<th>Survey Year</th>
<th>CMS Certification Number</th>
<th>Certification</th>
<th>Affiliation</th>
<th>Average Daily Census</th>
<th>Number of Beds</th>
<th>Primary C. diff Test Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>39455</td>
<td>2014</td>
<td>324569</td>
<td>DUAL</td>
<td>IFS</td>
<td>50</td>
<td>34</td>
<td>EIA</td>
</tr>
<tr>
<td>39455</td>
<td>2015</td>
<td>324569</td>
<td>MCARE</td>
<td>IFS</td>
<td>77</td>
<td>92</td>
<td>NAAT</td>
</tr>
<tr>
<td>39455</td>
<td>2016</td>
<td>324569</td>
<td>MCARE</td>
<td>IFS</td>
<td>100</td>
<td>200</td>
<td>ToxCol</td>
</tr>
<tr>
<td>41141</td>
<td>2014</td>
<td></td>
<td>DUAL</td>
<td>IFS</td>
<td>88</td>
<td>100</td>
<td>EIA</td>
</tr>
<tr>
<td>41141</td>
<td>2016</td>
<td></td>
<td>DUAL</td>
<td>IFS</td>
<td>95</td>
<td>100</td>
<td>NAAT</td>
</tr>
</tbody>
</table>

2. Plug in the number of beds in the calculation above:

   Estimating resident days for a facility at full capacity:  
   \[200 \times 1 \times 30 \text{ days} = 6000 \text{ resident days}\]

   Estimating resident days for a facility at half capacity:  
   \[200 \times 0.5 \times 30 \text{ days} = 3000 \text{ resident days}\]

3. The estimated resident days is between 3000 – 6000 days. We can use this estimate to assess the number of resident days reported by the facility.

   **Line Listing for All Summary Data**  
   As of: October 26, 2017 at 3:35 PM  
   Data Range: LTC SUMMARY summary YM 2017M01 to 2017M04

<table>
<thead>
<tr>
<th>Facility Org ID</th>
<th>Summary Year/Month</th>
<th>Event Type</th>
<th>Number of Resident Days</th>
<th>Admissions on C. diff Treatment</th>
<th>No Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>39455</td>
<td>2017M01</td>
<td>CDIF</td>
<td>5,700</td>
<td>1</td>
<td>N</td>
</tr>
<tr>
<td>39455</td>
<td>2017M02</td>
<td>CDIF</td>
<td>5,450</td>
<td>3</td>
<td>N</td>
</tr>
<tr>
<td>39455</td>
<td>2017M03</td>
<td>CDIF</td>
<td>5,350</td>
<td>3</td>
<td>N</td>
</tr>
<tr>
<td>39455</td>
<td>2017M04</td>
<td>CDIF</td>
<td>5,895</td>
<td>8</td>
<td>N</td>
</tr>
<tr>
<td>41141</td>
<td>2017M01</td>
<td>CDIF</td>
<td>2,800</td>
<td>10</td>
<td>Y</td>
</tr>
<tr>
<td>41141</td>
<td>2017M02</td>
<td>CDIF</td>
<td>2,800</td>
<td>8</td>
<td>Y</td>
</tr>
<tr>
<td>41141</td>
<td>2017M03</td>
<td>CDIF</td>
<td>0</td>
<td>5</td>
<td>N</td>
</tr>
<tr>
<td>41141</td>
<td>2017M04</td>
<td>CDIF</td>
<td>3,000</td>
<td>3</td>
<td>N</td>
</tr>
</tbody>
</table>

The number of resident days reported by the facility (orgID 39455) from January – April 2017 is 5,350 – 5,895 resident days. This is within the range for our estimate of 3,000 – 6,000 resident days.
HOW TO ACCESS FACILITY NAME AND ORGID?

**Facility Users Line list.** At this time, the tables listed above only display the facility orgID and it can be difficult to look up the orgID for each facility in a group. The facility-user line list allows users to create a table that includes the facility orgID, CCN, Name and address information.

To access the facility user line list:

1. Open the “Facility level data” folder
2. Select “Line listing-Facility Users”
3. Select “run report”
DEFINITIONS

Users are required to enter data about the total number of “short-stay” and “long-stay” residents when completing the Annual Facility Survey. The respective NHSN definitions may differ from your facility’s billing data. Therefore, we encourage you to adhere to the outlined NHSN definitions, which are described in the Table of Instructions for completing the LTCF Component Annual Facility Survey and below. If you have questions about collecting this information, consider consulting the facility administrator, MDS coordinator, business office, or billing personnel at your facility.

**NHSN definition of total number of long stay residents**

Total number of unique residents who stayed > 100 days in the previous calendar year.

**NHSN definition of total number of short stay residents**

Total number of unique residents who stayed ≤ 100 days in the previous calendar year. NOTE: If a resident starts as short stay but converts to long-stay, then count the resident in the total number of long-stay.