

New York—Reducing Avoidable Hospitalizations Data Resource Guide

Contents

Introduction	3
Data Collection and Storage	4
Identifying Actionable Measures	4
Standardization and Minimizing Subjectivity	5
Collecting and Storing Data	6
Exploring and Analyzing Data	7
Aggregating and Summarizing	7
Common Types of Analysis	7
Reporting and Visualization	8
Types of Visualization	8
Visualization Techniques	9
External Dataset and Quality Improvement Links	10
NY-RAH Excel Reporting Tool	11

Introduction

Data can be defined as any bit of information that is collected and stored. Data collection, analysis, and reporting are familiar terms to most organizations. As technology advanced in the second half of the 20th century, data use increased, and now, it serves as an integral part of many organization's business models. In addition to relying on data analysis to make importance decisions, organizations also use data to demonstrate the value of their services or products to the public and consumers.

The entire data analysis lifecycle can be broken down into stages. This resource guide will focus on three particular stages. First, **data collection** is the process of gathering and measuring information on targeted variables in an established systematic fashion, which then enables one to answer relevant questions and evaluate outcomes.¹ **Data analysis** is inspecting or examining data in order to discover useful information that can lead to valuable conclusions and support decision-making.² **Data reporting** or visualization serves to communicate of these conclusions to a specific audience that will ultimately be making a decision.³ Sometimes, the data analysis can occur immediately after the data collection in order to inform what will be reported. In other scenarios, organizations can create reports immediately after the data collection in order to inform the analysis of the data, and ultimately make specific conclusions, which help inform decisions.

Recently, Federal and state organizations began publicly reporting Skilled Nursing Facility (SNF) performance on a variety of health care measures to inform local consumers about SNF care within their geographic region. The publicly reported data includes CMS Nursing Home

Compare Star Rating and New York State's Nursing Home Quality Initiative, which use data-driven metrics that display skilled nursing facilities' performance on a variety of outcome and process measures. Due to this new data-driven landscape, it is increasingly apparent that SNFs need to become more comfortable with these data concepts. These three stages of the data analysis cycle, in particular, are important because they can help support major decisions that will be made at your organization. However, along with a proper understanding, it is important that each of these stages be executed efficiently and resourcefully in order to decrease any wasted effort or time that may affect your organizational progress.

As a participant in the New York–Reducing Avoidable Hospitalizations (NY–RAH) project, a Registered Nurse Care Coordinator (RNCC) collected a significant amount of resident-level data for your facility. Data collected included hospital transfer, change of condition, and advance directive information. This data was entered into the NY–RAH data portal, which is a web application that serves as a data collection tool for the project. Based on the data that was provided, the NY–RAH Data Team created facility-specific reports to analyze and summarize the information that the RNCCs entered. The measures from these reports will be discussed throughout this guide and serve as examples, but you can and should modify and customize any examples in this guide to best meet your facility needs.

Ultimately, this guide serves as a resource that allows your facility to better understand how to collect, analyze, and act on data that is available to you.

1. "Data collection" A Dictionary of Computing. Encyclopedia.com. 20 Oct. 2016. <http://www.encyclopedia.com>

2. "Data analysis" Cambridge Dictionary. Cambridge.org. 20 Oct. 2016. <http://dictionary.cambridge.org>

3. Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000.

Data Collection and Storage

Identifying Actionable Measures

Prior to analyzing and reporting data, it is important to understand what your facility is trying to measure. The answer to this question can help inform what data your facility should collect. In order to figure this out, some facilities start by categorizing areas of interest in order to help decide what category should be prioritized for data collection.

As previously mentioned, NY-RAH RNCCs collected resident-level data related to advance directives, changes of condition, and hospital transfers. However, each category includes a number of different data elements that were observed and analyzed. When selecting what measures would be most appropriate to analyze and measure, it is important to select measures that can be acted upon after summarizing data results. And in order to act upon a measure, a facility should also understand what processes or procedures impact a facility's performance on that measure. For our purposes, we will examine a process measure from the quarterly NY-RAH Palliative Care Report, the percentage of residents that had a quarterly advance directive discussion. In order to improve this measure, the facility must understand what additional processes must be monitored that affect the frequency of quarterly advance directive discussions. Simply measuring the frequency of the discussions will not necessarily lead to improvement in this measure, but identifying influential processes may lead to improvement. For example, collecting information for residents related to advance directives or family visits may help with identifying important or convenient times to schedule these discussions. In addition, there may be certain staff that must be present to hold the discussions, so understanding staff availability and how that affects discussions is also important and can be acted on. The same can be said for a measure related to an outcome like hospital transfers. In NY-RAH's Facility Progress Reports, each facility's hospital transfer rate was trended over four quarters. All facilities understand the number of processes that can be attributed to im-

proving hospital transfer rates in a facility. So, after understanding the question or problem you are aiming to solve, your facility should also aim to understand the associated processes that should be measured and monitored to inform the solution.

Selecting these associated measures before collecting data allows your facility to recognize what data should be collected in order to return the most accurate assessment. Below, we discuss in more detail some topics to consider when selecting measures for assessment.

Process vs. Outcome

In health care, two types of measures are often analyzed, process and outcome. A process measure assesses steps or activities that are carried out in order to deliver care.⁴ The example described above, **the percentage of residents that had a quarterly advance directive discussion**, serves as a process measure for the NY-RAH project. This process can help facility staff understand the end-of-life wishes of a resident. An outcome measure focuses on the product of a process or system, and tends to identify different or more complex underlying causes.⁵ As previously mentioned, **a facility's hospital transfer rate per 1000 resident days** is an example of an outcome measure. In addition, we previously mentioned how many different factors affect a facility's hospital transfer rate. Both process and outcome measures are useful when analyzing data. When implementing a new process, one would expect it to eventually impact specific outcomes. However, it is usually easier to move a process measure in the right direction in the short-term. It may take longer to move an outcome measure. Additionally, organizations will want to see if a positive change in a process measure has the expected and desired change on an outcome measure.

Data Source

Where will your facility get the data that will be used to calculate this measure? In the example of capturing ad-

4. Centers for Medicare & Medicaid Services (CMS), "Measure Indicator/Development Worksheet," Quality Assurance Performance Improvement, web. 5 Oct. 2016. <https://www.cms.gov/medicare/provider-enrollment-and-certification/qapi/downloads/measindicatdevwksdebedits.pdf>

5. See Note 5.

vance directive discussions, the facility may capture this information in the EMR. Therefore, this information can be quickly pulled and monitored through one system. However, when measuring hospital transfer rates, certain information may be available in a facility's EMR system, but there may be additional information needed that is only available through an Admission, Discharge and Transfer system. Ensuring that the data sources use consistent logic and methodology is important to understand how often the data can be reported. Some additional data sources that may be used for data measurement and collection include the Minimum Data Set and individual chart reviews. Each of these systems may be updated in different ways by different staff, so it is important to know the source of the data because the accuracy, timeliness, and completeness of data may vary among the different data sources. If two different data sources come to different conclusions, one is not necessarily right or wrong. Rather, it's important to understand each source's specifications to see why they may differ. In addition, if your facility needs to set up new processes or applications to collect data, this can take substantial time, which may slow your facility's ability to calculate certain measures.

Data Availability/Frequency

How often is your desired data collected, reported, and updated? Can you easily access your desired data? If a facility's goal is to make a quick decision, but the data that informs a measure is lagged and will not be updated in the near future, then the facility may want to reconsider using the selected measure to make an organizational decision. For example, the NY-RAH Data Team measured Length of Stay on the quarterly Hospital Transfer Reports. However, in order to calculate length of stay, an admission date and a discharge date are required. The length of stay will vary for each resident's transfer, therefore, in order to provide complete length of stay data, there must be some "run-off" time to account for a resident to be discharged. Data availability and update frequency should always be taken into consideration before selecting a measure because out-of-date or inconsistent data may mislead the audience.

Standardizing Data and Measures to Minimize Subjectivity

When choosing measures to calculate, observe, or analyze, a facility should work to standardize this data to the best of its ability. Data standardization utilizes a common format for data that can be expressed in different forms, and by selecting a common format, facilities can make more meaningful comparisons.⁶ In order to make informed decisions, organizations should do their best to compare data that are standardized or standardize the data that is being compared. By comparing standardized data, a facility will have a more accurate representation of the processes or outcomes that are being compared. It is also more likely that if a facility observes a change in the process or outcome data, it is reflective of an actual change in performance, as opposed to a change in how data is input or coded.

Data should be collected in a standardized way in order to reduce the subjectivity of those staff tasked with reviewing and analyzing the data at a later date. As part of the NY-RAH project, RNCCs completed and reviewed specific NY-RAH Quality Improvement (QI) Tools with facility staff in order to take a more detailed look at specific hospital transfers. In order to reduce the subjectivity of the individual completing the QI Tool or the individual reviewing it, the NY-RAH team worked to increase the objectivity of the questions. This ensures that, regardless of the individual who reviews or analyzes the results, the conclusion is clear and consistent.

Providing lists is an example of standardizing data. In the NY-RAH QI Tool, the reviewer is asked what symptoms are associated with the hospital transfer. The NY-RAH Project Management Team could have included a textbox in the QI Tool that would allow an RNCC to include any and all symptoms that accompanied the resident's hospital transfer, such as the example below.




Please list all symptoms that led to the resident being transferred.

6. IBM. "Initiate Glossary- Definition – Standardized Data," Initiate Glossary, web. 11 Nov. 2016. http://www.ibm.com/support/knowledgecenter/SSLVY3_10.0.0/com.ibm.mdshs.initiateglossary.doc/topics/r_glossary_standardized_data.html

Using the question above, you may get a number of answers that vary in length, specificity, spelling, etc. This variation makes it difficult for any analyst to summarize the information and deduce a valuable conclusion.

However, in order to reduce some of the varying factors, the NY-RAH PMT included a list of symptoms on the QI Tool that the RNCCs were able to select from when completing the associated question. This list enables an analyst to compare results regardless of the RNCC completing the QI tool. In addition, RNCCs were guided to categorize symptoms to the best of their ability and try to reduce the number of times the response 'Other' was selected. As with the previous example, when 'Other' is selected, it is difficult to aggregate and summarize specified results because each RNCC uses different language, acronyms, etc. So, we were able to ask the following question with a list of symptoms, which made it easier to analyze and summarize QI Tool results:

 Check the specific change, symptom, or sign that led to the transfer. (Check no more than 3 items)

[List of symptoms]

In addition to standardizing the results, the question limits the number of items that can be selected for a resident. This ensures that the individual completing the QI Tool prioritizes the most significant symptoms in an effort to focus on the precipitating symptom(s).

Using this second example above allows your facility to aggregate all respondents' answers and make a more accurate summarization of the symptoms that are leading to residents' transfers to the hospital.

Collecting and Storing Data

Data can be collected in a multitude of ways. Unless your data is collected and stored through some type of electronic device, you will need some form of manual data collection or data entry necessary to compile data into your database. Whether you collect data through a survey/questionnaire or from a resident assessment following an acute change of condition, the data will be captured manually and stored manually, or entered into some type of automated application that enables you to track that data over time. The table below highlights insights into areas that should be considered before determining what mode or application will be used to collect data.

Data Storage Type	Low Cost	Significant Data Constraint and Logic Capabilities	Significant Reporting Capabilities	Minimal Staff Education and Training	Minimal Technical Maintenance	Significant Long-Term Data Collection & Sustainability Capabilities
Manual/paper system	✓	✗	✗	✓	✓	✗
Data analysis software (i.e. Microsoft Excel)	✓	✗	✓	✗	✓	✗
Data collection software (i.e. Microsoft Access)	✓	✓	✗	✗	✗	✓
Fully-integrated application (i.e. EMR, MDS)	✗	✓	✓	✗	✗	✓

Exploring and Analyzing Data

Aggregating and Summarizing

Data aggregation combines gathered data for the purpose of summarizing the information and deducing general trends or information from the specific data. Aggregated data enables an analyst to combine individually collected data in order to observe specific trends, relative comparisons, and/or performance that would not be possible if the analyst observed each individually collected data point alone. In Phase One of the NY-RAH Project, RNCCs were tasked with summarizing a number of completed QITools each quarter based on an area of focus that was selected by the facility staff from their Hospital Transfer Report. The NY-RAH Data Team would review each suggested area of focus and would make alternate suggestions if the area of focus selected had too few hospital transfers. It is important when aggregating information that you ag-

gregate a large enough number of cases in order to reduce the potential to make generalizations from outliers or a non-representative sample.

Aggregation of data is an essential first step after collecting individual data elements for a specific metric or measure. Aggregating can enhance further data analysis and provide insight into what future analysis should be completed in order to take meaningful action. Analysts also can aggregate data into groups, such as age, staff, etc. By aggregating data into groups, an organization can see how certain groups perform versus others.

Common Types of Analysis

The table below highlights common types of analysis that can be done when observing and investigating data.

Analysis Type	Example	Description																										
Frequency Table	<table border="1"> <thead> <tr> <th>Symptom</th> <th>Frequency</th> </tr> </thead> <tbody> <tr> <td>Abnormal lab</td> <td>6</td> </tr> <tr> <td>Altered mental status</td> <td>8</td> </tr> <tr> <td>Bloody stool</td> <td>9</td> </tr> <tr> <td>Shortness of breath</td> <td>10</td> </tr> </tbody> </table>	Symptom	Frequency	Abnormal lab	6	Altered mental status	8	Bloody stool	9	Shortness of breath	10	Returns a numeric total for how often each value occurs in the data set. This can be repeated for each variable or category of data that is collected to determine how often the variable or category of data occurred during the specified time period or assessment.																
Symptom	Frequency																											
Abnormal lab	6																											
Altered mental status	8																											
Bloody stool	9																											
Shortness of breath	10																											
Descriptive Statistics	<table border="1"> <thead> <tr> <th></th> <th>n</th> <th>Mean</th> <th>Median</th> <th>Min</th> <th>Max</th> </tr> </thead> <tbody> <tr> <td>Length of Stay</td> <td>15</td> <td>5</td> <td>4</td> <td>1</td> <td>8</td> </tr> </tbody> </table>		n	Mean	Median	Min	Max	Length of Stay	15	5	4	1	8	Calculates basic statistical data: population size (n), minimum value, maximum value, average (mean and median), and in some cases the standard deviation for your variable or data set. This will provide a summarization of the data in a meaningful way. It can also serve as comparison information against the individual data.														
	n	Mean	Median	Min	Max																							
Length of Stay	15	5	4	1	8																							
Cross Tabulation Analysis	<table border="1"> <thead> <tr> <th colspan="5">Gender ^ Hospital Transfer Cross-Tabulation</th> </tr> <tr> <th colspan="2" rowspan="2"></th> <th colspan="2">Hospital Transfer</th> <th rowspan="2">Total</th> </tr> <tr> <th>Yes</th> <th>No</th> </tr> </thead> <tbody> <tr> <th rowspan="2">Gender</th> <th>Female</th> <td>45</td> <td>65</td> <td>110</td> </tr> <tr> <th>Male</th> <td>42</td> <td>62</td> <td>104</td> </tr> <tr> <th colspan="2">Total</th> <td>87</td> <td>127</td> <td>214</td> </tr> </tbody> </table>	Gender ^ Hospital Transfer Cross-Tabulation							Hospital Transfer		Total	Yes	No	Gender	Female	45	65	110	Male	42	62	104	Total		87	127	214	Compares two variables in order to understand whether the variables are similar or are there differences between the variables or groups being analyzed. For example, if you collect data on hospital transfers, you may also want to understand if a resident's gender plays a role in whether they are transferred to the hospital or not. A cross-tabs analysis helps determine whether the groups are similar as they relate to hospital transfers, or if there are differences between the groups.
Gender ^ Hospital Transfer Cross-Tabulation																												
		Hospital Transfer		Total																								
		Yes	No																									
Gender	Female	45	65	110																								
	Male	42	62	104																								
Total		87	127	214																								

Reporting and Visualization

Importance of Visualization

For the purpose of this section, the terms, reports and reporting will refer to any table or visualization (e.g., chart, graph) that is used to express results and share conclusions of a data analysis. Using visualization provides the audience with a quick and intuitive way to interpret analyzed data. This is especially important if the audience is unfamiliar with the data that is being presented. Good

visualization techniques can make it simple for users to understand desired conclusions and to further inform decision making.

There are various types of visualization techniques and specific-use cases that each should be used for. Below are examples of basic visualizations and common uses for each.

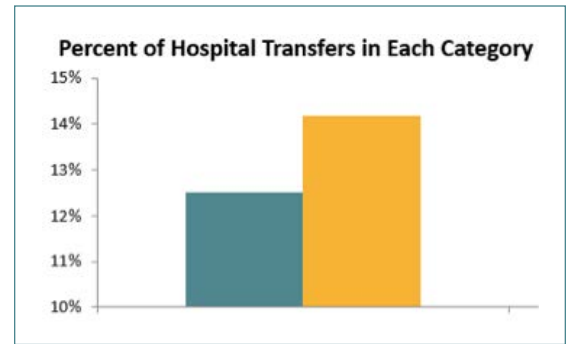
Graph Type	Visual Example	Use Case Example
Line Chart	 <p>The chart shows two data series: 'Your facility' (teal line with circles) and 'NY-RAH' (orange line with circles). The y-axis represents the transfer rate per 1000 resident days, ranging from 0.00 to 3.00. The x-axis shows quarters from Q4 2015 to Q3 2016**. 'Your facility' starts at approximately 1.0 in Q4 2015, drops to 0.5 in Q1 2016, rises to 1.2 in Q2 2016, and ends at 0.8 in Q3 2016**. 'NY-RAH' starts at 2.2 in Q4 2015, rises slightly to 2.4 in Q1 2016, drops to 2.2 in Q2 2016, and ends at 2.3 in Q3 2016**.</p>	<p>In the NY-RAH Facility Progress Reports, each facility's hospital transfer rate and the overall NY-RAH hospital transfer rates are displayed over a 12-month period in a line chart. For these measures, the data team chose to present them as a line chart because it is important to see how a facility's and the project's performance on these measures changes over time. A line chart provides a visual for audiences to quickly see whether the rate declined, stayed the same, or rose over the year. You can also include data for the same measure for a previous time period as a comparison.</p>
Histogram/Bar Chart	 <p>The bar chart compares 'Your Facility' (teal bars) and 'NY-RAH' (orange bars) on two metrics. The y-axis shows percentages from 0% to 100%. For '% of ACOCs that result in transfer', 'Your Facility' is at 49% and 'NY-RAH' is at 10%. For '% of ACOCs with any NY-RAH Tool Used*', 'Your Facility' is at 91% and 'NY-RAH' is at 55%.</p>	<p>In NY-RAH's monthly ACOC Reports, each facility's performance on ACOC communication tool usage is compared with the overall project's performance for the same period in a bar chart. Using a bar chart makes for easy interpretation of comparisons for the specific time period, but are not as effective for expressing trends over long periods of time.</p>
Scatterplot	 <p>The scatterplot shows the distribution of time between noticing an ACOC and hospital transfer. The x-axis is labeled 'Number of hours between noticing ACOC and hospital transfer' and ranges from 0.0 to 24.0. The y-axis represents the number of transfers. The data points are densely packed between 0 and 8 hours, indicating that most transfers occur within this time frame.</p>	<p>In a presentation made on hospital transfers, the NY-RAH team focused on how often residents were transferred after facility staff first noticed an ACOC. In order to express the distribution and variation of the time between the ACOC first being noticed and the hospital transfer, the data team created a scatterplot. Scatterplots make it easy to express potential relationships between variables or groups to your audience.</p>

Visualization Techniques

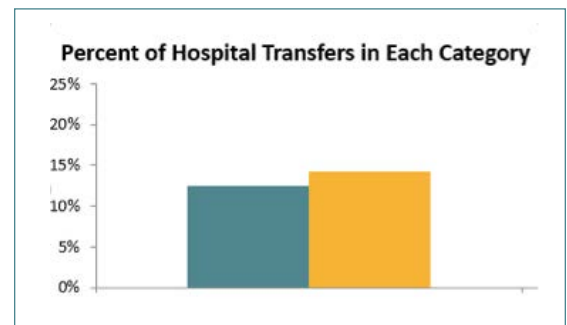
After deciding the best visualization type for your data, there are some other tips that are helpful when preparing your visualization for your audience. Check out the tips below for useful information when designing your charts.

Designing Tips

- Color selection
 - In general, many color-vision deficient individuals have trouble distinguishing a difference between **red** and **green**.⁷ Therefore, minimize the use of these colors in your reporting and visualizations.
 - Contrast can be very useful when trying to distinguish between different elements within a visualization. In addition, color-vision deficient individuals can determine contrast.
 - Similar to contrast, using different shades of the same color (monochromatic) is a useful way to present data. It allows individuals to easily recognize a variation between two items and eliminates any potential color impairment issues.
 - For more ideas on color schemes and style, please visit: <http://colorbrewer2.org>
- Scaling
 - Always be aware of the scales in your charts and graphs. Overstating and understating scales can greatly change how your chart or graph is interpreted by your audience. In the example below, the scale or y-axis range is 10%-15%, thus making the 1% (13% vs. 14%) difference between these two groups appear more drastic than it actually is.



- In the next example, we use the same data as the previous example, but we make the scale or y-axis range 0%-25%. Now, you can still see the difference between the two groups, but it greatly reduces the chances that someone may overreact to the 1% difference because it is less drastic visually when the scale is proper.



- Specify Direction
 - If you are presenting data where measure performance can be indicated by the data's direction, indicate the desired direction on the chart with a statement, for example, "Higher is better." This is an easy way for your audience to understand whether a chart or visualization represents poor or good performance.

7. National Eye Institute (NEI), "Facts about color blindness," National Institute of Health, web. 5 Oct. 2016. https://nei.nih.gov/health/color_blindness/facts_about

External Dataset and Useful Data Resource Links

Quality Assurance Performance Improvement (QAPI) Documents

- QAPI Process Tool Framework: <https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/QAPI/Downloads/ProcessToolFramework.pdf>

Nursing Home Compare Quality Measures

- Nursing Home Compare Dataset Downloads:
<https://data.medicare.gov/data/nursing-home-compare>
- Quality Improvement Tools for Nursing Home Compare Quality Measures:
<https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/QAPI/Downloads/Resource-Guide-QI-Tools.pdf>

NY-RAH Excel Reporting Tool

Below, please find a preview of NY-RAH's Excel Reporting Tool. Your facility can use this tool to generate monthly and quarterly hospital transfer rate reports. Using manual data entry, a user can enter resident day counts and hospital transfer counts, and the Excel Reporting Tool will document and create a line chart visualizing the data that

has been entered. Please feel free to work within the tool, while also expanding and improving upon it as you please.



[Download the Excel Reporting Tool here.](#)

DIRECTIONS: Please add appropriate information to highlighted cells and the accompanying rates will calculate automatically.

Total Resident Days												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Resident Days												
1000 Resident Days												
Hospital Transfers By Month												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Hospital Transfers												
Inpatient Admission												
ED Only/Observation Stay												
Hospital Transfers Per 1000 Resident Days												
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Hospital Transfers												
Inpatient Admission												
ED Only/Observation												

Monthly Hospital Transfers Per 1000 Resident Days

Monthly Inpatient Admissions and ED Only Visits/Observation Stays Per 1000 Resident Days