Step 6.10 – Change and Measure

A Measurement Guide for Long Term Care

Introduction

Stratis Health, in partnership with the Minnesota Department of Health, is pleased to present "A Measurement Guide for Long Term Care: How to Use Measurement for Root Cause Analysis" as part of the Root Cause Analysis Toolkit for Long Term Care. This document will assist nursing homes and other long term care settings in defining measures that will support and demonstrate the outcomes they have defined. The Root Cause Analysis toolkit is located on the Stratis Health website at http://www.stratishealth.org/providers/nursinghomes.html.

For more information about the use of RCA and how to conduct an RCA, contact Kristi Wergin, kwergin@stratishealth.org or Kathie Nichols, knichols@stratishealth.org.

Purpose of Measurement

Measurement for Quality Improvement

Measurement is an essential component in helping an organization determine:

- Did implemented interventions or actions lead to an improvement?
- Were the interventions implemented as expected?
- Did the implemented changes/actions lead to the expected outcomes?
- Have the changes been sustained and are they embedded into staff practice as expected?
- Have the changes resulted in improvement in care and services over time?
- How do the outcomes compare to and benchmark against performance at state and national levels?

Measurement data can also be used to inform individuals, family members, staff, board members and other stakeholders of the progress and success of organizational safety and quality improvement initiatives. Without data, organizations cannot know whether they are making progress toward the goals of increasing individual satisfaction and improving the quality of care and services provided in nursing homes. Measurement used for quality improvement does not need to be as complex or rigorous as methods used in a research study. Large samples for measurement and complex analyses are not necessary. Data collection should not be so complex, or the amount of data collected so large, that it impedes improvement efforts. Measures should be developed that will show the success or failure of changes implemented.

Root Cause Analysis, Action Plans and Measurement

Root cause analysis (RCA) is a problem solving method or process for conducting an investigation into an incident, failure, actual or potential problem or concern. RCA can be helpful to understand the factors that led to a serious event such as a fall with injury or harm resulting; reoccurring, unexpected or undesirable outcomes such as a reoccurring complaint, or development of pressure ulcers.

Using the RCA toolkit, once the team identifies the root cause(s) and contributing factors, an action plan is developed to alter the systems or processes identified as being at the root cause and/or contributing to the event. The action plan outlines the interventions to be taken to improve the systems, processes, or structural issues that are related to the root cause. An important element of the action plan is the measurement plan which monitors the impact of the interventions.

A measurement plan should evaluate whether the action plan was 1) implemented as intended, and 2) resulted in the intended changes in practice to the system, or a process of care. A measurement plan should not be limited to measuring the completion of the actions only. For example, the measurement plan should measure that the new process is occurring, not simply that staff have been trained on the new process or that the new process has been communicated.

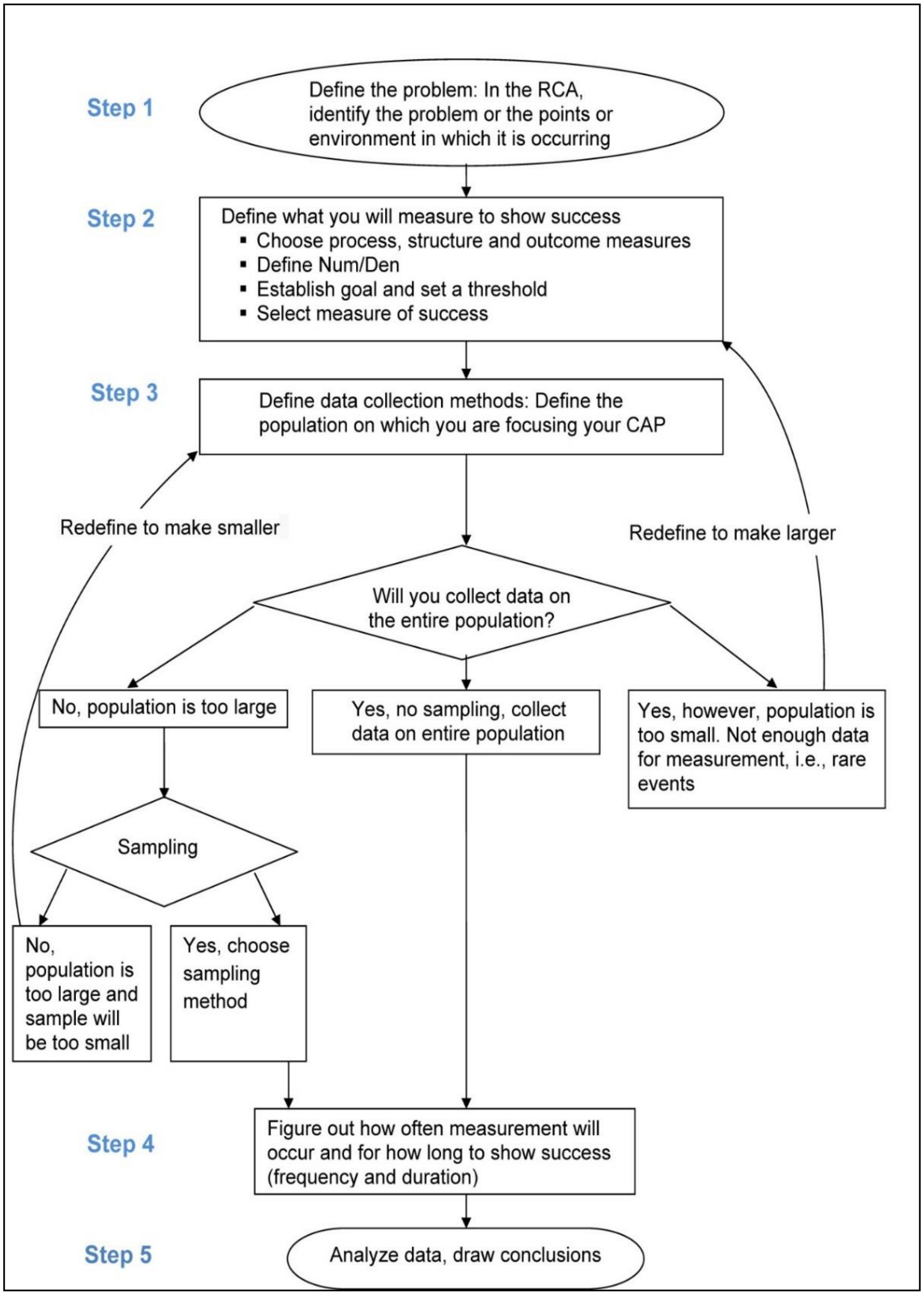
Ultimately, measurement plays a key role in quality improvement initiatives as well as advancing individual safety. Measurement findings are used to identify best practices and knowledge. Many are shared across the state and nationally to benchmark quality efforts.

Steps for Creating Measures

This section outlines the five steps required to create measures. (See Figure 1)

1. Define the root cause and identify the desired changes
2. Define what to measure to show success
    a. Determine type of measures to use - structural, process and/or outcome
    b. Define the numerator and denominator
    c. Define the baseline
    d. Establish a goal
    e. Set a threshold
3. Determine data collection methods
    a. Define population
    b. Determine sampling methodology and size
4. Define frequency and duration of measurement
5. Draw conclusions

Figure 1: Creating Measures



Step 1. Define the problem and identify the desired changes

The root cause is identified and defined in the RCA process. The action plan is created based on root cause finding and lays out specific changes to be made in the processes that are expected to prevent another similar problem from occurring.

Example

Event: An individual missed a scheduled therapy appointment.

Root Cause: The RCA team determined the therapy appointment was missed due to the nursing assistant not having access to an accurate and current therapy schedule.

Action plan: The actions are aimed at increasing the communication between the therapy department and the nursing assistants.The team develops a communication process that includes having accurate, current schedules and nursing staff and individuals attending therapy having access to them.

Step 2. Define what to measure to show success

Types of Measures

Three types of measures are relevant to quality improvement work: structural, process, and outcome measures. In root cause analysis, root causes and contributing factors of a problem or issue are identified. An action plan is developed to address the root causes and contributing factors, including a strategy to make changes in the organization which will prevent the problem from happening again. Depending on the nature of the problem, these actions can be a physical change to the environment or can be focused on changes to a process or system.

To demonstrate success, the organization must collect and monitor data over time to determine whether the actions proposed for the environment (structural measures) or the process or system (process measures) were implemented as expected, and whether they had the intended effect (outcome measures).

Structural measures

Structural measures are related to changes in the physical aspects of the environment or equipment. A need to monitor permanent structural changes, such as changing a type of door hardware, is not needed. However, evaluate whether the change is providing the intended outcome. Certain structural changes warrant periodic spot checks. For example, if properly fitting footrests are provided on all wheelchairs, an audit to assure they are in place can be done periodically.

Examples of Structural Changes

- Changing the color of all shower curtains to a color that contrasts with the wall color to prevent falls
- Changing the type of door to permit easier individual access
- Changing light bulb wattage to increase visibility and decrease individual harm.
- Changing tube connections so misconnection cannot occur

Process measures

Process measures provide data about a system or process. Process measures are used to indicate whether a change has been embedded into practice and has been sustained as expected. For example, when the process measure relates to ensuring the therapy appointment schedule is available to everyone, the process change would be monitored to assure therapy appointments were not being missed and the practice continued over time. Sources of data for process measures can be floor checks, observational audits, chart reviews, surveys, etc.

Examples of Process Measures

- Documentation following a completion of an incontinence risk assessment
- Completion of an incident report following a fall
- Consistent use of a tool for hand-off communication

Outcome measures

An outcome is an indicator of individual health status or change in health status that can be attributed to the care being provided. In RCA, outcomes are the events or conditions that the actions are intended to affect or change. Outcome measures provide information on whether the implemented action plans achieved the intended goal: Is individual care safer? Has better care is been provided? Have further problems been avoided?

Examples of Outcome Measures

- Fall rates
- Staff turnover rates
- Pressure ulcer rates

Sources of outcome measures can be data that is monitored as part of an organization's quality or safety program, including incident reports, chart reviews, satisfaction surveys, MDS data, quality measure/quality indicator reports, etc. Monitoring outcomes over time can show the impact of action plans on achieving broader goals related to individual quality of care and quality of life.

Guiding principles for determining the type of measurement indicated

Ideally, every action plan has a structural or process measure as well as an outcome measure. See Table 1. Process measure data collected and monitored over time identifies if the change has been sustained. Used alone, a process measure will not describe the impact the action plan had on preventing another incident. Using both process and outcome measures allows an organization to analyze whether the change has occurred and to know whether it has changed the system and will prevent recurring incidents. Using only one type of measure gives only part of the story; the lack of a recurrence of the event (outcome measure) could be coincidental and not attributable to the process change.

Table 1: How and when to use measures for action plans

Measure	When used	Companion measures	Example
Structural measure	The action plan calls for the removal or replacement of equipment or physical change to the environment	Outcome measure	**Structural measure** Toilet seats in individual bathrooms will be changed from white to black. **Outcome measure** Number of falls occurring in the bathroom
Process measure	The action plan calls for a system/process change	Outcome measure	**Process measure** Hourly rounding by nursing assistants will occur as expected **Outcome measure** Fall rate
Outcome measure	Extremely rare process where occurrence is difficult to predict; a way to monitor if a process or structural change has had the desired impact	Structural or process measure	**Process measure** Skin inspection is conducted consistently under and around a specific arm brace that is rarely used **Outcome measure** Pressure ulcer rates on individuals with this type of arm brace

Define the numerator and denominator

Once the problem and the interventions to make changes are identified, measures to monitor the progress of the action plan must be created. Effective measures will demonstrate if the change in the structure or process has occurred, and if the changes made are having an effect on improving the outcome. A measurement should be defined for each action identified in the action plan.

At least one measure should be created for each process or structural change made to show whether or not the changes have been implemented and sustained. One outcome measure should be created to show that the changes are having the desired effect.

Process measures are usually calculated by counting the number of cases or number of times a process occurs (numerator) and dividing it by the number of cases in which the problem or process could have occurred (denominator). The calculated rate is usually expressed as a percentage. For example: a numerator of 15 and a denominator of 30 (15/30) is expressed as 50%. **Outcome measures** are calculated in a similar fashion but instead count the number of times the problem or outcome occurs (numerator) and divide by the number of times the event could have occurred (denominator).

Both the numerator and denominator should be carefully defined to include only the cases to be counted in the numerator and those cases with the opportunity for the problem to occur in the denominator. Whatever is expected to be measured must be very clear – is it all medication errors, or just medication errors involving anticoagulants? Choose the numerator/denominator accordingly. Other methods are available to calculate outcome measures such as falls and pressure ulcer rates per patient days.

Example of a Measure

Measure = Numerator/Denominator X 100 = Rate

Measure: Percentage of individuals that have been scheduled for therapy and missed an appointment within the past month

Denominator: Number of therapy appointments that were scheduled in the past month. (200)

Numerator: Number of therapy appointments that were scheduled and missed in the past month. (25)

Calculated Rate: 25/200 X 100 = 12.5%

Result: 12.5% missed their therapy appointments in January.

Establish a goal

A goal is a level of expected compliance or outcome with a planned action and usually is expressed as a percentage. If compliance is critical to preventing another incident from occurring, the goal may be set at 100% compliance. However, in most cases, expecting 100% compliance over time or for most processes is unrealistic – errors may occur even when working within a stable system with well implemented processes.

Lack of compliance may be justified and appropriate in certain instances if it does not occur frequently and if there is a strong rationale behind the lack of compliance. For example, the skin safety policy calls for a weekly full skin inspection to identify any areas for potential breakdown, but an individual becomes critically ill and cannot tolerate repositioning to allow full skin inspection.

A goal should be identified for each measure created for the action plan. Goals should be written in the **SMART** format; *specific, measureable, attainable, realistic, and timely:*

Specific: A specific goal clearly defines what staff members are going to do and what they want to happen. A straightforward, specific goal is more likely to be met than a general goal. To help create a specific goal, answer the "W" questions (Who, What, When, Where, Why, How) using the example below:

Who: Individuals who meet the criteria of being assessed as high risk for falls and being selected as part of the sample.

What: The number of individuals in the sample with hourly rounding by nursing assistants.

When: The next six months starting (date), monitored monthly, individuals will be monitored during each shift.

Where: Identified individuals on neighborhood X.

Why: To assure individuals identified at high risk for falls have hourly rounding consistently by nursing assistants.

How: Individuals in the sample identified to be at high risk for falls will be observed by the neighborhood manager for hourly rounding by nursing assistants. For those individuals where hourly rounding is indicated, documentation will be audited to assure hourly rounding is documented in the plan of care.

Measurable: A goal should be measurable. Establish concrete criteria for measuring success and monitoring progress toward each goal set. When staff measures their progress, they stay on track. Visualizing success helps to continue putting in the effort required to reach the goal.

Attainable: Make sure the goal is attainable. Do not set the goal higher than can be attained in the allotted time frame.

Realistic: To be realistic, a goal must be something staff is both *willing* and *able* to work toward.

Timely: Set a timeframe for the goal, e.g., next week, within three months, by a certain date. Set an end point for the goal to be achieved to provide a clear target to work toward.

Example of a Goal

To confirm that hourly rounding by nursing assistants is being used appropriately for individuals who meet the criteria, a sample population of those identified as being at high risk for falls on Neighborhood X will be observed once each month for the next six months. The goal: 95% of all sample populations will have hourly rounding by nursing assistants implemented when indicated.

Set a Threshold

When a goal is the level of expected compliance with a planned action, a threshold is the minimum acceptable level of performance for that planned action – the level below which the planned action has not been adopted as expected. Falling below the threshold is an indicator or early warning sign that identifies problems that need immediate attention.

If the measure falls below the threshold, additional action is needed to increase compliance (e.g., a better process, or a change to the process), or analysis is needed to determine why the process has not been sustained or embedded. Consistently falling below a threshold indicates that a process change has not been embedded and sustained as expected, and that continuing with the same approach is unlikely to be effective.

Like a goal, a threshold usually is expressed as a percentage or rate. If the process change is thought to be a critical component within the system related to the event – meaning its failure is highly likely to result in another event – the threshold may be the same as the goal. For example, failure of hourly rounding by nursing assistants is highly likely to lead to another high risk fall. In this case, a high threshold should be set. In contrast, failure to document daily skin inspections as part of the safe skin procedures in a limited number of instances may be less likely to lead to another pressure ulcer developing by itself. In this case, the threshold could be set lower. Though both processes are important and should be done consistently, the first example may leave less room for error and may be more likely to result in another event if not completed every time. Therefore, the threshold for the first example may be set high and be the same as the goal.

In some instances, the threshold for a particular change may be set below 90%. For example, if a new, complex process is being introduced, moving the threshold up over time may be appropriate, such as setting the threshold at 70% in three months, and 90% in six months. However, in general, setting a threshold below 90% should only be done in rare circumstances with a specific purpose and rationale to support it.

One threshold should be applied to each measure created for the action plan.

Example of a Threshold

Goal: 95% of individuals identified as being at high risk for falls will have hourly rounding by nursing assistants.

Threshold: Neighborhood X will achieve a minimum of 95% of high risk individuals observed with hourly rounding.

Step 3. Determine data collection methods

This section will provide information on the key components to data collection: population, sampling, frequency, and duration. The goal of measurement is to be able to evaluate the processes that are in place and determine if changes made to those processes were successful. Measurement for quality improvement is not research; data collection should not be so rigorous that it impedes activities. However, data collection does need to be sufficiently rigorous to demonstrate that the intervention worked.

Population

In the context of measurement, population refers to the group of individuals impacted by the problem and its action. The population can be broad or narrow depending on the outcome and on the action or change being implemented. See Figures 2 and 3 below. Defining a population establishes parameters that clarify which individuals or processes should be included in the measurement. A population should be defined for each measure in the action plan and should only include individuals or processes that could have the outcome or root cause occur, or that are eligible to receive the process or structure change proposed in the action plan. The populations for the process measure and the outcome measures may not be the same, but large differences should be avoided. See Table 2. The data for measurement (the numerator and denominator) will be drawn from the population; so the population must always correspond with the action plan. Defining the population is important because it will help clarify what processes or individuals should be included in or excluded from the data collection.

Figure 2: Population for Action Plan Scenario A1

PROBLEM

- Population for the data collection is all individuals in the organization
- Population is too broad
- Cannot detect changes in population
- Cumbersome measurement and data collection
- Recommend focusing the population targeted for the data collection

Figure 3: Population for Action Plan Scenario A2

PROBLEM

- Population for the data collection is a small subset of individuals in the organization.
- Population is too narrow (e.g., rare events).
- Problematic if population targeted for the data collection is too small, resulting in not enough data for measurement.
- Recommend changing definition of population or expanding the population targeted for data collection

Population for process measures. A population for a process measure consists of the processes or group of individuals that are targeted in the action plan to receive an intervention or process change. The population for the process measure may be the same as the population for the outcome measure, a subset of the outcome measure, or a completely different population. See Table 2 for examples.

Population for outcome measures. A population for an outcome measure consists of the individuals for whom the outcome could occur. The outcome population can be determined broadly (e.g., every admission into the organization in a given year) or it can be narrowed to a specific population (e.g., admissions on one neighborhood, every person having a certain diagnosis, individuals at high risk for falls). Outcomes that occur in the population are counted, such as the number of falls or incidence of pressure ulcers. See Table 2 for examples.

Table 2: Population Examples for Outcome and Process Measures

	Process measure with identified population	**Outcome measure with identified population**	**Summary of population selection**
Root Cause Analysis found assessments for fall risk were not completed on admission. This pattern was noted on the neighborhood X. The action plan is aimed at increasing the consistency of completing a fall risk assessment on admission.	Process population: all individuals admitted to the neighborhood X Process measure: risk assessment completed upon admission for individuals admitted to the neighborhood X	Outcome population: all individuals admitted to the neighborhood X Outcome measure: fall rate for individuals admitted to the neighborhood X	Population is the same for outcome and process measure When the outcome and process population are the same, the risk for misinterpretation of the data is less likely
Root Cause Analysis found hourly rounding was not used as expected because the individual was near the nurses' station, which staff felt was intervention enough to deter the individual from getting up without help. Action plan is aimed at increasing the use of hourly rounding for all high-risk individuals by creating staff reminders and prompts, and thereby reducing the fall rate.	Process population: all individuals identified to be at high risk for falls Process measure: hourly rounding in place for high risk individuals	Outcome population: <u>all</u> individuals admitted to the organization in one year Outcome measure: fall rate for individuals admitted to the organization in one year	Population for process measure is a subset of the population for the outcome measure. One limitation of using a broad outcome with a more focused process measure: improvements made to the process that would affect the outcome will not be apparent (a broad outcome rate will dilute any effect on the specific population). Consider focusing the population targeted for the action plan.
Root Cause Analysis found fall risk assessments were not consistently conducted on neighborhood Y. Action plan is aimed at increasing the use of admission assessments to identify high-risk individuals and thereby reduce the fall rate. The action plan is rolled out to all staff on all neighborhoods. .	Process population: all individuals admitted to the organization Process measure: risk assessment completed upon admission.	Outcome population: all individuals admitted to neighborhood Y over the next six months Outcome measure: fall rate for individuals admitted to the neighborhood Y in the next six months	Population for outcome measure is a subset of the population for the process measure. The outcome measure is specific to the neighborhood with the problem identified, but the process is rolled out to all individuals in the organization. One limitation: when the process measure is broad and the outcome is specific, it will be difficult to determine if the process measure was adopted by the population with the problem. Recommend keeping a broad process measure to monitor if the process has been adopted organization wide and creating an additional measure to monitor the specific neighborhood with the problem.
Root Cause Analysis found a lack of clarity about the ability and expectation of staff to do a skin inspection after removing a certain brace that is rarely used.	Process population: individuals with the particular brace - used	Outcome population: individuals with the particular brace that is used	Populations for outcome and process measures are very small (processes that occur rarely). Expand both populations proportionately to increase sample sizes for measurement, but highly

	Process measure with identified population	Outcome measure with identified population	Summary of population selection
The action plan is aimed at developing a clear policy to address skin inspection for individuals with this particular brace, but also will expand the population to assure clarity for the range of all braces or devices used.	infrequently (rare event) Expand the population to individuals with any device or brace Process measure: skin inspections completed for individuals with any device or brace	infrequently (rare event) Expand the population to individuals with any device or brace Outcome measure: pressure ulcer rate for individuals with any type of brace	recommend monitoring the process and outcome for every case that occurs or monitoring every individual with the device or brace

Sampling

Often, it is not possible to measure every instance (the whole population) in which a process is supposed to occur or every individual that could have the outcome. If the population to be measured is large, collecting data for every individual is not feasible. In these cases, sampling can be used to reduce the data collection burden. When data are collected on a sample or subset of individuals, measures are calculated only for the sample. Any conclusions based on that sample are then applied to the remainder of the population. Because data assumptions are made when calculating measures from a sample, it is very important that this subset is an accurate representation of the population. One consequence of not including an accurate sample of the population in the action plan can be incorrectly concluding that a process has changed when the process has not actually changed. This incorrect conclusion may result in future reoccurrences.

The following can help assure the sample better represents the population:

- Appropriate sampling methodologies (e.g., random sampling or stratified sampling) and unbiased data collection (e.g., if a process occurs on all shifts, the sampling should include data from all shifts)
- Adequate sample sizes. The larger the sample size, the more likely the sample will accurately reflect the entire population; however, smaller sample sizes can be used as long as good data collection and sampling techniques are used.

Several proven methods for selecting samples help assure a reliable sample. When determining which sampling method will be most appropriate to use, consider the characteristics of the population, such as:

- Specific diagnosis
- Condition
- Procedure
- When the process being measured occurs
- When the teams being observed work

Table 3: Sampling Methodology Examples

Sample method	When used	Pros and cons of sample method	Examples
Random sampling Involves creating a list of the entire population from which the sample will be drawn, selecting a set number of cases randomly from that list, and collecting data on those cases	Typically used for rigorous research—when the stakes of the outcome are high	Pro: Most reliable method of sampling. Eliminates unintentional tendency to choose cases that are thought to be "typical" or "representative" of the population. Without a random sample, the cases are not necessarily a true representation of the population. Cases may have been selected because they happened to look particularly good or bad. Con: Can be difficult to create a complete population list. This method lends itself to retrospective data collection (such as chart reviews) and is not a good method with real-time or concurrent data collection (such as collecting data from assessments or the MDS).	Randomly select 30 charts from a list of all individuals admitted to the organization in the last week to verify if fall risks assessments have been conducted.
Stratified sampling Involves identifying subgroups (strata) of interest and collecting data from a random sample of cases within each group	When multiple factors (i.e., time of day, sex, race, type of treatments) need to be included in the sample	Pro: Helpful for evaluating if the process change has occurred and when and where the process is performed. Note: Cases should be selected randomly within each subgroup applicable to the population. Con: Can be time consuming to identify and select from each subgroup.	Randomly select 10 individuals from each shift (day, afternoon, and night) to observe whether fall prevention measures are in place (total of 30 individuals observed).
Systematic sampling Selects cases according to a simple, systematic rule, such as all persons whose names begin with specified letters, are born on certain dates (excluding year), or are located at specified points on a master list (every nth individual)	When the population is unknown and for cases or processes that occur infrequently	Pro: Possible to perform systematic sampling concurrently. The sample can be selected at the same time the list of individuals in the population is being compiled. This feature makes systematic sampling the most widely used of all sampling procedures. Con: Prone to bias depending on how the sample is collected and/or sorted.	Select every third individual from the list of individuals on the neighborhood to observe whether fall prevention assessments were completed (total of 30 individuals observed).
Convenience sampling Allows for the use of any available cases	When resources are limited and it is not possible to use random sampling. When validity of data is not an important factor (e.g., pilot testing)	Pro: Convenient—simple, easy design (a computer or a statistician is not required to randomly select the sample). Con: Since the sample is not random, the cases selected may not be typical of the population targeted for improvement.	Select all individuals on the neighborhood on the last day of the month to observe whether fall prevention measures are in place for those with high fall risk

Sample method	When used	Pros and cons of sample method	Examples
Quota Sample Involves selecting cases until the desired sample size is reached. Usually involves cases selected to assure data are collected for those with certain characteristics	When population size is unknown or when it is not possible to predict how many cases will occur in a given timeframe (e.g., certain surgeries performed or falls). Data collected until the desired number of cases has been reached	Pro: Ease of sample selection from a large population. Data collection can stop before the desired sample size is reached if the data indicate that the goal will not be met. Data collection stops, the problem is solved, or the process is changed and data collection is resumed Con: A judgment is made about the characteristics of the sample to be included with the hope that it will be as representative as possible of the population being targeted for improvement. Not a random sample so it has the same disadvantage as convenience sampling—risk of biased data. Prone to bias from selecting only a small window of time (e.g., collecting cases as they occur may result in only a sample of cases that occurred Monday morning vs. a sample of cases from the entire week, including the weekend). May use other sampling techniques with this method to reduce bias (e.g., add systematic sampling or systematic selection of cases, selecting every nth case).	Select 30 individuals as they are admitted to observe if fall prevention measures are in place for those with defined fall risks Or: Select 15 high-risk individuals and 15 low - risk individuals as they are admitted to observe whether fall prevention measures are in place.

The next step is to determine how large the sample should be. As in the case of selecting an appropriate sampling method, determining sample sizes involves tradeoffs between validity and practicality.

When the population targeted by the action plan is large, often it is not feasible to collect data on the entire population. Sampling reduces the amount of data to be collected by providing an estimation of what is occurring in the population. For example, records are reviewed for the entire population and a rate is calculated. The rate is 100/600=16.67%. However, it is likely not feasible to collect data from this many records for multiple measurements. So sampling is used to produce an estimate of the rate. A sample of records is chosen from the population, reviewed, and a rate is calculated. The rate for the sample is 5/30=16.67%. In this example, the sample produced a rate that is exactly the same as the rate calculated for the population. The sample provided a good estimate of what is actually occurring in the population.

However, this is not always the case. For example, a sample is drawn from this population five more times. Each time a sample is drawn, different records are selected by chance. The rate that is calculated for each will vary from sample to sample, referred to as sampling variability. The rates calculated will range, for example, from 5% to 30%.

The smaller the sample or the less data collected (e.g., fewer than 30 cases), the more variability in the rates calculated (larger range between each rate calculated). The larger the sample or the more data collected, the less sampling variability will occur (smaller range between rates). Larger sample sizes increase the likelihood that the rate calculated is accurate.

Note: When collecting data on the entire population, there is no estimation. The measurement includes all individuals or records so there is no variability in the data due to sampling. So collecting data for the entire population is ideal because it is the most accurate method; however, again, it is often not feasible.

Statistical methods are available to quantify how much variability exists in the data and measurement. But taking frequent measurements over time is a simpler method for understanding the variability that occurs. Monitoring frequent measurements over time can allow an organization to see the range of rates and can point out what is normal for its organization. Changes in the range and noticeable patterns can be reviewed to determine the reasons.

The example below shows data collected for skin inspections. In Figure 4, three measurements from a sample of 30 records were taken in April, May, and June. It appears as if the number of skin inspections has increased dramatically over time. But if this measurement were expanded to include more data points over a longer period of time, the organization would see that the data collected in these three months just shows variability in the data. See Figure 5.

Figure 4: Skin Inspection Rates for Organization A for Three Months

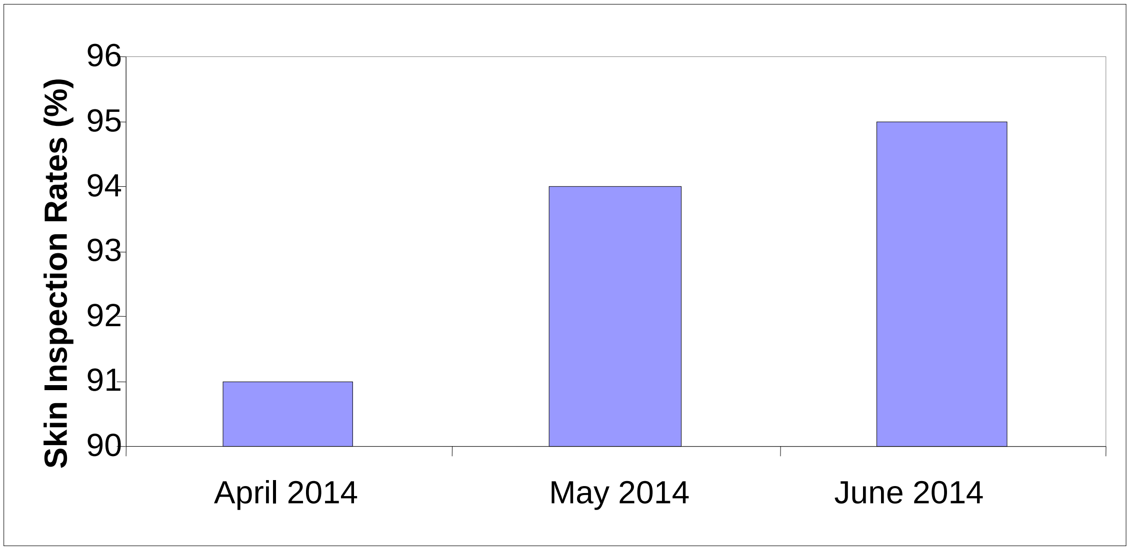
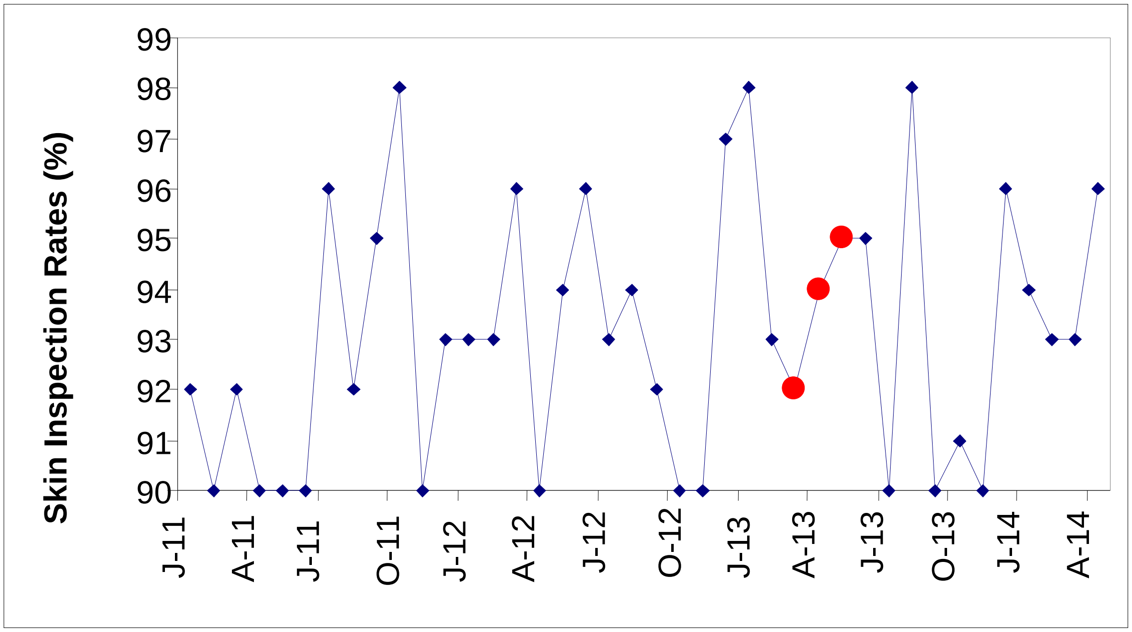


Figure 5: Skin Inspection Rates for Organization A by Month



A large sample size means more data will have to be collected, but more data can be helpful because there will be less variation, which increases the ability to draw good conclusions. However, many times large sample sizes are not practical or feasible. In those cases, smaller samples with frequent measurements can be used as a way to obtain a representative sample of the intended population. When small samples are used, frequent measurement will help illustrate variation in the data, which will increase the accuracy of interpretation. The size of a sample should be driven by the size of the population during the time frame of interest. See Table 6 for guidance in determining sample size.

Table 4: Determining Sample Size

Population size in the allotted data collection time frame	Recommended sample size
30 or fewer	Data should be collected on every case that occurs. Consider whether to broaden the population size or extend the time frame for the measurement to determine whether the action plan was successful. Results based on fewer than 10 cases are deemed “questionable,” and therefore difficult to show the effect of the change and whether it has been sustained and embedded as expected.
Greater than 30	In cases where the population is greater than 30, a sample can be drawn. Sample size calculations are used by statisticians to determine an adequate percentage of the total number of cases in the population that should be observed. In general, a sample of 30 or more observations or audits will have less variability, so the calculated measures will be more valid and conclusions about the success of the process change will be more accurate.

Small samples due to rare events. Because adverse events are usually rare, it may take a long time to collect enough data to draw conclusions about the effectiveness of the process changes through the use of outcome measures. To address this situation, pair the outcome measure with one or more process measures. For rare events, organizations can use alternative methodologies. See Table 7.

Table 5: Alternative Methodologies for Measuring Very Rare Events or Outcomes

Methodology	When to use	Example
Time between events is calculated and monitored	Changes that occur between events indicate how well the action plan or changes to the process is working. If the time between events increases (the event is occurring less frequently), the process change may be working. If the time between events decreases (event is occurring more frequently), the process change may not be working or there may be other root causes that led to the event recurring. Root cause analysis would be required to confirm what led to the event recurring.	The number of successful uses of a specific brace before pressure ulcers develop.
Combine data for similar cases or events	Particularly useful if the system or process found to be a root cause could result in a variety of adverse/undesired events. Some processes actually contribute to, or prevent multiple adverse/undesired events. For example, hourly rounding is conducted to prevent a variety of undesired events (e.g., falls, incontinence, complaints of pain, pressure ulcers). Combining data for all falls in this example will increase sample sizes.	In the case of falls that occur with a serious injury, the organization may consider combining all types of falls and monitoring whether hourly rounding is taking place as expected, rather than looking only at the falls with serious injury.

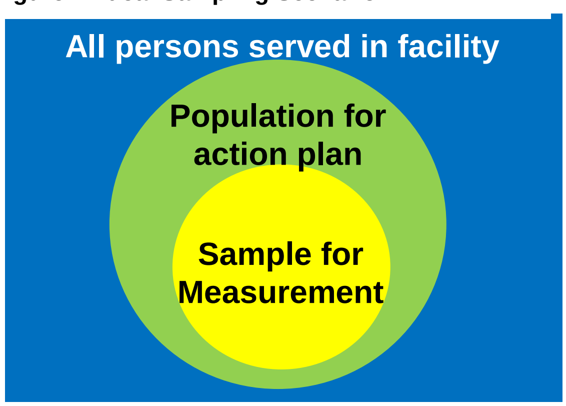
See Figures 6, 7, 8, and 9 below for illustrated ideal sampling and sampling pitfalls scenarios.

Figure 6: Ideal Sampling Scenario

IDEAL

- Population for testing the interventions is a selected number of individuals from the organization (not all individuals).
- Measurement is on the entire population targeted for the interventions (sample = entire population).
- Collecting data on the entire population for an intervention is a valid measurement.

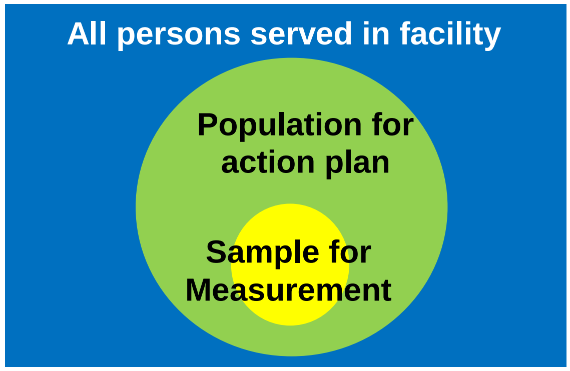
Figure 7: Ideal Sampling Scenario



IDEAL

- Population for the action plan is a selected number of individuals from the facility.
- Measurement is on a subset of the population targeted for the action plan (sample).
- Collecting data on a sample from the entire population for the action plan is a valid measurement if good sampling techniques are used.

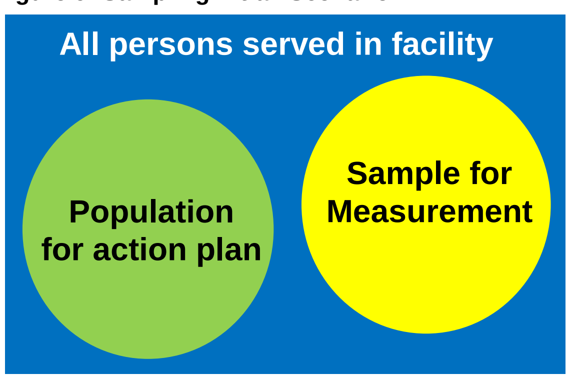
Figure 8: Sampling Pitfall Scenario



SAMPLING PITFALLS TO AVOID

- If it becomes evident when determining the sample size that the population targeted for the action plan is too large in relation to the desired sample size, the measurement may not be accurate.
- Recommend evaluating if the definition of the population targeted for the action plan is appropriate, and refining if necessary. Or additional data collection will be necessary to ensure accuracy of the measurement.
- Conversely, if the population targeted for the action plan is adequate, but the sample size proposed is too small in relation to the population, measurement may not be accurate.
- Recommend increasing the sample size, or conducting additional data collection of the smaller sample size over a longer period of time.

Figure 9: Sampling Pitfall Scenario



SAMPLING PITFALLS TO AVOID

If the sample selected is patients or records that did not receive the action plan intervention, the measurement will not be accurate. Recommend reviewing sampling methodology to include only patients or records that received the action plan intervention.

Step 4. Determine frequency and duration of measurement

Frequency: Frequency refers to how often data are collected for a measure, such as daily, weekly, monthly, quarterly, or annually.

Duration: Duration refers to the timeframe over which the data will be collected, such as the total number of weeks, months, or quarters.

Frequency and duration go hand in hand and are used together to monitor changes in the process and improvements in outcomes. Determining the appropriate frequency and duration for data collection depends on the size of the population being measured, the frequency with which the process or event occurs, and the characteristics of the population.

Size of the Population Being Measured

- If the size of the population (number of cases) is small, sampling may not be necessary or feasible. All records or cases will be audited for measurement. As a result, frequent measurement cannot occur, and duration for data collection will likely be longer because it will need to continue until enough data is collected.
- If the size of the population is too large to collect data on all cases, sampling should be conducted. Data collection will be less frequent to allow for an adequate sample size to be gathered (e.g., quarterly or annually).
- When the population is large, it is possible to collect all necessary data in a short period of time (e.g., in one day). However, collecting the data in a short period of time should be avoided. Smaller, more frequent measurement should occur (e.g., weekly, monthly, or over a period of several months).

Frequency with Which the Process or Problem Occurs

If the process to be measured occurs frequently, measurement should occur frequently (weekly or monthly) because the potential exists to miss capturing the true characteristics of the population and draw incorrect conclusions from the data.

Characteristics of the Population

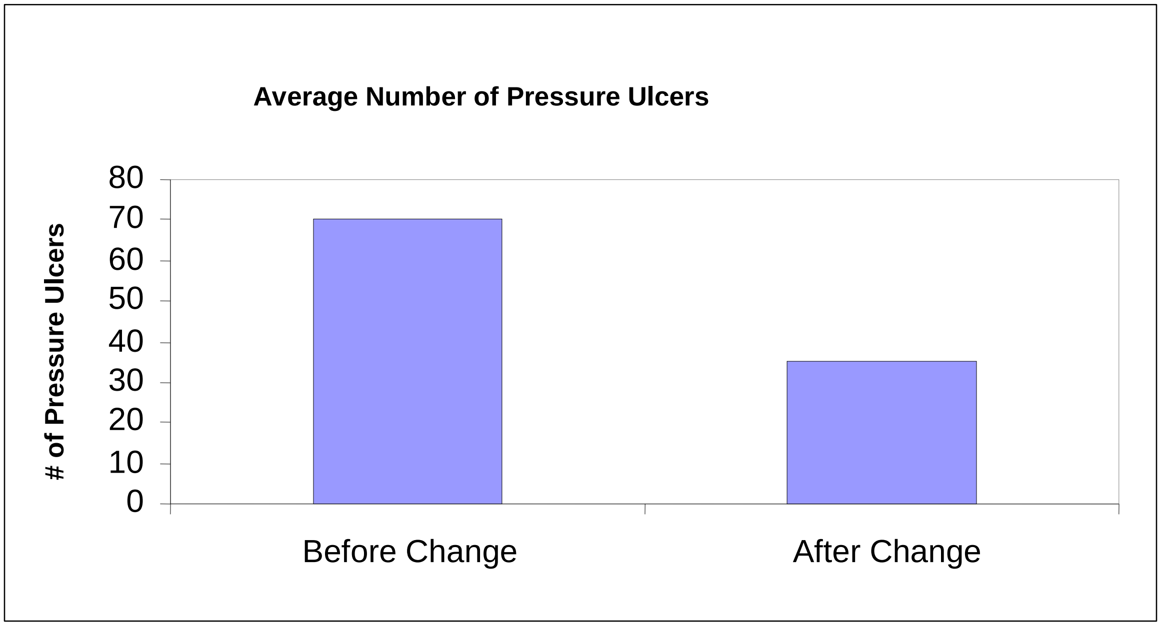
If the population being measured has seasonal considerations, such as procedures that are more common at certain times of the year, e.g., flu vaccines this must be taken into consideration for determining duration. In this case, the duration should cover a full year to determine if process change happens consistently throughout the year.

Frequency and duration are used to determine if a change is sustained over time. No clear formula exists for determining the appropriate frequency or duration for data collection because it is dependent on the sample size and characteristics of the population being measured. Smaller, more frequent data collection over a longer period of time is preferable to less frequent data collection. Smaller, more frequent measurement helps illustrate variability in the data and will improve the accuracy of the inferences drawn from the data.

Making a change to a core process or system can be a challenge to maintain over time. As more time passes after any training or intentional communication about the process change, practice can drift or slide back to old habits—“the way we have always done it.” Building a plan that allows an adequate length of time for data collection to monitor whether the changes stick, as well as to determine if the changes had a positive effect, is key. Experienced patient safety and quality improvement experts refer to the early period when staff can maintain the change in practice more easily as the honeymoon period. This period is often up to three months following implementation of a practice change. Continuing to collect data four to six months after implementation is the more accurate test of whether changes have been maintained and may allow a better assessment of whether the changes resulted in an improvement.

Multiple measurements over time help show if the intervention has resulted in the process being adopted and sustained. Collecting data for three months may only show if the process was adopted, but will not show if it was sustained over time. For example, the next set of figures shows the average number of pressure ulcers before and after a change was made (e.g., educational session on admission risk assessments). Figure 10 shows data collected before and after the change was made. A conclusion can be drawn that the number of pressure ulcers decreased after the educational session.

Figure 10: Pressure Ulcer Rates Before and After the Intervention



Collecting data for a longer period of time before and after the change was made provides additional information about what is happening, and can dramatically affect the conclusions made. It is helpful to understand the variability of the data used for measurement before and after a change is made to help determine the impact of the intervention. Figures 11 and 12 show the same data illustrated in Figure 10. But instead of summarizing the data in two data points (before and after the change), additional data points were collected and plotted over time. The conclusions drawn from these figures are very different. The decreasing trend in Figure 11 shows that pressure ulcers were decreasing before and after the educational session. Therefore, the educational session was not the only factor that contributed to the decrease in pressure ulcers. Figure 12 shows that the data are fluctuating, but at a higher rate before the educational session, then decreasing after the session.

This example illustrates the differences between collecting large amounts of data with less variability and more accurate measurement, and using smaller, more frequent measurements over time that show variation but can be very beneficial at detecting trends.

Figure 11: Pressure Ulcer Rates by Month Before and After the Intervention

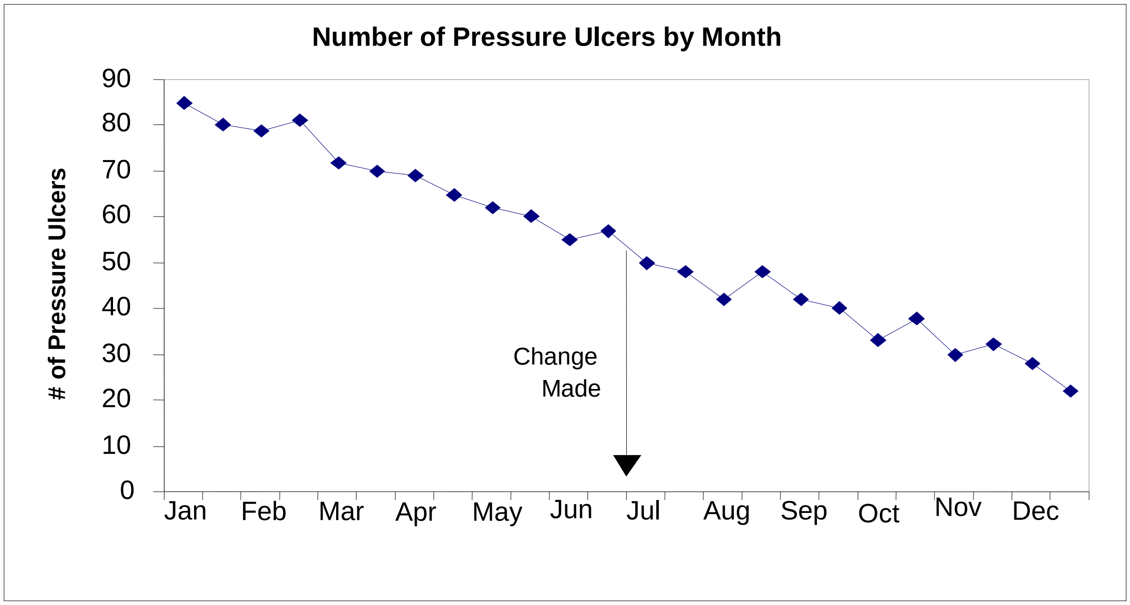
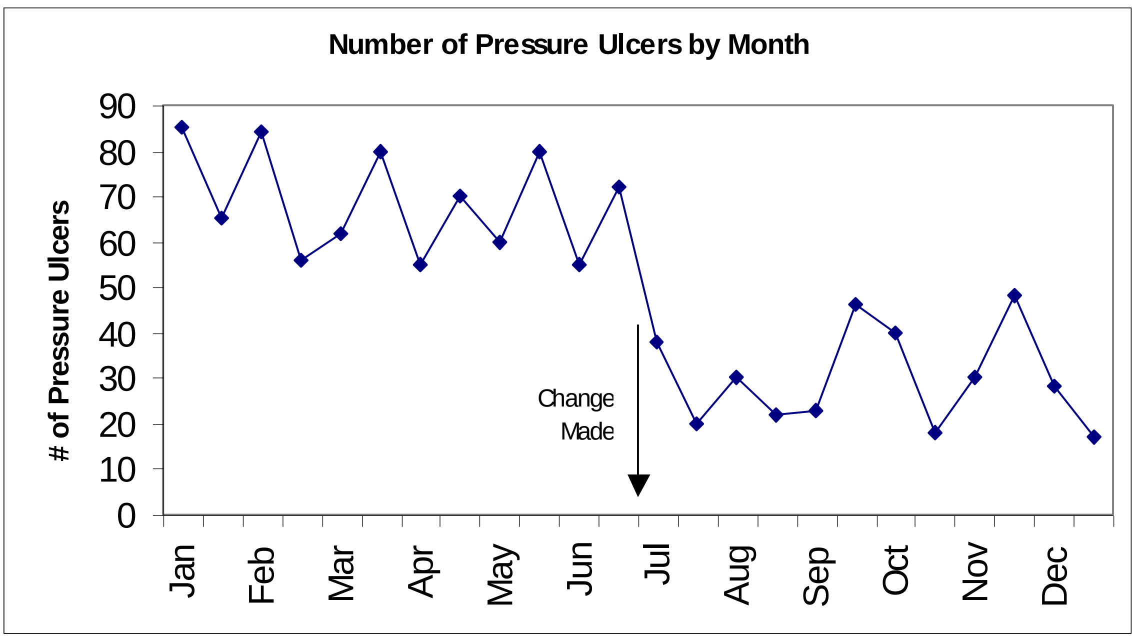


Figure 12: Pressure Ulcer Rates by Month Before and After the Intervention



Other Data Collection Considerations

In some cases, stopping data collection and measurement before the desired sample size is collected is reasonable if a failure trend is discovered. A failure trend is identified when it is mathematically impossible to meet the threshold. Data collection should only be stopped due to a failure trend if the sample has been defined and properly identified. Measurement cannot stop when measuring cases as they occur or when the population size is not known and data are being collected on the entire population.

Example of when measurement can stop due to failure trend

For a measurement of completed skin assessments, the sample size is set at 10 records each month for three months with a threshold of 80%. Three records are audited and none have documentation of a skin assessment. In this example, it is reasonable to stop auditing and investigate why there is no documentation. It is not worth collecting data from seven more records because the trend shows the threshold will not be met.

Example of when measurement should not stop

For a measurement of all completed physical therapy sessions as they occur during the month of March, a failure trend cannot be accurately calculated until the last physical therapy session is completed in March—when the whole population is known.

Step 5. Drawing Conclusions

If measurement data were collected before and after the action plan was implemented, the conclusion of whether the interventions led to a change is strengthened. Collecting data prior to the implementation of the strategies or interventions is ideal—both as a way of ensuring that the process change is addressing a real problem or gap and as a way of developing a baseline against which to measure progress. However, this is not always possible. If a formal baseline was not obtained, but the following criteria are met, you can reasonably conclude that the root cause analysis accurately identified the root cause and the linked interventions to make changes were appropriate.

In the absence of baseline data, all of the following criteria must be met to conclude successful action plan implementation:

- Data for the process measure were monitored over time
- Goal was attained (process and outcome)
- You are confident that the change is permanent
- Event is not repeated

If the problem occurs again, the newly launched root cause analysis should include collecting data on that process or structural measure again to verify whether the process change was sustained. If it has been sustained, there is another cause. If the process change was not sustained, that is a reasonable place to start the root cause analysis.

The following is an example that illustrates the use of measurement in the Root cause analysis action plan process:

Event: An individual identified as high risk for falling fell, resulting in a broken hip.

RCA: The team determined hourly rounding was not used because the individual was near the nurses' station and staff felt that would be enough to deter the individual from getting up without help.

Action plan: The action plan is aimed at increasing the use of hourly rounding for high-risk patients on this specific neighborhood by creating staff reminders and prompts. The team developed an awareness campaign to include fall prevention posters in rooms and added documentation prompts for hourly rounding in the medical record.

Process measure: Hourly rounding will be used in 95% of all individual identified as high risk for falls. The estimated population size is 20 individuals per month.

20 records every month (100% sample size due to the small population) will be audited for 6 months. The goal of 95% of audited records will show hourly rounding use with a threshold of 90%.

Outcome measure: The organization will measure falls rates every quarter with a goal of zero falls.

Analysis: At six months, the organization has collected enough data and is ready to analyze the results. Below are three possible scenarios their data may show.

Table 6: Analysis Scenarios

Scenario	Process measure	Outcome measure	Analysis
1	**Goal met:** 95% of audited records showed hourly rounding	**No change or goal not met:** Falls did not decrease	Increasing hourly rounding did not result in a decreased number of falls. The root cause was not correctly identified—or other factors contributed to the fall rate. Further analysis is needed. Perhaps the risk assessments were not conducted correctly, the assessment tool did not adequately capture all risks for falls, or other issues prevented staff from hourly rounding to prevent the fall.
2	**Goal not met:** Audited records show that hourly rounding is not being used consistently	**Goal met:** Falls decreased	Fewer falls occurred, but implementing hourly rounding on all high-risk patients has not been sustained or embedded into practice as expected. The root cause was not correctly identified, or other factors need to be addressed. Review of the event and systems involved is needed.
3	**Goal not met:** Audited records show that hourly rounding is not being used	**No change or goal not met:** Falls did not decrease	Because the action has not been implemented as expected, it is not possible to conclude if the true root cause is identified. Further investigation is needed to understand why the change was not sustained.

Scenario	Process measure	Outcome measure	Analysis
4	**Goal met:** 95% of audited records showed hourly rounding in use	**Goal met:** Falls decreased	The use of hourly rounding is embedded into practice and no falls have occurred since the change was implemented. The root cause was correctly identified and the action was successful in reducing falls.
5	One goal not met One goal met **Goal not met:** Hourly rounding not documented consistently on care plan **Goal met**: Observation audit showed hourly rounding in use	**Goal met:** Falls decreased	FOR CASES WITH MORE THAN ONE PROCESS MEASURE: No falls have occurred. One of the changes is embedded. The other change is not embedded and may not be as critical to preventing a fall as the change with the goal met.
6	One goal not met One goal met **Goal not met:** Need for hourly rounding not documented consistently on care plan **Goal met**: Observation audit showed hourly rounding in use	**No change or goal not met:** Falls did not decrease	FOR CASES WITH MORE THAN ONE PROCESS MEASURE: Falls are still occurring. One change is embedded and one change is not. The change embedded is not critical to preventing falls or the root cause has not been correctly identified.

By monitoring the process measure and the outcome measure over time, conclusions can be drawn as to whether a change is successful in preventing future adverse events. Continuing measurement over time can actually save time and money by alerting an organization when processes are not being followed and can potentially prevent future events.

Case Studies

The following case studies are examples of Root cause analysis measurement plans that could be strengthened.

Case study 1: Misalignment between action and measure

Event: An individual falls during a transfer using a new mechanical lift.

RCA: The team determined the individual fell because the support straps were not properly placed due to staff lack of knowledge of how to use the new lift.

Action plan: The action plan is aimed at holding mandatory training sessions for all staff on the how to use the new lift.

Process measure: 100% of staff will attend a training session.

Outcome measure: Zero individual falls when using the new lifts. .

Comments/analysis: The process measure is set up to calculate attendance at the training sessions. While it is important that staff attend, it is the information provided at the training session that is expected to change the practice and is the key intervention—the training session is the method of sharing the information.

A stronger process measure would be observational audits of the lift being used by staff. . In situations where the process change happens infrequently or cannot be predicted when it will occur, observational audits may not be feasible. It is possible to evaluate the effectiveness of the education through a demonstration of knowledge. This can be accomplished with use of a posttest or return demonstration of the expected process steps. In this example all staff has to do a return demonstration of how to transfer with the lift during training.

Case study 2: Missing measurement components

Event: Stage 2 pressure ulcer

RCA: The team determined the correct wheelchair cushion was not used due to unclear policy.

Action plan: The action is aimed at 1) implementing a clarified policy, and 2) implementing a process to evaluate all policies for effectiveness and clarity for staff (not examined here).

Process measure: Random chart audits of at-risk patients for implementation of policy. Expect 90% compliance.

Outcome measure: After goal met, monitor implementation of policy with a random audit of five records every quarter.

Comments/analysis: The process measure would benefit from the addition of several elements. In addition to the method and goal, a complete measurement would include sample size, frequency, and duration. In addition to or an alternative to documentation audit is observational audits for the use of the appropriate wheelchair cushions. A measurement strategy that includes all the elements is needed to assure good data collection and the ability to draw inferences and conclusions; for example, weekly random audits of 25 charts for one quarter of at-risk patients on Units Y and Z for use of the indicated interventions according to the policy. The audit will begin on (date) and be conducted by the wound prevention team. Expect 90% compliance with a threshold of 90%. In addition, the proposed outcome measure measures a process not an outcome. An appropriate outcome measure could be monitoring the pressure ulcer rate over time.

Case Study 3: Unclear measurement over time

Event: Hypoglycemia.

RCA: The team determined the error occurred due to lack of staff communication during shift change report.

Action plan: The action is aimed at using a structured way to communicate information to staff across shifts.

Process measure:	Process measure: The new nurse to nurse documentation for shift report will be audited 10 times each month for use of the structured communication process for eight months. Expect 90% compliance with a threshold of 90%.
Outcome measure:	Ongoing.
Comments/analysis:	The measurement over time information would be strengthened if it were more specific and defined. For a stronger outcome measure, after the goal is met, monitor the number of hypoglycemic events.

Conclusion

"We can't manage what we don't measure."
Brent James, MD

This guide is intended to be a resource for root cause analysis action plans or quality improvement efforts with the need for a robust measurement plan. Solid measurement is an essential component of quality improvement and patient safety work. It helps answer the question: How will we know that a change is an improvement?

Making our health care system safer is enormous, important work. We hope this guide can be a useful resource for your organization's patient safety and quality improvement efforts.

Appendix A: Resources

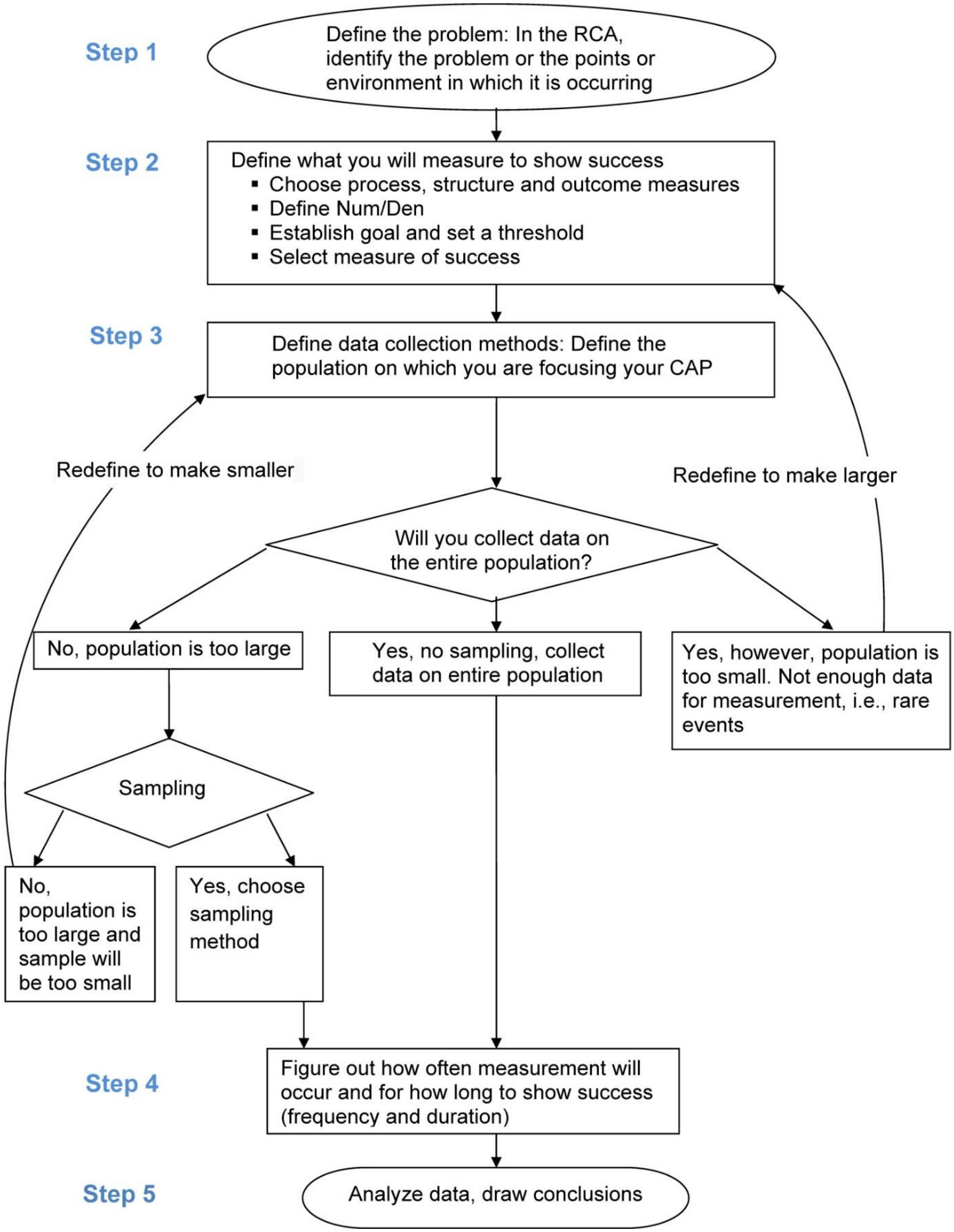
Minnesota Hospital Association has information on Call to Action programs, patient safety news, and updates.
http://www.mnhospitals.org/index/patient1

Stat Trek: Teach Yourself Statistics. This site provides a statistics tutorial to help solve common statistical problems.
http://stattrek.com/AP-Statistics-2/Data-Collection-Methods.aspx?Tutorial=AP

Stratis Health has a series of recorded Webinars on the basics of quality improvement. These sessions allow provider organizations to hone a specific quality improvement skill set, orient new staff, or offer in-service workshops for teams. Visit the Stratis Health website at www.stratishealth.org for recorded trainings and webinars.

Appendix B: Steps for Creating Measures

Figure 13: Steps to Create Measures



Produced under contract with the Minnesota Department of Health

Produced with the use of Federal Nursing Home Civil Money Penalty Funds