Minnesota Adverse Health Events Measurement Guide

Prepared for the Minnesota Department of Health
Revised December 2, 2015

Stratis Health is a nonprofit organization that leads collaboration and innovation in health care quality and safety, and serves as a trusted expert in facilitating improvement for people and communities.

Bloomington, Minnesota ● 952-854-3306 ● www.stratishealth.org

Contents

Introduction

Stratis Health, with the Minnesota Department of Health (MDH), is pleased to present the Minnesota Adverse Health Events Measurement Guide. The guide provides instruction on the components required for adverse event measurement plans submitted to the Patient Safety Registry, including examples of commonly missing elements, and clarification on confusing topics regarding measurement.

Under contract with MDH, Stratis Health reviews all root cause analyses and corrective action plans—including measurement plans—that are submitted under Minnesota's Adverse Health Events Reporting Law, and provides technical assistance to Patient Safety Registry users. Through this work, Stratis Health has detected common areas of confusion and missing elements in measurement plans. With its expertise and knowledge of events in the Patient Safety Registry and with the skills of its analytic and epidemiology staff, Stratis Health has created a practical guide based on sound analytic theory and relevant to adverse event reporting requirements.

The guide's primary intent is to serve as a how-to measurement guide for those new to the Minnesota Adverse Health Events Reporting Law and its reporting requirements. It is intended as a tool for use by root cause analysis and corrective action teams that are struggling with questions related to measuring the success of their interventions as well as a resource for events and situations that fall outside the 29 reportable events required to be reported under Minnesota's Adverse Health Events Reporting Law. The guide also can serve as a resource for more experienced users and for other patient safety or quality improvement efforts that require a robust measurement plan. For information on entering data into the Patient Safety Registry, see Resources listed in Appendix A.

Solid measurement is an essential component of quality improvement work. At a minimum, quality improvement measurement allows organizations to know if an intervention has been implemented as expected and if that intervention resulted in the intended improvement. Measurement data can be used to inform staff, administration, and board members of the progress and success of patient safety and quality improvement initiatives, and to illustrate improvement needs.

Without data, organizations cannot know whether they are making progress toward the goal of making the health care delivery system safer. Stratis Health and MDH intend for this guide to be a resource for your organization's patient safety and quality improvement efforts.

Purpose of Measurement

Measurement for quality improvement

Measurement is essential in helping an organization make the case for quality improvement efforts, communicate with staff, and gain staff buy-in for process changes and quality initiatives. Measurement is used to determine if a change has been sustained and embedded into staff practice as expected and if the change has resulted in improvement in care over time. It provides a reference point to compare and benchmark an organization's performance at state and national levels.

> Measurement is used to determine if a change has been sustained and embedded into staff practice as expected and if the change has resulted in improvement in care over time.

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 for this type of measurement. 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. Smaller numbers can be used with a well-developed measure.

> Data collection should not be so complex or the amount of data collected so large that it impedes improvement efforts.

Adverse events and measurement

Minnesota state law requires hospitals, ambulatory surgical centers, and community health hospitals to report 29 specific adverse events into the Patient Safety Registry. Root cause analysis (RCA) is the standardized method that all reporting organizations use to help identify one or more human factors or systematic causes that led to an adverse health event (AHE).

Once the root causes and/or contributing factors are identified, a corrective action plan (CAP) is developed to address the systems or processes identified as being at the root cause and/or contributing to the event. The CAP outlines the actions to be taken to improve the systems, processes, or structural issues that are related to the root cause. An important element of the CAP is the measurement plan which monitors the impact of the actions taken.

A measurement plan should evaluate whether the CAP was 1) implemented as intended, and 2) resulted in the intended changes in practice, in the system, or in 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 rolled out.

> A measurement plan should evaluate whether the CAP was implemented as intended and resulted in the intended changes in practice, the system, or a process of care.

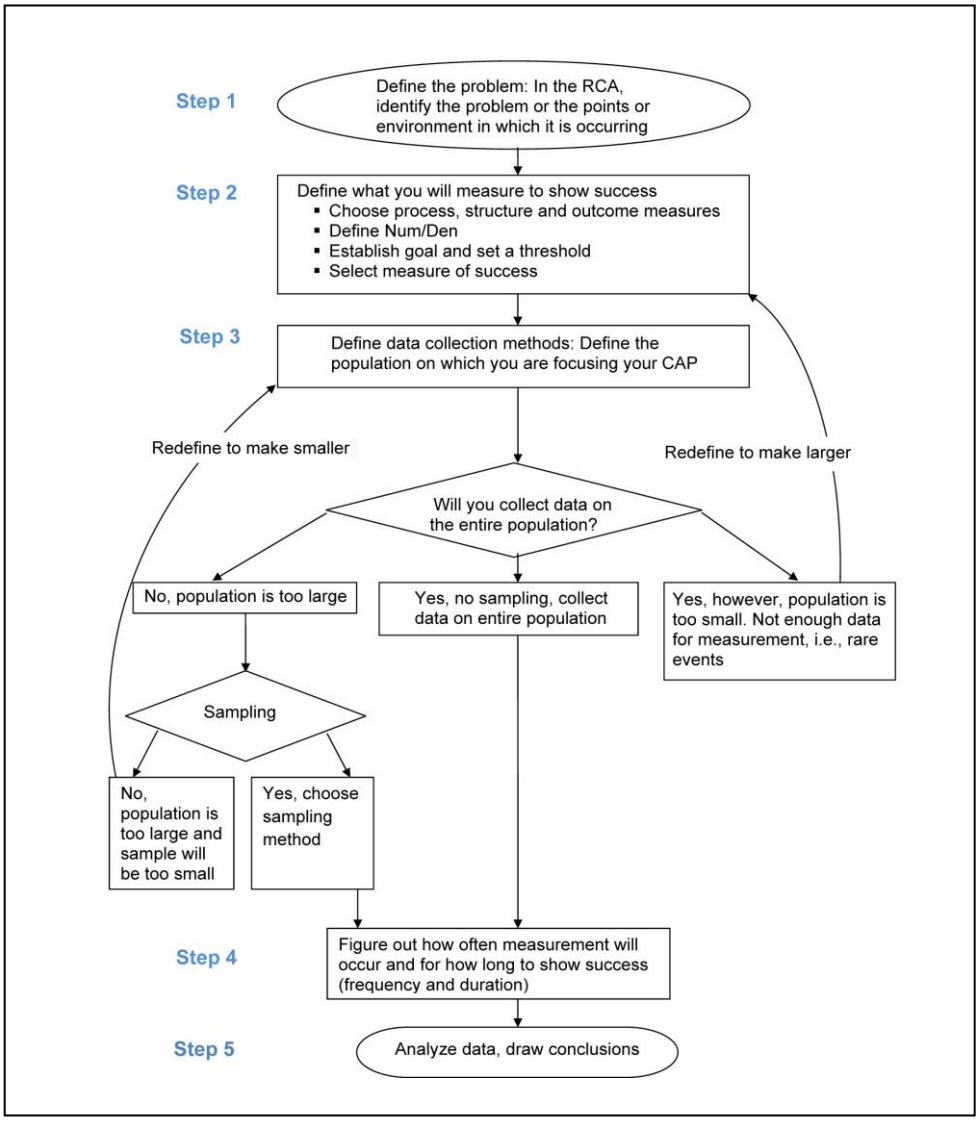
Ultimately, measurement plays a key role in advancing safety as part of the Minnesota Adverse Health Events Reporting Law. Measurement findings are used to identify best practices and knowledge, and are shared across the state to help prevent adverse events and make health care delivery in Minnesota safer.

Steps for Creating Measures

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

1. Define the problem and identify the desired changes
2. Define what to measure to show success
    a. Determine type of measures to use (structural, process and outcome)
    b. Define the numerator and denominator
    c. Establish a goal
    d. Set a threshold
    e. Select a measure of success
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 Flowchart



Step 1. Define the problem and identify the desired changes

The suspected cause or causes of an AHE are identified and defined in the RCA process. The CAP is created based on root cause findings, links directly to the root cause findings and lays out specific changes to be made in the processes that are expected to prevent another similar AHE from occurring.

Example

Event. A patient fell, resulting in a broken hip. The patient had previously been identified as high risk for falling.

RCA. The RCA team determined the "within–arms-reach" policy was not followed as expected because the patient requested privacy while using the bathroom.

CAP. The CAP is aimed at creating a script to help staff explain to patients the reason for staying with-in-arms reach. According to the CAP, the team develops an awareness campaign that provides scripting to all nursing staff.

Step 2. Define what to measure to show success

Types of Measures

Three types of measures are relevant to AHE work: structural, process, and outcome measures. In the RCA process, root causes and contributing factors of an AHE are identified. A corrective action plan is developed to address the root causes and contributing factors to the AHE, including a strategy to make changes in the facility which will prevent the event from happening again. Depending on the nature of the event, these actions can be a physical change to the environment or can be focused on a process or system.

To demonstrate success, the facility must collect and monitor data over time to determine whether the corrective 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, may not be apparent. Evaluate whether the change is providing the safeguard intended. Certain structural changes warrant periodic spot checks. For example, if the type of dressings used on a surgical set up is changed to allow only tailed sponges, periodic monitoring is recommended to confirm that other types of sponges do not return to the surgical set up trays.

Structural measures are related to changes in the physical aspects of the environment or equipment.

Examples of Structural Changes

- Equipment that malfunctioned removed from use and removed from reorder/purchasing procedure
- Changing the type of door hardware to prevent patient self-harm
- Adding windows to increase the ability of staff members to observe patients
- A hard stop in the EHR which will force the ordering practioner to specify discontinue date on certain medications

Process measures

Process measures provide information about a system or process.

Process measures provide information 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, in the case where the process measure relates to staff staying "within arm's reach" when indicated, the process change would be monitored to assure staff are staying within the reach of the patient when indicated and that the practice continued over time. Sources of data for process measures can be observational audits and patient surveys.

Examples of Process Measures

- Frequency of OR debriefings which include accounting for all specimens
- Frequency of surgical sites correctly marked
- Consistent use of a tool for hand-off communication

Outcome measures

An outcome is an indicator of health status or change in health status that can be attributed to the care being provided.

An outcome is an indicator of health status or change in health status that can be attributed to the care being provided. In the case of adverse health events, outcomes may be the events or conditions that the corrective actions are intended to affect or change. Outcome measures provide information on whether the corrective actions implemented achieved the intended goal: care is safer and further adverse health events are avoided.

Examples of Outcome Measures

- Number of lost specimens
- Number of wrong site surgeries
- Number of unacted upon critical lab results

Sources of outcome measures can be data that is monitored as part of an organization's quality/safety program, claims data, incident reports, chart reviews, and electronic health record data collection.

Monitoring outcomes over time can show the impact of corrective actions on achieving broader goals related to adverse health events or health status.

Guiding principles for determining the type of measurement indicated

Ideally, every AHE corrective action plan has a process or structural measure as well as an outcome measure.

Ideally, every AHE corrective action plan has a structural or process measure as well as an outcome measure. (See Table 1 below.) 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 corrective action had on preventing another adverse event. Using both process and outcome measures as companion measures allows an organization to analyze whether the change has occurred and to know whether it has made the system safer and will prevent further adverse events. Conversely, using only one type of measure gives only part of the story; the lack of a recurrence of the event (outcome measure) may be coincidental and not attributable to the process change.

Table 1. How and when to use measures for AHE reporting

Measure	When used	Companion measures	Example
Structural measure	The corrective action plan calls for the removal or replacement of equipment or physical change to the environment	Outcome measure	**Structural measure** Clamp with detachable parts to be removed from stock **Outcome measure** Number of retained objects
Process measure	The corrective action plan calls for a system/process change	Outcome measure	**Process measure** Outpatient fall risk assessments 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** Patients admitted to the ED with suicidal thoughts are roomed immediately **Outcome measure** Elopement rate of patients with suicidal thoughts

Define the numerator and denominator

Once the problem is identified and the changes to be made are identified, measures to monitor the progress of the CAP 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 identified corrective action or process change in the CAP.

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. And 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 event 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 event 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 event 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 medication X? Choose the numerator/denominator accordingly. Other methods are available to calculate outcome measures such as fall and pressure ulcer rates per patient days.

Example of a Measure

Measure = Numerator/Denominator X 100 = Rate

Measure: Percentage of procedural Time Outs where all activity in the room stops during the Time Out

Denominator: Number of procedural Time Outs observed for all activity in the room stopping during the time out (53)

Numerator: Number of procedural Time Outs observed where all activity in the room stopped during the Time Out (32)

Calculated Rate: 32/53x100=60%

Result: Only 60% of procedural Time Outs had all activity in the room stop during the time out

Establish a goal

A goal is a level of expected compliance with a planned action and usually is expressed as a percentage. If compliance is critical to preventing another AHE from occurring, the goal may be set at 100% compliance. In most cases, expecting 100% compliance over time 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 daily full skin inspection to identify early any potential for breakdown, but for a patient in ICU who only tolerates micro-turning due to becoming critically hypotensive with repositioning, full skin inspections may not be possible.

A goal should be identified for each measure created for the CAP. Goals should be specific, measureable, attainable, realistic, and timely (SMART):

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

What: The number of patients in the sample with hourly rounding.

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

Where: Patients on Unit X.

Why: To assure patients identified at risk for falls have consistent hourly rounding.

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

M: 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.

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

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

T: 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 is being conducted for patients who meet the criteria, a sample population of patients identified as being at high risk on Unit X will be observed once each month for the next six months. The goal: 95% of all sample populations will have hourly rounding conducted when indicated.

The registry only allows entry of the rate or number that is set as the goal. In the goal example above, the tool used to capture results of the observation should specify what is considered high risk, what conditions or findings would constitute an affirmative finding.

Set a threshold

While 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., additional cognitive aids, 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, the failure to use two independent source documents when verifying surgical procedures is highly likely to lead to another surgical event. 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 event 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 CAP.

Example of a Threshold

Goal: 100% of debriefings after a case include accounting for all specimens
Threshold: 95% of unused labels and unused labeled containers are discarded before the next case.

Select a measure of success

Under the Minnesota Adverse Events Reporting Law, a measure of success (MOS) is required for all reported adverse events except pressure ulcers. Per the Joint Commission, a MOS is a quantifiable measure that demonstrates whether an action was effective and sustained. The Minnesota Department of Health uses MOS as a way for all facilities to report on the success of their CAP. Each event has one reported root cause, and one reported intervention. One measure reported for the CAP is also used as the MOS to evaluate the action plan. The MOS should be a process or structural measure, not an outcome measure. In general, the minimum acceptable threshold for an MOS is 90%.

During the three months after the process or structural changes are implemented, the facility must continue to collect data on the MOS to show how well a proposed change has been sustained and embedded into practice. If the threshold was not met by the third month after change implementation of the CAP, the MOS must continue to be monitored and reported into the registry at the sixth month after change implementation.

Step 3. Determine data collection methods

This section will provide information on the key components to data collection: population, sampling, frequency, and duration as they relate to AHE reporting. The goal of measurement for AHE 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 quality improvement activities—however, it does need to be sufficiently rigorous to demonstrate that the intervention worked.

Measurement for quality improvement is not research; data collection should not be so rigorous that it impedes quality improvement activities—however, it does need to be sufficiently rigorous to demonstrate that the intervention worked.

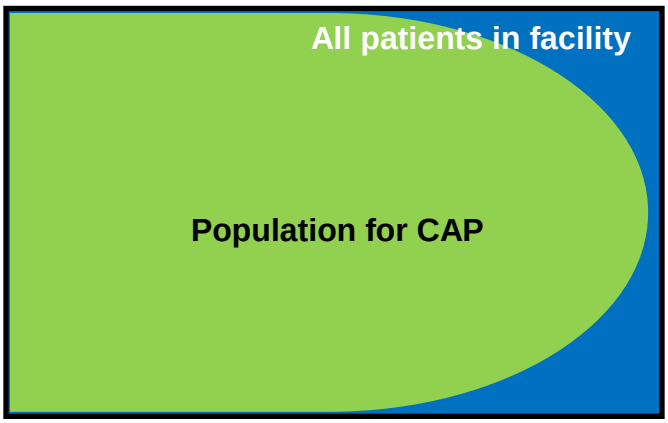
Population

In the context of measurement, population refers to the group of patients impacted by the AHE and its corrective 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 cases or events should be included in the measurement. A population should be defined for each measure in the CAP and should only include patients, events, or cases that could have the outcome or AHE, or that are eligible to receive the process or structure change proposed in the CAP. The populations for the process measure and the outcome measures may not be the same, but large differences should be avoided. (See Table 2 below.) The data for measurement (the numerator and denominator) will be drawn

from the population; so the population must always correspond with the CAP. Defining the population is important because it will help clarify what processes or patient types (cases) should be included in or excluded from the data collection.

Consequences of not properly identifying the population are an incorrectly targeted CAP, inaccurate data, and incorrect assumptions. (See Figures 2 and 3 for examples.)

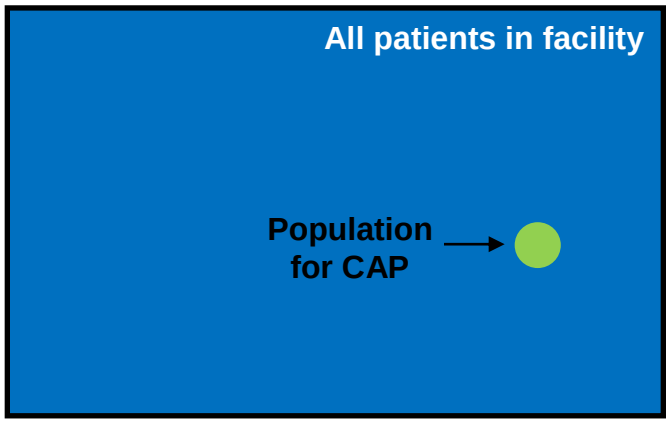
Figure 2. Population for CAP, Scenario A1



PROBLEMATIC

- Population for the CAP is all patients in the facility.
- Population is too broad.
- Cannot detect changes in CAP.
- Cumbersome measurement and data collection.
- Recommend focusing the population targeted for the CAP.

Figure 3. Population for CAP, Scenario A2



PROBLEMATIC

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

Population for process measures. A population for a process measure consists of the processes or group of patients or cases that are targeted in the CAP 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 below for examples.)

Population for outcome measures. A population for an outcome measure consists of the patients for whom the outcome or adverse event could occur. The outcome population can be determined broadly (e.g., every admission into the facility in a given year, every surgical patient) or it can be narrowed to a specific population (e.g., admissions on one unit, every person having a certain type of procedure, patients with critical lab results). Outcomes that occur in the population are counted, such as the number of falls, number of lost specimens, or number of wrong-site surgeries. (See Table 2 below for examples.)

Table 2. Population examples for outcome and process measures

RCA and CAP	Process measure population	Outcome measure population	Summary of population selection
RCA found assessment for fall risk was not completed on admission. This pattern was noted on the unit. CAP is aimed at increasing the consistency of completing fall risk assessment on admission.	Process population: all patients admitted to the unit Process measure: risk assessment completed upon admission for patients admitted to the unit	Outcome population: all patients admitted to the unit Outcome measure: fall rate for patients admitted to the unit	Population 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.
RCA found a critical lab result was not acted on because of miscommunication between staff. CAP is aimed at increasing effective communication between staff by teaching reporting staff to expect and receiving staff to perform a "read back" of critical lab values.	Process population: all critical lab results Process measure: critical lab values are read back to reporting staff	Outcome population: all lab results reported in the facility in one year Outcome measure: miscommunication of critical lab results in the facility 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 CAP.
RCA found a particular drill bit was not consistently inspected for being intact after use during procedures. CAP is aimed at increasing the inspection of all instruments for all procedures in the facility.	Process population: all procedures performed in the facility Process measure: procedural equipment inspected for being intact after use .	Outcome population: procedures which require the particular drill bit over the next six months Outcome measure: Retained object rates for procedures which require the particular drill bit over the next six months	Population for outcome measure is a subset of the population for the process measure. The outcome measure is specific to one type of equipment but the process is rolled out to all equipment. 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 facility-wide, and creating an additional process measure to monitor the specific procedure with the problem.
RCA found a lack of clarity about the ability and expectation of staff to remove a certain brace that is rarely used to do skin inspections. CAP is aimed at developing a clear policy to address skin inspection for patients with this particular brace, but also will expand the population to assure clarity for the range of all braces or devices used.	Process population: patients with the particular brace used infrequently (rare event) Expand the population to patients with any device or brace	Outcome population: patients with the particular brace that is used infrequently (rare event) Expand the population to patients with any device or brace	Populations for outcome and process measures are very small (rare events). Expand both populations proportionately to increase sample sizes for measurement, but highly recommend monitoring the process and outcome for every rare event that occurs.

RCA and CAP	Process measure population	Outcome measure population	Summary of population selection
	Process measure: skin inspections completed for patients with any device or brace	Outcome measure: pressure ulcer rate for patients with any type of brace	

Sampling

Often, it is not possible to measure every instance (the whole population) in which a process is supposed to occur or on every patient that could have the outcome or AHE. 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 CAP can be incorrectly concluding that a process has changed when the process has not actually changed. This incorrect conclusion may result in future AHEs. (See Table 3 below for examples of sampling methodologies.)

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-	Randomly select 30 charts from a list of all patients admitted to the facility in the last week to verify if fall risk assessments have been conducted.

Sample method	When used	Pros and cons of sample method	Examples
		time or concurrent data collection (such as collecting data from surgeries or other cases as they occur).	
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 surgery) 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 6 procedures from each OR and Interventional Radiology rooms (five rooms) to observe whether Time Out processes are conducted as expected (total of 30 cases 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 OR case on the OR schedule to observe whether specimen transportation protocols are in place (total of 30 cases or all if less than 30 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.	On the last day of the month, observe that all surgical cases are set-up for inspection of equipment and/or supplies
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. Popular in AHE because 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	Select 30 patients as they are admitted to observe fall prevention measures are in place. Or: Select 15 high-risk patients and 15 low- risk patients as they are admitted to observe whether fall prevention measures are in place.

Sample method	When used	Pros and cons of sample method	Examples
		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).	

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 CAP 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 patients 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 facility. Changes in the range and noticeable patterns can be reviewed to determine the reasons.

The example below shows data collected for reading back critical lab results. 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 critical lab result read backs has increased dramatically over time. But if this measurement were expanded to include more data points over a longer period of time, the facility would see that the data collected in these three months just shows variability in the data (Figure 5).

Figure 4: Critical lab result read back rates for Hospital A for three months

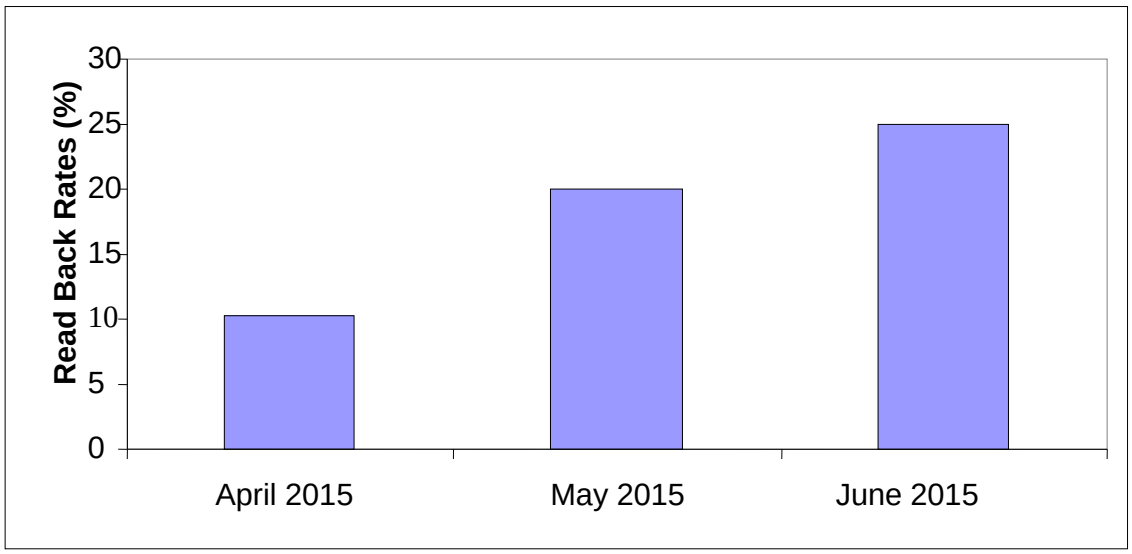
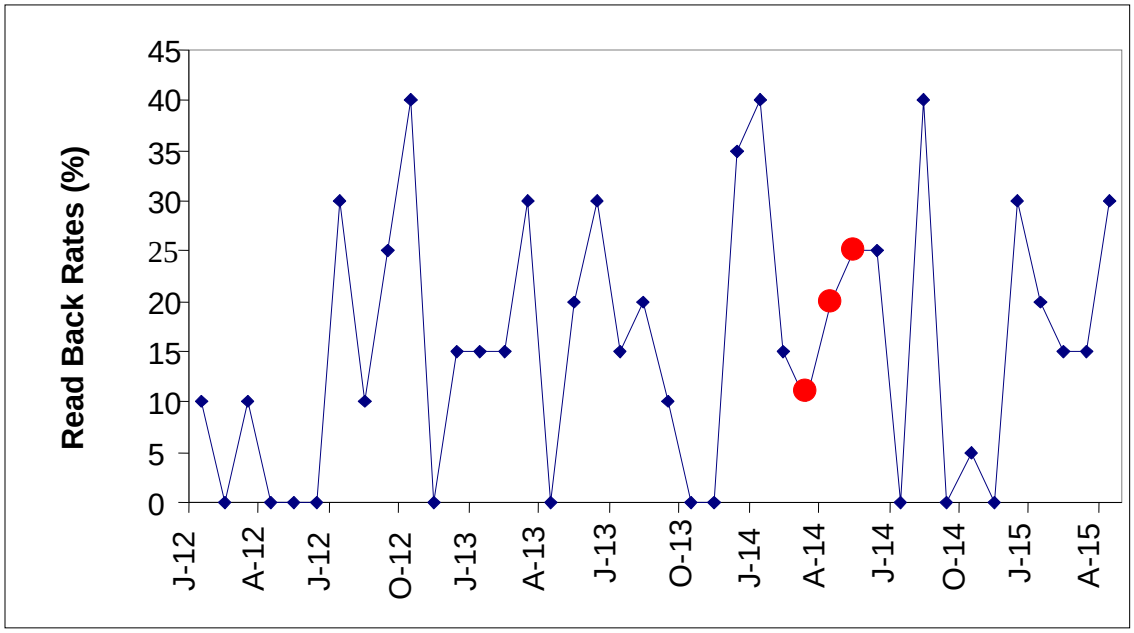


Figure 5: Critical lab result read back rates for Hospital A by month



In summary, 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, whether due to cost constraints, timing, or the rarity of the process or event. 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 the interpretation of the data. The size of a sample should be driven by the size of the population during the time frame of interest. (See Table 4 below 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 corrective action 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.

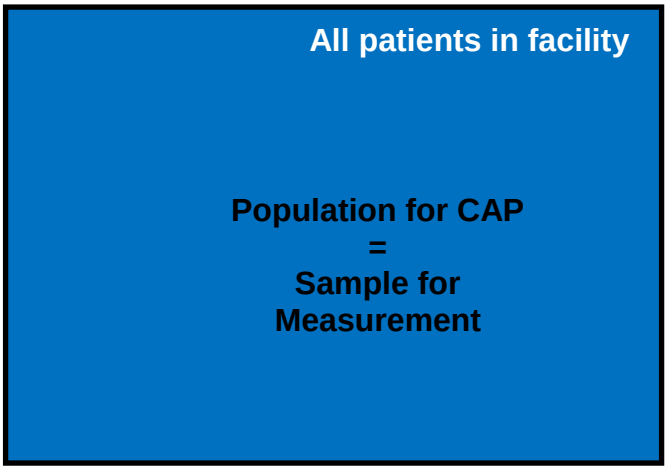
Small samples due to rare events. Because adverse events are usually rare, it may take a long time to collect enough data to draw conclusions about the effectiveness of the process changes through the use of outcome measures. To address this situation, pair the outcome measure with one or more process measures. For rare events, facilities can use alternative methodologies. (See Table 5 below.)

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 corrective action or change 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 events. Some processes actually contribute to, or prevent multiple adverse events. For example, timeouts are conducted to prevent a variety of adverse events (e.g., wrong-site surgeries, incorrect patients, and wrong surgical procedures). Combining data for all surgeries in this example will increase sample sizes.	In the case of a wrong-site surgery that occurred during a rare procedure, the facility may consider combining all types of surgeries and monitoring whether the time-out process is taking place as expected, rather than looking only at the type of surgery during which the event occurred.

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

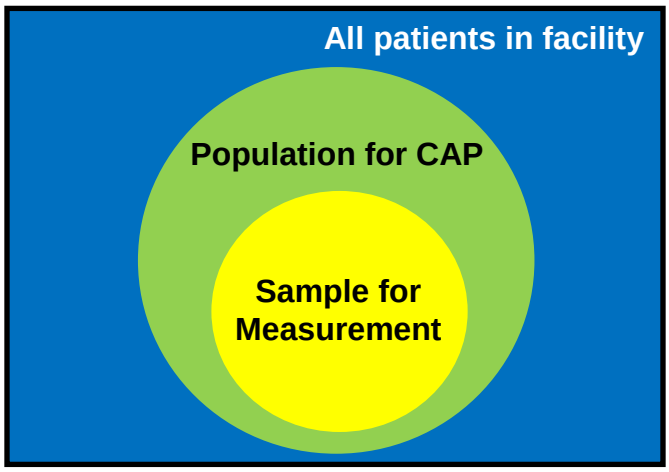
Figure 6. Ideal Sampling Scenario



IDEAL

- Population for the CAP is a selected number of patients from the facility (not all patients).
- Measurement is on the entire population targeted for the CAP (sample = entire population).
- Collecting data on the entire population for a CAP is a valid measurement.

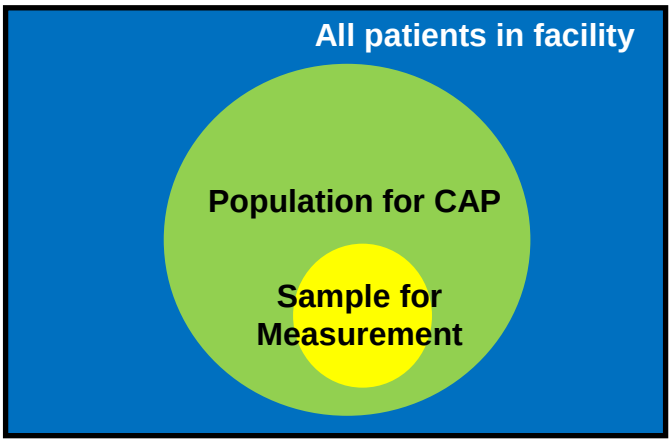
Figure 7. Ideal Sampling Scenario



IDEAL

- Population for the CAP is a selected number of patients from the facility (not all patients).
- Measurement is on a subset of the population targeted for the CAP (sample).
- Collecting data on a sample from the entire population for the CAP is a valid measurement if good sampling techniques are used.

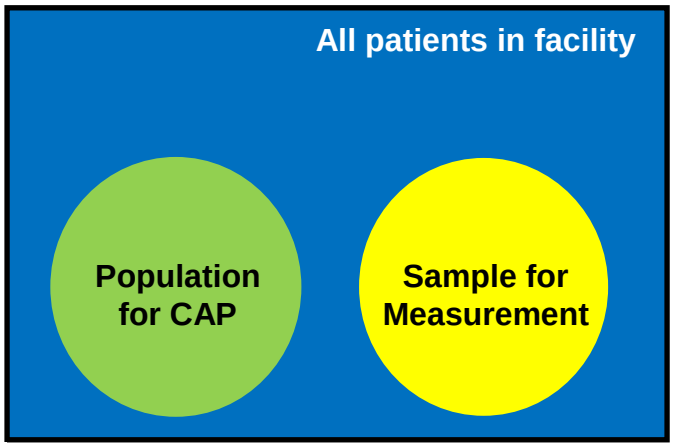
Figure 8. Sampling Pitfall Scenario



SAMPLING PITFALLS TO AVOID

- If it becomes evident when determining the sample size that the population targeted for the CAP 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 CAP 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 CAP is adequate, but the sample size proposed is too small in relation to the population, measurement may not be accurate.
- Recommend increasing 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 CAP intervention, the measurement will not be accurate. Recommend reviewing sampling methodology to include only patients or records that received the CAP 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 event occurs

  If the process or event 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, 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.

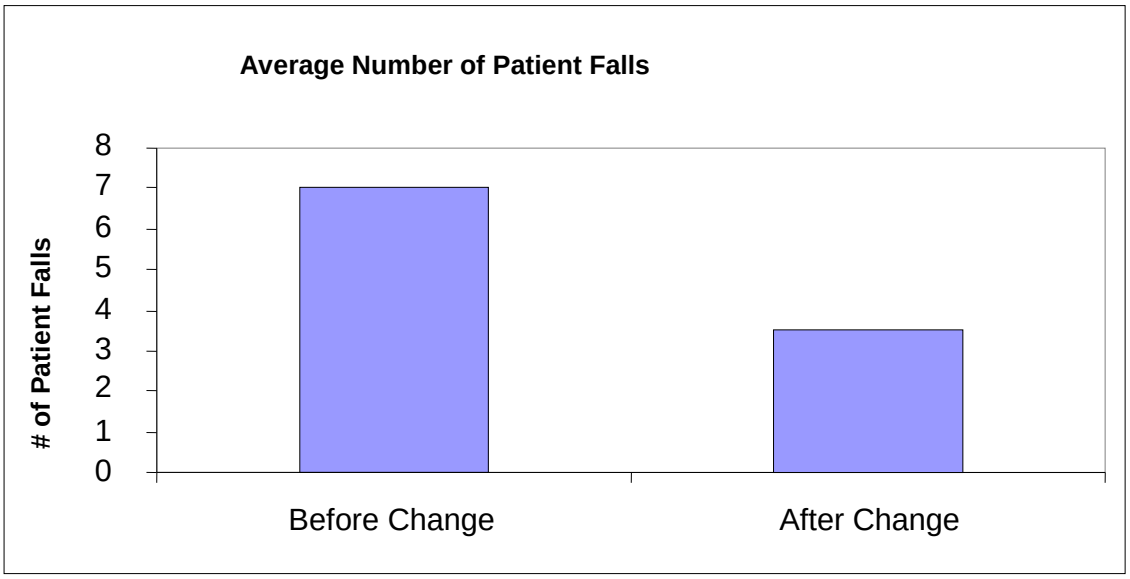
Frequency and duration are used to determine if a change is sustained over time. No clear formula exists for determining the appropriate frequency or duration for data collection because it is dependent on the sample size and characteristics of the population being measured. Smaller, more frequent data 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 patient falls before and after a change was made (e.g., implement hourly rounding). Figure 10 below shows data collected before and after the change was made. A conclusion can be drawn that the number of patient falls decreased after hourly rounding started.

Figure 10. Fall rates before and after the CAP 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 below 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 patient falls were decreasing before and after implementing hourly rounding. Therefore, the new process was not the only factor that contributed to the decrease in patient falls. Figure 12 shows that the data are fluctuating, but at a higher rate before, then decreasing after the start of hourly rounding.

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. Patient fall rates by month before and after the CAP intervention

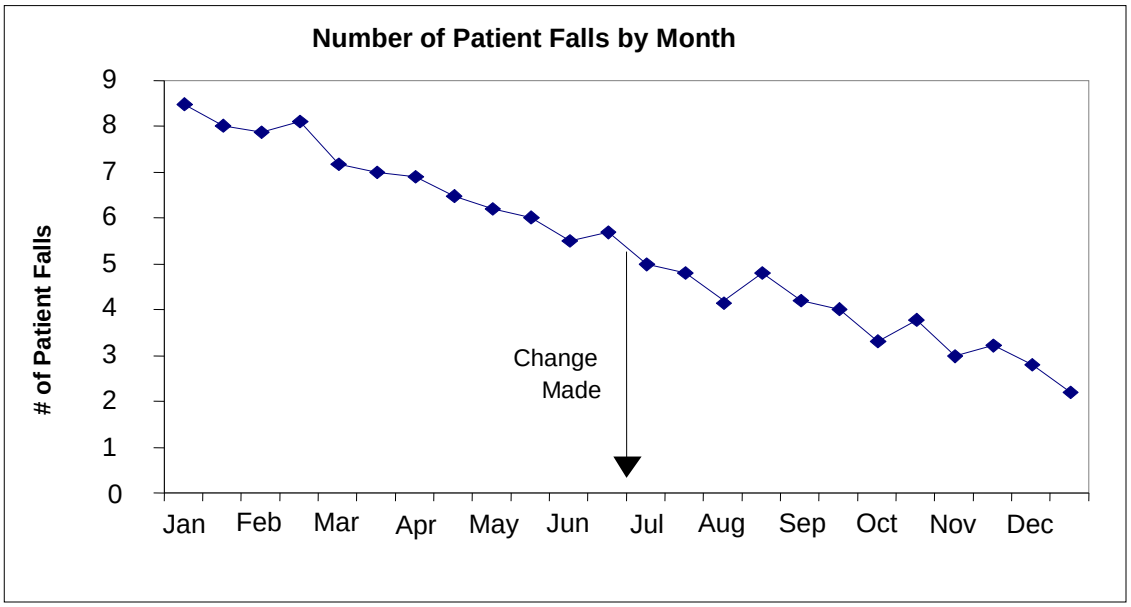
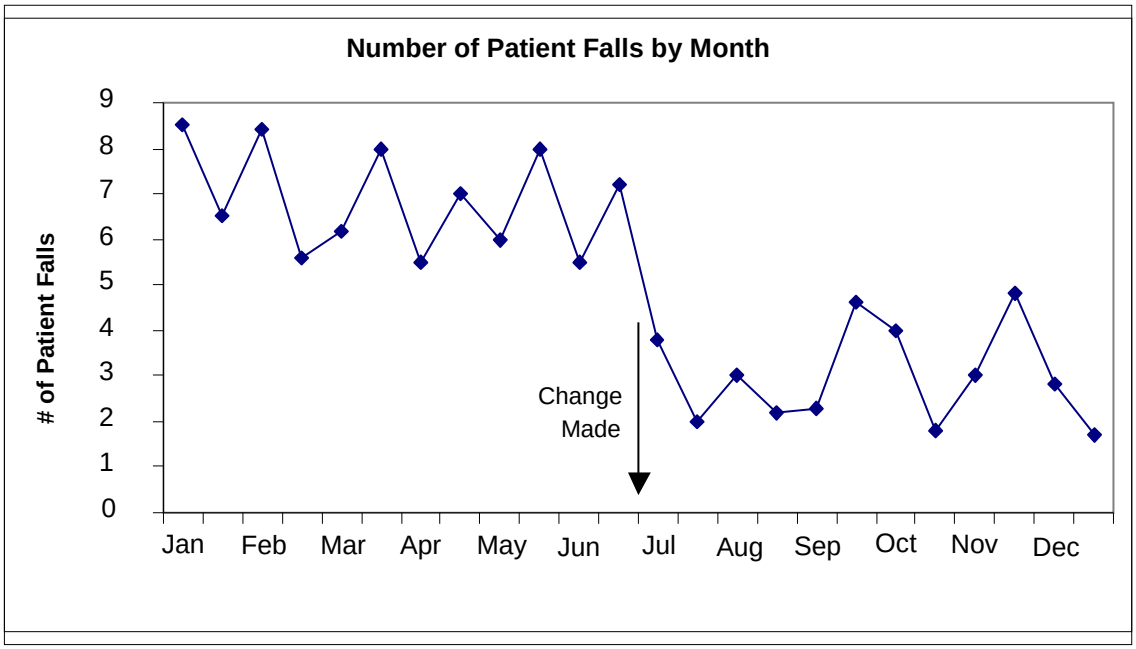


Figure 12. Patient falls by month



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 all activity stopping during a Time Out, the sample size is set at 10 procedural observations each month for three months with a threshold of 90%. Two observations are conducted and there is no observed stop in activity. In this example, it is reasonable to stop auditing and investigate why activity is not stopping during Time Outs. It is not worth collecting data from eight more records because the trend shows the threshold will not be met.

Example of when measurement should not stop

For a measurement of all surgical case Time Outs as they occur during the month of March, a failure trend cannot be accurately calculated until the last surgical case is completed in March—when the whole population is known.

Step 5. Drawing Conclusions

If measurement data were collected before and after the corrective action was implemented, the conclusion of whether the corrective action led to a change is strengthened. Collecting data prior to the implementation of the corrective action 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 corrective action was appropriate.

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

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

If the event occurs again, the newly launched RCA 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.

Following is an example that illustrates the use of measurement in the RCA/CAP process:

Event: Retained central line guidewire visible on CT three days after insertion at bedside.

RCA: The team determined there was a lack of a cognitive aide to remind staff to account for guidewire removal of bedside central line insertions

CAP: The CAP is aimed at developing a prompt on the central line form for two staff to verify the removal of the guidewire.

Process measure: Verification will be documented by two staff of the removal of the guidewire in 95% of central line placements at bedside.

The estimated population size is 5 patients per month.

5 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 documented verification by two staff of the removal of the guidewire with a threshold of 90%.

Outcome measure: The organization will measure retained guidewires of bedside central line insertions every quarter with a goal of zero.

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 two staff verification of guidewire removal	**No change or goal not met:** retained guidewires did not decrease	Two staff verification of guidewire removal did not result in a decreased retained guidewires. The root cause was not correctly identified—or other factors contributed to the retained guidewire rate. Further analysis is needed.
2	**Goal not met:** Audited records showed two staff verification of guidewire removal is not done	**Goal met:** retained guidewires decreased	Fewer retained guidewires occurred, but implementing the two staff verification of guidewire removal has not been 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 showed two staff verification of guidewire removal is not done	**No change or goal not met:** retained guidewires did not decrease	Because the two staff verification of guidewire removal has not been embedded into practice 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.
4	**Goal met:** 95% of audited records showed two staff verification of guidewire removal	**Goal met:** retained guidewires decreased	Two staff verification of guidewire removal is embedded into practice and no retained guidewires have occurred since the change was implemented. The root cause was correctly identified and the CAP was successful in reducing the AHE.
5	One goal not met One goal met **Goal not met:** Chart audit showed two staff verification of guidewire removal **Goal met**: Staff education on potential complications of guidewire retention	**Goal met:** retained guidewires decreased	FOR CASES WITH MORE THAN ONE PROCESS MEASURE: No retained guidewires have occurred. One of the changes is embedded. The other change is not embedded and may not be as critical to preventing retained guidewires as the change with the goal met.

Scenario	Process measure	Outcome measure	Analysis
6	One goal not met One goal met **Goal not met:** Chart audit showed two staff verification of guidewire removal **Goal met**: Staff education on potential complications of guidewire retention	**No change or goal not met:** retained guidewires did not decrease	FOR CASES WITH MORE THAN ONE PROCESS MEASURE: Retained guidewires are still occurring. One change is embedded and one change is not. The change embedded is not critical to preventing retained guidewires 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 RCA/CAP measurement plans that could be strengthened.

Case study 1: Misalignment between CAP and measure

Event: A sponge was retained after a bowel resection.

RCA: The team determined the sponge was retained due to staff lack of knowledge of accounting for sponges.

CAP: The CAP is aimed at holding mandatory training sessions for all surgical staff on the expected process for accounting for sponges.

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

Outcome measure: Zero retained sponges the next three quarters.

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 sponge accounting process. In situations where the process changed 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 post test or return demonstration of the expected process steps.

Case study 2: Missing measurement components

Event: Medication error – anticoagulation dosing was not included in discharge instructions from day surgery.

RCA: The team determined medication reconciliation did not occur as expected due to unclear policy.

CAP: The CAP 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 discharged surgical 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. A measurement strategy that includes all of 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 discharged surgical patients from day surgery for evidence that medication reconciliation occurred according to the policy. The audit will begin on (date) and be conducted by the day surgery safety team. Expect 95% 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 medication error 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 when transferring patient from the OR to recovery unit.

CAP: The CAP is aimed at using a structured way to communicate information to staff across shifts and units.

Process measure: Process measure: 15 patient transfers a month will be reviewed for documentation of 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 always hope for the easy fix: the one simple change that will erase a problem in a stroke. But few things in life work this way. Instead, success requires making a hundred small steps go right - one after the other, no slipups, no goofs, everyone pitching in." *Atul Gawande*

This guide is intended to be a resource for patient safety 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 Adverse Healthcare Events: MDH Reporting Evaluation Tool. This document in the Patient Safety Registry toolkit provides the evaluation criteria Stratis Health uses when reviewing root cause analyses and corrective action plans. Requests for more information and additional comments are based on this document.
https://www.patientsafetyregistry.com/includes/AHEReviewCriteria.pdf

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

Patient Safety Registry User Guide provides practical how-to information for those who use the Web site. Topics include managing personal information, obtaining reports, and the three step cycle to reporting an event.
http://www.mnhospitals.org/Portals/0/Documents/ptsafety/User's-Guide-Aug09.docx

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

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.
http://www.stratishealth.org/expertise/quality/QIBasics.html

The Department of Veterans Affairs National Center for Patient Safety (NCPS) supports and leads patient safety activities for all VA medical centers. This site contains tools, training, and software to facilitate patient safety and root cause analysis investigations.
http://www.patientsafety.va.gov/professionals/onthejob/cognitive.asp

The Department of Veterans Affairs National Center for Patient Safety (NCPS) Web site provides information on the hierarchy of corrective action plans. This page provides examples of actions that are labeled as weak, intermediate, or strong as determined by their ability to cause change.
http://www.patientsafety.va.gov/docs/joe/rca_tools_2_15.pdf

Appendix B: Steps for Creating Measures