CHAPTER 8

Leveraging Analytics in Quality Improvement Activities

Knowing is not enough, we must act.

—Johann Wolfgang von Goethe

Data and information alone are not sufficient to achieve transformation in healthcare. Information and insight need to operate within a framework or methodology for quality and performance improvement decision making. Such a framework is necessary to identify priorities for improvement and evaluating outcomes. This chapter will focus on how to leverage analytics within a quality improvement (QI) environment to assist the healthcare organization (HCO) in achieving its quality and performance goals.

Moving from Analytics Insight to Healthcare Improvement

When used in concert with QI methodologies such as PDSA, Lean, or Six Sigma, analytics helps to identify the most pressing quality issues facing the HCO based on needs defined by patient safety, the quality goals of the HCO, national standards, and legislative requirements. These improvement methodologies leverage the insights gained from analytics, within their respective structured approaches, to develop interventions and solutions for healthcare quality and performance issues and evaluate outcomes to ensure the long-term sustainability of improvements.

FIGURE 8.1 Five Quality Improvement Phases

To gain maximum value from analytics, QI projects need to be able to:

- Retrieve and analyze baseline data to document current performance and/or quality measures against which to compare future performance;
- Utilize or develop well-defined performance indicators (PIs) that accurately reflect the processes, procedures, policies, or treatments being changed; and
- Perform ongoing evaluation and reporting of relevant PIs to quantify the impact of implemented changes and to identify if further revisions to processes or policies are required.

Chapter 4 outlines the necessity of clearly defining what quality means in the context of the HCO, and also provides an overview of several common QI methodologies, such as PDSA, Lean, and Six Sigma. Regardless of which methodology is chosen, analytics can be incorporated at many decision and analysis points throughout a QI project, and in most decision-making processes within the HCO. This chapter will focus on the key role that analytics plays throughout quality and performance improvement initiatives. Following are five phases of a QI project during which analytics can be leveraged to help move from analytics insight to healthcare innovation and improvement. The five phases are illustrated in Figure 8.1 and described in more detail here.

Analytics is important in almost every phase of healthcare quality and performance improvement. Regardless of the QI framework chosen, the analytics needs of QI initiatives depend on:

- The phase of the initiative;
- Who is using analytics;

- What information and insight is required; and
- How that insight and information is being used.

Accurate information is necessary to understand the scope of the problem(s), identify the best possible solutions, evaluate those solutions once implemented, and monitor ongoing performance to help ensure that the improvements have been sustained.

The quality and performance problems that HCOs endeavor to address should, where possible, be in alignment with the quality and performance goals of the organization as stated in its quality and performance strategy. Those quality and performance goals serve as a kind of "true north" for keeping the organization on track as competing interests and requirements detract from strategic improvements. On occasion, problems will emerge that are not strictly aligned with the quality strategy but are nonetheless necessary to deal with. Because HCOs may have limited resources for undertaking multiple projects, the competing priorities need to be ranked in some way so that the highest-priority issues come first. Depending on the HCO's needs, the importance of problems can be ranked in different ways; three common bases for ranking issues facing an HCO are:

1. **Clinical.** Clinical concerns are perhaps the most important reason to undertake an improvement project. An HCO's clinical performance directly impacts the satisfaction and safety of patients whose care has been entrusted to the HCO. Clinically related improvement initiatives have as their goal to reduce adverse clinical outcomes and to ensure that patient care is delivered as per best practice guidelines. An example of an opportunity with a clinical focus is working on emergency department processes for stroke patients so that they receive the recommended imaging tests and medications within the proper window to minimize the loss of brain tissue.
2. **Financial.** As HCOs strive to become more financially healthy, financial considerations are likely to be a priority when selecting improvement initiatives. In this case, the costs of inefficiency, errors, or simply of doing business are likely to factor highly. For example, a healthcare payer may realize that doctors at a certain hospital are ordering more, expensive diagnostics than at other, similar hospitals. In this case, the payer would work with the hospital in question to identify new processes, guidelines, procedures, and even training to reduce these unnecessary diagnostics and associated costs.
3. **Regulatory/legislative.** Another prime driver of improvement activities is regulatory and legislative changes and/or incentives. For example, the U.S. Centers for Medicare and Medicaid Services introduced an incentive program rewarding qualifying HCOs and providers for

implementing or upgrading electronic medical records and demonstrating meaningful use.[1] This incentive program resulted in many HCOs and providers updating their processes and policies around the use of this technology in order to qualify for the incentive payment.

When classifying problems into each group listed previously (or other types of issues not listed here), HCOs can calculate the cost (whether measured in financial, clinical outcomes, or other terms) of poor quality and performance and begin to estimate the potential value or benefit of addressing their root causes. Those with higher associated costs and/or greatest potential benefit or value are considered higher-priority than those with lower associated costs. Using this type of ranking rationale helps organizations to become more transparent and quantitative in their decision making so that a decision to address one or a few problems over others can be quantitatively supported.

Analytics in the Problem Definition Stage

Once a problem has been identified and selected as a priority for the organization to address, the first step is to start with a clear and detailed description of the quality or performance problem (or other issue) that must be improved. The important considerations of this step are:

- What are the goals/objectives to which this problem relates?
- What are the relevant indicators and metrics?
- What baseline data is available, and what data will be available moving forward for monitoring and evaluation?

One of the first steps is to quantify and measure the magnitude of the problem. The problem should relate to the strategic quality and performance improvement objectives of the organization and/or the tactical-level improvement goals of units, departments, and programs, and wherever possible be described and quantified in terms of the appropriate metrics and indicators. The results of this process help to filter all the possible metrics and indicators, based on possibly hundreds of available data points, down to the critical few indicators required for the success of a quality and performance improvement project.

The information-gathering and benchmarking phase typically sees QI team members obtaining data regarding the issue in question. Once metrics and indicators are decided upon, critical to any QI initiative is effective baseline performance information. Useful baseline information, however, is more than just the collection of historical data. Baseline performance is a quantitative description of some aspect of the HCO's performance measured

Baseline Data

Baseline performance data is a quantitative description of some aspect of the organization's performance measured prior to undertaking an improvement initiative.

prior to the undertaking of an improvement initiative. (If no quantitative data is available, a qualitative description is sometimes helpful.) An accurate baseline is necessary for determining whether any actual change in performance has taken place, and what the magnitude of that change is.

Every QI project will require baseline data. In fact, baseline performance data is helpful in the everyday operations of the HCO. Baseline data can put current performance in perspective; for example, are the number of visits, lengths of stay, admission rates, or bed turnaround times today better or worse than typical? This has an important impact on dashboards and other similar tools, where baseline quality and performance information can add very valuable context to current real-time performance values.

In some cases, baseline data may not be available (for example, with the opening of a new unit or clinic, or with the adoption of a brand-new technology). In these cases, baseline data should be established as soon as possible into the improvement activity in order to gauge performance changes. The information gathered in this phase of the QI process is typically more static and historical in nature, may require some basic statistics to tease out actual performance values, and can be visualized in various types of statistical process control (SPC) charts to determine how "in control" a process is.

Knowing the true magnitude of a change requires both an accurate starting point and ending point. For example, almost any medical procedure performed on patients at a healthcare facility requires a complete baseline (including height, weight, lab results, and diagnostic imaging as necessary)—no healthcare provider would even consider performing a procedure without knowing as much about the patient as possible. Yet healthcare QI projects are undertaken too often without the benefit of clearly knowing current and/or historical performance, and this is why many of those projects fail.

Tip

Baseline data should be based on and measured in the exact same way as the indicators developed for the QI project.

Baseline data should be based on and measured in the exact same way as the indicators developed for the QI project. In order to ensure that baseline data is a true measure of existing performance, the following considerations should be made:

- **Data source.** Wherever possible, the source of baseline data should be the same as that used for ongoing measurement and evaluation; if the sources of data are different (for example, before and after the implementation of a new system), differences in the way data were recorded, processes were interpreted, and data was analyzed must be taken into account or there is a risk of performing an invalid comparison and perhaps reaching inaccurate conclusions.
- **Data quality.** If baseline data is drawn from legacy systems, or even collected from paper sources, data quality may be an issue. Ensure that the quality of baseline data is as high as possible prior to performing any analysis or comparisons.
- **Time period.** Baseline data must cover a long enough time to be an accurate reflection of performance, and must be recent enough to be a valid comparison. If baseline data is too old or does not cover a sufficient time period, processes, performance, and quality may have changed since the time encompassed by the baseline data.
- **Indicators.** Comparisons with baseline data work best if the baseline data is analyzed and reported using the same indicators that will be used moving forward for monitoring and evaluation.

Data used for the baseline should be reliable—it should be trusted as a true measure of performance, as future performance will be compared against this. If no data is available, it may be obtained manually (for example, via chart reviews), but this may have implications for what indicators are chosen for the actual improvement initiatives.

Manually Collected Data

Just because data is collected manually doesn't mean that it must always live on paper; results of audits and other manual checks can be stored electronically and made available for analysis. Several business intelligence tools will allow users to import data from external spreadsheets into existing frameworks. In my experience, we have successfully integrated external data, such as from process audits, that was collected manually into our BI tools, and used that data as part of analysis of data collected from clinical systems.

When reporting baseline data, one important consideration is how to aggregate the data. For example, when looking at a year's worth of baseline data, the temptation might be to group the data by month. When grouping by month, however, it is possible that certain details are being lost that might be helpful to determine if a change is in fact occurring.

Consider Chart "A" in Figure 8.2, in which average clinic patient length-of-stay data is grouped by month over a 12-month period. During that time, the performance of the clinic appears to be relatively static. Chart "B" hints at a slightly different story—aggregating the data by week instead of month shows that midway through the baseline data, variation in the data seems to have decreased. A decrease in variation in performance is considered to be a step toward improvement. So even though on a monthly scale performance may appear to be static, when looked at through a weekly lens, it can be

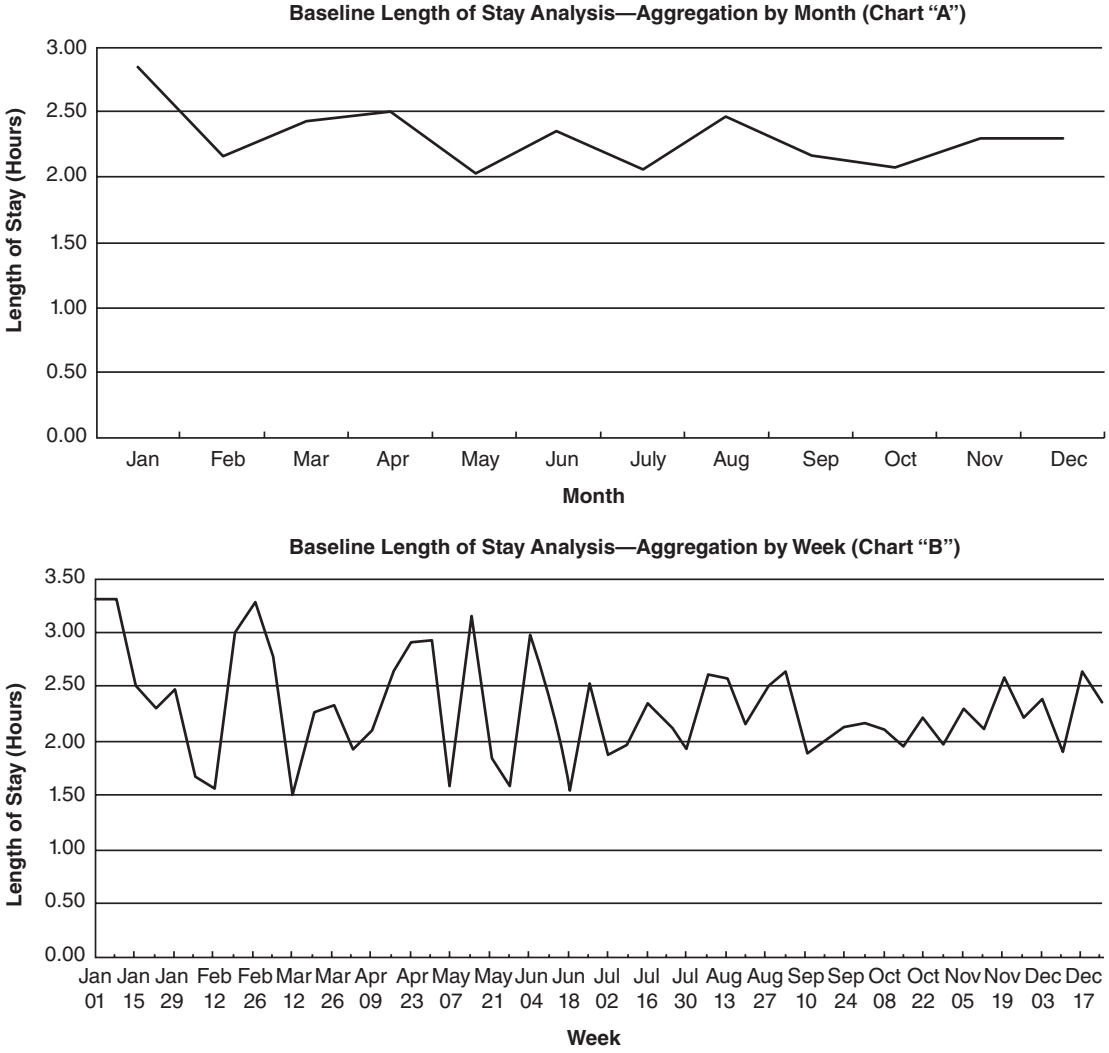

FIGURE 8.2 Baseline Data Showing a Change in Performance When Aggregated Differently

suggested that some improvement has already started to take place. The reverse is true as well; if baseline data shows increasing variability in a set of data, it is possible that capacity to perform at expected levels is deteriorating.

Using Analytics to Identify Improvement Opportunities

Once a problem has been identified and quantified, the next step is to identify improvement opportunities—that is, what steps an HCO can take to achieve the desired outcomes and levels of performance. To achieve these improvements requires specific actions and interventions on the part of QI teams. Of course, not all identified opportunities and resultant improvement activities will have the same impact, so HCOs need to establish ranking criteria with which to rank, evaluate, and select opportunities. At the problem definition stage, a cost (in financial, clinical, or other terms) associated with the quality or performance problem would have been calculated. The cost of actual improvement efforts can be compared to the cost of the problem, and a decision can be made to proceed with specific improvement initiatives that will have the greatest impact and require the least resources and effort—in other words, achieve maximum value.

The three steps in identifying and selecting improvement opportunities are as follows:

1. Determining likely root cause(s) of quality/performance problems.
2. Identifying possible countermeasures to address root causes.
3. Estimating countermeasure impact and effort to achieve goals.

DETERMINING ROOT CAUSE Healthcare QI teams often fall victim to addressing *symptoms* of problems, not the actual problems themselves. This will likely only result in the introduction of a workaround, not a solution to the problem. For example, if a medication cabinet in an observation area is regularly understocked, the "solution" may be for a healthcare aide or other staff member to raid (for lack of a better term) a medication cabinet of another area of the department. This not only results in additional inefficiency by requiring staff to move unnecessarily, but it may also cause problems downstream with other staff who rely on medications from the raided cabinet.

True QI requires healthcare professionals to move beyond "who" and "what" into "why" errors, defects, or waste occur. For real change to occur, we must change from being a blaming culture and into a solution-finding culture in healthcare by moving beyond responding to symptoms and start addressing actual root causes of problems.

Because blaming seems to come more naturally than finding root causes, there are tools that can be employed within QI methodologies to help find

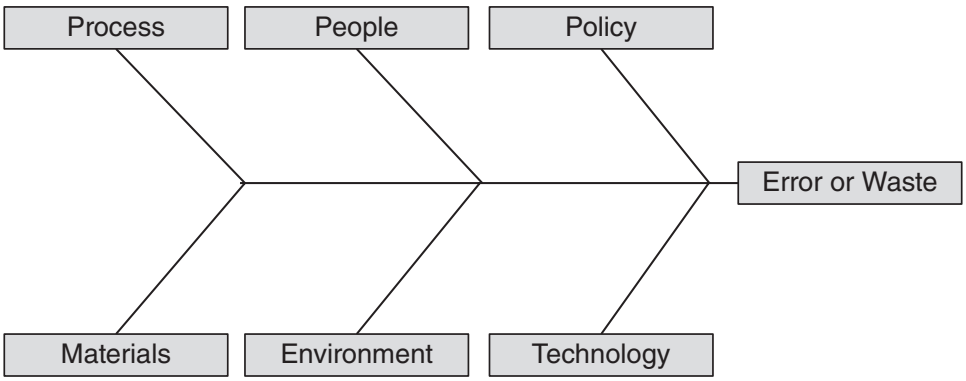

FIGURE 8.3 Sample Ishikawa (Fishbone) Diagram Used for Identifying Causes of Problems, Errors, or Waste

root causes. One approach to identify root causes using an Ishikawa (or fishbone) diagram, as illustrated in Figure 8.3. A problem, error, or waste is identified and all possible causes for it are listed under various categories. (Typically, the categories are process, people, policy, materials, environment, and technology, although variations exist and oftentimes the categories are changed to fit the particular problem under investigation.)

Once possible root causes are identified, the contribution of each root cause can be mapped in a Pareto chart (see Figure 8.4). Pareto charts are very useful to highlight the most important contributing factors to a problem or issue. The main components of a Pareto chart are the identified causes of a quality or performance problem, vertical bars to represent the number (or percentage) of times the problem occurred as a result of that cause, and a line plotting the cumulative frequency of the causes (which should add up to 100 percent).

Suppose an emergency department has implemented a new computerized triage tool and is experiencing an unacceptably high rate of triage overrides, which occur when the triage nurse does not agree with the triage score determined by the algorithm in the triage tool. During an investigation into the root causes of the overrides, triage nurses are asked to identify the reason they overrode the computer. The nurses identify five key issues that caused them to enter a score other than what the algorithm determined:

1. Too many clicks on the form; nurses are bypassing certain form fields to save time.
2. The design of the triage form is confusing, and triage nurses are missing important fields on the form.
3. Triage nurses do not feel they received enough training.
4. Triage is too busy in general to complete the form properly.
5. Some triage staff don't trust the algorithms.

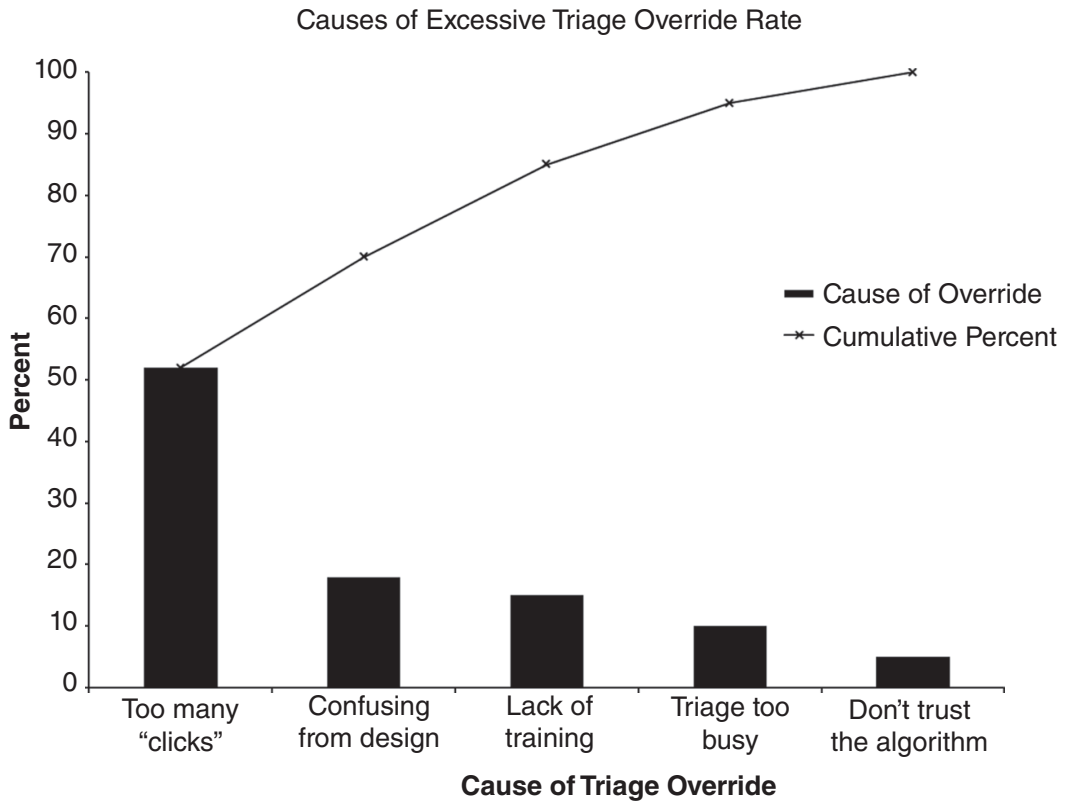

FIGURE 8.4 Sample Pareto Chart Used to Determine Important Contributing Factors to Quality Problems

These five causes of triage overrides are then plotted on a Pareto chart (see Figure 8.4). By looking at the chart, the QI team is able to determine that approximately 70 percent of overrides are the result of two main issues: too many clicks, and confusing form design. With this information, the QI team can be reasonably sure that a significant percent of overrides can be prevented by addressing these two issues, which in fact would bundle nicely into a single triage form improvement project.

Analyzing the root causes of problems in a quantitative manner such as a Pareto chart can provide QI teams the insight they require to make a transparent, evidence-informed choice on which QI initiatives to undertake.

ESTIMATING IMPACT AND EFFORT There may be some process changes that HCOs can make that would result in a large impact to the department or HCO but would require relatively little effort. Is there a way that we can quantify these high-impact changes? Consider Figure 8.5, which illustrates an impact/effort grid that can be used to help select QI projects. I have yet to encounter an HCO that has unlimited resources to dedicate to QI projects. Because of this, even though there may be many opportunities and needs for improvement, only a few projects are feasible to undertake at any one time.

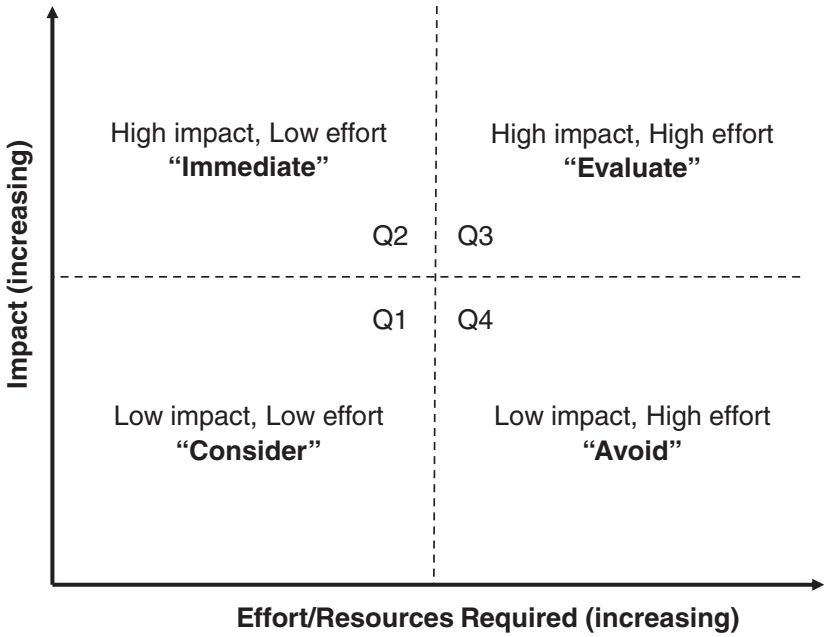

FIGURE 8.5 Sample Impact/Effort Grid for Selecting Quality Improvement Projects

The impact/effort grid helps QI teams map projects according to estimated impact (that is, how much change or benefit can be expected) and what the anticipated effort to achieve that change might be. Projects that fall into quadrant 2 (high impact, low effort) are generally favored by HCOs, as these represent potential game changers, where big wins are possible with relatively little effort. Projects falling into quadrant 4 (low impact, high effort) should be avoided, as they can be a major drain on HCO resources and QI team morale, without any definite long-term positive impact on quality. Projects that fall into quadrant 1 (low impact, low effort) can be considered, but may detract resources from other projects that may have a bigger impact. Quadrant 3 projects, those that are high impact and high effort, should be evaluated carefully to determine whether the potential benefit of the project is worth the time and resources applied to achieve that benefit.

Analytics can be very helpful when mapping projects on an impact/effort grid, especially if relevant, detailed, and high-quality baseline data is available. For example, the anticipated impact of a project can be estimated by examining baseline levels of performance against anticipated targets, or, better yet, by looking at comparable performance metrics for other programs or sites that are performing better.

Estimating the impact of a change can be challenging; without any quantitative basis, impact estimates are little more than a mere guess. Various analytical approaches, including regression modeling and other predictive approaches, as well as simulation, can be helpful in better quantifying the effect of a change. See Chapter 11 for a discussion of various advanced analytics approaches.

Analytics in the Project Execution Phase

Executing improvement activities involves finding new ways of doing the required work to achieve the desired improvements (innovation), working with staff to implement the changes (intervention), and examining interim results to make any midcourse corrections (experimentation). During project execution, QI teams use analytics to closely monitor the processes that have been changed (or that have been impacted by changes to equipment, layout, or staffing levels, for example) to quantify differences in performance and quality.

Innovation and Experimentation in Healthcare

All change in healthcare should be treated as an experiment. If the results are positive, adopt (and continue to tweak) the changes. If the results are negative, reject the changes or identify what additional changes must be made to obtain the desired results.

During the execution phase, quality teams require detailed data with a rapid turnaround to make quick adjustments to their efforts to maximize the amount of positive change (or to mitigate any negative effects the changes might have introduced). This sees a shift in the type of analytics required. During this phase, QI teams will be much more directed in the information they are seeking. The QI teams are likely dissecting existing processes and workflows and developing new processes. Depending on the methodology used and the time frame of the initiative, the QI teams will likely be changing the actual workflows and processes that staff members are performing. It is therefore likely that QI teams will require data that is closer to real-time, and much more specific to the desired processes. As processes and workflows are changed, QI teams will need to be able to see if the changes have actually led to a change in performance.

QI projects (especially ones that employ a PDSA methodology) usually begin with small-scale, localized changes as part of an initial evaluation. During the execution phase of quality group initiative, the data requirements of the team become very specific. QI teams will often break down processes into very minute detail, and will seek available metrics to measure the performance of these process components. For that reason, data and utilization requirements on QI projects are different from higher-level monitoring.

Using indicators that are monitored at a departmental or organizational level may not be sufficiently granular to detect a localized change over a period of time. For example, a new process resulting in a reduction in

hospital discharge times achieved over several days within a single unit may not even register on a more global indicator. For this reason, the output of analytics for a quality or performance improvement project in this stage should be:

- Relevant to the process or other change that is subject of the improvement effort.
- Focused locally on the department, unit, or other region where the change occurred.
- Available in near real time (or "short cycle") to allow for rapid adjustments.
- Presented in appropriate formats (such as run charts or statistical process control [SPC] charts) to evaluate both variability in the data and the magnitude of change in performance. (Remember that a reduction in process variability is a key step toward improvement.)

It is possible, and indeed likely, that a lot of information during this phase will not be available from existing sources. Don't be surprised if some required data is not even available in computerized form. For example, many improvement initiatives rely on audit data (that may not be available in electronic medical records or other systems) such as:

- Number of times a computer system (e.g., RIS/PACS) is not available (on downtime).
- Number of times a new process was followed correctly (as defined in standard work).
- Number of times a medical admission form was not completed properly.

Just because information may not be easily obtained or currently available does not mean that it is not important; in fact, very often these process components are overlooked. There are a few options to obtain this information. One option is to manually collect the data on audit forms placed throughout the unit and compile it in something like a spreadsheet. Although this may seem like a lot of work, often there is no other way to obtain such information. What may be possible is to link this manually collected and entered data to your current BI platform so that it can be integrated into existing dashboards or other performance reports for use by the team.

Using Analytics to Evaluate Outcomes and Maintain Sustainability

After the project execution phase and a change in performance or quality has been successfully implemented and tested on a smaller scale, the

project team naturally will deploy the improvement throughout other applicable areas within the HCO. Once the project has been deployed, continued monitoring and evaluation are necessary to ensure that the desired changes in quality and performance are occurring on the new, larger scale.

There are several possible approaches for evaluating the impact of an improvement project. For example, the team might want to compare quality or performance before and after the implementation of a new process or other innovation to determine if a change occurred, and what the magnitude of that change was. Or they may compare the impact of two different changes to determine which has the greater magnitude. For example, a department manager may wish to evaluate the impact of a new type of staff member (such as nurse practitioner or physician assistant) and compare baseline department performance to performance since the new role was added to determine the overall impact.

Once the desired changes are in place, analytics can be used to quantify the impact of the changes based on the initial indicators, and comparisons with benchmarks can be used to gauge progress toward meeting the designated targets. There are many ways to monitor performance—usually in

PROCESSES VERSUS OUTCOMES

When working on quality and performance improvement, there may be a tendency to focus on indicators relating to patient flow (interval times, lengths of stay, etc.). It is important to not lose sight of the indicators that are truly important—those that relate to patient outcomes. The bias toward patient flow indicators may be related to what is most conveniently obtained from electronic medical records systems. However, nobody would argue that a shorter length of stay is a benefit to the patient when that same patient is readmitted to hospital a few days later. Having said that, it is also *not* acceptable to keep a patient in hospital or in the emergency department longer than necessary simply to prevent a possible readmission (when the risk of readmission for a particular patient is not even quantified).

One of the goals of quality and performance improvement initiatives must be to balance the overall flow of an organization with the outcomes that are important to the patient. This is where the concept of value comes into play: How can HCOs maximize the value to patients (outcomes) while improving quality, efficiency, and safety? If value is increased, the likely outcome will be happier and healthier patients and HCOs that are more efficient and more profitable.

the form of performance dashboards or other reports. SPC charts are, once again, a very valuable tool to monitor the ongoing performance relating to changes being made. Just as is necessary during the project execution phase, evaluation results must be available quickly enough to take meaningful corrective action if necessary.

Sustaining Changes and Improvements

Careful analysis and redesign of healthcare processes can be successful at improving quality and performance. Without looking at the right data, or analyzing it correctly, it may be difficult to evaluate the impact of a change. And when a change does occur, it is less likely to be sustained without ongoing monitoring and corrective action.

Healthcare QI projects typically start off with a flourish, but all too often end with a whimper. This is because the excitement generated with a new project, new opportunities, and great expectations elevates QI teams with a feeling that anything is possible. This enthusiasm usually dies down by the end of a project, when team members look to other problems to address.

It is important not to lose sight of improvements that are made, so performance must continually be monitored to ensure that things don't revert to previous (undesired) levels of quality or performance.

This final stage is critical for QI initiatives, because sustainability is in fact one of the most challenging objectives to achieve. Many initiatives appear initially successful, but the desired performance begins to tail off after a few weeks or a few months. Monitoring during this phase must allow QI teams (and the HCO's leadership) to monitor ongoing performance of the improved processes. The actual metrics being followed during this phase may be fewer in number than during the execution phase, but the metrics chosen for monitoring and evaluation must be the most relevant to the performance desired. In addition to a performance dashboard highlighting the key indicators of a newly implemented improvement initiative, regular reports and automatic alerts that are e-mailed to key stakeholders draw the attention of QI team members when performance begins to deteriorate or otherwise deviate from desired parameters.

IMPROVING RETURN ON INVESTMENT

Many information technology (IT) projects are initiated, implemented, and deployed without ever defining or measuring the return on investment (ROI); and when an ROI is claimed, the values are often unclear at best and dubious at worst. In other words, many HCOs are unclear

as to what value many of their IT solutions in fact provide. Of course, analytics is not an exclusively IT undertaking (and should in fact be a strong partnership between the business and IT). But the sole purpose of BI and analytics within an HCO is to improve quality and performance. That would imply, then, that an ROI is in fact necessary—if no measurable improvement occurs, there is no return on investment.

The ROI of healthcare analytics should not be measured in terms of outputs of the system. For example, the number of reports, analytical applications, predictive models, and other analytical products is not a valid measure of ROI, since there is no indication of the value of these efforts. (In fact, it may be argued that more reports actually means less value!)

There are other types of value generated through analytics, such as in the areas of research and education. For example, the analytics teams that I work with have provided much value to clinical research efforts. The analytics infrastructure has made the extraction and analysis of data for research projects much more efficient; much data for research can be extracted and analyzed in the span of several hours with the tools now available (as opposed to the days and weeks necessary before much of the clinical data was available electronically). So ROI can be measured in terms of increased productivity in addition to real money saved.

Of course, quality and performance improvements cannot be entirely attributed to the use of analytics, but such metrics do provide a compelling measure of return on investment for QI efforts as a whole.

Note

1. Centers for Medicare and Medicaid Services, "EHR Incentive Programs," www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/index.html.

CHAPTER 9

Basic Statistical Methods and Control Chart Principles

If your experiment needs a statistician, you need a better experiment.

—Ernest Rutherford

There are complementary methods to measure the impact of a change or innovation on quality and performance—statistically and using control charts. Statistical methods to determine changes in performance rely on the performance of statistical tests to determine if changes in quality, performance, or other metrics are "statistically significant." Graphical approaches, on the other hand, use specialized charts known as statistical process control (SPC) charts (and specific rules to aid the interpretation of those graphs) to determine if a change in quality or performance is in fact occurring. This chapter discusses how both of these methods can be employed for quality and performance improvement.

Statistical Methods for Detecting Changes in Quality or Performance

I chose the epigraph at the start of this chapter rather tongue-in-cheek. My intent with the quotation isn't to say that statistics (and statisticians) should be avoided, but rather that the job of analytics professionals (including statisticians) is to make statistics more *accessible* and easily understood to all users of information through the use of the right tools in addressing the right problems (those of the quality and performance issues of the organization).

Statistics offers a wide range of methods with which to analyze quality and performance data. In this section, it is my intention only to introduce some basic statistical tests and terminology as they pertain to quality and performance improvement projects. I strongly encourage the reader to explore additional resources for more in-depth coverage of additional statistical topics.

Analytics is comprised of the tools, techniques, and systems necessary for obtaining deeper insight into the performance of an organization. Statistics is but one of those tools—an important tool to be sure, but not the only tool. I often say that analytics for healthcare quality improvement (QI) projects does not require *fancy* statistics, but rather *appropriate* statistics. A large part of the discussion in this book is about ensuring that high-quality data is available, that it is compiled into relevant indicators, and that it is made available to QI teams using structured methodologies. I have seen many analysis efforts become derailed because an analyst was overly concerned with applying complex statistical analysis, when this level of analysis was not necessary. Hopefully this section will demonstrate how statistics can be but one valuable tool in the quality and performance improvement toolbox.

Most QI methodologies do not require extensive statistical knowledge and the use of exotic statistical methods. Rather, statistical methods can best be used to identify any unusual variations in performance and to help pinpoint the causes of this variation.[1] This is more common in some methodologies (such as Six Sigma) that require more statistical analysis than others (such as Lean). Most QI projects can benefit from some basic statistical summaries such as average, median, and percentiles to report baseline information and current performance (as discussed in Chapter 6). Six Sigma delves deeper into more statistics, such as analysis of variance (ANOVA) and other tests, to either detect differences in performance or to increase the certainty of a result. In my experience, I have always tried to use statistics where appropriate and necessary to clarify and strengthen conclusions, not to search for a needle in a haystack.

If extreme statistical analyses are necessary to detect a change in performance, then a few scenarios are likely:

- The quality of data being used is poor.
- The wrong indicators were developed or used.
- There is no change in performance, or it is too small to be relevant.

Statistics are typically applied in one of two ways. *Descriptive statistics* are used to describe a large number of values or observations (representing an entire population or a sample thereof). There is a wide variety of descriptive statistics; the most commonly used ones include the mean (for example, the average weight of patients visiting a clinic) and the median (for example, length of stay [LOS] of emergency department patients). *Inferential statistics*,

on the other hand, analyze a sample of data to help evaluate and draw conclusions about a population. See Chapter 6 for a discussion on measures of central tendency and the use of descriptive statistics. This section will focus more on inferential statistics used to confirm the statistical significance of a change in performance.

Hypothesis Testing

Consider a facility that was observing longer than desired lengths of hospital stay and decided to implement a new streamlined patient discharge protocol. Prior to the implementation of the protocol, three months of baseline data showed an average length of hospital stay of 4.54 days. Following the implementation of the new protocol, the results were evaluated and the three-month post-implementation average length of hospital stay was 3.56 days. See Figure 9.1 for a graph illustrating the results. The difference pre- and post-implementation was 0.98 days. QI teams needed to determine whether this difference is the result of the new protocols, or whether the protocols made no difference and the observed difference is entirely by chance.

The process of determining whether this difference in values is due to natural variation and chance or the result of the change in process is known as *hypothesis testing* and typically involves a test of *statistical significance.* Because of natural variations in performance, no two sets of randomly selected data will ever be exactly the same even if the two samples are drawn from the same population of patients. Hypothesis testing and tests of statistical significance will help to determine if any observed differences between two (or more) groups are likely due to *actual* differences in the populations being studied (the result of a process change or other intervention), or if the observed differences are due to random variation and chance.[2]

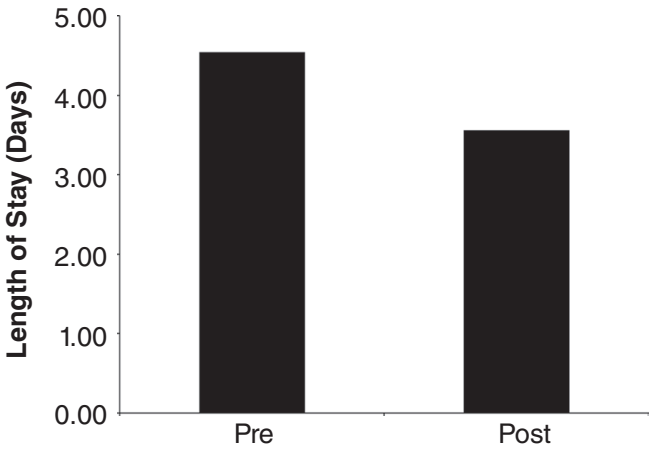

FIGURE 9.1 Sample Length of Hospital Stay before and after Implementation of New Discharge Protocols

Hypothesis testing starts with the assumption that there actually is no difference between the groups (that is, any observed differences are caused by random variation) unless there is compelling evidence to demonstrate otherwise. This is called the *null hypothesis*, and is expressed as:

$$H_0: \mu_1 = \mu_2 \text{ or } H_0: \mu_1 - \mu_2 = 0$$

The null hypothesis states that the means of data sets 1 and 2 are equivalent (that is, subtracting the mean of one data set from the mean of the other would return zero). In the case of the streamlined discharge protocol, H_0 states that there is no difference in the mean hospital LOS before and after the implementation of the new protocol—or that the new protocol had no effect on patient LOS. In the event that the null hypothesis is demonstrated to be false, the *alternative hypothesis* is then assumed to be true (that is, that the means of data sets 1 and 2 are not equal, and the differences observed between two data sets are likely *not* due to chance).

$$H_a: \mu_1 \neq \mu_1 \text{ or } H_a: \mu_1 - \mu_2 \neq 0$$

In our discharge protocol example, the alternative hypothesis is that there is a true difference in the means of discharge times measured before and after the protocol was introduced, suggesting that the new protocol *did* have an effect on patient LOS in hospital.

Comparing Performance between Two Groups

One common statistical test to evaluate situations like the pre-post evaluation of the protocol implementation is the *t*-test. The *t*-test is a statistical method that can be used to help determine if a statistical parameter (such as the mean, or average) is the same when compared between two groups (the null hypothesis) or different (the alternative hypothesis).[3]

LEARNING MORE ABOUT STATISTICS

If you are interested in learning more about the scientific and statistical basis behind hypothesis testing and statistical significance, there are many good statistical textbooks that cover these topics. I would also encourage you to review the resources listed in this chapter for more information. You can also visit this book's web site, http://HealthcareAnalyticsBook.com, for links to relevant resources.

A *t*-test can be used in two situations, depending on the number of samples. The *one-sample t*-test is used to compare one point of interest to a sample. For example, the one-sample *t*-test can be used to compare a sample's average performance to the target value of an indicator. If an emergency department's average left without being seen (LWBS) is 3.4 percent and the target LWBS rate is 2.5 percent, a one-sample *t*-test could be used to determine if the difference between actual performance and the target value is statistically significant.

A *two-sample t*-test is used to compare the performance of two groups. There are two varieties of this type of *t*-test; the best one to use depends on the two samples being tested. For example, if you are testing the hospital LOS at two different hospitals—Hospital A versus Hospital B—then the test to use is the *independent t*-test. The independent *t*-test assumes that the two populations are indeed independent, and are normally distributed. The independent *t*-test would not be appropriate in our example of the pre-post analysis of the streamlined discharge protocols. A pre-post study evaluation, also known as a *repeated measures* design, requires use of the *dependent t*-test variant.

A *t*-test can be applied to the "pre-change" and "post-change" groups in the example highlighted in Figure 9.1 to see if the difference of 0.98 days is statistically significant. Normally, statistical tests such as the *t*-test would be performed in a statistical software package or a spreadsheet with statistical capabilities. In the case of our discharge protocol example, running a dependent *t*-test on the two groups generates the following output from the statistical software used to run the test, which in this case is R:

```
t = -33.3139, df = 89, p-value < 2.2e-16
alternative hypothesis: true difference in means is not
  equal to 0
95 percent confidence interval:
 -1.0385635 -0.9216485
sample estimates:
mean of the differences
               -0.980106
```

What do these results mean? Consider if we repeated the discharge protocol pre-post test a second time and the results were similar, with a difference of 0.92 days; chances are our confidence in the results would improve, with two repeated tests demonstrating the same trend. Now, if we repeated the pre-post test 100 times and found that 95 out of these 100 times produced similar results, our confidence would be pretty high that the discharge protocols actually did decrease LOS for patients. If we repeated the test 100 times and found that the LOS of the protocol patients

was shorter than nonprotocol patients for only 60 of the trials, we would be less confident in the results. Finally, if each group had the shorter LOS 50 percent of the time over 100 repetitions of the pre-post test, we would likely deem that the protocols did not result in shorter lengths of stay.

Of course, it would be extremely time consuming and expensive to repeat trials such as our discharge protocol evaluation the necessary number of times to fully gauge confidence in results. This is where statistical tests are very useful, to determine how confident we can be that any differences observed are the result of a process change, or whether the observed difference likely occurred by chance.

More formally, the statistical significance is the probability of obtaining the observed (or more extreme) results if the null hypothesis were in fact true.[4] This chance or probability is calculated on the basis that the null hypothesis is correct; the smaller this chance, the stronger the evidence against the null hypothesis. Statistical analyses such as the *t*-test provides a quantitative assessment of this confidence with the *p*-value. A common *p*-value target often used in scientific research and QI is 0.05 or less, which means that there would be less than a 5 percent chance of obtaining the observed results if the null hypothesis was true.

In the previous example, the *p*-value is estimated to be less than $2.2e^{-16}$ by the computer software, suggesting that there is an extremely small chance of observing an LOS difference of 0.98 days if there was in fact no difference between the protocol and nonprotocol groups. This small *p*-value can provide the QI team with confidence that the discharge protocols actually are making a difference in lengths of stay of patients.

Another value reported by the computer software on the example above is the 95 percent confidence interval (CI). The CI is a computed range of numbers within which the true value is expected to lie.[5] In this case, the *t*-test calculated that the 95 percent CI, or the range of values in which the difference in LOS for protocol and nonprotocol patients can be expected to lie, is most likely between −1.039 and −0.922 (with rounding). In other words, we can say there is only a 1 in 20 chance that the true difference between the groups is *not* within that range.

If the CI included zero (for example, if the CI was between −0.5 and +0.5), it would imply that "no difference" in means, or a difference of zero, was as likely as other values within the CI. Because zero is not within the CI, however, it is likely that there is in fact a true difference.

The description of the *t*-test and test of significance earlier is to provide a flavor of how tests of statistical significance can help determine whether an actual change is occurring in a process. The *t*-test is ideal for comparing two groups, but what if more than two groups need to be tested at once, or you needed to test categorical data, or if other assumptions required of the data to perform a *t*-test are not met?

Comparing Performance of More Than Two Groups

What if an analyst needs to compare the performance of more than two groups? For example, consider the case where an analyst needs to compare the hospital LOS between three different facilities for patients who undergo a coronary artery bypass graft (CABG) procedure. Table 9.1 illustrates sample CABG patient LOS for three different facilities—A, B, and C. What would be the best approach to determine if there is a statistically significant difference between groups, or if the differences observed are simply the result of random variation?

The first instinct might be to perform pairwise comparisons—that is, compare A to B, A to C, and B to C using standard *t*-tests. This approach has two drawbacks. First, as the number of groups to compare grows, the number of pairwise comparisons that are required becomes unwieldy very quickly; after just seven groups, the number of comparisons required would be 21. Technically, we can get computers to run multiple *t*-tests quite simply, so the number of comparisons is not really a concern. However, as more *t*-tests are performed, the risk of obtaining a statistically significant difference *purely by chance* increases. Although there are corrections (such as the Bonferroni correction) that can be made for this when performing multiple tests (such as *t*-tests) on the same set of data, other statistical options are available.

In this case, and other cases when the *t*-test is not appropriate to use, there are other tests that can be used. The ANOVA (*an*alysis *o*f *va*riance) test is helpful when you need to compare more than two samples to each other to determine whether any of the sample means is statistically different from the other sample means.[6] A *one-way* ANOVA is valid if the groups are independent (as three sites would be), the data is normally distributed, and the variance in the populations is similar. Without going into the formulas, ANOVA works by comparing the variance *within* each group with the variance *between* the groups, and comparing the ratio of the within-group and between-group variance; the ratio is known as the *F-statistic*. If the variation between groups is much higher than the variation within groups, and the F-statistic exceeds a critical value, then a difference observed between the groups can be considered statistically significant. (Note that the critical F-statistic value can be looked up on a specially designed table of critical values, but more often than not, this will be performed by computer.)

TABLE 9.1 Sample Hospital LOS for CABG Patients for Three Sites

Facility	Hospital Length of Stay (days) for CABG Patients
Hospital A	8.5
Hospital B	9.8
Hospital C	8.9

Comparing Observations of Normal and Ordinal Values

What if the data that needs to be compared between two (or more) groups is nominal or ordinal, that is, data for which a mean cannot be generated for a test like a *t*-test or an ANOVA? A *chi-square test* is useful for determining if there is in fact a relationship between two categorical variables,[7] and would be appropriate in this situation. Rather than comparing a statistic such as the mean of two or more groups, the chi-square sums the squared differences observed and expected frequency of observations within each category.[8]

The different scenarios for which you may consider using the different types of ANOVA tests include:

- **One-way between groups.** Use the one-way between-groups ANOVA when the performance of three or more groups needs to be compared (as in the above example).
- **One-way repeated measures.** When performance has been measured a few times (for example, prior to a QI project, during the execution of the QI project, and after the QI project), the one-way repeated measures ANOVA can test for a statistically significant change in performance.
- **Two-way between groups.** This is used when looking for more complex interactions. For example, if comparing hospital LOS for CABG procedures, QI teams may be interested in understanding the interaction between whether the site is a teaching hospital or community hospital, and the overall hospital LOS.
- **Two-way repeated measures.** This is similar to the one-way repeated measure, but includes an interaction effect (for example, if you wanted to test whether type of X-ray had any impact on changes in the processing time of diagnostic imaging patients).

Lessons Learned

Many software packages include statistical tests built in, so it is relatively simple to perform a *t*-test, ANOVA, or other statistical test. Keep in mind that although software makes it easy, applying a statistical test to a set of numbers on a spreadsheet may not achieve accurate results. Before proceeding with a statistical test, always ensure that the basic assumptions required of the test are met (for example, does the test require normally distributed data?), and that the test can provide the type of answer being sought.

Graphical Methods for Detecting Changes in Quality or Performance

Hypothesis testing and tests of statistical significance are one method of determining if any change is occurring in quality and performance within a healthcare organization (HCO). The challenge with statistical tests, however, is that most require large samples of data to be accurate, and can be cumbersome to run every time the performance or quality associated with a process needs to be measured. Another issue is that they tend to utilize aggregated data (for example, determining if the mean of two samples is statistically significant). If all analysis is done in aggregate, it is possible to lose sight of variations in the way that processes are performed and in the outcomes of those processes. One danger of solely relying on aggregate data and statistical analysis is that although *average* values of data sets might be meeting a target value, individual performance and quality may vary so widely that the inconsistency poses a risk to patient safety.

Control charts are a very common visual approach to evaluate performance and quality with associated rules to determine if a process is in control and improving (or getting worse). Graphical analysis is a highly regarded approach in healthcare QI. It has been recommended that "methods for the analysis of data should be almost exclusively graphical (always including a run order plot), with minimum of aggregation of the data before the initial graphical display."[9]

Graphical analysis of performance data provides visual evidence of the variability inherent in a process. Measuring and understanding the variation in a process is merited because it is "important to eliminate extraneous process variation wherever possible, while moving well-defined metrics toward their target values."[10]

Variation in Performance

There are many different causes of variation in performance. Causes can range from differences in the way individuals perform tasks to calibration differences in equipment. All the different causes of variation, however, can be divided into two categories:

1. **Common (or random).** These are causes of variation that are inherent in the work being performed, affect everyone who performs the work, and affects all outcomes of the process.[11] Common cause variation is generally predictable and expected and can be caused by myriad reasons ranging from complexity of patient needs to materials available. An example is the natural variations in the time it takes to triage an emergency patient; although every triage is different because each

patient presentation is unique, there is a typical range in the time it normally takes to complete a patient triage.

2. **Special (or assignable).** These are causes of variation that are *external* to the system or the work being performed, and do not occur all the time; they arise due to special circumstances.[12] An example of special-cause variation would be a nurse who takes significantly longer to triage patients than is typical. This may be caused, for example, by a nurse who is improperly trained on the use of the triage system.

Quality not only means that a process is able to meet target performance on average, but it must accomplish this within certain tolerances and consistency; that is, it must be considered *stable.* A stable process refers to one that is free of special-cause variation. The term "in control" is also used when variations in data are present and exhibit a pattern that is random.[13] (Note that "in control" does not mean an *absence* of variation, since even the best processes will demonstrate some variability.) In addition, a statistically "in control" process may still not be acceptable if the variation falls outside a range that is deemed safe or otherwise acceptable by the HCO, clinical experts, or governing bodies.

One of the tenets of process improvement is that a process must be stable before it can be improved. Strictly speaking, even the act of changing a process from one that is out of statistical control to one that is within statistical control (i.e., with reduced variability in the output of the process) can be considered an improvement.

Almost every report showing any metric will display some variability in the performance of a process. No process in healthcare is so stable that it is able to produce the same results every single time. The question is how to determine how much variability in a process is too much, and how much is acceptable. Statistical process control (SPC) is a technique that QI teams use to improve, evaluate, predict, and control process through control charts.[14] In essence, an SPC chart is the chronological time series plot of an indicator, metric, or other important variable and is used for, among other things, analyzing the occurrence of variations within a process. Many statistics can be plotted on an SPC chart, including averages, proportions, rates, or other quantities of interest.[15]

Rather than simply plotting values on a graph, one of the unique components of SPC charts is the addition of upper and lower reference thresholds, which are called *control limits*. The control limits are calculated based on the process data itself; the plotted points of data must almost always fall within the control limit boundaries, as the control limits specify the natural range of variation within the data. Points falling outside of the control limit boundaries "may indicate that all data were not produced by the same process, either because of a lack of standardization or because a change in the

process may have occurred."[16] When looking for changes in performance, then, a reduction in variation and/or a deliberate and consistent shift to values near (or outside of) the control limits may signal that changes in a process are occurring.

Statistical Process Control Chart Basics

Many analytics tools with even basic visualization capabilities can be used to generate SPC and run charts. There are some stand-alone software tools (as well as plug-ins for Microsoft Excel) that can generate excellent SPC and run charts (and provide other visualization tools for quality and performance improvement). Even without dedicated SPC generation capabilities, very useful charts similar to SPC charts can be generated with the basic graphics capabilities of most analytics and business intelligence software provided that the basics of SPC charts are understood (and a little creativity is applied).

Tip

For a list of software tools that can be used to generate SPC and run charts, and examples on how to build them, please visit this book's web site at http://HealthcareAnalyticsBook.com.

See Figure 9.2 for a sample control chart. The important features of control charts are:

- **Data points** that represent a quality or performance indicator associated with a process (and may be a statistic such as mean or proportion).
- A **centerline (CL)** that is drawn at the mean value of the statistic.
- An **upper control limit (UCL)** and **lower control limit (LCL)**, which represent the values outside which performance of the process is considered statistically unlikely.

The centerline of a control chart is drawn at the mean ($\bar{x}$) or average value of the observations being plotted. Upper and lower control limits are typically drawn at $+3\sigma$ and -3σ (where σ is one standard deviation) from the centerline. The sample SPC chart in Figure 9.2 demonstrates a process that would generally be considered to be in control. All the data points are randomly scattered around the mean ($\bar{x} = 9.20$) and all fall within the upper control limit (UCL = CL$+3\sigma$ = 9.47) and the lower control limit (LCL = CL -3σ = 8.94).

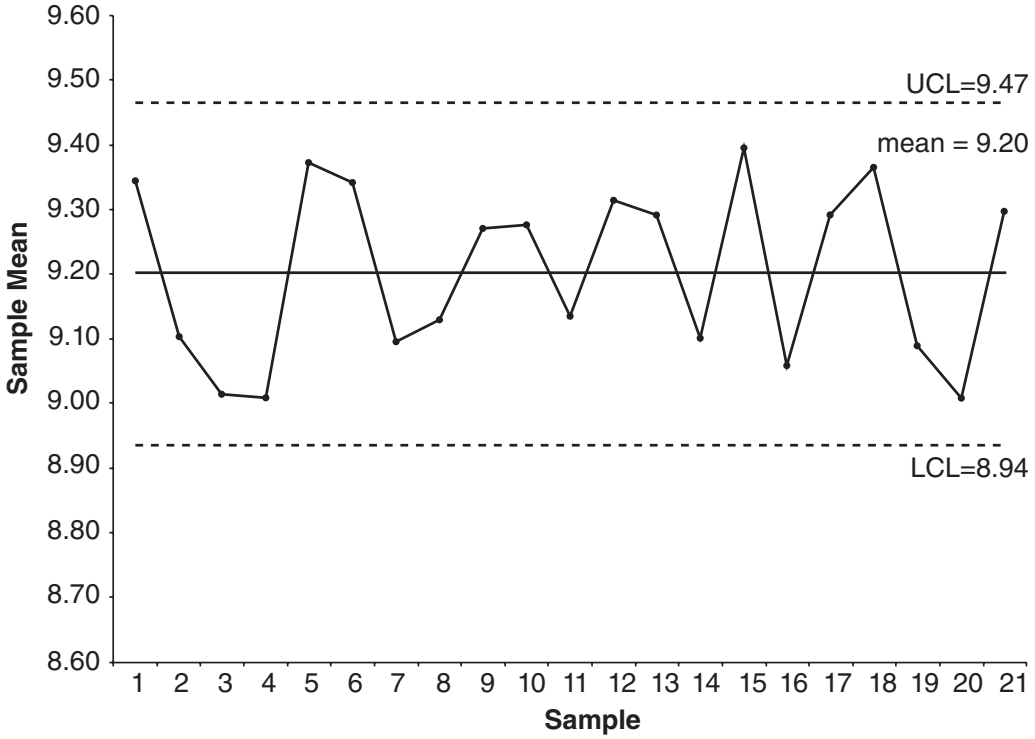

FIGURE 9.2 Sample Control Chart

When a change in process occurs as the result of a QI activity (or some other cause of change), the SPC chart can be used to monitor if a change in performance and/or outcomes has occurred. The limits of an SPC chart should be revised when the existing limits are no longer relevant or useful. When a shift in process occurs, it is helpful to reset the mean and control limit lines to better isolate the new process from the old process in the chart. If new mean and control limits are not reset, the existing mean and control limits will expand (or otherwise adjust) as new data is added to the calculations. This may make it more difficult to identify any actual change in outcomes or performance.

Figure 9.3 illustrates an updated version of the Figure 9.2 chart, this time with new data points added after a process change, and with the new mean and control limits added. With this chart, the baseline performance is shown, the time at which the new process was introduced is clearly evident, and the performance of the new process stands out from the baseline data. When the SPC chart is drawn in this way, the new performance can be evaluated not only to see if the desired target performance is being met but also to investigate the stability of the new process and whether it is in control.

Data Considerations for Statistical Process Control Charts

As mentioned earlier, when developing analytics for quality and performance improvement, it is important to use the right data, and that the

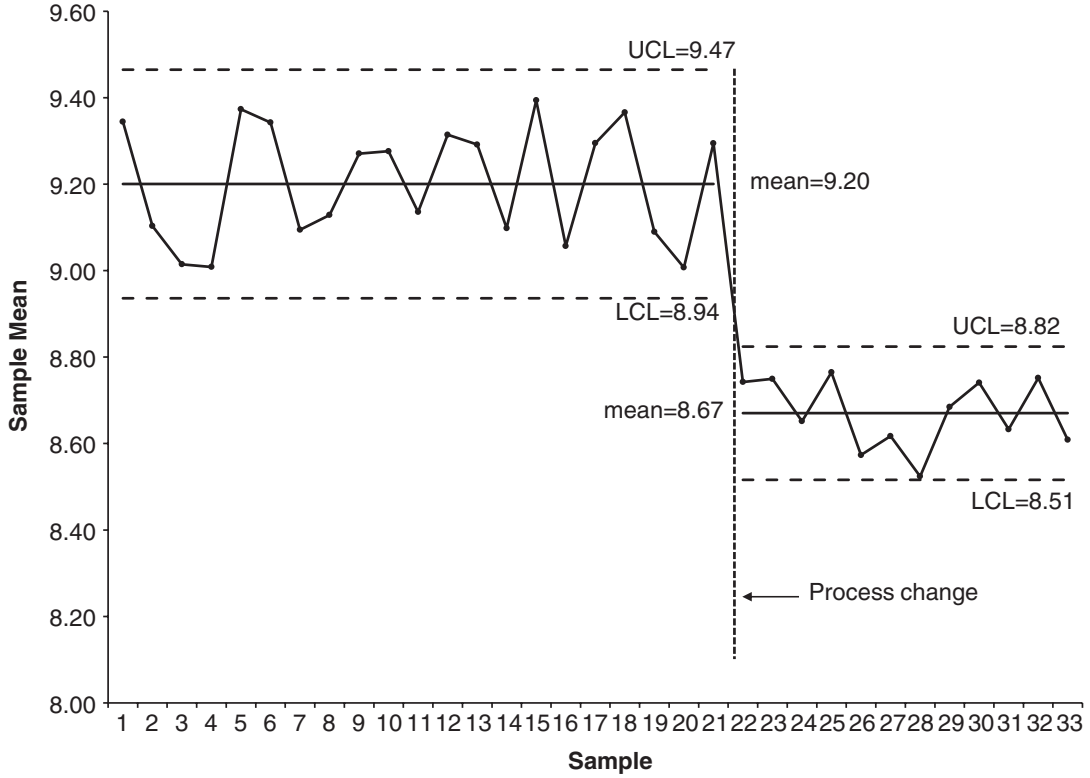

FIGURE 9.3 Sample Control Chart Highlighting Performance before and after a Change in Process

underlying assumptions of any statistical test or other tool are being met, otherwise inaccurate results are possible. SPC charts are no different; there are a few data considerations to ensure that SPC charts are accurate and that any conclusions drawn from them are valid.

When obtaining data for SPC charts, it is recommended that there be a minimum of 20 to 30 consecutive subgroups,[17] which are comprised of at least 100 consecutive observations.[18] For example, if the sample SPC chart in Figure 9.2 was evaluating the emergency department LOS for patients to be admitted, it would be ideal to plot at least 100 admissions over a period of 20 days for optimal validity of the control chart. In this example, assuming that there are at least five admissions from the emergency department every day, each subgroup would be the average LOS for admitted patients over a 20-day period. The mean for each subgroup would be plotted on the *y* axis, and each day would be plotted *in chronological order* along the *x* axis. This number of observations is necessary because, as in most evaluation methods, insufficient data may lead to inaccurate results.

Graphically Displaying the Stability of a Process

As long as the basic data requirements are met, a change in process can be quite clearly identified on an SPC chart. It takes more than simply

"eyeballing" it, however, to determine a change in performance or to detect undesirable variations and trends in the data. Figure 9.4 outlines a set of rules that can be used to determine the stability of a process based on data plotted on a control chart. The rules help quality teams to interpret the process patterns on the charts, specifically to special causes of variation. Figure 9.5 is a visual representation of the rules specified in Figure 9.4.

Different patterns that manifest on control charts may signal different issues or different causes of variation. In manufacturing and other industries, many of the rules help detect problems with machinery and other manufacturing issues. Healthcare is in many ways much more complex than manufacturing, so changes in control charts may be caused by any number of reasons. For example, a single point above the UCL (or below the LCL) may indicate that a single abnormality occurred that day (with possible causes ranging from a multicasualty incident causing a surge of patients at the emergency department to a lab equipment glitch requiring all blood work to be redone).

Some of the indications in Figure 9.4 and Figure 9.5 that a significant change has occurred or a process may not be in control include:

- Eight (or more) points in a row above or below the centerline
- Four of five points between $+1\sigma$ and $+2\sigma$ (or -1σ and -2σ)
- Two of three points between $+2\sigma$ and the UCL (or -2σ and the LCL)
- Any one point above the UCL or below the LCL

When reviewing SPC charts, the important point to remember is that any time the chart stops exhibiting random variation and patterns begin to manifest in one or more of the ways described, it is an indication that *something* is causing a process to change, whether as the result of deliberate intention or due to inconsistent practices, performance, or other causes.

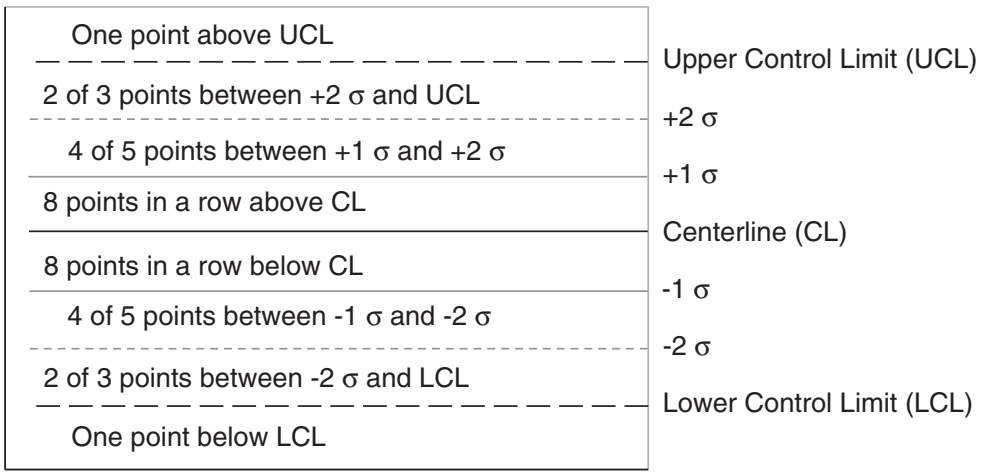

FIGURE 9.4 Detecting Stability in a Process Using Control Charts

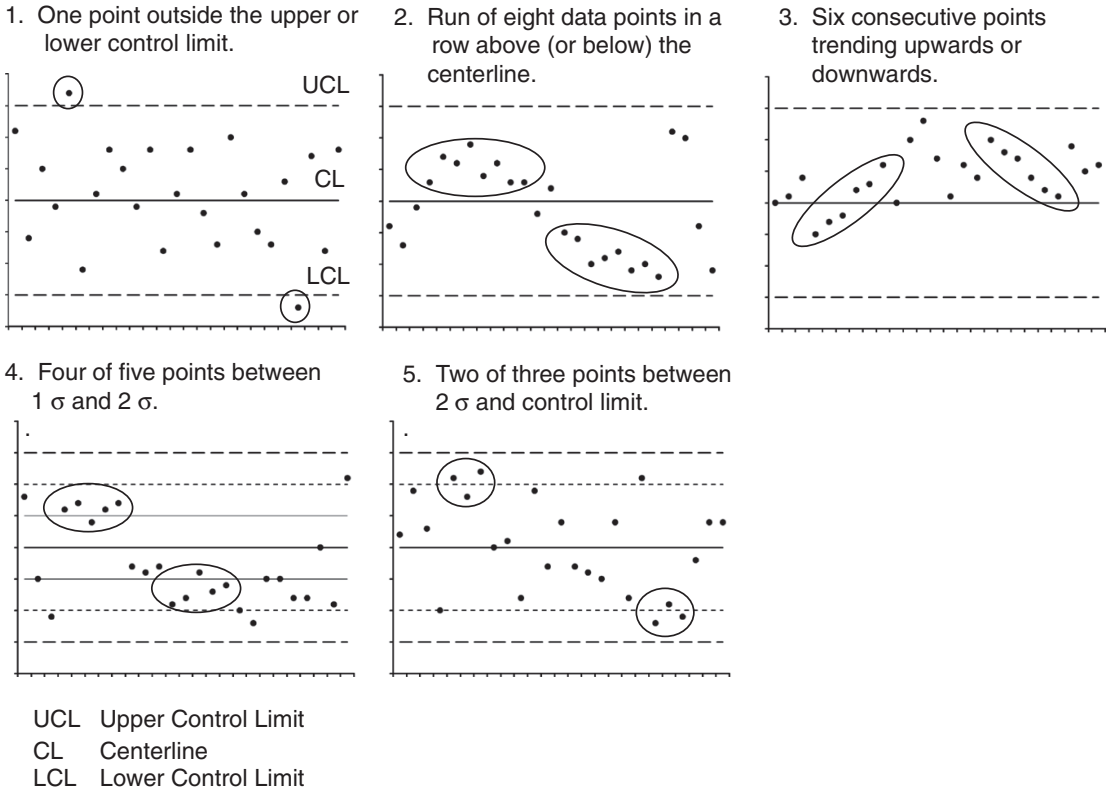

FIGURE 9.5 Sample SPC Chart Rule Violations

Keep in mind that even though a value is within the UCL and the LCL, it might not be acceptable from a clinical perspective. In addition to the σ and control limit values on an SPC, often a *specification limit* will be added. The specification limit is the range of values that is acceptable to the customer (or, in this case, to the patient and/or best clinical practice guidelines).

Consider, for example, an HCO that is improving its care of patients who experience a stroke. To achieve acceptable clinical standards, the HCO might identify a target duration of three hours from the time a patient experiences a stroke to the time rt-PA is administered. If rt-PA administration times for patients are plotted on an SPC, a specification limit of three hours would be added as a visual indicator. In this case, rt-PA administration times that were within the UCL (meaning within statistical control) but outside of the specification limit would still be a cause for concern. As the assessment and treatment of stroke patients was improved and variation in performance decreased, it is likely that UCL and LCL would tighten to the point where they were inside (or very close to) the specification limit.

Types of Statistical Control Charts

This chapter discusses the fundamentals of SPC charts, and there are actually several different kinds of SPC charts that can be used.

TABLE 9.2 Examples of Common Control Chart Types and How They Are Used

Data Type	Chart Type	Usage
Discrete	P-chart	Percentages
	C-chart	Counts
	U-chart	Rates
	T-chart	Time between rare events
Continuous	I-Chart (Sometimes called X-MR, where MR = moving range)	Individually measured data points
	X-Bar	Subgroups of data at the same point in time

Source: www.qihub.scot.nhs.uk/knowledge-centre/quality-improvement-tools/shewhart-control-charts.aspx.

The selection of a type of control chart depends on several factors,[19] including:

- The type of data being used (continuous versus discrete).
- Sample size available.
- What is being plotted (such as percentages, counts, rates, or time between rare events).

Table 9.2 shows a collection of common control chart types, what type of data they are appropriate for, and how they are used.

A summary of the different types of SPC charts, and a guide to selecting the best one for your particular needs, is downloadable from the book's web site at http://HealthcareAnalyticsBook.com.

Putting It Together

Critical to the development of analytics is the knowledge of who needs to use information, and how they can best make use of it. For example, QI experts working with Six Sigma and other methodologies often use SPC charts and statistical analysis in raw form to study the performance of a process and its resultant quality. In doing so, they are using SPC charts and statistics to analyze one (or a few) process changes in depth to uncover opportunities for further performance improvements and changes in quality. Healthcare executives, managers, and other healthcare leaders, who are usually concerned with the operations of an entire unit, hospital, or system,

on the other hand, are not likely to benefit from SPC charts and *t*-tests outlining every performance indicator relevant to a QI activity.

Knowing who needs what kind of information is important to developing effective analytics. Analytics is able to provide deeper insight into the performance of an HCO, and is designed to make decision making easier for QI teams and healthcare leaders. Statistical analysis and SPC charts were invented long before modern analytics. The power of analytics is in the synthesis of information and insight from statistical and graphical analysis into more meaningful and easier-to-interpret formats where appropriate, and presentation of more detailed information when necessary.

Rather than simply displaying a collection of graphs and charts, dashboards and reports can be made more analytical by embedding insights gained from statistical analyses and graphical analysis. Figure 9.6 illustrates a sample dashboard for a diagnostic imaging department displaying several key performance indicators for that unit, baseline performance for the previous six months, the indicators' respective targets, the current month's performance data, the performance for the previous month, and a trendline of performance over the last eight weeks. (See Chapter 10 for a discussion on dashboard and data visualization design.)

The dashboard is a simple representation of these indicators, provides an overview of performance, and also includes embedded insight from both statistical analysis and SPC rules. Using superscript values next to the current month's data and descriptive text in the "Notes" section, the dashboard in Figure 9.6 indicates to the user that there was: (1) a statistically significant difference in performance on indicator A between last month and the current month, and (2) an SPC rule violation for indicator D. The statistically significant ($p < 0.05$) decrease in X-ray order to patient pickup times (indicator A) from the last month to the current month may suggest to QI teams and DI managers that efforts to improve processes associated with indicator A might be having a positive effect.

Diagnostic Imaging Department - Performance Dashboard **December**

Indicator	Target	Six-Month Baseline	Last Month	Current Month	8-Week Trend
A. X-Ray order to patient pickup	10.0	33.2	22.5	13.4 [1]	
B. Pickup to X-Ray start	10.0	12.2	11.9	11.8	
C. X-Ray duration	15.0	14.5	15.2	15.1	
D. X-Ray stop to patient return	10.0	12.1	11.9	17.5 [2]	

(All values are in minutes)

Notes:
1) p<0.05 - month over month performance
2) SPC rule violation: two of three points between +2σ and UCL (**Link** to SPC chart)

FIGURE 9.6 Sample Dashboard Including Embedded Analytics

Although dashboards are often designed to be printed out, they are most useful when designed to be viewed on a screen (such as a computer display, smartphone, or tablet) and interacted with. Such interaction allows users to drill down into more detail (such as to view a control chart for performance indicators), select additional or other indicators to view, and even manipulate date ranges and other dashboard parameters. The example in Figure 9.6 shows that the SPC rule violation alert also includes a link to the actual SPC chart that triggered the violation, so an interested user of this dashboard would be able to launch an additional view that contains the desired additional material. The design objective is that the most important features and insight of this dashboard are immediately visible, with additional detail (links to an SPC chart) available via a simple click.

Statistical and nonstatistical approaches to evaluating quality and performance are not at odds but are entirely complementary. When used in concert, and synthesized on interactive information displays such as dashboards and other analytical tools, users of the information can quickly identify where and when performance needs to be improved, and perhaps even what actions need to be taken.

Developing effective analytical tools does require effort and expertise. Analytical teams must understand the context of the data, know control charts (and their associated performance variation rules are used), and be comfortable in enabling the basic statistical analyses required. Although the mechanics of performing the required statistical analyses and building the appropriate charts, graphs, and other data displays are possible in many analytical tools, it still requires a knowledgeable analytics team working closely in concert with quality and performance improvement experts to design concise and effective analytical information displays that provide insight about quality and performance issues, help suggest appropriate mitigation steps, and monitor ongoing results.

Notes

1. Thomas P. Ryan, *Statistical Methods for Quality Improvement, 3rd ed.* (Hoboken, NJ: John Wiley & Sons, 2011), 9.
2. Michael L. George et al., *The Lean Six Sigma Pocket Toolbook: A Quick Reference Guide to 100 Tools for Improving Speed and Quality* (New York: McGraw-Hill, 2005), 156.
3. Ibid, 161.
4. David Freedman et al., *Statistics*, 2nd ed. (New York: W.W. Norton & Company, 1991), 433–435.
5. Ibid., 348–350.