RadioGraphics

Quality Initiatives

Statistical Control Charts: Simplifying the Analysis of Data for Quality Improvement[1]

Yvonne Y. Cheung, MD, MS • Boyoun Jung, BS • Jae Ho Sohn, MS
Greg Ogrinc, MD, MS

Scan this code for access to supplemental material on our website.

TEACHING POINTS
See last page

Quality improvement (QI) projects are an integral part of today's radiology practice, helping identify opportunities for improving outcomes by refining work processes. QI projects are typically driven by outcome measures, but the data can be difficult to interpret: The numbers tend to fluctuate even before a process is altered, and after a QI intervention takes place, it may be even more difficult to determine the cause of such vacillations. Control chart analysis helps the QI project team identify variations that should be targeted for intervention and avoid tampering in processes in which variation is random or harmless. Statistical control charts make it possible to distinguish among random variation or noise in the data, outlying tendencies that should be targeted for future intervention, and changes that signify the success of previous intervention. The data on control charts are plotted over time and integrated with various graphic devices that represent statistical reasoning (eg, control limits) to allow visualization of the intensity and overall effect—negative or positive—of variability. Even when variability has no substantial negative effect, appropriate intervention based on the results of control chart analysis can help increase the efficiency of a process by optimizing the central tendency of the outcome measure. Different types of control charts may be used to analyze the same outcome dataset: For example, paired charts of individual values (x) and the moving range (mR) allow robust and reliable analyses of most types of data from radiology QI projects. Many spreadsheet programs and templates are available for use in creating x-mR charts and other types of control charts. Supplemental material available at *http://radiographics.rsna.org/lookup/suppl/doi:10.1148/rg.327125713/-/DC1*.

Abbreviations: LCL = lower control limit, QI = quality improvement, UCL = upper control limit

RadioGraphics 2012; 32:2113–2126 • **Published online** 10.1148/rg.327125713 • **Content Codes:** QA RS

[1]From the Department of Radiology, Dartmouth Hitchcock Medical Center, One Medical Center Dr, Lebanon, NH 03756 (Y.Y.C.); Dartmouth Medical School, Hanover, NH (B.J., J.H.S.); and Department of Internal Medicine, VA Medical Center, White River Junction, Vt (G.O.). Presented as an education exhibit at the 2011 RSNA Annual Meeting. Received February 23, 2012; revision requested June 22 and final revision received August 5; accepted August 8. All authors have no financial relationships to disclose. **Address correspondence to** Y.Y.C. (e-mail: *yvonne.cheung@hitchcock.org*).

Introduction

Quality improvement (QI) projects are an integral part of contemporary radiology practice. Radiologists conduct QI projects to add value and safety to their practice, as well as to satisfy regulatory requirements: The completion of QI projects is a requirement for maintenance of certification by the American Board of Radiology; in addition, the Joint Commission, an independent organization that certifies hospitals and healthcare programs, requires that hospitals conduct QI projects for accreditation (1–5).

QI projects commonly involve the study of processes, analysis of data, and introduction of process change through intervention. A major challenge to the accurate analysis of outcome data is the random variation that is inherent in virtually all processes. Since random variation causes the observed results to fluctuate even without any intervention, how can we be certain that divergence in outcome data represents true change and not random variation due to common causes? Control charts provide a means for answering that question. Control charts are an important component in the statistical process control methods used in manufacturing to monitor, improve, and predict performance and reduce random variations (6). Control charts are analytic tools that allow a visual distinction between meaningful change and random variation or "noise" in a process by comparing the actual distribution patterns of outcome data with standardized distribution patterns derived from probability statistics (6).

Teaching Point

Although control charts are practical, easy to interpret, and ideally suited for use in the kinds of QI projects commonly undertaken by radiologists, they are still generally underused in radiology. Their use is seldom reported in the literature (7,8) and rarely described in published guidelines for performing radiology QI projects (9). This underuse could be due in part to an overall lack of familiarity with the principles underlying their creation and interpretation.

This article provides guidance for preparing and using control charts to analyze and manage key processes affecting the quality of radiology services. First, the conceptual basis, historical origins, and anatomy of control charts are outlined, and the advantages of their use are briefly described. Next, the type of variation present in a process, the type of process, and the type of outcome data are discussed with regard to their bearing on the creation of particular types of control charts for specific analyses. The principles that govern control chart creation and analysis are explained and illustrated with examples from QI projects relevant to radiology.

Conceptual Basis for Control Chart Analysis

Among the many different methods of statistical analysis available, radiologists are probably most familiar with classic hypothesis-testing methods learned in statistics courses, such as the χ^2 and Student *t* tests, which are performed to determine whether the averages for two compared groups are similar or different. Although these hypothesis-testing methods are standard in scientific research, they are of limited use when analyzing QI processes because the time element is not captured in the analysis, and the conclusions drawn by comparing the values from one point in time with those from another time point may be misleading. Control charts show changes in outcome measures over time and thereby offer a clear advantage over classic methods of statistical analysis. For example, hypothesis testing of data collected during a QI project in pediatric radiology showed a substantial difference between the numbers of computed tomographic (CT) scans performed in children before and after implementation of Image Gently™ guidelines (Fig 1a), a finding that might have led to the conclusion that the QI intervention was responsible for the difference. However, when the project team reviewed control charts on which the data were visually linked with the time points at which they were obtained, it became obvious that the decline in the number of pediatric CT scans actually preceded the Image Gently implementation (Fig 1b).

Teaching Point

Walter Shewhart (1891–1967), a physicist working for Bell Telephone Laboratories (now AT&T), invented control charts to help engineers produce telephone components that were of uniform quality. Shewhart theorized that to improve the quality of a process, one must control variation. His approach to statistical analysis was pragmatic and simple: He reduced all process variations to two types, and he created charts that would allow the detection of the type of variation occurring in a process. Shewhart further categorized processes into two funda-

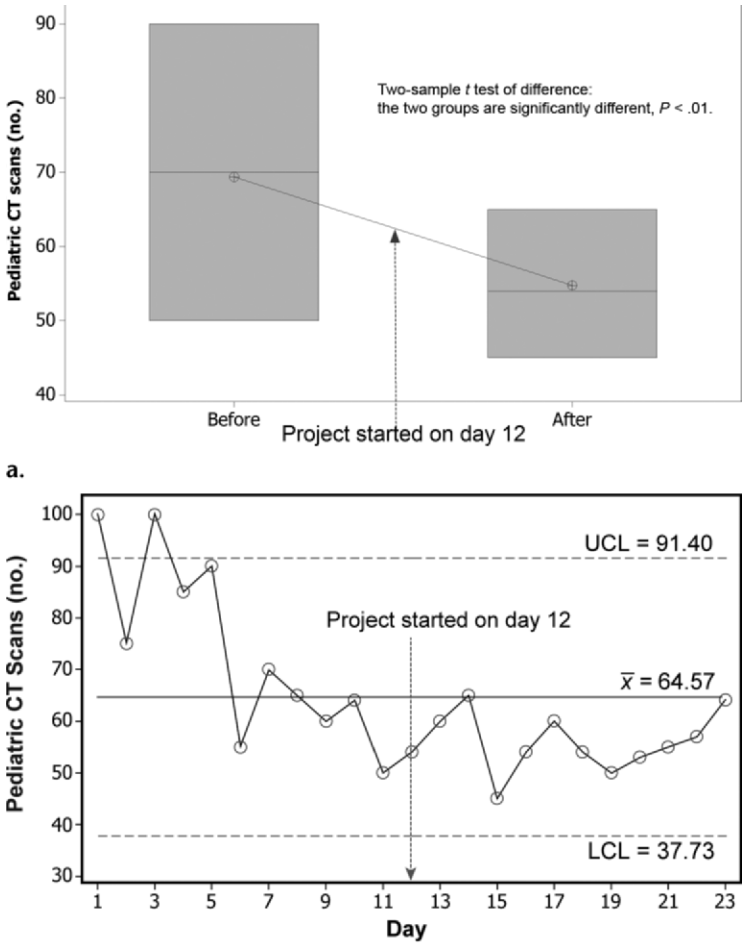


Figure 1. Charts of statistical results from an Image Gently QI project show the importance of including the time element in the outcome analysis. **(a)** Box plots show the numbers of pediatric CT scans performed during two periods of data collection, before and after the implementation of the QI initiative. Results of standard hypothesis testing showed a significant difference in the average values (data points on center lines) collected during the two periods. **(b)** Control chart incorporating the time of data collection shows that the decline in the number of pediatric CT scans actually began before day 12, when the QI initiative was implemented. The effect of the project was therefore difficult to determine. *LCL* = lower control limit, *UCL* = upper control limit.

mental types based on the type of variation occurring in them. Effective management of process improvement is simplified by understanding and responding to the fundamental type of process represented by the outcome data on a control chart. This conceptual framework constitutes the central theoretical basis of what is now known as statistical process control (6,10,11). Widely accepted in manufacturing, the use of control charts has expanded in recent years to the field of healthcare (12,13). For example, the Joint Commission (1) uses control charts to analyze performance in hospitals. Most QI processes that require data analysis, including those in radiology, are amenable to Shewhart's simple but powerful charting techniques.

There are many types of control charts from which to choose. The selection of an appropriate chart template for a specific QI analysis is facilitated by a straightforward characterization of key elements in a process: the type of variation present (either random and due to common causes, or due to "special" causes), the current state of the process (either in or out of statistical control), and the process outcome that is being measured (either a variable or an attribute).

Identifying the Cause of Variation: Common or Special?

Variation is inherent in all processes, and understanding and managing variation is key for achieving control over a process. Shewhart distinguished between two types of variation: variation due to common causes and variation due to special causes (6,13). Common cause–related variations are random and lead to outcome measures that are within a predictable distance from the mean. An example of a common cause variation is interobserver variability, which is encountered in routine radiology practices such as repeat measurements of the size of a lymph node: No two such measurements are ever exactly the same. By contrast, special cause–related variations reflect extraneous or assignable effects that are unlikely to be due to chance alone. For instance, a sudden increase in

Teaching Point

Table 1
Glossary of Statistical Terms Relating to Process Control

Term	Synonyms	Definition
Autocorrelation	…	Interdependence of observations in a time series, with resultant clustering of data (tendency for high values to follow high values and low values to follow low values)
Common cause–related variation	*Routine variation, random variation, noise, unassignable variation, predictable variation*	Inherent variation due to random fluctuations in a process; all variation in a process that is stable and predictable is due to common causes
Control limits	*Natural process limits*	Limits computed from collected data to allow the differentiation of predictable variation from unpredictable variation
Process in statistical control	*Predictable process, process in control, stable process*	A process in which only common cause–related variation occurs and in which 99% of the process measures are within control limits
Process out of statistical control	*Unpredictable process, process out of control, unstable process*	A process in which both special cause– and common cause–related variations are detectable by using Shewhart's rules
Special cause–related variation	*Exceptional variation, assignable variation*	New, unexpected, or previously overlooked signal not explained by chance alone
Statistical control chart	*Process behavior chart, Shewhart chart, control chart*	Shows trends in outcome measures over time; helps distinguish between variations due to common causes and those due to special causes
Statistical process control	*Method of continuous improvement*	Method of process improvement based on analyses of variation and control charts
Tampering	…	Changing a process in which only common cause–related variation occurs

the number of missed appointments ("no-shows") on a heavy snow day is unlikely to be due to chance. Many other terms can be used to describe these two types of variation (Table 1).

Neither type of variation is intrinsically good or bad. However, recognition of the type of variation occurring in a process can provide insight into its origin and proper management. Control charts combine outcome data with information about the time points of measurement and the statistical framework in a simple visual display that facilitates the detection of variations that are related to special causes. A set of "special cause rules" developed by Shewhart and others is helpful for distinguishing between special and common causes when analyzing control charts (see the section on "Guidelines for Analyzing Charts").

Assessing the Stability of the Process: In or Out of Statistical Control?

Shewhart described two categories of processes: those that are within statistical control and those that are outside statistical control. The distinction is based on the presence or absence of special cause signals. When special cause signals are detected in outcome data, the process is unpredictable, unstable, and out of statistical control. By contrast, only common cause–related variations are found in a process that is predictable, stable, and in statistical control. A baseline assessment of the stability of a process is helpful not only for understanding and managing present variability but also for predicting the future performance of the process.

Categorizing the Data: Variable or Attribute?

Outcome measures analyzed in a QI project can be categorized as either variable data or attribute data. The type of data determines which control chart is most appropriate for the analysis (Tables 2, 3).

Table 2
Control Charts for Analyzing Variable Data

Control Chart Type	Description of Uses and Limitations	Sample Size	Examples of Appropriate Outcome Measures
x (*I*)	Is used to assess variability of individual values, commonly in combination with the moving range; all special cause rules are applicable	1	Number of MR imaging examinations performed per shift, patients' wait time in minutes
mR (*m*oving *r*ange)	Is used to monitor variation by plotting the difference between two consecutive measurements over time; is never used in isolation; only special cause rule 1 is applicable	1	Number of MR imaging examinations performed per shift, patients' wait time in minutes
x-mR	Is used to assess process stability by visualizing the difference between individual values and that between ranges of values; all special cause rules are applicable to *x*, but only rule 1 is applicable to mR	1	Number of MR imaging examinations performed per shift, patients' wait time in minutes
$\bar{x}$ (average)	Is used to plot the mean value for a sample with more than one measurement; is never used in isolation; all special cause rules are applicable	>1	Daily average turnaround time for emergency department "wet reads," weekly average access times (interval between time of appointment request and time of appointment) for MR imaging examinations
R (*r*ange)	Is used to plot the range for a sample; is never used in isolation; all special cause rules are applicable	2–9	Daily average turnaround time for emergency department wet reads, weekly average access times (interval between time of appointment request and time of appointment) for MR imaging examinations
$\bar{x}$-*R*	Is used to assess process stability by visualizing the difference between sample averages and that between sample ranges; all special cause rules are applicable to both $\bar{x}$ and *R*	2–9	Daily average turnaround time for emergency department wet reads, weekly average access times (interval between time of appointment request and time of appointment) for MR imaging examinations
S (*s*igma, *s*tandard deviation)	Is used to show the square root of variance (ie, σ or SD) for a sample; is never used in isolation; all special cause rules are applicable	>10	Daily average turnaround time for emergency department wet reads, weekly average access times (interval between time of appointment request and time of appointment) for MR imaging examinations
$\bar{x}$-*S*	Is used to visualize the difference between averages and SD or σ levels for multiple samples; all special cause rules are applicable to both $\bar{x}$ and *S*	>10	Daily average turnaround time for emergency department wet reads, weekly average access times (interval between time of appointment request and time of appointment) for MR imaging examinations
G (*g*eometric distribution)	Is used for the analysis of rare events separated by time intervals that have a geometric distribution	>1	Days between infections after imaging-guided steroid injections into joints

Note.—Variable data are quantitative values expressed on a continuous scale with a single unit of measure. MR = magnetic resonance, SD = standard deviation.

Table 3
Control Charts for Analyzing Attribute Data

Control Chart Type	Description of Uses and Limitations	Sample Size	Examples of Appropriate Outcome Measures
p (*p*roportion or percent defective)	Is used to plot counts of defective outcomes (ie, defectives*) as a proportion or percentage of total outcomes, or to plot dichotomous outcomes that either meet or fail to meet the standard, within a specified time period of measurement; the formulas used to calculate control limits are based on the assumption of a binomial distribution of the data	Unequal	Proportion of inappropriate imaging examinations per 100 imaging examinations performed each month
NP (*n*umber defective)	Is used to plot the number of defective outcomes (ie, defectives*); assumes a binomial distribution of the data	Equal	Count of late arrivals per 100 appointments
U (*u*nequal areas of opportunity)	Is used to plot counts of defects† or events with a specific outcome that is negative or substandard; counts may be given as a proportion or percentage; assumes a Poisson distribution of the data	Unequal	Proportion of patient identification errors per calendar month
C (*c*onstant area of opportunity)	Is used to plot counts of defects† as a proportion or percentage of total outcome; assumes a Poisson distribution of the data; is similar to the *U* chart, except that the denominator is constant	Equal	Proportion of findings that are nonroutine per 100 CT examinations

*Defectives are counts of dichotomized outcomes (outcomes that either meet or fail to meet the desired standard).
†Defects are counts of substandard events contributing to a specific undesirable outcome. More than one defect may occur in the same sample or affect the same patient. When counts of defects are expressed as a proportion or percentage, the numerator (eg, number of identification errors of any type—wrong patient, wrong side, or wrong body part—occurring each month) is not necessarily a subset of the denominator (eg, the number of imaging examinations per month) and therefore may exceed it.

Variable data are measures that can be represented on a continuous scale representing all possible values. Examples of variable data are weight, height, CT radiation dose, wait time, and patient satisfaction scores. These data are expressed as numbers of units, such as milligrays for CT dose index (ie, CTDI) or minutes for wait time.

Attribute data are counts of similar or dichotomized outcomes or events. Examples of attribute data include the number of late arrivals versus that of on-time arrivals, cases with allergic reaction versus those without allergic reaction, cases with complications versus uncomplicated cases, and inappropriate examinations versus appropriate examinations. Attribute data that are counts of similar outcomes or events are expressed as whole numbers; those that are counts of dichotomized events are expressed as proportions or percentages.

Attribute data can be further subdivided into two categories: defectives (ie, counts of dichotomized outcomes) and defects (ie, counts of a specific outcome or event that is negative or substandard). An example of a defective is a missed appointment: The patient either shows up or misses the appointment. By contrast, more than one defect may contribute to an undesirable outcome or event: For example, imaging of the wrong patient, wrong body part, and wrong side are all defects resulting from identification errors that might occur in the same hospital admission, affecting the same patient. For this reason, when defects are expressed as a proportion or percentage, the numerator (eg, the number of all identification errors) is not always a subset of the denominator (eg, the total number of patients admitted) and may exceed it.

Anatomy of Control Charts

A control chart is a graphic display of data points (plotted on the x-axis) collected at regular intervals in time (plotted on the y-axis) and includes statistics-based components allowing a visual assessment of the relations between individual data points (Fig 2). The center line on a control chart indicates the average of the collected data points. The UCL and LCL denote the boundaries within which 99% of the data points will be found when the limits are set at a distance of 3σ (ie, 3 standard deviations) from the

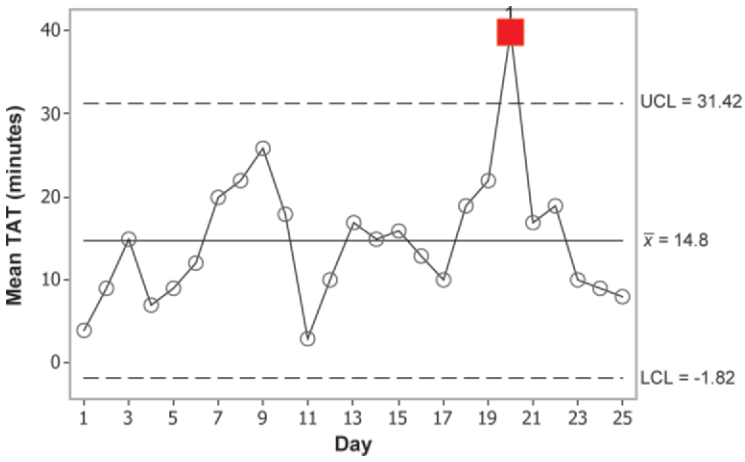


Figure 2. Control chart from a QI project to improve turnaround time *(TAT)* in the reading of emergency department radiographs shows the five components of all statistical control charts: the outcome measure on the y-axis; the time point of data collection on the x-axis; the average of all collected measurements, which is represented by the center line; the UCL, which is set at 3σ above the mean; and the LCL, which is set at 3σ below the mean. Note the special cause signal (red square), a single outlier beyond the UCL, on day 20.

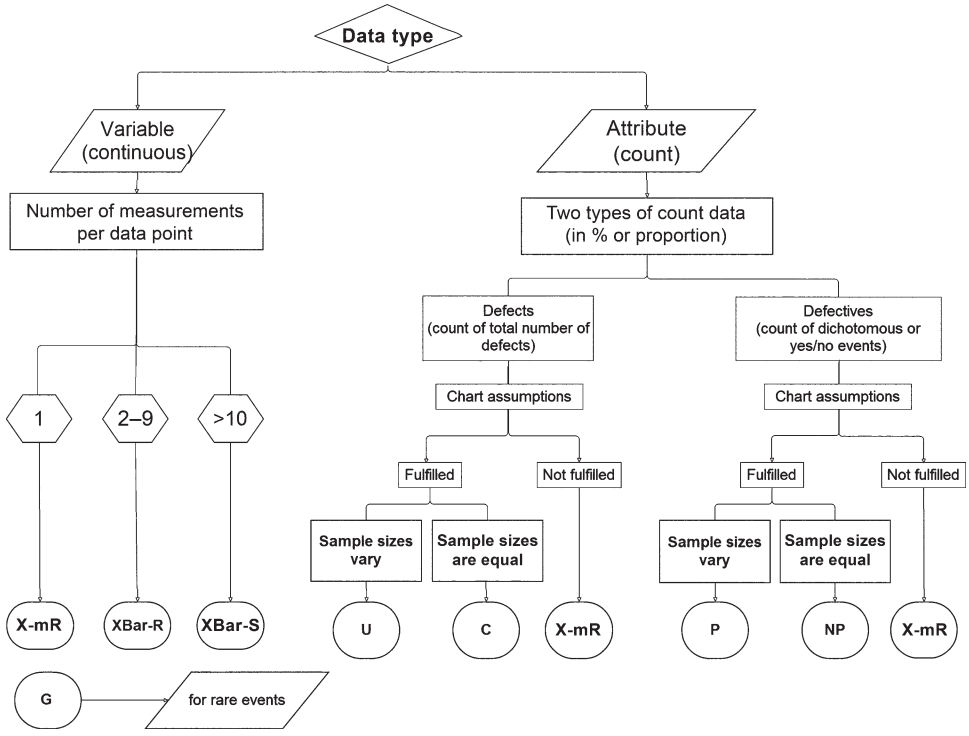


Figure 3. Flowchart shows decision-making steps in the selection of appropriate chart types for data analysis in a QI project. The term *sample size* refers to the denominator when the outcome measure is a proportion. (Adapted and reprinted, with permission, from reference 13.)

mean and when the data are normally distributed. The choice of a 3σ limit is a practical one that is designed to strike a balance between the likelihood of a type I error (ie, concluding that there is a special cause when none is present) and that of a type II error (ie, concluding that there is no special cause when one is present). Furthermore, the UCL and LCL represent the limits of expected performance measures in a stable process, before any change is introduced.

The formulas used to calculate the UCL and LCL are specific to the chart type. If the calculated limits are unrealistic, they must be reset: For example, in an assessment of radiology report turnaround times, a calculated LCL with a negative value must be reset to zero.

Guidelines for Selecting Appropriate Chart Types

Control charts are used both to perform an initial assessment of the state of a process and to monitor the effect of subsequent QI interventions in the process. Different types of control charts may be needed, depending on the specific goal of the analysis and the outcome measures selected for assessment. Different types of charts are constructed to represent different types of data and different subgroup (sample) sizes or numbers of observations represented by each data point. Each type of control chart is based on specific statistical formulas and assumptions about the distribution of data points (1,6,11,13) (Tables 2, 3; Fig 3).

Other considerations also may influence chart selection (Fig 4). In general, charts of continuous data (ie, variable data) are more powerful than charts of attribute data and therefore are preferred for process analysis. For example, although counts of "late arrival" are considered attribute data because the patient is either on time or late, "late arrivals" could be quantified not only as an absolute number or proportion but also in relation to a continuous variable such as time. Recording patients' arrival times so as to allow calculation of the time lost to arrival delays could result in an outcome measure substantially different from that produced by recording the mere number or percentage of late arrivals; for example, the effect on the process when 50% of patients arrive 2 minutes late can be expected to differ greatly from the effect when 50% of patients arrive 30 minutes late.

It is sometimes necessary to use more than one chart type to unearth special cause–related signal in a dataset. However, time may be needed to master the nuances of different chart types so as to select the optimal combination for analyzing a specific outcome measure. In the interim, it is appropriate to rely primarily on the *x*-mR chart for most analyses. This general-purpose chart is easy to use and robust to departures from normality in data distribution (14).

1. Data type is only one determinant in chart selection.
2. More than one chart type can be applicable in the analysis of one type of data.
3. Count data are weaker than continuous measurements.
4. The *x*–mR chart is versatile and commonly used as a general-purpose chart. It can be applied to analyze attribute data (count can be considered as a continuous variable).
5. Specialized charts are based on specific probability models of data distribution. For example, *p* and NP charts are based on binomial distribution of the collected data, *U* and *C* charts are based on Poisson distribution, and *G* charts are based on geometric distribution. If the normality of the data distribution is unknown, these specialized charts may not be the best choice. When in doubt, use the *x* chart.

Figure 4. These principles help guide the selection of appropriate chart types for analyzing outcome measures in a QI project.

Chart Types for Analyzing Continuous Data

Charts of x, $\bar{x}$, mR, R, and S are used to monitor both the average for, and variability across, sampled values. The x (or individual values) chart is used to analyze samples that cannot be subdivided, such as the number of MR imaging cases per month ($n = 1$). The $\bar{x}$ (or averages) chart facilitates monitoring of the outcome when each collected data point represents the average of multiple measurements. For instance, during a study of the daily report turnaround time for a period of 14 days, five random measurements might be collected daily; in this case, the total number of data points on the chart is 14, and the sample size (ie, n) is five. The mR (or moving range) chart allows monitoring of variability in a process by plotting the difference between two successive observations over time. The S (standard deviation, or sigma) chart allows visualization of the extent of the data spread or deviation from the mean. The R (range) chart depicts the difference between the largest and smallest observations in a subgroup or sample over time.

Control charts of different types are often paired for process analysis: For example, an x chart might be combined with an mR chart to allow a comparison of variability in individual measurements ($n = 1$) at different time points. An $\bar{x}$ chart might be paired with an S chart to facilitate analyses in which each data point represents a large number of measurements ($n \geq 10$); or for data subgroups of two to nine, an $\bar{x}$ chart might be coupled with an R chart.

Chart Types for Analyzing Attribute Data

The type of chart selected for the analysis of attribute data is determined first by the type of data collected (eg, counts of defectives or defects) and then by the constancy of the subgroup size. The

1. Calculate control limits by entering data onto spreadsheet templates and using the appropriate formula from statistical textbooks.
2. Create charts.
3. Examine patterns on charts.
4. Detect special cause variations by matching patterns on charts to special cause rules (Figs 6, 7).
5. Classify behavior of process as in control or out of control.
6. Manage process according to behavior of process (Fig 8).

Figure 5. These steps are important when building and analyzing control charts.

p chart, in which the proportion is plotted on the y-axis against time on the x-axis, is the attribute chart most commonly used to analyze defectives. It is used when the subgroup size (the denominator used to calculate the proportion) varies over time. Thus, the *p* chart might be used to analyze the proportion of missed appointments when the total number of appointments (the denominator or subgroup) varies daily; by contrast, the NP chart would be selected for use when the outcome measure is the number of missed appointments per 100 examinations. *U* and *C* charts are used in a similar way to analyze counts of specific events (ie, defects): The *U* chart is used to analyze data for which the subgroup size varies over time, whereas the *C* chart is predicated on the assumption of a constant subgroup size. For example, the *C* chart might be used to plot the proportion of all types of identification errors per 100 examinations performed during the period of analysis.

The statistical assumptions underlying the charting of attribute data (ie, binomial distribution for *p* and NP charts and Poisson distribution for *U* and *C* charts) could make the data analysis burdensome. By contrast, chart types that are designed for the analysis of variable data (eg, *x* charts) are more versatile and robust to departures from normal data distribution. If the data distribution is unknown and there is no time in which to test for normality and transform skewed data to make the distribution more normal, the *x*-mR chart is an easy-to-use alternative to attribute charts.

Guidelines for Building Charts

Once the appropriate chart type is selected, the data are entered on a spreadsheet and control limits are calculated (Fig 5). The formulas for calculating control limits can be found in textbooks (14). A number of easy-to-use software tools (eg, Chartrunner, PQ Systems, Dayton, Ohio; Matlab, MathWorks, Natick, Mass; and Minitab Statistical Software, Minitab, State College, Pa) are available for use in creating charts. The authors have used Excel add-ins (Microsoft, Redmond, Wash) to create their own chart templates. (Interested readers may obtain these templates from Y.Y.C.)

The commonly recommended number of data points to allow confident detection of special causes on an *x*-mR chart is 20 to 25 (10). Plots with fewer than 20 points might be somewhat useful, but the outcome of an analysis based on such charts could be subject to an increased risk of a type II error, or false-negative result (ie, a special cause signal is present but overlooked). For *p* charts, the American Society for Testing and Materials has developed formulas for calculating optimal subgroup sizes (ie, *n*) on the basis of the average percentage (ie, $\overline{p}$). The formula for calculating the minimal subgroup size is $n > 4/\overline{p}$ (15).

Guidelines for Analyzing Charts

Teaching Point

Control chart analysis helps focus scarce resources on variations for which there is evidence of special causes and helps avoid tampering in systems where random variation explains the observed difference. The goal of control chart analysis is to gain insight into the chain of events that caused the observed results and to use that insight to improve future system performance. Charts are analyzed visually for signals of special causes. The presence of a special cause signal characterizes the process as out of statistical control; the absence of such a signal means that the process is in statistical control.

Shewhart described a set of data distribution patterns that are unlikely to be seen in processes where variation is random and due to chance alone (6). Observation of any of these four patterns or rules (Fig 6, rules 1–4) signals the likelihood that a special cause is present and further

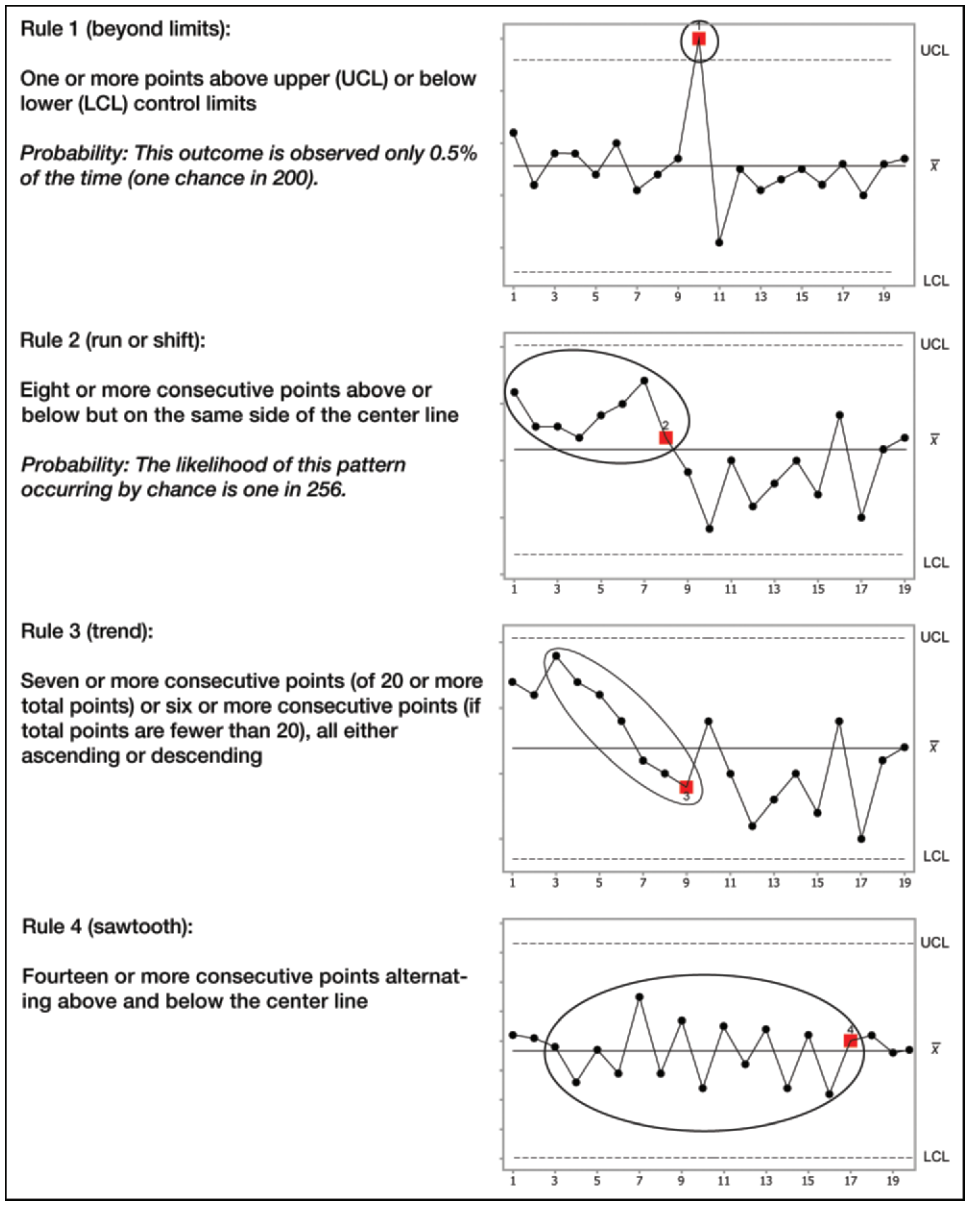
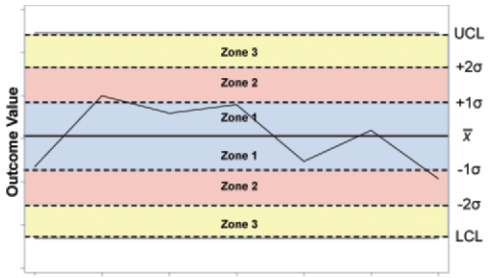


Figure 6. Shewhart's rules for identifying special cause signals. These four rules may be used to analyze all control charts, with the exception of the moving range (ie, mR). To analyze mR charts, only rule 1 should be applied. (Adapted and reprinted, with permission, from reference 12.)

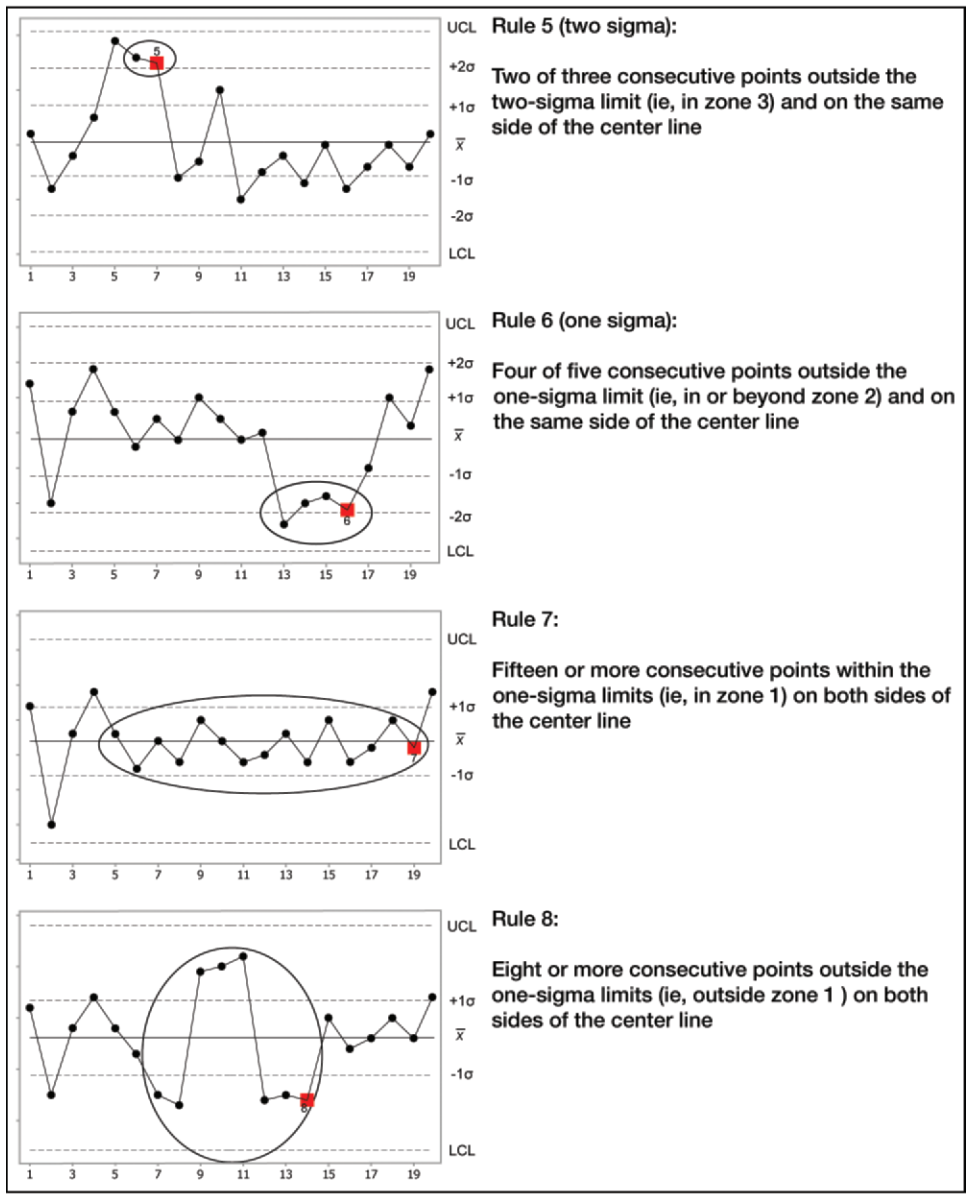
investigation into that cause is in order. In the first pattern (rule 1), an abrupt spike or drop-off is seen in the data, with one or more points beyond the UCL or LCL. In the second pattern (rule 2), small runs or shifts in direction are seen; and in the third (rule 3), a more consistent upward or downward trend is discernible. In the fourth pattern (rule 4), the data points fluctuate, sawtooth-like, between values above and below the mean.

Since these four patterns were first described by Shewhart, four additional patterns (Fig 7b) have been proposed, the observation of which may allow increased sensitivity in the detection of special causes (16). Rule 5 represents an intermediate degree of process change in which two of three consecutive data points are outside the 2σ limit and on the same side of the center line. Rule 6 describes a minimal to moderate nonrandom disturbance in the process, with four of five consecutive data points outside the 1σ limit and on the same side of the center line. Rule 7 represents an extended sawtooth-like pattern of fluctuation, with 15 or more consecutive points within the 1σ limits on both sides of the center line. Rule 8 defines a pattern of wider divergence, with eight or more consecutive data points beyond the 1σ limits on both sides of the center line.

All eight of these patterns or rules may be useful for assessing patterns of variation on control charts designed for analysis of continuous data



a.



b.

Figure 7. Additional rules for identifying special cause signals. Control chart "skeleton" **(a)** shows the common structural elements of charts used to identify special cause signals when applying the four rules described in **b**: the outcome measure on the x-axis, the day or event number on the y-axis, and the σ control limits above and below the average (center line). These four rules may be used to analyze charts of x, $\bar{x}$, R, and S. (Adapted and reprinted, with permission, from reference 12.)

(ie, x, $\bar{x}$, $\bar{x}$-R, and $\bar{x}$-S). However, as the number of rules applied increases, so does the number of false-positive special cause signals, and the use of just four or five rules (rules 1, 2, 5, and 6) may suffice (14). Current guidelines recommend that only rule 1 be used in analyses of the moving range (mR) and that only rules 1–4 be used in analyses of attribute data (R, S, G, p, NP, U, and C).

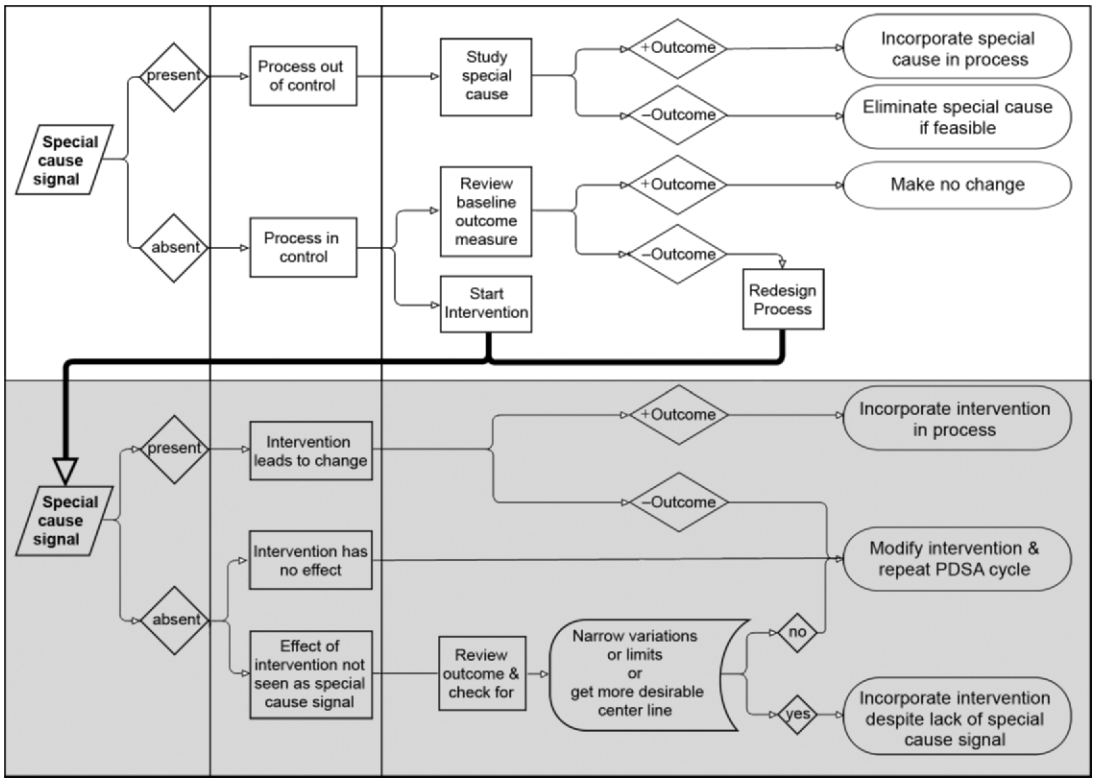


Figure 8. Flowchart shows the steps involved in process analysis and management cycles completed at QI project baseline (white area at top) and after intervention (gray-shaded area). From left to right, these steps include control chart analysis, investigation of special cause signals, and appropriate action. Process change should be attempted only when the process is stable. − = Negative, + = positive, *PDSA* = plan-do-study-act, an improvement strategy for retooling a process to achieve the desired outcome.

When charts are analyzed in pairs, the same special cause rules are applied to both charts. The only exception is in the analysis of paired *x* and mR charts: All eight rules are applicable to the analysis of *x* charts, but only rule 1 is applicable to the analysis of mR charts.

Management Strategies Based on the Type of Variation

Teaching Point

Processes with outcome variability due to special causes are managed differently from those in which variations are due to common causes alone. Since control charts help differentiate between these two types of variation, the analysis of control charts is central to the management of all stages of a QI project (Fig 8).

Management Options at Project Baseline (Precontrol Phase)

Before any change is introduced into a process, the baseline process must be stable so that the effect of the change can be readily separated from noise. All baseline data (optimally, 20 data points or more) should be analyzed to uncover special cause signals.

If special cause signals are seen on baseline control charts, the process is considered to be out of statistical control. The first step toward bringing the baseline process into statistical control is to investigate the events leading to the special cause signal. In the data analysis for a QI project to decrease no-shows (see "Case 1" in Appendix E1 [online]), our team discovered special cause signals on control charts indicating that the baseline process was unstable (Fig E1a [online]). We

then compared these signals with variables such as the day of the week and the referring clinic, uncovered a correlation, and devised a strategy for eliminating the apparent cause of the undesirable outcome. After intervening to stabilize the process, we collected and analyzed a new dataset. When the baseline process was in control as evidenced by a lack of special cause signal (Fig E1b [online]), the process was ready for improvement.

The goal of our intervention in this case was to lower the center line (ie, the average number of no-shows) and narrow the distance between the UCL and LCL. An intervention in a process can be targeted at a single undesirable outcome (eg, implementation of automated reminder technology to decrease no-shows); alternatively, the entire process can be redesigned with the objective of increasing the reliability and predictability of a desirable outcome.

Management Options after Intervention

If the baseline process is in control, the effect of intervention can be detected on the control chart as a special cause signal (Fig E2a [online]). If the special cause signal indicates the desired outcome, the intervention is successful and should be replicated and incorporated into the process. If the special cause signal indicates a negative outcome or inefficiency, the special cause must be investigated and the intervention revised appropriately.

The absence of a special cause signal also has more than one possible implication: The most obvious conclusion is that the intervention has had no effect. However, if the intervention has resulted in a more optimal center line and narrower range between the UCL and LCL, then it has increased the reliability and predictability of the outcome, a beneficial effect (Fig E3 [online]).

Limitations of Control Chart Analysis

The special cause rules invented by Shewhart to allow the detection of nonrandom variations have their basis in probability distribution theory. This theory, which is embodied graphically in the familiar bell curve, predicts that 68.26% of collected data points will be within the limits at 1σ from the mean, 95.44% will be within the 2σ limits, and 99.73% will be within the 3σ limits. Shewhart empirically selected 3σ for calculating control limits because in the ideal situation only 13.5 of 10,000 data points exceed those limits, and investigation into the causes of rare outliers is more efficient. However, these statistical predictions are based on the assumption of an ideal situation in which the data are normally distributed and the subgroups (samples) are independent, with no relation or autocorrelation between adjacent data points; in reality, such assumptions rarely prove true. Shewhart recognized these constraints but argued that the control limits could still be useful, finding support in his own accumulated experience as well as in Chebyshev's inequality theorem, which specifies that at least 89% of observations will fall within 3σ of the mean irrespective of the normality of their distribution, with that proportion increasing to more than 99% when the distribution of observations approaches normality. Thus, Shewhart's charts are robust to departures from normality.

This does not necessarily mean that QI teams should be unconcerned about whether the distribution of their data approximates normality. The normality of distribution can easily be checked with graphic plots such as histograms and normal probability plots, the preparation of which is simplified by the inclusion of the necessary tools in most statistical software programs; when the distribution is skewed, the same software tools facilitate correction of nonnormality. To control for autocorrelation, one can vary the sampling interval (17).

In summary, it is not essential that the assumptions of statistical models be met in order for nonrandom variations to be charted and detected. Statistical theory merely provides information that is useful for the interpretation of data in ideal conditions. Effective charting depends on practical experience, a reasonable choice of parameters, and an understanding of the process that is to be assessed (13). Some types of control charts are particularly robust and therefore are useful even when ideal conditions are not met.

Conclusions

Statistical control charts are powerful visual tools that combine graphic and statistical methods to facilitate monitoring and management of process outcomes. Visualization of the temporal relationship between interventions and outcomes facilitates the differentiation of real effects from confounding random variations. Insights gained from the appropriate use of control charts can provide valuable guidance for process management. There are many different kinds of charts with a specialized purpose among which one may choose, but the general-purpose *x*-mR chart is robust and allows reliable analysis of most data types collected in radiology QI projects. The management benefits long enjoyed by those in manufacturing who have made extensive use of control chart analysis are readily available to all clinical radiologists who are interested in conducting more efficient and more results-oriented QI projects.

Acknowledgment.—We thank Daniel R. Deneen, Department of Radiology at Dartmouth Hitchcock Medical Center, Lebanon, NH, for his editorial assistance during manuscript preparation.

References

1. Lee K, McGreevey C. Using control charts to assess performance measurement data. Jt Comm J Qual Improv 2002;28(2):90–101.
2. Blackmore CC. Defining quality in radiology. J Am Coll Radiol 2007;4(4):217–223.
3. Donnelly LF, Gessner KE, Dickerson JM, et al. Department scorecard: a tool to help drive imaging care delivery performance. RadioGraphics 2010;30 (7):2029–2038.
4. Duszak R Jr, Saunders WM. Medicare's physician quality reporting initiative: incentives, physician work, and perceived impact on patient care. J Am Coll Radiol 2010;7(6):419–424.
5. Johnson CD, Miranda R, Aakre KT, Roberts CC, Patel MD, Krecke KN. Process improvement: what is it, why is it important, and how is it done? AJR Am J Roentgenol 2010;194(2):461–468.
6. Shewhart WA. Statistical method from the viewpoint of quality control. Mineola, NY: Courier Dover, 1986; 155.
7. Kruskal JB, Eisenberg R, Sosna J, Yam CS, Kruskal JD, Boiselle PM. Quality improvement in radiology: basic principles and tools required to achieve success. RadioGraphics 2011;31(6):1499–1509.
8. Bluth EI, Havrilla M, Blakeman C. Quality improvement techniques: value to improve the timeliness of preoperative chest radiographic reports. AJR Am J Roentgenol 1993;160(5):995–998.
9. Nagy P, Safdar N. Online radiology quality resources. J Am Coll Radiol 2010;7(6):459–460.
10. Carey RG. Improving healthcare with control charts: basic and advanced SPC methods and case studies. Milwaukee, Wis: ASQ Quality Press, 2003; 194.
11. Wheeler DJ. Making sense of data. Knoxville, Tenn: SPC Press, 2003; 395.
12. Pujar S, Calvert S, Cortina-Borja M, et al. Statistical process control (SPC)—a simple objective method for monitoring seizure frequency and evaluating effectiveness of drug interventions in refractory childhood epilepsy. Epilepsy Res 2010;91(2-3):205–213.
13. Amin SG. Control charts 101: a guide to health care applications. Qual Manag Health Care 2001;9(3): 1–27.
14. Wheeler DJ, Chambers DS. Understanding statistical process control. 3rd ed. Knoxville, Tenn: SPC Press, 2010.
15. American Society for Testing and Materials. ASTM manual on presentation of data and control chart analysis, MNL 7. 6th ed. Philadelphia, Pa: American Society for Testing and Materials, 1990.
16. Nelson LS. The Shewhart control chart: tests for special causes. J Qual Technol 1984;16(4):237–239.
17. Provost LP, Murray SK. The health care data guide. San Francisco, Calif: Jossey-Bass, 2011; 264.

RadioGraphics

Quality Initiatives
Statistical Control Charts: Simplifying the Analysis of Data for Quality Improvement

Yvonne Y. Cheung, MD, MS • Boyoun Jung, BS • Jae Ho Sohn, MS • Greg Ogrinc, MD, MS

RadioGraphics 2012; 32:2113–2126 • **Published online** 10.1148/rg.327125713 • **Content Codes:** QA RS

Page 2114
Control charts are analytic tools that allow a visual distinction between meaningful change and random variation or "noise" in a process by comparing the actual distribution patterns of outcome data with standardized distribution patterns derived from probability statistics (6).

Page 2114
Control charts show changes in outcome measures over time and thereby offer a clear advantage over classic methods of statistical analysis.

Pages 2115
Variation is inherent in all processes, and understanding and managing variation is key for achieving control over a process. Shewhart distinguished between two types of variation: variation due to common causes and variation due to special causes (6,13). Common cause–related variations are random and lead to outcome measures that are within a predictable distance from the mean. [...] By contrast, special cause–related variations reflect extraneous or assignable effects that are unlikely to be due to chance alone.

Page 2121
Control chart analysis helps focus scarce resources on variations for which there is evidence of special causes and helps avoid tampering in systems where random variation explains the observed difference.

Page 2124
Processes with outcome variability due to special causes are managed differently from those in which variations are due to common causes alone.