Journal of the Pediatric Infectious Diseases Society

INVITED REVIEW

Displaying and Analyzing Quality Improvement Data

Patrick W. Brady,[1,2] Michael J. Tchou,[1,2] Lilliam Ambroggio,[1,3] Amanda C. Schondelmeyer,[1,2] and Erin E. Shaughnessy[1]

[1]Division of Hospital Medicine, Department of Pediatrics, [2]James M. Anderson Center for Health Systems Excellence, Department of Pediatrics, and [3]Division of Biostatistics and Epidemiology, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Ohio

In this second article in the quality improvement (QI) methods series, we discuss how data are best displayed and analyzed in QI projects while focusing on some similarities with and differences from traditional clinical research. We demonstrate why displaying data over time on a run or control chart is superior to using pre–post analysis for QI studies. We introduce several types of statistical process control charts for data commonly collected during QI programs and provide guidance on how to use the proper chart. Last, we present solutions to several common data challenges in QI projects.

Keywords. quality improvement; measurement; statistical process control; run chart; control chart.

The first article in this quality improvement (QI) series introduced essential QI methods, including the importance of iterative testing cycles and plotting data over time [1]. In this second article, we discuss how data are best displayed and analyzed in QI projects. First, we show why displaying data over time on a run chart is superior to using a pre–post analysis. Second, we introduce several types of statistical process control charts for data commonly collected during QI projects and how they assist in measuring variability. Last, we briefly discuss common data challenges in QI projects and their potential solutions.

PROBLEMS WITH PRE–POST ANALYSIS

Time is a tremendously important variable when judging the success of a QI initiative. A pre–post analysis assumes a static time point at which the intervention was implemented, but the complex and dynamic nature of care systems and iterative QI interventions make this approach ill advised [2]. Interventions take time to adopt and scale; sustaining change is difficult. Care providers change from shift to shift and/or month to month. To illustrate the limitations of pre–post analysis for QI interventions, we share the example of an improvement team that is excited about early evidence of success in improving the reliability of narrow-spectrum antibiotic use for children hospitalized with uncomplicated community-acquired pneumonia. The team describes their improvement intervention (a pocket-sized version of the Pediatric Infectious Diseases Society/Infectious Diseases Society of America pneumonia guidelines given to each resident), and the encouraging pre (40% narrow-spectrum antibiotic use) and post (60% narrow-spectrum antibiotic use) results. The team performed their work at a busy pediatric center, so they had sufficient numbers to meet statistical significance ($P < .05$) in χ^2 testing. However, for many reasons, the team is unwise to be confident that their intervention led to improved performance. Each panel of Figure 1 displays a 40% preintervention and 60% postintervention use of narrow-spectrum antibiotics, but the time-series data tell quite different versions of the story.

Biostatistical techniques, such as an interrupted time-series approach, are used for the analysis of longitudinal data [3]. These approaches confer some important advantages, particularly when a secular trend exists before the QI intervention and when a single-packaged or bundled QI intervention (as opposed to series of small tests of change) is performed. However, during most improvement projects, run charts and control charts are easier to use and a preferred way to gauge the ongoing success of interventions.

COMMON-CAUSE AND SPECIAL-CAUSE VARIATION

In any process, inherent outcome variability exists because of the variations at each step. For example, variations occur in the turnaround time for laboratory results at a hospital because of small variations in the time to draw the blood sample, the time to send the sample to the laboratory, and the time to process the sample in the laboratory. This type is known as common-cause variation, defined as the variation expected by chance in a stable process [4, 5]. Reacting to common-cause (or expected) variation is similar to a type I error, or inferring a significant difference when one does not truly exist. Alternatively, special-cause variation is variation that would not be expected by chance in a stable process and is an indication of a changed or unstable process [4, 5]. Examples of special-cause variation in laboratory turnaround

Received 27 February 2017; editorial decision 15 August 2017; accepted 21 August 2017.

Correspondence: P. W. Brady, MD, MSc, Department of Pediatrics, Division of Hospital Medicine, James M. Anderson Center for Health Systems Excellence, Department of Pediatrics, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, ML 9016, Cincinnati, OH 45229 (patrick.brady@cchmc.org).

Journal of the Pediatric Infectious Diseases Society 2017;00(00):1–4

DOI: 10.1093/jpids/pix077

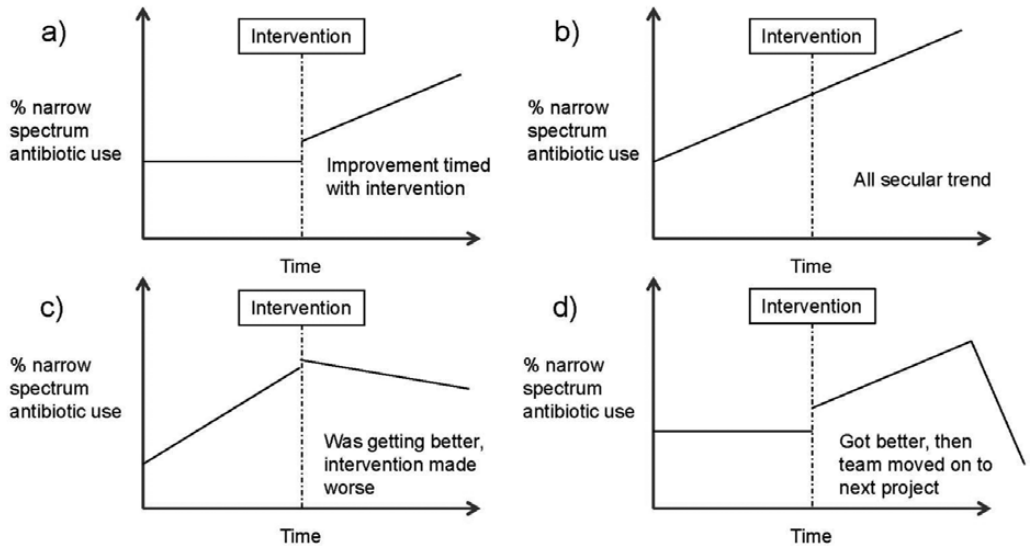

Figure 1. Four improvement stories shown over time. Each chart looks the same in a pre–post analysis. (a) Modest success at the project start, which increases with additional education, awareness, and reminders. (b) Ongoing, gradual improvement in the outcome, and the intervention was not associated with any change in the trajectory of narrow-spectrum antibiotic use. (c) Timing of the intervention was associated with a decrease in narrow-spectrum antibiotic use, but averaging the pre–post data falsely indicated improvement. (d) The team's intervention was associated with a quick and large increase, but this success was not sustained when the team lost focus.

time would include a power outage that prolongs turnaround time or an improvement initiative that reduces turnaround time by changing the laboratory technician workflow. In both of these examples, the underlying process has changed fundamentally.

RUN AND CONTROL CHARTS

Run charts and other statistical process control (SPC) charts present data over time and enable the improvement team to identify quickly when variation that is unlikely due to chance (special-cause variation) has occurred. SPC charts (also called Shewhart charts) were developed by Walter Shewhart, a young engineer, physicist, and statistician working at Western Electric Company in the early 20th century [6, 7]. Run charts are simple displays of data over time with a median line that indicates the central tendency [2, 4, 8]. The most common run-chart rules for identifying special-cause variation are presented in Table 1 and displayed visually in the first article in this series [1].

Control charts have advantages over run charts in that they define expected variation in a process [4, 7, 9–11]. The first goal of many QI interventions is to reduce this variation [5]. The centerline in SPC charts (Figure 2) is most commonly the mean of the data points (versus a median used in run charts). The upper and lower control limits (usually shown visually as dotted lines) are defined on the basis of the distribution of the data with each approximately 3 standard deviations, or σ, above and below the centerline. For normally distributed data, it might be helpful to think of a control chart as a Bell curve turned on its side, with the centerline at the peak of the Bell curve and the upper and lower control limits located 3σ out such that approximately 99% of the data points (if they come from the same distribution) would be expected to fall inside control limits. A properly annotated run or control chart conveys a tremendous amount of data and can greatly improve a QI report, as discussed in more detail in the third article in this series [12].

CHOOSING THE CORRECT SPC CHART

The proper SPC chart is chosen on the basis of the type of data and their underlying distribution [4, 9]. A fixed or varying subgroup size (eg, number of patients per week with the outcome of interest) influences which chart would be the best choice, although we have found that a fixed subgroup size is uncommon in healthcare QI projects. Table 2 lists the most common types of control charts used in healthcare QI reports, the types of data for each, and example scenarios in which this type of control

Table 1. Most Commonly Used Run and Control Chart Rules for Identifying Likely Special-Cause Variation

Type of Chart	Rule	Definition
Run chart	Shift	Six consecutive points, all above or below the median line; points on line do not break the run but are not counted in the run
	Trend	Five or more consecutive points, with each going all up or all down
Control chart	Control limits	Any single data point outside of the 3-σ control limit
	Run	Eight consecutive points, all above or below the mean line; points on line do not break the run but are not counted in the run

Data source, Provost and Murray's *The Health Care Data Guide* [4].

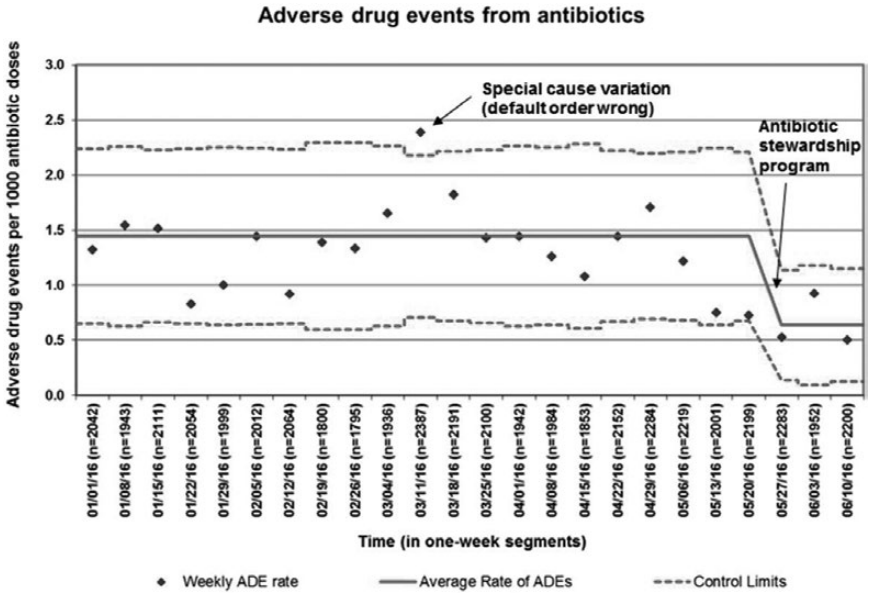

Figure 2. Control chart showing adverse drug events resulting from antibiotic use per 1000 antibiotic doses. The first 10 points show modest variation that is common-cause variation, because it all occurs within the control limits. The 11th point shows special-cause variation, related in this case to an error in the default antibiotic choice in an order set in the electronic health record. Once this error was corrected, common-cause variation continued until a new antibiotic stewardship program reduced the rate of prescribing of antibiotics with higher rates of adverse drug events (ADEs). With 2 of 3 points outside of previous control limits, the improvement team readjusted the centerline downward. These count data are displayed on a u chart.

chart might be used. In the case of control charts with a varying subgroup size (as in Figure 2), the control limits will vary on the basis of sample size, being more narrow with larger subgroups.

Table 1 lists the most commonly used run and control rules for identifying likely special-cause variation. It is important to remember that this variation might still be a result of chance or common-cause variation, akin to ~5% of studies of equivalent treatments that still produce a *P* value of .05 or less. Special-cause variation that occurs closer to when a QI intervention is first implemented is more likely related to that QI intervention. Special-cause variation that occurs several months after the intervention should be viewed more skeptically, because it might be a result of other contextual or extraneous factors. For instance, if a data point was special cause in the desired direction only during the week in which the QI team lead was on the consult service, it might not be right to conclude that the underlying process has changed and adjust the centerline. In that situation, more data points should be obtained.

Table 2. Types of Statistical Process Control Charts Commonly Used in Healthcare Quality Improvement

Type of Chart	Type of Data	Example Healthcare Scenario
p	Proportion (attribute data classified as yes or no)	Percentage of all discharges from the hospital that were associated with a readmission within 30 days
u	Count (attribute data classified only as present)	Adverse drug events from antibiotics
XmR	Continuous (variable data for an individual)	Vancomycin serum levels for 1 patient with oxacillin-resistant *Staphylococcus aureus* meningitis
XbarS	Continuous (variable data for groups/subgroups)	Minutes from arrival in emergency department to first antibiotic in febrile, neutropenic patients with a central venous catheter

COMMON CHALLENGES AND THEIR SOLUTIONS

The following common challenges arise with QI data, but some potential solutions exist.

Uncommon Outcomes

Uncommon outcomes, such as surgical site infections (SSIs), are a common issue for safety measures. It can be difficult to measure improvement, because many units of time will have no SSI events, and the lower control limit is almost certain to be 0. In such a case, g charts can be helpful. Using SSIs as an example for a g chart, the team would include event date of each SSI on the x-axis and the number of surgeries between each SSI on the y-axis. A t chart (Figure 3) is a special type of g chart in which time (usually a date) is on the x-axis; these charts are generally less preferable when event data are available because surgical procedure counts vary according to time of year and day of week.

Large Seasonal Variations

An additional challenge in infectious-disease–related QI projects is the large seasonal variations in the count of children with

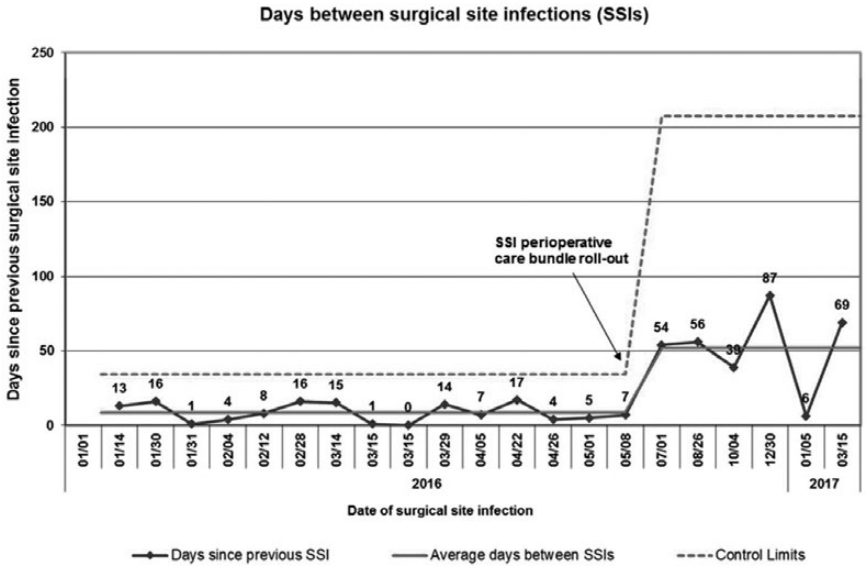

Figure 3. t chart illustrating days between surgical site infections (SSIs) on the y-axis and dates of the SSIs on the y-axis. A substantial increase in days between was seen immediately after roll-out of a perioperative care bundle; the average number of days between SSIs went from <10 to >50 days.

endemic diseases over the year. If a QI team has a goal to improve narrow-spectrum antibiotic prescribing for children hospitalized with pneumonia, it can be challenging—and arguably misleading—to give equal prominence on an SPC chart to July (in which 2 children were admitted) and January (in which 45 were admitted). Fixed subgroup sizes (eg, 10 consecutive pneumonia admissions per point on the x-axis) helps with this issue, although it might be important to better show time of year in the figure or figure legend.

Too Few Baseline Data

Minimal preintervention data can be a substantial problem for both the team that judges the success of the project and the potential for publishing the data in the peer-reviewed literature. At least 10 and ideally 20 preintervention data points are needed to establish the average and baseline variations in performance [4, 10]. If fewer data points are available, the problem sometimes can be addressed by making the subgroups smaller (eg, using monthly instead of biweekly data), but there also can be a tension here between having the subgroups then be too small (subgroups of 10 help limit the "zig-zag" of small sample sizes).

CONCLUSIONS

Finally, and perhaps most importantly, SPC and run charts are used to their utmost value and validity during a QI project. "Reverse-engineering" a retrospective study (or even retrospectively collected outcome data from a prospective educational intervention) is both inefficient QI and potentially misleading. An SPC chart used optimally is both a statistical tool for understanding the performance of a process and a resource for informing the team of which interventions might be effective and worth implementing on a larger scale. A proper SPC chart can be an invaluable resource in driving great local QI programs, and it gives the greater scientific community a better understanding of effective QI interventions that might be best spread or evaluated more rigorously in an experimental design [13].

Note

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Shaughnessy E, Shah A, Ambroggio L, et al. Introduction to quality improvement. J Pediatric Infect Dis Soc **2017**; in press.
2. Langley GJ. The Improvement Guide: A Practical Approach to Enhancing Organizational Performance, 2nd ed. San Francisco: Jossey-Bass; **2009**.
3. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin; **2001**.
4. Provost LP, Murray SK. The Health Care Data Guide: Learning From Data for Improvement. San Francisco, CA: Jossey-Bass; **2011**.
5. Wheeler DJ. Understanding Variation: The Key to Managing Chaos, 2nd ed. Knoxville, Tenn: SPC Press; **2000**.
6. Shewhart WA. Economic Control of Quality of Manufactured Product. New York: D. Van Nostrand Company, Inc; **1931**.
7. Shewhart WA, Deming WE. Statistical Method From the Viewpoint of Quality Control. Washington, DC: Department of Agriculture, The Graduate School; **1939**.
8. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. BMJ Qual Saf **2011**; 20:46–51.
9. Benneyan JC. Statistical quality control methods in infection control and hospital epidemiology, part II: chart use, statistical properties, and research issues. Infect Control Hosp Epidemiol **1998**; 19:265–83.
10. Benneyan JC. Statistical quality control methods in infection control and hospital epidemiology, part I: introduction and basic theory. Infect Control Hosp Epidemiol **1998**; 19:194–214.
11. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. Qual Saf Health Care **2003**; 12:458–64.
12. Schondelmeyer AC, Brower LH, Statile AM, et al. Writing and reviewing quality improvement manuscripts. J Pediatric Infect Dis Soc **2017**; in press.
13. Ambroggio L, Schondelmeyer AC, Hoefgen E, et al. Advanced designs for quality improvement studies. **2017**; in press.