Counterpoint

‘Counterpoint’ is an occasional feature presenting discussion of a topic that is currently under debate in quality of care circles. We invite readers to submit Letters to the Editor adding their opinion to the topic.

Risk-adjusted sequential probability ratio tests and longitudinal surveillance methods

Use of statistical process control in health care

Statistical methods can play an important role in detecting changes in many hospital processes, including mortality and adverse event rates. Several surveillance methods have an established history of use within health care, while others such as standard and modified statistical process control (SPC) are experiencing growing interest. The retrospective use of SPC by Spiegelhalter et al. [1] is an excellent example of the potential role that risk-adjusted control charts could have played in earlier detection of higher mortality rates in the Bristol Royal Infirmary and of general practitioner Harold Shipman.

Statistical control charts were first developed in the 1920s by Walter Shewhart at Bell Laboratories [2], popularized by physicist and statistician W. Edwards Deming [3], and extended over the years to be more powerful or more appropriate for various applications. Although initially intended for use in industrial and chemical processes, both Shewhart and Deming recognized the value of these methods for detecting statistical changes in just about any application, and as early as 1942 Deming proposed their potential value for disease surveillance and rare events [4]. Important health care concerns in which control charts have been shown to be effective include surgical site infections, adverse drug events, needle sticks, ventilator-associated pneumonia, and other iatrogenic injuries.

Given the staggering incident and cost estimates of medical mistakes, care-induced injury, and hospital-acquired infections in the US alone—including 770 000–2 million injured patients per year, 44 000–180 000 deaths, and $8.8 billion annually [5–8]—the growing use of SPC in the health care literature is not surprising. As with the Shipman data, however, more advanced or special-purpose SPC methods may need to be used in applications for which standard assumptions are not met.

Special-purpose control charts and statistical issues

The simplest types of statistical control charts, called Shewhart charts, perform fairly well for detecting moderate-to-large rate changes. In some industrial applications, more advanced tools such as cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) charts are used to detect smaller changes, to monitor low rates, or in situations where sufficiently large sample sizes are not available. Examples of health care CUSUM and EWMA applications include surveillance of seasonal influenza, community Salmonella, and fever curves in neutropenic patients. In other applications, new SPC methods have been developed for applications dealing with rare events, infectious diseases and other events that naturally occur in clusters, overdispersion, naturally cyclic behavior, and risk adjustment, such as in the case of the Shipman data. Related SPC methods have also been developed to handle non-homogeneous events in manufacturing, such as for different production lines, small batches, and job shops.

Risk-adjustment

In many (if not most) applications, patients do not have identical likelihoods of an adverse event, mortality, or other outcome. Various risk-adjustment methods have been developed to stratify patients into homogeneous categories or to estimate patients’ a priori survival probabilities, including logistic regression models and others [9–10]. For example, the analysis by Spiegelhalter et al. [1] risk-adjusts at the provider macro-level essentially by estimating each annual mortality rate using local general practitioner and national annual data as norms. Other methods that incorporate risk-adjustment at the patient level into Shewhart, EWMA, CUSUM, and binary sequential probability ratio test (SPRT) charts could have greater sensitivity if detailed data are available, although this frequently is not possible.

Chart sensitivity

It is interesting, however, that even simple $p$ charts risk-adjusted at the macro-level, with control limits set so that $\alpha = 0.000001$, signal a rate increase for Shipman’s male patients in 1995 and 1996 and in 10 of 22 years for female patients (1979, 1985, 1988–1989, and 1993–1998), with one year (1997) being almost 40 standard deviations above the mean. For rate increases of such large magnitudes, simple charts therefore may be good enough for annual data. Mathematically, for example, it can be shown that a binary SPRT and a conventional Shewhart $p$ chart both will detect a rate
change from \( p = 0.009 \) to \( p = 0.045 \) within the first 1.67 years on average (using Shipman’s average annual sample size of 235 patients and \( \alpha = 0.000001 \) for the Shewhart chart and \( \alpha = \beta = 0.000001 \) for the SPRT).

**Rare events**

Another motivation for cumulative monitoring approaches is to accommodate rare events that otherwise would require large samples to yield adequate statistical sensitivity. CUSUMs, SPRTs, and EWMAs are excellent for this purpose, and several other SPC methods also have been developed for rare events. Many of these are based on some variation of the idea of monitoring the number of cases or time between adverse events rather than the more traditional approach of monitoring the number of events or deaths within a fixed time period or sample size [11–12].

Applying these methods to risk-adjusted data could lead to even better (faster) detection performance, although in many cases collecting data and risk-adjusting at the individual patient level may not be practical or even possible. Number-between types of SPRTs also could be used in such cases and in theory should yield optimal detection performance, although their use has not been explored in detail. Binary SPRT and moving average charts based on each dichotomous case also have been investigated for detecting infectious disease outbreaks and rate increases [13].

**Other statistical issues**

Although a methodological review is not our focus here, it is useful to comment briefly on several important statistical issues inherent in these types of studies. Many risk-adjusted analysis methods utilize some type of comparison of the expected and observed number of outcomes, most commonly an Observed/Expected (O/E) ratio or difference. Once computed, the O/E results typically are treated and analyzed as independent Gaussian (i.e. normal) statistics, although in many applications they are neither Gaussian nor independent. This issue typically has not received much attention, although it is easy to construct examples in which results can be significantly incorrect.

The observed number of events within a heterogeneous population also classically is not a simple random variable, even (especially) after being divided by a constant (i.e. the expected number). A third important issue in some studies is the convention of basing the expected number of events for monitoring adverse events.

**Summary**

Regardless of the exact method employed, the application of statistical process control, SPRTs, or related longitudinal analysis methods can significantly improve the ability to monitor clinical processes and outcomes. Incorporation and adaptation of risk-adjustment and rare events into these methods represent important contributions to their use in health care. Fostering greater and more widespread use of these methods, however, remains a significant challenge. Hopefully studies such as those by Spiegelhalter et al. [1] will lead to more awareness of their value for contributing to a safer health care system.

James C. Benneyan, Alicia D. Borgman
Northeastern University
Boston, MA, USA

**References**