

Letter to the Editor

The use of statistical process control methods in monitoring clinical performance

To the Editor: The article by Spiegelhalter and colleagues [1] and the Counterpoint papers by Benneyan and Borgman [2], Lim [3], and Bolsin and Colson [4] deserve further comment. At least two issues should be raised, the first of which, the primacy of systems, is of crucial importance.

Benneyan and Borgman state that 'Fostering greater and more widespread use of these methods remains a significant challenge'. In Australia, statistical process control (SPC) methods were implemented with enthusiasm in the mid-1980s, as the Australian Council on Healthcare Standards (ACHS) embarked on widespread hospital accreditation. They quickly fell into disuse because they were found not to be useful. This occurred because the role of SPC was not understood. Thus we have waited nearly 20 years for their resurrection in hospitals, albeit with much improved methods. Unless we learn from the mistakes of the mid-1980s, these very valuable methods will once again be found wanting and they will once again fall into disuse.

The problem is that processes must be brought into statistical control before SPC is useful; the message that control limits are useless unless the process is in control is so fundamental that it seems easily to be forgotten. It can, of course, be argued that the risk adjustment imposes the required statistical control. I believe, however, that this problem still exists, certainly at a philosophical and quality improvement level, but also in all probability at a scientific and statistical one too. Shewhart, Deming and others have repeatedly pointed out that control limits are meaningless unless the system has first been brought into statistical control. This is eloquently described in the wonderful book by David Salsburg, *The Lady Tasting Tea. How Statistics Revolutionized Science in the Twentieth Century* [5]. Bringing processes into statistical control in hospitals implies that hospital systems are first carefully analysed and optimized [6–9].

All substandard performance is underpinned by unsatisfactory systems; even an incompetent operator is a system problem because the system allows him or her to operate. In Australia in the early 1980s, hospital 'reform' was in full swing under the influence of a managerialist management system obsessed with very short-term financial objectives and individual performance. The system was one of reducing costs in the short term and judging individuals, not of providing a quality service. SPC was bound to fail. Analysing and optimizing systems is a management problem, therefore institutions

wishing to avoid substandard performance must ensure that managers are trained to analyse and optimize systems, and that they have the resources and authority to see that the systems in their hospitals are sound. The practice of judging individuals who have no authority to fix systems damages morale, encourages substandard performance, and guarantees increased medical error. In addition, data are used to justify rather than to learn how to do better, and data can easily be rendered meaningless by subtle gaming so that they become of no use in improving quality.

Lim [3] would like to see greater sensitivity and specificity in monitoring clinical performance. However, if we wish to be certain of removing all the bad apples from a barrel, we are likely to remove good ones as well. Conversely, if we wish to remove only some bad apples, it is probable that some bad ones will remain. We have to make up our minds about what we wish to do. In a judgmental environment, a high level of specificity is required, but this makes it likely that problems will either be missed or will take such a long time to be identified that preventable patient injury will occur. In a learning environment, high sensitivity is possible as occasional false positives can be tolerated. In fact, with most hospital adverse events, false positives are not difficult to identify, but 'false negative states' where an unidentified problem exists can be very dangerous. Surely decision making will be improved if we can foster an environment in which surveillance and data analysis are used to learn how to improve rather than to judge and to blame.

Bolsin and Colson [4] reported the optimistic message that their trainee anaesthetists found SPC very useful and that they embraced it enthusiastically. This occurred because they had a good system for anaesthetic training and trainees were able to use these valuable tools to learn how to improve. They emphasized the essential components of the Deming cycle: good feedback and a supportive training environment. SPC worked for them because they had a good system.

If attention is first directed towards fixing systems, the methods described by Spiegelhalter *et al.* [1] have the wonderful potential to enrich and improve the quality of patient care. If they are employed merely to judge without first attending to the underlying systems, they will be seen to fail when the sort of problems that occurred at Bristol recur, and will fall into disuse, as happened in Australia nearly 20 years ago. It is worth our while to strive to prevent this from happening.

A more minor matter is that most count data adverse events in hospitals can display too great a variability for the methods described by Spiegelhalter *et al.* [1] to be applicable. This can occur with colonizations due to multiple antibiotic resistant organisms (MROs), needlestick injuries, pressure ulcers, hospital readmissions, and patient falls. It is often due to a breakdown in the independence that is required for methods based on the Poisson distribution to be appropriate. When this occurs, it is often useful to employ a negative binomial distribution. The required weights then become: $w = F \times \log_e[(1 - P_1)/(1 - P_0)] - S \times \log_e[P_0/P_1]$, where F is the observed count (e.g. the number of new MRO isolates for the current month); $S = M^2/(V - M)$, where M and V are the mean and variance, respectively, derived from monthly counts when the monthly new MRO isolate rate is stable and endemic; and $P_0 = S/(S + M)$ and $P_1 = S/(S + R \times M)$, where R is the increase in the rate to be identified, for example for a doubling of the rate $R = 2$.

Anthony P. Morton

Queensland University of Technology
Tarragindi, Queensland, Australia

References

1. Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003; **15**: 7–13.
2. Benneyan J, Borgman A. Risk-adjusted sequential probability ratio tests and longitudinal surveillance methods. *Int J Qual Health Care* 2003; **15**: 5–6.
3. Lim T. Statistical process control tools for monitoring clinical performance. *Int J Qual Health Care* 2003; **15**: 3–4.
4. Bolsin S, Colson M. Making the case for personal professional monitoring in health care. *Int J Qual Health Care* 2003; **15**: 1–2.
5. Salsburg D. *The Lady Tasting Tea. How Statistics Revolutionized Science in the Twentieth Century*. New York: W. H. Freeman and Co., 2001.
6. Reason J. Human error: models and management. *Br Med J* 2000; **320**: 768–770.
7. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *Br Med J* 1998; **316**: 1154–1157.
8. Plsek P. Innovative thinking for the improvement of medical systems. *Ann Intern Med* 1999; **131**: 438–444.
9. Morton A. Dealing with mortality rates after Shipman. Electronic response to Baker R, Jones D, Goldblatt P, Monitoring mortality rates in general practice after Shipman. *Br Med J* 2003; **326**: 274–276.