Letter to the Editor

Annals of Surgery Open

OPEN

Comment on A Method for Continuous Surgeon Improvement in Rectal Cancer

Observed Minus Expected (O–E) Chart Versus CUSUM Chart

Quentin Cordier, MPH,*† and Antoine Duclos, MD PhD*†‡

We read with great interest the manuscript by Ferrari et al[1] entitled "A Method for Continuous Surgeon Improvement in Rectal Cancer: Risk-Adjusted Cumulative Sum" and recently published in Annals of Surgery. In this article, the authors present a chart to monitor the routine performance of individual surgeons in rectal cancer surgery. We agree with authors that real-time evaluation of surgical outcomes is relevant for the quality of care improvement, and commend the author's efforts to propose a tool designed for this purpose. However, the presence of certain semantic and methodological confusions, as introduced by the authors, highlights the challenges surgeons encounter when selecting the appropriate tools to monitor their outcomes.

Various tools are available to monitor surgical indicators in real time. The family of statistical process control tools—to which belongs the Cumulative Sum (CUSUM) chart—is large and includes a great variety of control charts. The first of those tools were designed a century ago in the manufacturing industry. Today they are widely implemented for monitoring surgical outcomes[2,3] and have also proven impactful in preventing major adverse events.[4] Specific charts, such as Shewhart control charts,[5] are designed to monitor the measurement of surgical outcomes aggregated over consecutive periods. Alternatively, other charts rely on procedure-by-procedure measures to offer real-time feedback, such as the observed minus expected (O–E) chart, the exponentially weighted moving average chart, the sequentially probability ratio tests chart, or the CUSUM chart.

The abundance of different tools makes it challenging for surgeons to choose the best one and accurately name it for monitoring their outcomes. The name of the tool proposed by the authors here is misleading. Instead of a risk-adjusted CUSUM chart, they designed a risk-adjusted O–E chart, which is also known synonymously as the Variable Life-Adjusted Display (VLAD)[6] or Cumulative Risk-Adjusted Mortality (CRAM) chart.[7] O–E and CUSUM charts differ significantly in their conception and interpretation, much like the gastric sleeve and gastric bypass differ in bariatric surgery. On one hand, the risk-adjusted O–E chart was proposed by Lovegrove et al.,[6] allowing users to intuitively visualize variations in surgical outcomes over time. It calculates the cumulated difference between observed and expected events on a procedure-by-procedure basis (see Fig. 1). The resulting curve rises in cases of poor performance and falls in cases of good performance, indicating the deviation of the number of observed events from what would have been expected. The final value represents the number of avoidable events if greater than 0 and the number of avoided events if lower than 0. However, the absence of control limits in the risk-adjusted O–E chart precludes the statistical interpretation of observed variations to detect significant improvements or deteriorations in the safety of care. For instance, in Ferrari et al's[1] manuscript, although authors observe a seeming decrease in the rate of complications, they cannot formally conclude that this decrease was statistically significant, nor can they identify the full sequence of intervention in which this reduction would indicate significant improvement. On the other hand, the CUSUM chart was initially developed by Page[9] in 1954 and later applied for monitoring surgical performance. The CUSUM chart is much more complex to design than the O–E chart but enables the identification of statistically significant changes in surgical performance trends with great sensitivity. It assesses, at each intervention, through a hypothesis test whether the statistical process is under control (null hypothesis: the rate remains constant over time) or out of control (alternate hypothesis: a significant change in the rate is observed).[10] A CUSUM score is calculated by summing weighted deviations between the process and an expected target, and compared with lower and upper horizontal lines known as control limits. If the score exceeds the control limits, the process is considered out-of-control, indicating a signal toward either improvement or deterioration (see Fig. 1). The theory behind control limits relies on average run lengths, which represent the average number of interventions needed to detect a signal in both out of control process (true positive signal) and in-control process (false positive signal). By setting the control limits, users seek a balance in the tool's sensitivity to detect a maximum of true positive signals while avoiding an excessive detection of false positives. After detecting an unusual variation in surgical outcomes, the surgeon can investigate related causes and consider decisions to improve patient care. While the CUSUM chart is particularly efficient to detect small changes in daily performance,[11] it requires more statistical knowledge and may pose interpretation difficulties for surgeons.

In conclusion, although the authors constructed a risk-adjusted chart to track surgical outcome variations in real-time,

*From the *Research on Healthcare Performance RESHAPE, INSERM U1290, Université Claude Bernard Lyon 1, Lyon, France; †Health Data Department, Hospices Civils de Lyon, Lyon, France; and ‡Center for Surgery and Public Health, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.*

Disclosure: The authors declare that they have nothing to disclose.

Quentin Cordier and Antoine Duclos contributed equally to this work.

Reprints: Quentin Cordier, MPH, Service des Données de Santé, Hospices Civils de Lyon, Pôle de Santé Publique, 162 avenue Lacassagne 69424 Lyon cedex 03, France. Email: quentin.cordier@chu-lyon.fr.

Annals of Surgery Open (2024) 3:e473

Received: 6 June 2024; Accepted 15 June 2024

Published online 1 August 2024

DOI: 10.1097/AS9.0000000000000473

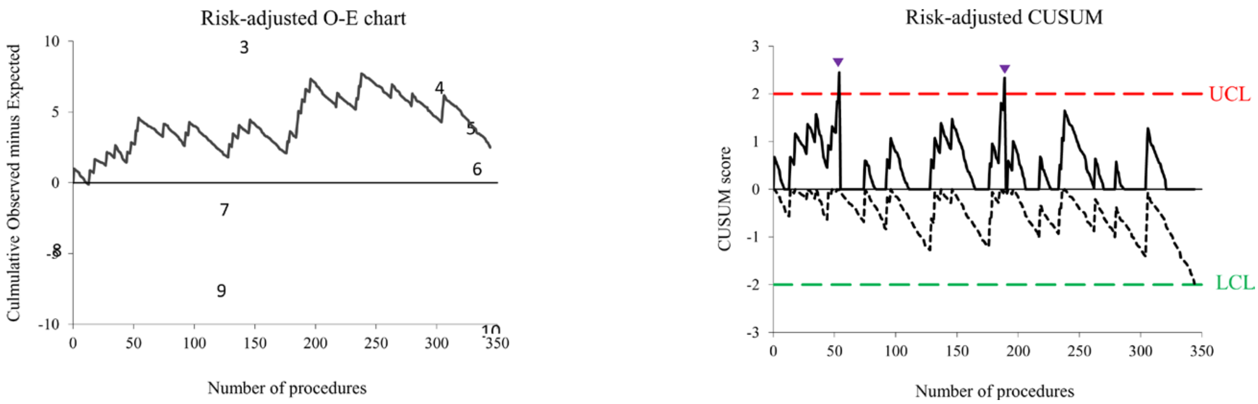

FIGURE 1. Examples of risk-adjusted observed minus expected (O–E) and cumulative sum (CUSUM) charts in thyroid surgery. Examples of a risk-adjusted O-E chart and a CUSUM chart displaying the occurrence of recurrent laryngeal nerve palsy after thyroidectomy for a single surgeon. In the O-E chart (A), the curve moves upward if the number of operations with recurrent laryngeal nerve palsy increased above that predicted by the risk model, and moves downward if the number decreases. In the CUSUM chart (B), the upper curve detects signals of deterioration in surgical performance, while the lower curve detects signals of improvement. When the curves cross over the control limits, a signal is detected (shown with circles), indicating a statistically significant change in surgical performance. The graph is subsequently reset to allow for further monitoring. In this example, deteriorations were detected at procedures 57 and 192. Data originated from Duclos et al.[8]

on a procedure-by-procedure basis, the tool they proposed is not a CUSUM chart but rather an O–E chart. Considering the existing challenges for surgeons in navigating through various types of statistical process control tools, we believe that terminology matters. Calling this tool a 'CUSUM' chart is misleading for readers.

REFERENCES

1. Ferrari D, Violante T, Merchea A, et al. A method for continuous surgeon improvement in rectal cancer: risk-adjusted cumulative sum. *Ann Surg*. doi: 10.1097/SLA.0000000000006330.
2. Thor J, Lundberg J, Ask J, et al. Application of statistical process control in healthcare improvement: systematic review. *Qual Saf Health Care*. 2007;16:387–399.
3. Le Thien MA, Cordier Q, Lifante JC, et al. Control charts usage for monitoring performance in surgery: a systematic review. *J Patient Saf*. 2023;19:110–116.
4. Duclos A, Chollet F, Pascal L, et al; SHEWHART Trial Group. Effect of monitoring surgical outcomes using control charts to reduce major adverse events in patients: cluster randomised trial. *BMJ*. 2020;371:m3840.
5. Shewhart WA. *Economic control of quality of manufactured product*. New York: D. Van Nostrand Company, Inc; 1931.
6. Lovegrove J, Valencia O, Treasure T, et al. Monitoring the results of cardiac surgery by variable life-adjusted display. *Lancet*. 1997;350:1128–1130.
7. Poloniecki J, Valencia O, Littlejohns P. Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. *BMJ*. 1998;316:1697–1700.
8. Duclos A, Lifante JC, Ducarroz S, Soardo P, Colin C, Peix JL. (2011), Influence of Intraoperative Neuromonitoring on Surgeons' Technique During Thyroidectomy. *World J Surg*. 2014;35:773–778.
9. Page ES. Continuous inspection schemes. *Biometrika*. 1954;41:100–115.
10. Steiner SH, Cook RJ, Farewell VT, et al. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics*. 2000;1:441–452.
11. Montgomery DC. *Introduction to Statistical Quality Control*. 8th Edition. Wiley; 2019.