

Use of cumulative summation (CUSUM) as a tool for early feedback and monitoring in robot-assisted radical prostatectomy outcomes and performance

Garson Chan, MD¹; Stephen E. Pautler, MD^{1,2}

¹Division of Urology, Department of Surgery, Western University, London, ON, Canada; ²Division of Surgical Oncology, Department of Oncology, Western University, London, ON, Canada

Cite as: *Can Urol Assoc J* 2019;13(2):53-8. <http://dx.doi.org/10.5489/auaj.5350>

Published online July 31, 2018

Abstract

Introduction: Today's surgical practice has evolved, with increasing emphasis on quality assurance. Many forms of quality-control monitoring have been suggested, but they are often impractical or difficult to implement. Cumulative summation (CUSUM) is a simple method to provide visual feedback before significant quality issues arise. We present our initial use and practical application of CUSUM in a surgical practice.

Methods: A retrospective analysis was applied to a prospectively collected database of 577 sequential patients who have undergone robot-assisted radical prostatectomy from a single surgeon over a 10-year period. Outcome measures were analyzed with CUSUM, which included a composite complication score, continence rates, length of hospital stay, biochemical recurrence, and need for adjuvant radiation. If any outcomes were out of control, they would cross the CUSUM failure line.

Results: CUSUM chart-plotting for incontinence demonstrated an initial upward slope followed by trending to a new safety limit. Additionally, outcomes in complications and biochemical recurrence did not reach the established safety boundaries. Length of stay and radiation outcomes did initially cross the safety line, but were improved over time.

Conclusions: The use of CUSUM in clinical practice can fulfill the need for quality assurance. CUSUM plotting in our practice reflected the initial learning curve, followed by ongoing maintenance and improvement in performance. These changes were consistent with the implementation of changes in surgical techniques. Although this tool was used retrospectively, this strengthens our argument to implement the tool prospectively and assess real-time refinement of surgeon skill. We have demonstrated that CUSUM can be appropriately used to assure quality control in a surgical practice.

Introduction

There is a shift in modern medicine, with increasing importance on evidence-based practice and quality assurance. This change has encouraged approaches to identify objective measures of quality and accountability. However, there is a practical issue in applying these standards and being able to evaluate clinical outcomes in real-world practice.¹ The use of cumulative summation (CUSUM) has been suggested for both surveillance and quality control.

A CUSUM chart or graph is a visual representation of a trend in the outcomes of a series of consecutive procedures over time. It is designed to quickly detect changes in performance. The measured outcome must be binary (i.e., did the event occur → yes or no?). This is then easily graphed and allows useful feedback. Every outcome is plotted and allows for early trends to be visualized. Other advantages of CUSUM charting include tracking learning curves, competency, and quality assurance. CUSUM allows early detection of small changes so that corrective action can be taken if necessary, thus it is a timely warning system. The binary system is easy to use, objective to record, and intuitive. It enables simple comparisons between current and past performance, or between trainees, or consultants.²

CUSUM has been used in many different industries, and more recently emphasis has been placed on its application in medicine.^{1,3-5} In fact, CUSUM has been recommended by the Royal Australasian College of Surgeons as a tool for self-analysis and audit.⁶ Since then, more recent publications of CUSUM in urology have followed, but implementation remains limited.⁷⁻¹⁰ We believe that CUSUM can be used in a surgical practice to assess quality outcomes. Our objective was to apply CUSUM charting to analyze and evaluate our performance in robot-assisted radical prostatectomy (RARP) patient outcomes.

Methods

A single fellowship-trained uro-oncologist surgeon performed all RARPs at Western University in London, Ontario. Patients treated between April 2005 and August 2015 were prospectively included in the study (REB#13086E). A prospective database of 577 patients was collected that included patient demographics, as well as postoperative factors and postoperative complications. We retrospectively looked at several outcomes, including a composite complication score (CCS), continence rates, length of stay (LOS), use of radiation, and biochemical recurrence (BCR). The CCS compiles all postoperative RARP outcomes within the first year. This includes transfusions, peritonitis, urinary tract complications, need for gastrointestinal stoma creation, strictures or contractures, fistula formation, hospital readmission, and death. The score is based on a recent 2014 Ontario health initiative working group looking at all prostatectomy outcomes in Ontario.¹¹

Statistical analysis was performed and CUSUM graphs were generated retrospectively. CUSUM charting was used to assess surgeon performance and outcomes in comparison to acceptable standards from the literature. These charts were constructed with the methodology described by Rogers.¹² If performance is unacceptable, the CUSUM slopes upward. In this way, CUSUM graphs represent cumulative failure charting. Any sustained slope changes are signals providing early identification that a quality issue might exist. Type I and II error rates were set at 0.05%, in keeping with previous literature. These rates are displayed as control lines, which act as boundaries for our curves. If the graph rises above the 'unacceptable' line, performance is said to be unacceptable, but if the graph remains between the two control lines, monitoring should continue.

The difficulty in constructing CUSUMs is in setting acceptable and unacceptable rates. However, this can also be seen as an advantage because what might be acceptable in one patient population might be unacceptable in another. We chose limits based on a combination of recent literature, expert opinion, and local consensus. It is important to realize that an acceptable rate is not solely a level that a surgeon aims to, but is also a level at which a surgeon has reached a standard the surgeon believes is safe. Consideration can then be given by surgeons to resetting the graph so they do not get into 'credit' for their previously low rates.

We chose boundaries for post RARP incontinence (PRI) at one year, with $\geq 10\%$ as an unacceptable range based on >1 pad per day.¹³ We defined BCR at $>25\%$ based on the American Urological Association (AUA) definition of having confirmatory prostate-specific antigen (PSA) value ≥ 0.2 ng/ml.^{14,15} We also looked at radiation with $>25\%$ post-RARP.¹⁶⁻¹⁸ We considered LOS based on <3 days, starting from patient entering the hospital to discharge, and a CCS $\geq 10\%$. Again, these were binary outcomes (i.e., yes or no).

Results

CUSUM graphs were generated for 577 patients on outcomes of LOS, CCS, BCR, radiation, and PRI. The graphs illustrate simply the various changes in surgical outcomes for our patient population over a consecutive 10-year period. Fig. 1 demonstrates the curves for LOS, with the grey lines acting as the upper failure boundary. Fig.1A illustrates the absolute curve, with Fig.1B plotted on a cumulative log-likelihood ratio test. Similarly, Figs. 2–5 are illustrated with grey acting as the failure line for the outcome. When the line is crossed, as in Figs. 1, 4, and 5, the process is out of control. This is the warning marker and signals the need for corrective action. The CUSUM graph in Fig.1 approached and crossed the upper boundary near patient 150 and stayed elevated until after patient 200. Figs. 2 and 3 for CCS and BCR did not cross the upper boundary. The radiation CUSUM graph shown in Fig. 4 did cross the boundary quite early, near patient 20, and did not decline until after patient 140. Most interestingly, the PRI CUSUM chart in Fig. 5 showed an upper boundary crossing near patient 120, with an abrupt drop near patient 150, and another sharp decline around patient 375.

Discussion

Surgical quality and quality assurance have become increasingly important in urology. The surgeon is responsible for balancing the needs of the individual patient and the hospital, as well as cost implications when introducing a new technique or surgical technology. The use of tools for self-appraisal allows for objective measurements to assess individual ability and progression or deterioration. Performance feedback is invaluable for physicians doing technically demanding and complex procedures. However, this is an area that receives little attention during surgical training and continued medical education. This area is becoming increasingly important in our professional development, as the system we serve demands more tangible methods to allocate scarce resources. In addition, this technique identifies areas for improvement, enhances self-esteem, and develops self-awareness for the surgeon. Importantly, it allows users to identify reasons for discrepancies and may become a positive platform for career promotion and progression. For regulatory bodies, this may supplement supervisory systems that lack regular direct supervision. In these numerous ways, CUSUM is a substantially beneficial tool that is underused.

The graphs from our CUSUM results clearly depict changes in rates over the surgical series and are excellent markers of changes in surgical technique. Although the curve did approach and cross the unacceptable rate for LOS, radiation, and PRI, ongoing technical improvements allowed the curves to once again return to acceptable rates. We believe part of

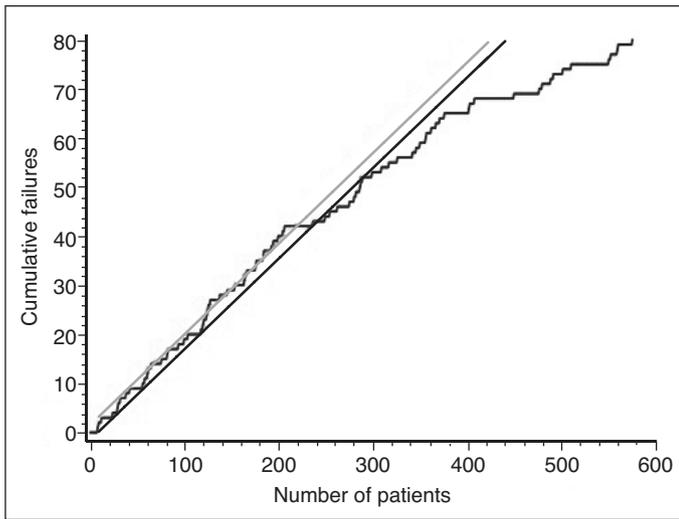


Fig. 1A. Plot of cumulative failures for the outcome of length of hospital stay greater than 3 days (n=577 patients). The plotted line (jagged black) is the cumulative failure. The acceptable and unacceptable failure rates were set at 10% (black) and 30% (grey), respectively.

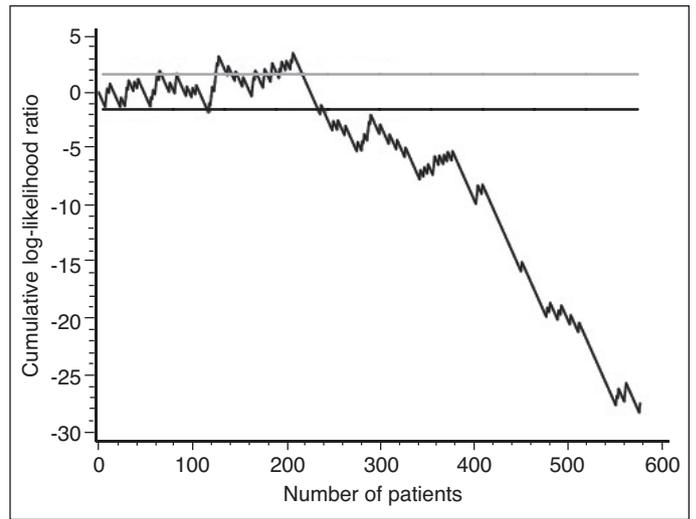


Fig. 1B. Plot of cumulative log-likelihood ratio test for the outcome of length of hospital stay greater than 3 days (n=577 patients). The plotted line (jagged black) is the cumulative failure. The acceptable and unacceptable failure rates were set at 10% (black) and 30% (grey), respectively.

the ‘unacceptable’ performance reflected a short learning curve for RARP. For instance, in the first 100 cases, there is an upward slope in LOS, radiation, and incontinence rates, seen in Figs. 1, 4, and 5, with improvement over time and with increased caseloads. The CCS CUSUM is within limits and even better over time, with an eventual plateau (Fig. 2).

In Fig. 1 LOS, the early upslope is likely related to the learning curve: robot and room setup time, operating time, and postoperative course, as patients were initially managed with the same postoperative pathway as patients who underwent open radical prostatectomy. However, over time, the pathway for RARP recovery changed, as reflected in the improvement seen in the CUSUM graph. There are also slope changes that we discovered correlate accurately with resident

and fellow rotation changes. This is not likely the only cause, but highlights how this interesting application alerts us that a change happened and allows us to investigate further.

When we consider the CUSUM graphs for BCR and radiation (Figs. 3 and 4) we notice an early rise in radiation (Fig. 4), which may correspond to different practice patterns, treatments, or patient preference over a decade ago. However, when compared to Fig. 3 for BCR, it is difficult to determine conclusively the effect radiation had on BCR.

Interestingly, the CUSUM for incontinence rates shows a marked decline around patients 150 and 375, seen in Fig. 5. This first change may be attributable to the introduction of a new technique — posterior reconstruction (“Rocco stitch”)

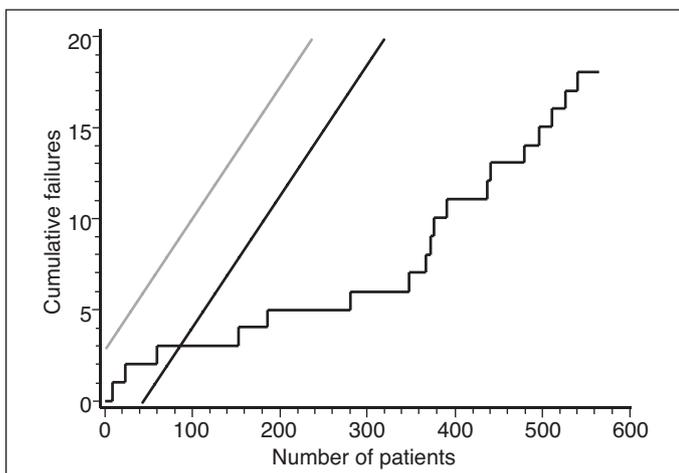


Fig. 2A. Plot of cumulative failures for the outcome of complications within 5 years of procedure (n=563 patients). The plotted line (jagged black) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 10% (grey), respectively.

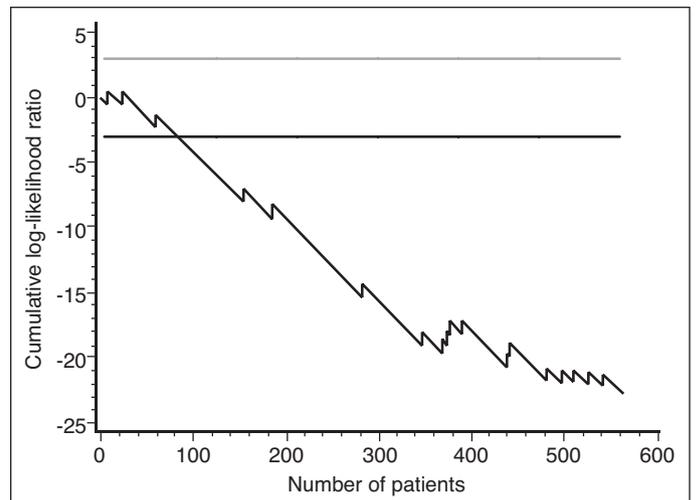


Fig. 2B. Plot of cumulative log-likelihood ratio test for the outcome of complications within 5 years of procedure (n=563 patients). The plotted line (jagged black) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 10% (grey), respectively.

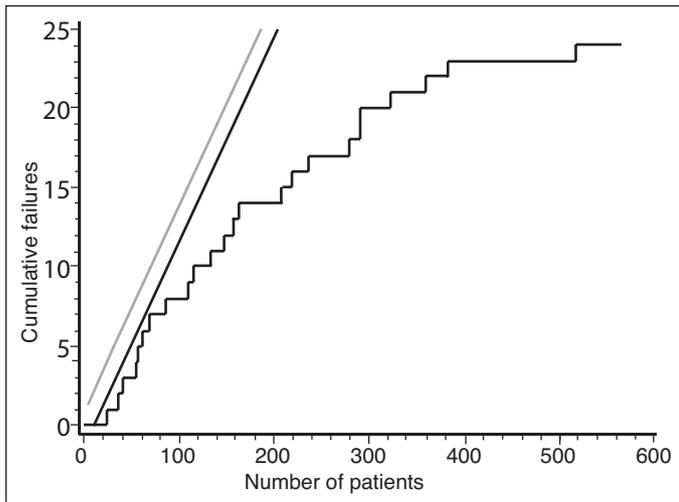


Fig. 3A. Plot of cumulative failures for the outcome of prostate-specific antigen (PSA) >0.2x2 over 10 years (n=563 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 25% (grey), respectively.

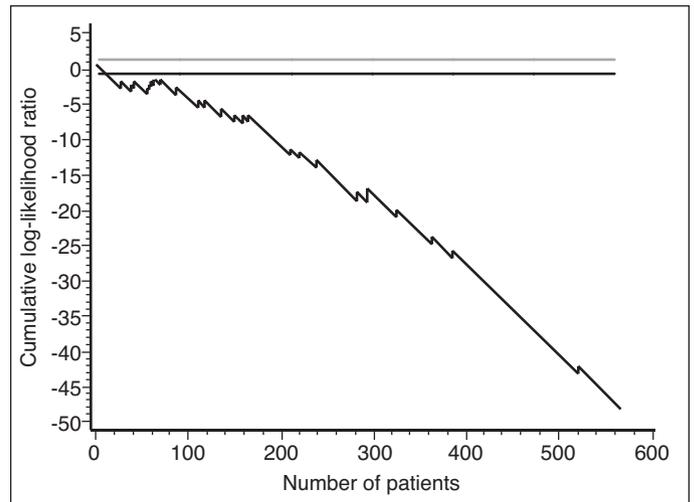


Fig. 3B. Plot of cumulative log-likelihood ratio test for the outcome of prostate-specific antigen (PSA) >0.2x2 over 10 years (n=563 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 25% (grey), respectively.

— and the positive impact on outcomes. A second downward slope correlates again with another modification in technique. In this case, the change was the introduction of a barbed suture for the running anastomosis and a posterior reconstruction with simultaneous puboperineoplasty to create an autologous sling. These are just some examples of how CUSUM can help in self-analysis and easy visualization to identify changes that may need attention.

Limitations include the retrospective nature of how the CUSUM values were applied. To be of greater value, a prospective analysis in real-time should be applied. However, this paper illustrates the important clinical role that may promote the acceptance of these tools into daily practice.

Additionally, these results validate the positive impact of subtle technical modifications to the RARP technique. Although this study was performed retrospectively, implementation of key changes was clearly evident during the CUSUM analysis. This further promotes the use of this tool prospectively, providing preventative and corrective analysis. With this quality framework, we will continue to prospectively collection information with confidence, using the CUSUM technique to analyze ongoing quality control and correct deviations early on if necessary. Our data affirms the next step in our use of this tool and implementing real-time monitoring and analysis. We encourage surgeons to join us and engage in this quality assurance measure.

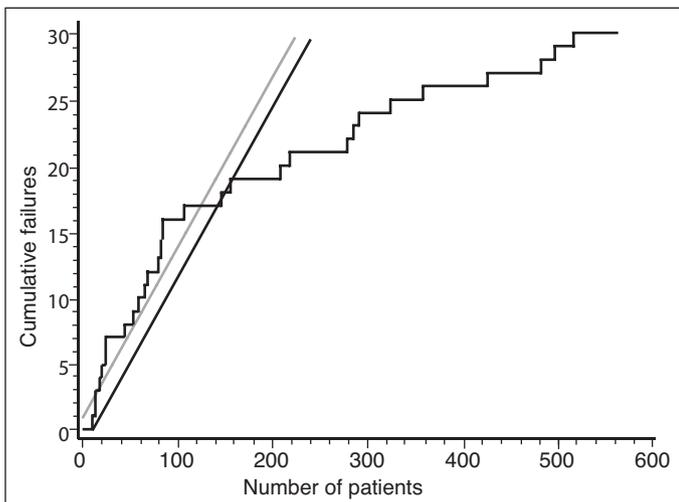


Fig. 4A. Plot of cumulative failures for the outcome of the need for adjuvant radiation therapy (n=563 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 25% (grey), respectively.

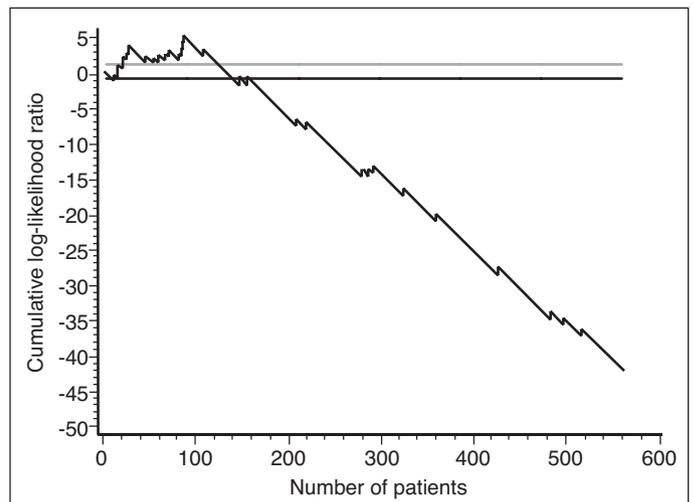


Fig. 4B. Plot of cumulative log-likelihood ratio test for the outcome of the need for adjuvant radiation therapy (n=563 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 25% (grey), respectively.

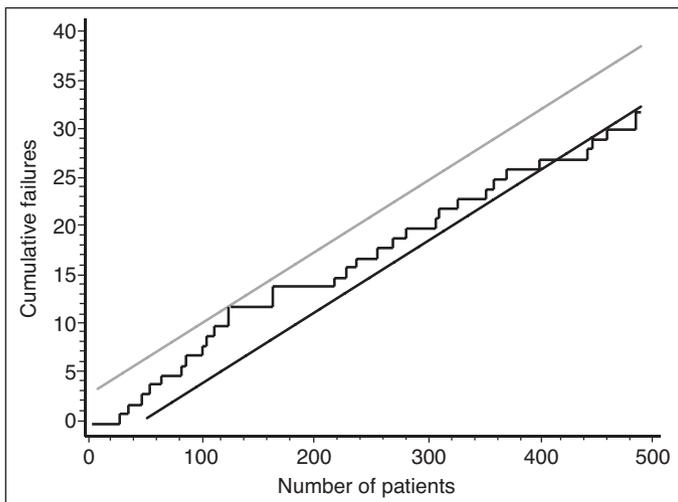


Fig. 5A. Plot of cumulative failures for the outcome of PPD ≤ 1 (n=488 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 10% (grey), respectively.

CUSUM is objective and easily understood, but is limited to only quantitative measures with binary outcomes. CUSUM will not show improvements that do not change the binary outcome. This discounts incalculable values, such as patient care, patient satisfaction, communication, and teamwork. It can be time-consuming and difficult to completely compare CUSUM from different surgeons due to multiple factors: a heterogeneous patient population, health system factors (hospitals, staffing, etc.), and other factors resulting in diverse outcomes. Thus, CUSUMs are best applied to individual surgeons' outcomes. CUSUMs do not directly measure the contributions by trainees, such as residents and fellows, where a new assistant might lead to significantly longer operating room time, blood loss, etc. While these contributions should even out over time, the CUSUM calculation might be affected. In addition, while crossing the lower boundary is considered to be positive, the graph should be reset to bring the CUSUM back to baseline so that surgeons will not get 'credit' for previous low rates of the measured outcome and they can continue to be accurately monitored. This 'reset' allows for continuous monitoring and improvement. As well, this tool was able to confirm desirable competency of surgeon performance and patient selection despite complex patients at a high-volume tertiary referral centre. Lastly, there is sparse literature on individual surgeon performance using CUSUM analysis to which to compare our outcomes in a meaningful manner.

Conclusion

CUSUM can be a simple yet dynamic and versatile technique that allows us to identify the trend of deteriorating performance early, prompting preventative and corrective measures. This is useful during the learning curve and for

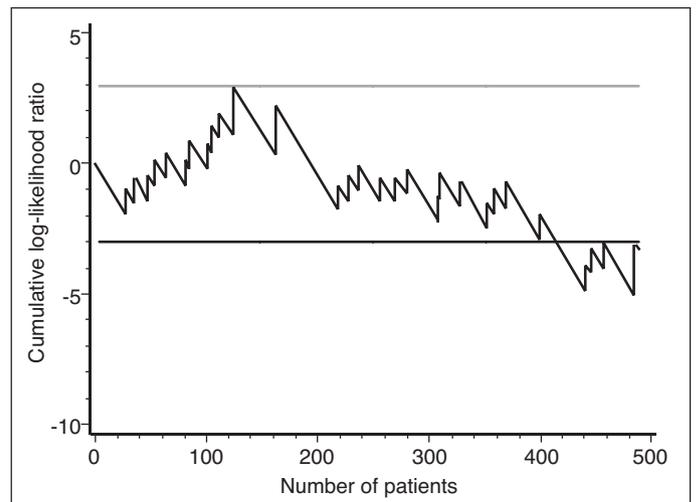


Fig. 5B. Plot of cumulative log-likelihood ratio test for the outcome of PPD ≤ 1 (n=488 patients). The plotted line (jagged back) is the cumulative failure. The acceptable and unacceptable failure rates were set at 5% (black) and 10% (grey), respectively.

quality control in ongoing surgical performance. It also allows accountability and reassurance to our patients and health system. CUSUM methodology can be applied to RARP, as the procedure is common, the outcome is well-defined, and it is closely related to surgeon skill. Future studies will include a real-time prospective CUSUM analysis of outcomes from heterogeneous groups of patients using a risk-adjusted or difficulty-adjusted CUSUM process, which is more time-consuming and complex. However, even a simple CUSUM application is able to detect important changes, thereby imploring us to investigate further and stay proactive with respect to surgical quality assurance.

Evidence-based medicine and quality self-assurance is guiding the future of healthcare. CUSUM is a valuable tool that is easy to use for a variety of outcomes. In this study, RARP outcomes were analyzed to identify quality variances and identified favourable outcome responses to technical changes. This instrument encourages accountability to our patients and for the use of health resources. There are challenges to CUSUM implementation, but this alone should not impede our usage of this valuable tool. Future real-time use of these instruments will guide our quality assurance and accountability.

Competing interests: The authors report no competing personal or financial interest related to this work.

Acknowledgements: The authors would like to thank Workhorse Consulting Group's Andrew McClure and Norine Foley for statistical advice and consulting, as well as Dr. Marie Dion and Linda Nott for compiling and assistance with the exhaustive RARP database.

This paper has been peer-reviewed.

References

1. Steiner SH, Woodall WH. Debate: What is the best method to monitor surgical performance? *BMC Surgery* 2016;16:15. <https://doi.org/10.1186/s12893-016-0131-8>
2. Young A, Miller J P, Azarow K. Establishing learning curves for surgical residents using Cumulative Summation (CUSUM) Analysis. *Curr Surg* 2005;62:330-4. <https://doi.org/10.1016/j.cursur.2004.09.016>
3. Bolson S, Colon M. The use of the CUSUM technique in the assessment of trainee competence in new procedures. *Int J Health Care Qual Assur* 2000;12:433-8. <https://doi.org/10.1093/intqhc/12.5.433>
4. Institute Of Medicine. Crossing the quality chasm: A new health system for the 21st century. Institute of Medicine's Quality of Health Care in America project 2001.
5. Biau DJ, Resche-Rigon M, Godiris-Petit G, et al. Quality control of surgical and interventional procedures: A review of the CUSUM. *Qual Saf Health Care* 2007;16:203-7. <https://doi.org/10.1136/qshc.2006.020776>
6. The Royal Australasian College of Surgeons. Surgical audit and peer review, c2008. Available at: https://www.surgeons.org/media/20671311/surgical_audit_and_peer_review_guide_2014.pdf. Accessed June 6, 2017.
7. Siddiqui KM, Izawa JJ. Systematic methods for measuring outcomes: How they may be used to improve outcomes after radical cystectomy. *Arab J Urol* 2015;13:122-7. <https://doi.org/10.1016/j.aju.2015.02.008>
8. Williams AK, Chalasani V, Martinez CH, et al. Cumulative summation graphs are a useful tool for monitoring positive surgical margin rates in robot-assisted radical prostatectomy. *BJU Int* 2011;107:1648-52. <https://doi.org/10.1111/j.1464-410X.2010.09634.x>
9. Maguire T, Mayne CJ, Terry T, et al. Analysis of the surgical learning curve using the cumulative sum (CUSUM) method. *NeuroUrol Urodyn* 2013;32:964-7. <https://doi.org/10.1002/nau.22375>
10. Nabi G, McLoman L, Cook J, et al. MP-16.14: Endoscopic extrapitoneal radical prostatectomy: CUSUM analysis of learning curve. *Urol J* 2009;74:S123-4. <https://doi.org/10.1016/j.urolgy.2009.07.1319>
11. Ontario Health Technology Advisory Committee (OHTAC). Robotic-assisted minimally invasive prostatectomy: OHTAC recommendation. Toronto: Queen's Printer for Ontario; 2014 January.
12. Rogers CA, Reeves BC, Caputo M, et al. Control chart methods for monitoring cardiac surgical performance and their interpretation. *J Thorac Cardiovasc Surg* 2004;128:811-9. <https://doi.org/10.1016/j.jtcvs.2004.03.011>
13. Ficarra V, Novara G, Rosen RC, et al. Systematic review and meta-analysis of studies reporting urinary continence recovery after robot-assisted radical prostatectomy. *Eur Urol* 2012;62:405-17. <https://doi.org/10.1016/j.eururo.2012.05.045>
14. Pound CR, Partin AW, Eisenberger MA, et al. Natural history of progression after PSA elevation following radical prostatectomy. *JAMA* 1999;281:1591-7. <https://doi.org/10.1001/jama.281.17.1591>
15. Stephenson AJ, Kattan MW, Eastham JA, et al. Defining biochemical recurrence of prostate cancer after radical prostatectomy: A proposal for a standardized definition. *J Clin Oncol* 2006; 24:3973-8. <https://doi.org/10.1200/JCO.2005.04.0756>
16. Zincke H, Oesterling JE, Blute ML, et al. Long-term (15 years) results after radical prostatectomy for clinically localized (stage T2c or lower) prostate cancer. *J Urol* 1994; 152:1850-7. [https://doi.org/10.1016/S0022-5347\(17\)32399-6](https://doi.org/10.1016/S0022-5347(17)32399-6)
17. Stephenson AJ, Scardino PT, Kattan MW, et al. Predicting the outcome of salvage radiation therapy for recurrent prostate cancer after radical prostatectomy. *J Clin Oncol* 2007;25:2035-41. <https://doi.org/10.1200/JCO.2006.08.9607>
18. Bottke D, Wiegel T. Adjuvant radiotherapy after radical prostatectomy: Indications, results, and side effects. *Urol Int* 2007;78:193-7. <https://doi.org/10.1159/000099336>

Correspondence: Dr. Stephen E. Pautler, Division of Urology, Department of Surgery and Division of Surgical Oncology, Department of Oncology, Western University, London, ON, Canada; stephen.pautler@sjhc.london.on.ca