Direct comparison of risk-adjusted and non–risk-adjusted CUSUM analyses of coronary artery bypass surgery outcomes

Richard J. Novick, MD,[a] Stephanie A. Fox, BSc, RRCP,[a] Larry W. Stitt, MSc,[b] Thomas L. Forbes, MD,[c] and Stefan Steiner, PhD[d]

Objective: We previously applied non–risk-adjusted cumulative sum methods to analyze coronary bypass outcomes. The objective of this study was to assess the incremental advantage of risk-adjusted cumulative sum methods in this setting.

Methods: Prospective data were collected in 793 consecutive patients who underwent coronary bypass grafting performed by a single surgeon during a period of 5 years. The composite occurrence of an "adverse outcome" included mortality or any of 10 major complications. An institutional logistic regression model for adverse outcome was developed by using 2608 contemporaneous patients undergoing coronary bypass. The predicted risk of adverse outcome in each of the surgeon's 793 patients was then calculated. A risk-adjusted cumulative sum curve was then generated after specifying control limits and odds ratio. This risk-adjusted curve was compared with the non–risk-adjusted cumulative sum curve, and the clinical significance of this difference was assessed.

Results: The surgeon's adverse outcome rate was 96 of 793 (12.1%) versus 270 of 1815 (14.9%) for all the other institution's surgeons combined ($P = .06$). The non–risk-adjusted curve reached below the lower control limit, signifying excellent outcomes between cases 164 and 313, 323 and 407, and 667 and 793, but transgressed the upper limit between cases 461 and 478. The risk-adjusted cumulative sum curve never transgressed the upper control limit, signifying that cases preceding and including 461 to 478 were at an increased predicted risk. Furthermore, if the risk-adjusted cumulative sum curve was reset to zero whenever a control limit was reached, it still signaled a decrease in adverse outcome at 166, 653, and 782 cases.

Conclusions: Risk-adjusted cumulative sum techniques provide incremental advantages over non–risk-adjusted methods by not signaling a decrement in performance when preoperative patient risk is high.

From the Divisions of Cardiac Surgery[a] and Vascular Surgery,[c] London Health Sciences Centre, and the Biostatistical Support Unit,[b] Department of Clinical Epidemiology and Biostatistics, the University of Western Ontario, London, Ontario, Canada; and the Department of Statistics and Actuarial Science, University of Waterloo,[d] Waterloo, Ontario, Canada.

An abstract of this work was presented in a poster format at the Canadian Cardiovascular Congress in Montreal on October 24, 2005.

Funded by a grant from the Lawson Health Research Institute Internal Research Fund, London, Ontario, Canada.

Received for publication Dec 2, 2005; accepted for publication Feb 21, 2006.

Address for reprints: Richard J. Novick, London Health Sciences Center, University Hospital, 339 Windermere Road, PO Box 5339, London, Ontario, Canada N6A 5A5.

J Thorac Cardiovasc Surg 2006;132:386-91

0022-5223/$32.00

doi:10.1016/j.jtcvs.2006.02.053

ACD

Coronary artery bypass grafting (CABG) is one of the most intensely scrutinized procedures in medicine, and its outcomes are often publicly disseminated. Therefore, it is important that highly sensitive statistical methods be used to assess CABG outcomes. We previously used cumulative sum (CUSUM) failure techniques to analyze the transition from on-pump to off-pump CABG[1] and the learning curve in robotic-assisted cardiac procedures,[2] and found that CUSUM methods were more sensitive than standard statistical techniques such as chi-square testing in detecting an improvement or decrement in performance. Increasing interest has been paid recently to the role of CUSUM and other sequential probability ratio tests in the analysis of the outcomes of cardiac surgery procedures,[3-6] and the limitations of our previously described method of CUSUM analysis have been highlighted.[4] In addition, it has been suggested that CUSUM and other sequential probability ratio methods should be risk-adjusted to "level the playing field" with

Abbreviations and Acronyms
ARL = average run length
CABG = coronary artery bypass grafting
CUSUM = cumulative sum
VLAD = variable life-adjusted display

respect to the analysis of surgical outcomes. Nonetheless, although the theoretic advantages of a risk-adjusted, as opposed to a non–risk-adjusted CUSUM analysis, are presumed, no direct comparison between the 2 techniques in a cohort of patients undergoing CABG has, to our knowledge, been published in the cardiac surgical literature. Furthermore, the clinical significance of differences between the 2 techniques has not been demonstrated in patients undergoing CABG.

During the past year we formed a multidisciplinary, multiinstitutional research group that presented risk-adjusted CUSUM analyses after different types of surgical procedures.[7-9] We used a risk-adjusted CUSUM technique that was developed by Steiner and associates[10,11] and that was used to assess risk-adjusted mortality in a busy tertiary referral intensive care unit.[12] In this study, we used this risk-adjusted CUSUM method to analyze a surgeon's CABG outcomes over a 5-year interval. We hypothesized that the risk-adjusted CUSUM technique would prove superior to non–risk-adjusted CUSUM methods in a direct "head-to-head" comparison and that this difference in outcomes would be clinically significant. Furthermore, we also explored the use of variable life-adjusted display (VLAD) methods[4,13] as a complementary technique to assess CABG outcomes.

Patients and Methods

The subjects of this study included all patients undergoing elective and urgent CABG by a single surgeon and a group of 6 colleagues at the London Health Sciences Centre, University Hospital, between July 1999 and December 2004. Detailed preoperative, operative, and postoperative data were entered prospectively into a cardiac surgery database that tracked postoperative outcomes. Although the date of onset of the study preceded a legislative requirement that individual patient consent be routinely obtained, the lead author asked the Research Ethics Board of the University of Western Ontario to review and determine individual patient consent requirements for this study. The Research Ethics Board issued an exemption from this requirement on May 27, 1999. Moreover, in a further review requested by the author the Research Ethics Board of the University of Western Ontario approved the study on June 20, 2005 (Review 11527E).

In this study, we used the composite occurrence of an "adverse outcome" as our primary outcome variable, which included mortality or any of 10 predefined major complications (stroke, reoperation for bleeding, mediastinitis, sepsis, sternal dehiscence, new intra-aortic balloon pump use, perioperative myocardial infarction, renal failure, respiratory insufficiency, and life-threatening arrhythmia). The operational definitions of each of these variables have been described.[1] Patients undergoing emergency or salvage CABG procedures, or CABG in addition to other surgical procedures, were excluded.

The first step in the analysis involved the development of an institutional logistic regression model for adverse outcome using 2608 contemporaneous, consecutive patients who underwent elective and urgent coronary artery bypass between July 1999 and December 2004. Variables considered for the model included age, gender, preoperative length of stay, ventricular grade, urgency status, primary or redo operation, body mass index, chronic obstructive pulmonary disease, myocardial infarction within 30 days before surgery, peripheral vascular disease, Canadian Cardiovascular Society angina class, diabetes (diet-controlled, treatment with oral agents, or insulin-dependent), history of transient ischemic attack or stroke, left main coronary artery disease, preoperative creatinine, congestive heart failure, palpable aortic atherosclerosis at surgery, and planned type of CABG procedure (on or off pump). By using stepwise logistic regression, allowing for entry and removal of variables at the 0.05 level, we determined the set of predictors of an adverse outcome in this institutional patient cohort, which included the first author's patients. This then permitted the calculation of the predicted risk of this composite occurrence in each of the surgeon's 793 patients. Both non–risk-adjusted and risk-adjusted CUSUM curves were then generated, after specifying the control limits and odds ratio.[7,8,10-12] In this study we chose a control limit of ± 2 and odds ratios of 3/2 (1.5, a 50% increase) and 2/3 (0.67, a 33% decrease). The rationale for these choices is highlighted in the "Discussion" and the Appendix; the Appendix includes the CUSUM statistic.

When both the risk-adjusted and non–risk-adjusted CUSUM curves are being plotted to detect changes in the adverse outcome rate, patient scores associated with an adverse outcome are positive, whereas those associated with the avoidance of an adverse outcome are negative. With the risk-adjusted analysis the "penalty" for an adverse event in a low-risk patient is larger than that of an adverse event in a high-risk patient.[8] The resulting graph includes 2 curves, with the chart designed to identify a decrement in surgical performance placed above that designed to identify an improvement in performance with a decreased complication rate. The risk-adjusted CUSUM procedure is designed to signal when the plot falls above the upper control limit or below the lower control limit, indicating evidence of a 3/2 (50%) increase or 2/3 (33%) decrease in the odds ratio, respectively (see CUSUM statistic and weights in Appendix). In addition, the risk-adjusted CUSUM plot can be reset to zero after each signal to continue the quality assurance process for subsequent patients.[8] To prove or disprove our a priori hypothesis, we compared the non–risk-adjusted CUSUM curve of the individual surgeon's performance with the risk-adjusted CUSUM curve. In addition, we further assessed the individual surgeon's risk-adjusted performance using VLAD methods,[4,13] which monitored the cumulative expected minus actual adverse outcome rate for the entire 793-patient experience. The theoretic relationship between CUSUM and VLAD methods has been reviewed in a recent publication by Grunkemeier and associates,[14] and 95% pointwise 2-sided prediction limits were plotted in our VLAD analysis using the method de-

TABLE 1. Predictors of an adverse outcome on multivariable analysis (n = 2608)

Variable	Odds ratio (95% CI)	*P* value
LVEF < 35%	1.74 (1.33-2.28)	<.001
Preoperative creatinine >120 μmol/L	1.73 (1.27-2.34)	<.001
Increasing age (per decade)	1.43 (1.26-1.62)	<.001
Increasing CCS angina class	1.36 (1.12-1.67)	.002
Cerebrovascular disease (previous CVA or TIA)	1.59 (1.15-2.20)	.006
Redo surgery	2.18 (1.17-4.07)	.015
COPD	1.41 (1.06-1.88)	.018

The intercept term in the multivariable logistic regression model was −5.60. *CI*, Confidence interval; *LVEF*, left ventricular ejection fraction; *CCS*, Canadian Cardiovascular Society; *CVA*, cerebrovascular accident; *TIA*, transient ischemic attack; *COPD*, chronic obstructive pulmonary disease.

scribed in that article. Although these prediction limits do not account for multiple testing and are thus difficult to interpret, they provide useful information concerning expected variation.[14]

Results

The overall predicted risk of an adverse outcome was 14.9% ± 8.1% (standard deviation) in the surgeon's patients versus 13.7% ± 8% (standard deviation) for all of the other surgeons' patients combined ($P < .001$). Nonetheless, during the course of the study the actual adverse event rate was 96/793 (12.1%) and 270/1815 (14.9%) in the surgeon's patients versus his colleagues' patients, respectively ($P = .06$). The predictors of an adverse outcome in the institutional series of 2608 patients are shown in Table 1. The Hosmer-Lemeshow goodness of fit statistic for the multivariable logistic regression model was $X_8^2 = 6.740$, $P = .565$, whereas the C statistic (area under the curve) was 0.688. The logistic regression equation generated by this analysis was then used to calculate the risk of an adverse outcome in each of the surgeon's individual cases, which ranged from 2.2% to 48.0%.

Results of the non–risk-adjusted CUSUM analysis are shown in Figure 1. The non–risk-adjusted CUSUM curve reached below the lower control limit, signifying excellent clinical outcomes, between cases 164 and 313, 323 and 407, and 667 and 793. However, the non–risk-adjusted CUSUM curve transgressed the upper control limit between cases 461 and 478, suggesting a decrement in performance.

As depicted in Figure 2, the intervals of improved performance on the risk-adjusted CUSUM analysis were similar to those in the non–risk-adjusted CUSUM analysis. However, the risk-adjusted CUSUM curve never transgressed the upper control limit, signifying that cases preceding and including 461 to 478 were at an increased predicted risk of an adverse outcome. Confirmatory evidence in this regard was provided by a post hoc analysis of the predicted risk of an adverse outcome in each 50-patient cohort, as shown in Table 2. The data in that table clearly show that the second highest predicted risk of death or major complications was in the 50-patient cohort preceding patients 461 to 478. This increased risk was adjusted for in the risk-adjusted CUSUM curves in Figure 2, but not in the non–risk-adjusted Figure 1 curves; thus, the risk-adjusted CUSUM curve never transgressed the upper control limit throughout the 793-patient cohort. Furthermore, if the risk-adjusted CUSUM curve was reset to zero whenever a control limit was reached, it still signaled a decrease in risk-adjusted adverse outcome at 166, 653, and 782 cases (Figure 3).

A VLAD analysis of the cumulative expected minus actual adverse outcome rate (Figure 4) showed that there were 23 fewer adverse events in the surgeon's 793-patient experience than predicted by the institutional multivariable logistic regression model. The pointwise prediction limits in Figure 4 were included as a useful way of assessing variability, but they do not account for multiple testing.[14] Similar to the pattern shown by the risk-adjusted CUSUM curves, the VLAD analysis demonstrated an increasing trend of adverse outcomes at the beginning of the cohort and from cases 400 to 478, versus a decreasing trend of adverse events in the remainder of the surgeon's experience.

Discussion

Increasing attention has been paid to the use of sequential probability ratio tests (including the CUSUM method) in the

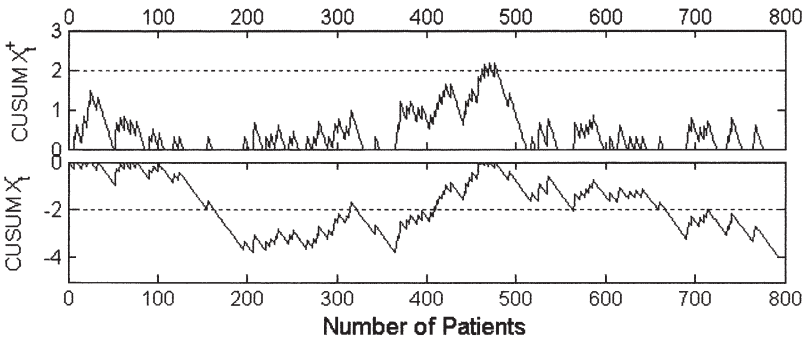

Figure 1. Non–risk-adjusted CUSUM analysis of the surgeon's adverse outcome rate. The curve designed to identify a decrement in surgical performance is placed above that designed to identify improvement in performance with a decreased complication rate. Pre-set control limits (*dashed horizontal lines*).

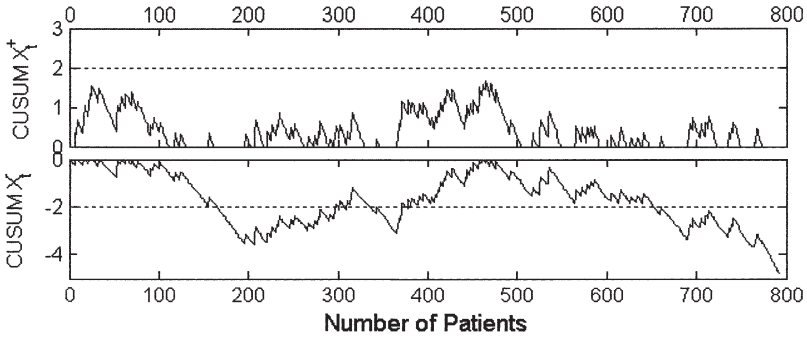

Figure 2. Risk-adjusted CUSUM analysis of the surgeon's adverse outcome rate. The curve designed to identify a decrement in surgical performance is placed above that designed to identify improvement in performance with a decreased complication rate. Pre-set control limits (*dashed horizontal lines*).

analysis of surgical results during the past few years,[1-9,15] after the initial application of this technique in neonatal heart surgery by de Leval and colleagues.[16] Our group has used non–risk-adjusted CUSUM methods to analyze the learning curve of a cardiac surgeon in the first 10 years of practice[15] to examine the impact of a policy change from on-pump to off-pump CABG,[1] to assess the learning curve in robotic-assisted beating heart CABG,[2] and to analyze a surgeon's early experience with elective open abdominal aortic aneurysm repair.[17] All of these studies showed that the non–risk-adjusted CUSUM method was more sensitive than standard statistical techniques such as the chi-square test in identifying a cluster of surgical failures or successes. Nonetheless, a recent article suggested possible methodologic limitations of our non–risk-adjusted method of CUSUM analysis.[4] Furthermore, it has been presumed that contemporary CUSUM analyses should ideally be risk-adjusted, because the non–risk-adjusted method treats all patients identically and does not take into account variations in preoperative risk. The practical importance of risk adjustment was highlighted in the current study, in which the predicted risk of an adverse outcome in the surgeon's patients ranged from 2.2% to 48.0%.

TABLE 2. Predicted risk of major complications or death in each 50-patient interval in the surgeon's cohort

Surgery no.	Risk Mean (SD)	No. of deaths or major complications (%)
1-50	0.126 (0.067)	8 (16.0%)
51-100	0.142 (0.092)	8 (16.0%)
101-150	0.146 (0.083)	3 (6.0%)
151-200	0.142 (0.074)	2 (4.0%)
201-250	0.130 (0.068)	7 (14.0%)
251-300	0.146 (0.077)	8 (16.0%)
301-350	0.137 (0.070)	5 (10.0%)
351-400	0.151 (0.071)	8 (16.0%)
401-450	0.166 (0.080)	10 (20.0%)
451-500	0.163 (0.087)	8 (16.0%)
501-550	0.135 (0.081)	5 (10.0%)
551-600	0.172 (0.092)	6 (12.0%)
601-650	0.156 (0.091)	6 (12.0%)
651-700	0.144 (0.073)	4 (8.0%)
701-750	0.164 (0.081)	6 (12.0%)
751-793	0.166 (0.093)	2 (4.7%)

SD, Standard deviation.

The method of risk-adjusted CUSUM analysis used in this study was developed by Steiner and associates[10,11] and was used to track risk-adjusted outcomes in a tertiary referral adult intensive care unit.[12] The latter study detected a decrease in risk-adjusted intensive care unit mortality, which was temporally related to increased senior staffing levels and quality improvement efforts.[12] The Vascular Surgery group at the London Health Sciences Centre recently used the same technique to analyze risk-adjusted outcomes after elective open abdominal aneurysm repair[7] and ruptured abdominal aortic aneurysm surgery.[8] The former study demonstrated improved results after the first year of experience, which coincided with the adoption of routine intraoperative cell-saver use.[7] In the second study, the risk-adjusted CUSUM curve was set to signal with a halving or doubling of the odds ratio; the CUSUM plot signaled at patient 45 by crossing the lower control limit, indicating improved risk-adjusted outcomes one third of the way through the patient cohort.[8]

Although the theoretic advantages of a risk-adjusted, as opposed to a non–risk-adjusted, CUSUM analysis of CABG outcomes are presumed, to our knowledge only 1 "head-to-head" comparison between the 2 methods has been published, in a biostatistical journal.[10] However, a recent publication compared risk-adjusted versus non–risk-adjusted CUSUM analyses in 2 cohorts of patients undergoing heart or lung transplantation.[18] The findings of our study indicated that despite wide variations in preoperative patient risk, the non–risk-adjusted (Figure 1) and risk-adjusted (Figure 2) CUSUM curves paralleled each other closely, signifying that the overall impact provided by risk adjustment was of modest clinical significance. Nonetheless, there were more extreme fluctuations of the CUSUM curves in the non–risk-adjusted analysis than in the risk-adjusted analysis. Furthermore, although improved performance was de-

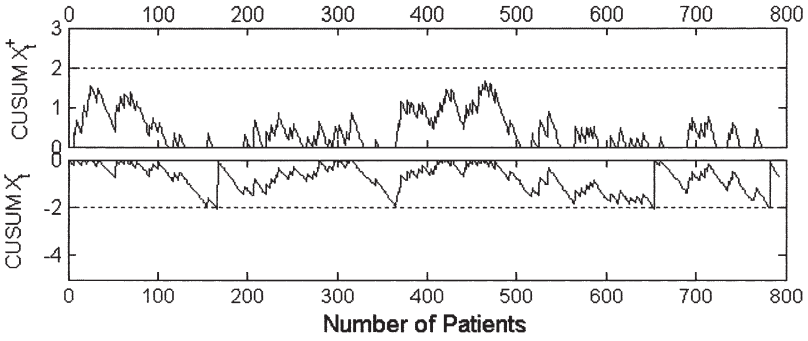

Figure 3. Risk-adjusted CUSUM analysis of the surgeon's adverse outcome rate, with resetting of the CUSUM curve to zero whenever a pre-set control limit was reached. The curve designed to identify a decrement in surgical performance is placed above that designed to identify improvement in performance with a decreased complication rate. Pre-set control limits (*dashed horizontal lines*).

tected at similar intervals in the non–risk-adjusted and risk-adjusted CUSUM analyses, the non–risk-adjusted curve showed an increased adverse outcome rate between patients 461 and 478. In the risk-adjusted analysis, the curve did not come close to reaching the upper control limit, implying that the patients preceding and including 461 to 478 were at a higher predicted risk. This qualitative impression was confirmed in a post hoc analysis of the predicted risk of major complications or death in each 50-patient interval in the surgeon's cohort, as presented in Table 2. Indeed, patients 400 to 500 had the highest predicted risk of an adverse outcome in the series to that point. It is interesting that patients 551 to 600 had an even higher predicted risk, yet the non–risk-adjusted CUSUM curve was relatively flat and the risk-adjusted CUSUM curve was downsloping during this interval.

Although this study showed that risk-adjusted CUSUM analyses have advantages over non–risk-adjusted CUSUM methods in a retrospective analysis of surgical performance, the greatest advantage of the former technique will likely be when it is used prospectively. An analytic technique that accounts for the patient's preoperative risk will facilitate an accurate analysis of factors that have an immediate and direct impact on clinical practice. Furthermore, the risk-adjusted CUSUM curve can be reset to zero whenever a boundary line is transgressed, thus continuing the quality assurance process for subsequent patients. This zero reset method could also facilitate an interpretation of the impact of new quality improvement initiatives on postoperative outcomes.

With a CUSUM analysis there is an inherent tradeoff in the choice of the control limit. A large limit provides protection against false signals resulting from chance, but will be slow to detect a change in the process performance (ie, odds ratio), whereas a small control limit will detect changes quickly but have more frequent false signals resulting from chance when the process is unchanged. We can quantify the CUSUM chart performance in terms of the average run length (ARL), which represents the average number of patients before a signal. Large ARLs are better when the process performance is unchanged, whereas small ARLs are more optimal for large process changes. Given the observed patient mix in our study, a control limit set at 2.0 resulted in an ARL of 604. Thus, even if the odds ratio is unchanged, one would expect a signal on average for every 604 patients. This value of ARL was deemed to provide

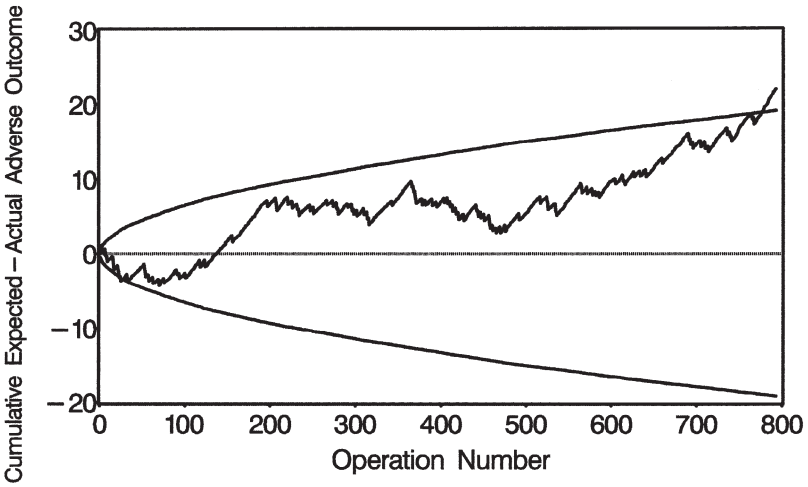

Figure 4. VLAD analysis of the surgeon's cumulative expected minus actual adverse outcome rate. The 95% pointwise 2-sided prediction limits are plotted as aids in interpretation.[14]

reasonable protection from frequent false alarms, while being sensitive to changes in the odds ratio.

The limitations of this study include the fact that it compares only a single surgeon's patient outcomes with those of his colleagues in a single institution; nonetheless, the 2 patient cohorts numbered 793 and 2608 patients, respectively, and the experience spanned 5 years. A second limitation is that the primary outcome variable used represented the composite occurrence of mortality and/or any of 10 predefined major complications. Some of these complications (eg, reoperation for bleeding) are less consequential than others, such as a perioperative stroke. A more targeted analysis incorporating only irrevocable adverse outcomes such as mortality and stroke could have been performed, but would have required a larger number of patients to generate the same statistical power. Another potential limitation is the choice of odds ratio of 1.5 (3/2) and 2/3 and a control limit of ± 2 in this study. Odds ratios of 3/2 and 2/3 were used because a relative increase of 3/2 or decrease of 2/3 in the composite outcome was deemed to be clinically significant. Furthermore, the control limit was set lower than in a recent risk-adjusted CUSUM analysis of early mortality after ruptured abdominal aortic aneurysm repair,[8] so that a decrement or improvement in surgical performance could be more rapidly detected. Use of a control limit less than 2.0 would have significantly decreased the ARL at which a positive or negative deviation in care would have been detected by chance alone, and would likely have been too sensitive an analysis for this study.

Conclusion

This study demonstrated that the results of the non–risk-adjusted and risk-adjusted CUSUM analyses paralleled each other; however, there were incremental advantages to the latter technique, which did not inappropriately signal a decrement in performance when the predicted risk was high. Future studies should use risk-adjusted CUSUM methods prospectively to assess the impact of new clinical protocols and quality assurance methods on surgical outcomes.

References

1. Novick RJ, Fox SA, Stitt LW, et al. Cumulative sum failure analysis of a policy change from on-pump to off-pump coronary artery bypass grafting. *Ann Thorac Surg*. 2001;72:S1016-S21.
2. Novick RJ, Fox SA, Kiaii BB, et al. Analysis of the learning curve in telerobotic, beating heart coronary artery bypass grafting: a 90 patient experience. *Ann Thorac Surg*. 2003;76:749-53.
3. Blackstone EH. Monitoring surgical performance. *J Thorac Cardiovasc Surg*. 2004;128:807-10.
4. Rogers CA, Reeves BC, Caputo M, Ganesh JS, Bonser RS, Angelini GD. Control chart methods for monitoring cardiac surgical performance and their interpretation. *J Thorac Cardiovasc Surg*. 2004;128:811-9.
5. Spiegelhalter DJ. Monitoring clinical performance: a commentary. *J Thorac Cardiovasc Surg*. 2004;128:820-2.
6. Treasure T, Gallivan S, Sherlaw-Johnson C. Monitoring cardiac surgical performance: a commentary. *J Thorac Cardiovasc Surg*. 2004;128:823-5.
7. Forbes TL, Steiner SH, Lawlor DK, DeRose G, Harris KA. Risk-adjusted analysis of outcomes following elective open abdominal aortic aneurysm repair. *Ann Vasc Surg*. 2005;19:142-8.
8. Harris JR, Forbes TL, Steiner SH, Lawlor DK, DeRose G, Harris KA. Risk-adjusted analysis of early mortality after ruptured abdominal aortic aneurysm repair. *J Vasc Surg*. 2005;42:387-91.
9. Novick RJ, Fox SA, Stitt LW, Forbes TL, Steiner SH. Risk-adjusted CUSUM analysis is superior to non risk-adjusted methods in CABG outcome analysis. *Can J Cardiol*. 2005;21:132C.
10. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics*. 2000;1:441-52.
11. Steiner SH, Cook RJ, Farewell VT. Risk-adjusted monitoring of binary surgical outcomes. *Med Decis Making*. 2001;21:163-9.
12. Cook DA, Steiner SH, Cook RJ, Farewell VT, Morton AP. Monitoring the evolutionary process of quality: risk-adjusted charting to track outcomes in intensive care. *Crit Care Med*. 2003;31:1676-82.
13. Lovegrove J, Valencia O, Treasure T, Sherlaw-Johnson C, Gallivan S. Monitoring the results of cardiac surgery by variable life-adjusted display. *Lancet*. 1997;350:1128-30.
14. Grunkemeier GL, Xing Y, Furnary AP. Cumulative sum techniques for assessing surgical results. *Ann Thorac Surg*. 2003;76:663-7.
15. Novick RJ, Stitt LW. The learning curve of an academic cardiac surgeon: use of the CUSUM method. *J Card Surg*. 1999;14:312-20.
16. de Leval MR, Francois K, Bull C, Brawn W, Spiegelhalter D. Analysis of a cluster of surgical failures. Application to a series of neonatal arterial switch operations. *J Thorac Cardiovasc Surg*. 1994;107:914-24.
17. Forbes TL. A cumulative analysis of an individual surgeon's early experience with elective open abdominal aortic aneurysm repair. *Am J Surg*. 2005;189:469-73.
18. Rogers CA, Ganesh JS, Banner NR, Bonser RS. Cumulative risk-adjusted monitoring of 30-day mortality after cardiothoracic transplantation: UK experience. *Eur J Cardiothorac Surg*. 2005;27:1022-9.

Appendix

Definition of CUSUM statistic and weights in this study

$X_t^+ = \max(0, X_{t-1}^+ + w_t)$ and $X_t^- = \min(0, X_{t-1}^- - w_t)$ are used for the CUSUM statistic to detect increases and decreases, respectively. In both expressions the weights are based on the formula below, where p_t is the predicted preoperative risk (as given by the logistic regression) and for X_t^+ and X_t^- OR_A is 3/2 and 2/3, respectively.

$$w_t = \begin{cases} \log\left[\dfrac{OR_A}{(1 - p_t + OR_A p_t)}\right] & \textit{if} \text{ adverse outcome for patient t} \\ \log\left[\dfrac{1}{1 - p_t + OR_A p_t}\right] & \textit{otherwise} \end{cases}$$

ACD