ISQua

International Journal for Quality in Health Care, 2017, 29(3), 343–348
doi: 10.1093/intqhc/mzx036
Advance Access Publication Date: 21 April 2017
Methodology Paper

OXFORD

Methodology Article

Improved implementation of the risk-adjusted Bernoulli CUSUM chart to monitor surgical outcome quality

MATTHEW J. KEEFE[1], JUSTIN B. LODA[1], AHMAD E. ELHABASHY[2,3], and WILLIAM H. WOODALL[1]

[1]Department of Statistics, Virginia Tech, 405 Hutcheson Hall (0439), 250 Drillfield Drive, Blacksburg, VA 24061, USA, [2]Grado Department of Industrial and Systems Engineering, Virginia Tech, 250 Perry Street, Blacksburg, VA 24061, USA, and [3]Production Engineering Department, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt

Address reprint requests to: Matthew James Keefe, Department of Statistics, Virginia Tech, Blacksburg, VA 24061, USA. Tel: +302-530-6553; E-mail: mjkeefe@vt.edu

Editorial Decision 22 February 2017; Accepted 2 March 2017

Abstract

Methodology issue: The traditional implementation of the risk-adjusted Bernoulli cumulative sum (CUSUM) chart for monitoring surgical outcome quality requires waiting a pre-specified period of time after surgery before incorporating patient outcome information.

Proposed solution: We propose a simple but powerful implementation of the risk-adjusted Bernoulli CUSUM chart that incorporates outcome information as soon as it is available, rather than waiting a pre-specified period of time after surgery.

Evaluation: A simulation study is presented that compares the performance of the traditional implementation of the risk-adjusted Bernoulli CUSUM chart to our improved implementation. We show that incorporating patient outcome information as soon as it is available leads to quicker detection of process deterioration.

Advice to practitioners: Deterioration of surgical performance could be detected much sooner using our proposed implementation, which could lead to the earlier identification of problems.

Key words: cumulative sum, risk-adjustment, statistical process control, surgical performance

Introduction

Statistical process monitoring techniques are becoming more widely used in healthcare applications. In particular, methods for monitoring surgical outcomes are used to detect deterioration in surgical performance as quickly as possible to avoid undesirable consequences. For a comprehensive review of monitoring techniques for surgical outcome quality, see Woodall *et al.* [1].

Frequently in healthcare applications, patients' pre-operative risks vary widely across the population. To account for the heterogeneity across patients, many risk-adjusted monitoring techniques have been developed. These techniques incorporate information about each patient's potential risk factor characteristics, such as age, gender, and health status, into the calculation of a statistic that is monitored. For example, if a patient who was old and unhealthy died shortly after surgery, this result would be more likely than an instance where a young, healthier patient died. Risk-adjusted monitoring techniques use this information about each individual to allow for meaningful monitoring of surgical outcomes. We propose an implementation scheme for the risk-adjusted Bernoulli cumulative sum (CUSUM) chart that significantly improves the time until detection of process deterioration.

The risk-adjusted Bernoulli CUSUM chart proposed by Steiner *et al.* [2] is a control chart that can be used to monitor 30-day mortality rates prospectively, where each patient has a predicted probability of 30-day mortality based on a risk-adjustment model. This approach can be used to monitor the rate of other adverse events, not just mortality. When discussing our proposed implementation for risk-adjusted monitoring, we focus exclusively on the risk-adjusted Bernoulli CUSUM chart since it has the strongest theoretical justification and is the most popular approach [3].

When monitoring surgical outcomes, the outcome of interest is usually based on some pre-specified period of time after surgery. For example, Steiner *et al.* [2] considered death within 30 days after surgery. When outcomes such as this one are used, there is a period of time during which the outcome of the patient may be unknown. Specifically, for patients who survive the entire time period (e.g. 30 days), their outcomes are unknown until the end of the time period. However, for patients who do experience the adverse event sooner than the end of the time period (e.g. death occurs within 30 days after surgery), the outcome is obtained earlier. The standard risk-adjusted Bernoulli CUSUM method monitors patients in the order in which they undergo surgery, despite the fact that information about many of their outcomes is known sooner than 30 days. For example, if we are monitoring 30-day mortality rate and a patient dies one day after surgery, the traditional implementation of the risk-adjusted Bernoulli CUSUM chart may not incorporate this outcome information into the chart until 29 days later. We propose an implementation scheme for the risk-adjusted Bernoulli CUSUM chart that incorporates patients' surgical outcomes as soon as they are available, rather than waiting the length of the pre-specified time window to incorporate their information. Our proposed implementation considers all adverse outcomes immediately.

Methods

Risk-adjusted Bernoulli CUSUM procedure

The risk-adjusted Bernoulli CUSUM chart proposed by Steiner *et al.* [2] is capable of monitoring binary outcomes while adjusting for prior risk of the adverse event occurring. A risk-adjustment model is fit to a Phase I sample so that predicted probabilities of the adverse event of interest (e.g. 30-day mortality) can be calculated. Phase I data is typically a historical sample collected that characterizes how the process being monitored operates under stable conditions. For a comprehensive overview of Phase I and its importance in statistical process monitoring, see Jones-Farmer *et al.* [4]. When monitoring surgical outcome quality, a logistic regression model is typically fit with covariate information about the patients in order to obtain the predicted probability of the adverse event of interest.

The risk-adjusted CUSUM chart is designed to detect a shift in an odds ratio R from R_0 to $R_1 > R_0$, where typically $R_0 = 1$. We let p_t represent the predicted probability of the event of interest for the tth observation. Thus, the odds of the event can be calculated as $p_t/(1-p_t)$. Therefore, under the in-control odds R_0, the odds of the event of interest are given by $R_0 p_t/(1-p_t)$. Likewise, under R_1, the odds of the event of interest are given by $R_1 p_t/(1-p_t)$. Thus, the corresponding in-control and out-of-control probabilities are given by the following equations:

$$p_{0t} = \frac{R_0 p_t}{1 - p_t + R_0 p_t} \quad \text{and} \quad p_{1t} = \frac{R_1 p_t}{1 - p_t + R_1 p_t}, \tag{1}$$

This leads directly to the calculation of the score for the risk-adjusted Bernoulli CUSUM given by the following equation:

$$W_t = \begin{cases} \log\left[\dfrac{(1 - p_t + R_0 p_t)R_1}{(1 - p_t + R_1 p_t)R_0}\right] & \text{if } y_t = 1 \\ \log\left[\dfrac{1 - p_t + R_0 p_t}{1 - p_t + R_1 p_t}\right] & \text{if } y_t = 0. \end{cases} \tag{2}$$

The CUSUM statistics are given by the following equation:

$$S_t = \max(0, S_{t-1} + W_t), \tag{3}$$

where, $S_0 = 0$ and a signal is given when $S_t > h$ for $h > 0$.

Currently, the observations would be indexed in the order in which the patients undergo surgery. However, when operating under this assumption, the outcomes for those patients who do not survive are observed earlier than 30 days, but for patients who survive, outcomes are not observed until the end of the 30-day time window.

Proposed approach

We show that it is beneficial to incorporate patient outcome information as soon as it is obtained, rather than waiting until the end of the waiting period associated with the outcome. To illustrate this result, we propose a simple implementation scheme for the risk-adjusted Bernoulli CUSUM chart that incorporates patients' surgical outcomes as soon as they are available.

To illustrate our proposed implementation scheme of the risk-adjusted Bernoulli CUSUM chart, we will use the same data set from a United Kingdom center for cardiac surgeries as was used by Steiner *et al.* [2]. The data set consists of 6994 patients from the years 1992 through 1998 and contains descriptive information such as surgery date, pre-operative Parsonnet score, and the number of days before any patient mortality. The Parsonnet score is a single value used to characterize a patient's overall health status [5]. The first two years of data (1992–93) were taken as Phase I data and were used to fit the following logistic regression model for risk-adjustment:

$$\text{logit}(p_t) = -3.68 + 0.077X_t, \tag{4}$$

where X_t is the Parsonnet score of patient t and p_t is the pre-operative risk of mortality within 30 days of surgery for this patient.

In this set of data, for those patients who died within 30 days after surgery, the distribution of days lived is heavily right-skewed. Figure 1 shows that the majority of patients from the first two years of data who died within 30 days, died within a week after surgery. Practically, if a patient dies 5 days after surgery, it does not seem reasonable to wait up to an additional 25 days to include the outcome in the monitoring scheme. Figure 1 illustrates why it could be very helpful to incorporate outcome information as soon as it is available. If deaths occur closer to the 30-day point, then use of our proposed method would not be as advantageous.

Proposed implementation

We propose an implementation scheme where we continue to consider the patients in the order of their surgical operations as is done in the traditional implementation [2]. Although there is no way to know in advance if a particular patient is going to live or die, we initially assume in advance that all patients are going to survive until the end of the 30-day time window, and then, as the operations occur and the actual adverse outcomes are obtained, the chart is updated accordingly, rather than waiting 30 days. Using this implementation, all patients would be assumed to survive until determined otherwise. Specifically, the chart would be updated when

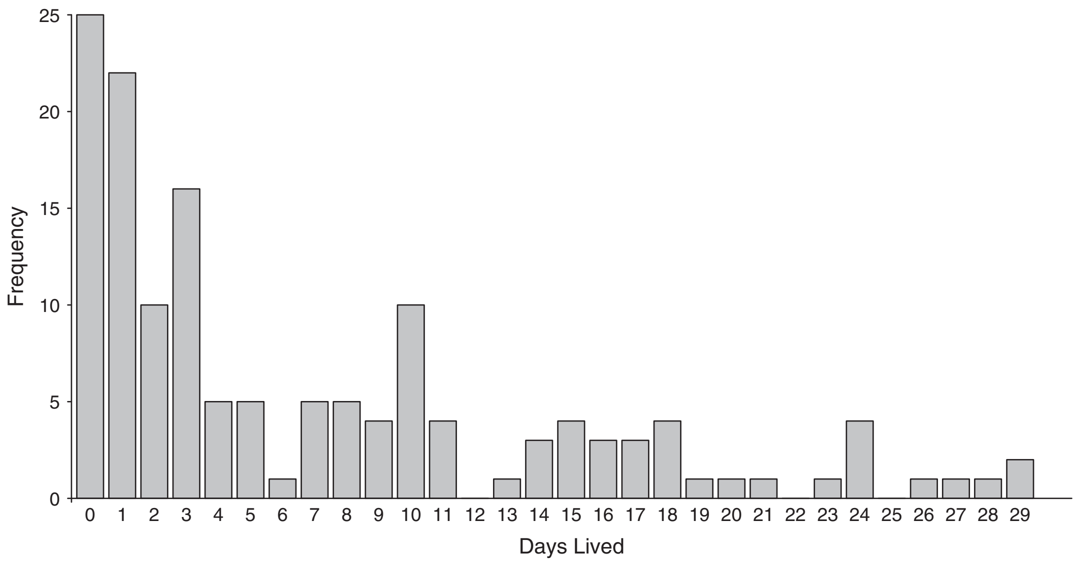

Figure 1 Distribution of the number of days lived for patients who died within 30 days of surgery (1992–93).

patients who were assumed to survive actually die within 30 days after surgery. Thus, as soon as a patient dies, we would incorporate this outcome into the chart immediately. In this manner, some previously charted CUSUM statistics would dynamically change throughout the monitoring process.

The proposed implementation is a recursive process that would be initiated every time an adverse event occurs. In this manner, there would be a 30-day moving window in which the CUSUM statistics could be updated. For instance, if we are now considering the outcomes obtained on the 45th day of monitoring, then this backtracking window would update all outcomes considered from the 16th day to the 45th day. Patients who were not operated on within the last 30 days no longer need to have their control chart values updated. It is important to realize that at the end of the monitoring process once all outcome information has been obtained, the CUSUM chart using the proposed implementation will be identical to that with the traditional implementation. The dynamic updating of the CUSUM statistics through time allows our proposed implementation to detect process deterioration sooner than the traditional implementation.

Illustration

As an illustration of the proposed implementation, consider the small, artificial data set of surgical outcomes with an associated 30-day mortality response given in Table 1. Each row corresponds to a patient. Information regarding whether or not each patient survived 30 days after surgery, as well as the number of days lived after surgery was recorded. The number of days lived for patients who survived 30 days after surgery was recorded as 30+. The day on which outcome information was obtained is calculated based on each patient's operation day and the number of days they lived after surgery.

Using the proposed implementation scheme, each plotted CUSUM statistic corresponds to a patient. However, the chart is updated through time for each day that new outcome information is obtained. Figure 2 illustrates how the chart updates information as it is obtained. On Day 1, all patients operated on thus far are assumed to survive 30 days, and thus all CUSUM statistics are zero.

Table 1 Example data set

Patient	Operation day	Survived 30 days?	Days lived	Outcome day
1	1	Yes	30+	31
2	1	No	3	4
3	1	No	2	3
4	2	Yes	30+	32
5	2	Yes	30+	32
6	3	No	1	4
7	3	No	4	7
8	3	Yes	30+	33
9	4	No	3	7
10	5	Yes	30+	35

The CUSUM statistics for the third patient and all subsequent patients are updated on Day 3 when it was learned that the third patient died. All patients after the third patient on Day 3 are still assumed to survive 30 days, and thus result in a decreasing trend of CUSUM statistics. On Day 4, patients 2 and 6 died and the CUSUM statistics are updated. Finally, on Day 7 outcome information for patients 7 and 9 was obtained and the CUSUM statistics were updated again. Hence, previously plotted CUSUM statistics are dynamically updated as outcome information is obtained, similar to reliability monitoring schemes involving dynamically changing observations [6]. Also, note that the resulting CUSUM chart for Day 7 is identical to the CUSUM chart obtained after Day 35 using the traditional implementation. For this illustration, a control limit of $h = 2.5$ is used, but another control limit could be used in practice. If the proposed implementation is used, the chart signals on Day 7, whereas if the traditional implementation is used, the chart would not signal until Day 34. Clearly, detection time can be reduced by using the proposed implementation.

Simulation study

A simulation study was conducted using the UK cardiac surgery data to compare the in-control and steady state out-of-control performance of the proposed implementation scheme with that of the traditional implementation of the risk-adjusted Bernoulli CUSUM

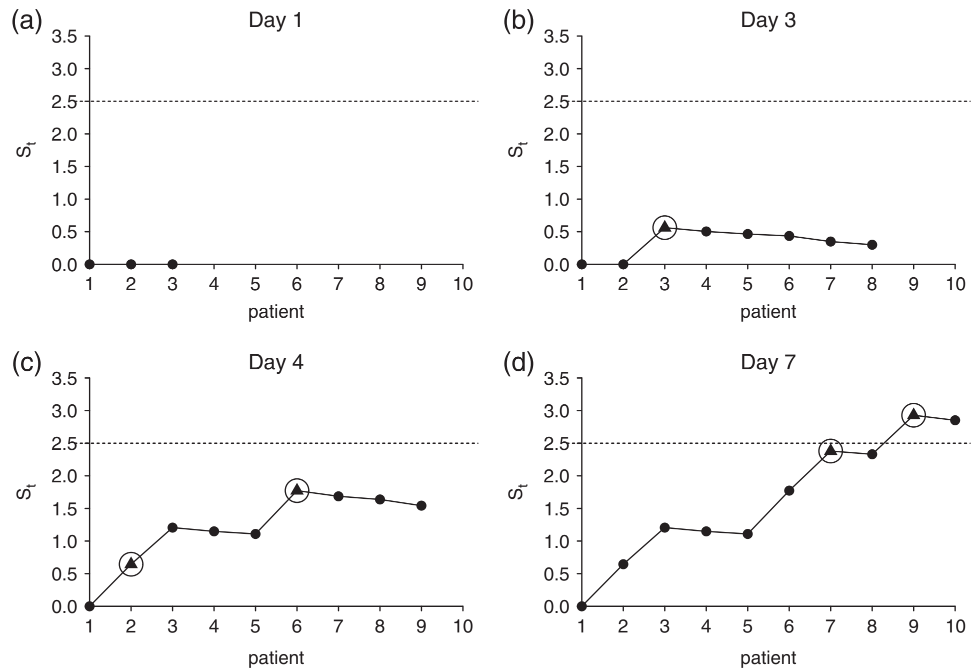

Figure 2 Illustration of proposed implementation of the risk-adjusted Bernoulli CUSUM chart. Circled triangles indicate patients who died resulting in CUSUM statistics that were updated on the given day.

chart. For both implementations, each CUSUM statistic plotted on the control chart corresponds to a patient, where patients are ordered by the date of their operation. It is more informative to consider the average run length (ARL) in number of days, rather than in number of patients. It is important to note that our proposed implementation will never signal before the traditional implementation in terms of the number of patients. When implementing the chart, however, the time until a signal can be determined by the number of days since monitoring began. The benefit in our implementation scheme is clearly seen in recognizing deterioration sooner, in terms of number of days, rather than number of patients, due to removing the 30-day wait time restriction. For our simulation study, the number of operations for a given day was drawn with replacement from the empirical distribution of the Phase I data and varies from one to eight. We also considered simulations where we fixed the number of operations per day to assess performance of the proposed implementation as the number of operations per day increases.

The simulation procedure used to compare our proposed implementation method to the traditional implementation method can be described in the following steps. For each simulated patient t,

1. Sample with replacement a Parsonnet score from the in-control empirical distribution.
2. Use Equation (4) to determine the predicted probability, p_t, of death within 30 days of surgery. Adjust p_t based on the assumed odds ratio R.
3. Generate a Bernoulli random variable with probability of 'success' p_t.
4. With the outcome obtained in Step 3, calculate the CUSUM statistic using Equation (3) with $R_1 = 2$.
5. Repeat steps 1–4 until $S_t > h$.

To be consistent with the work of Steiner *et al.* [2], we used an upper control limit of $h = 4.5$ which produces an in-control ARL, in terms of the number of patients, of ~7400 [7]. We set up the control chart so that it is designed to detect a shift of $R_1 = 2$. For the out-of-control simulations, the process was initially simulated as in-control under the baseline model with $R = 1$ for the first 50 patients to achieve steady state conditions and the odds of death within 30 days was shifted after patient 50. We considered various values of the odds ratio R between 1 and 10. Additionally, we considered values of $R = 2,3,4,5$ for fixed numbers of operations per day of 1 through 10. Each ARL simulation result is based on 1000 simulated control charts.

Results

The proposed implementation scheme shows improved detection time of process deterioration in terms of days. Typically, monitoring schemes are compared using ARL, where the in-control ARLs of the schemes are set to be equal and the out-of-control ARLs are observed for different size shifts. The results of the simulation study for various odds ratios R are provided in Table 2. We note that in our case, the in-control ARLs of the two implementation schemes are close, but not quite equal. With the proposed implementation, the in-control ARL is on average 15.8 days less than with the traditional implementation. Any false alarm obtained by the traditional method would likely be obtained at an earlier time by our proposed implementation scheme. As the size of the odds ratio R increases, the average time until detection for the proposed implementation improves. For this application, we can signal process deterioration up to 29 days sooner; however, if the waiting period for the response of interest were more than 30 days, more time could be saved. For example, in organ transplantation applications the monitoring of survival times is frequently used rather than the risk-adjusted Bernoulli CUSUM because the waiting period after transplantation is usually one year [8–10]. In this case, our proposed implementation makes the risk-adjusted Bernoulli CUSUM more

Table 2 ARL comparison (in days) for traditional and proposed implementation schemes

R	Traditional	Proposed	Difference (days)	% Reduction relative to traditional method
1	2080.9	2065.1	15.8	0.8
1.5	175.5	159.0	16.5	9.4
2	79.0	61.5	17.5	22.2
2.5	60.4	41.8	18.6	30.8
3	51.1	31.8	19.3	37.7
3.5	47.1	27.4	19.7	41.8
4	44.5	24.3	20.2	45.4
4.5	42.4	21.9	20.5	48.4
5	41.0	19.8	21.1	51.6
5.5	39.8	18.5	21.4	53.6
6	39.2	17.6	21.6	55.2
6.5	38.6	16.7	21.9	56.8
7	37.9	15.6	22.3	58.9
7.5	37.3	14.9	22.4	60.1
8	37.0	14.5	22.5	60.7
8.5	36.6	13.9	22.6	61.9
9	36.3	13.5	22.8	62.7
9.5	36.0	13.0	23.0	63.8
10	35.7	12.5	23.2	65.0

similar to methods with continuous updating schemes that consider time until event data and would lead to significantly improved performance in terms of days until detection of a process change.

Another important aspect to notice is that the time until detection of process deterioration will depend on the number of operations performed per day. Figure 3 shows that the out-of-control ARL improves for both the traditional and proposed implementations of the risk-adjusted Bernoulli CUSUM as the number of operations per day increases, yet the difference in days until detection between the two implementations does not change with the number of operations per day. However, with the proposed implementation one always detects the shift sooner than with the traditional implementation. As expected, larger shifts in the process result in lower ARLs for both methods. In the limiting case, if all adverse events occurred immediately (i.e. death on the first day) the improvement in detection for the proposed implementation would be exactly 29 days.

Discussion

With the traditional implementation of the risk-adjusted Bernoulli CUSUM chart, one monitors patient by patient in the order of operation with a waiting period to determine the outcome. Practically, it is inefficient to wait a specific time period, such as 30 days, if information

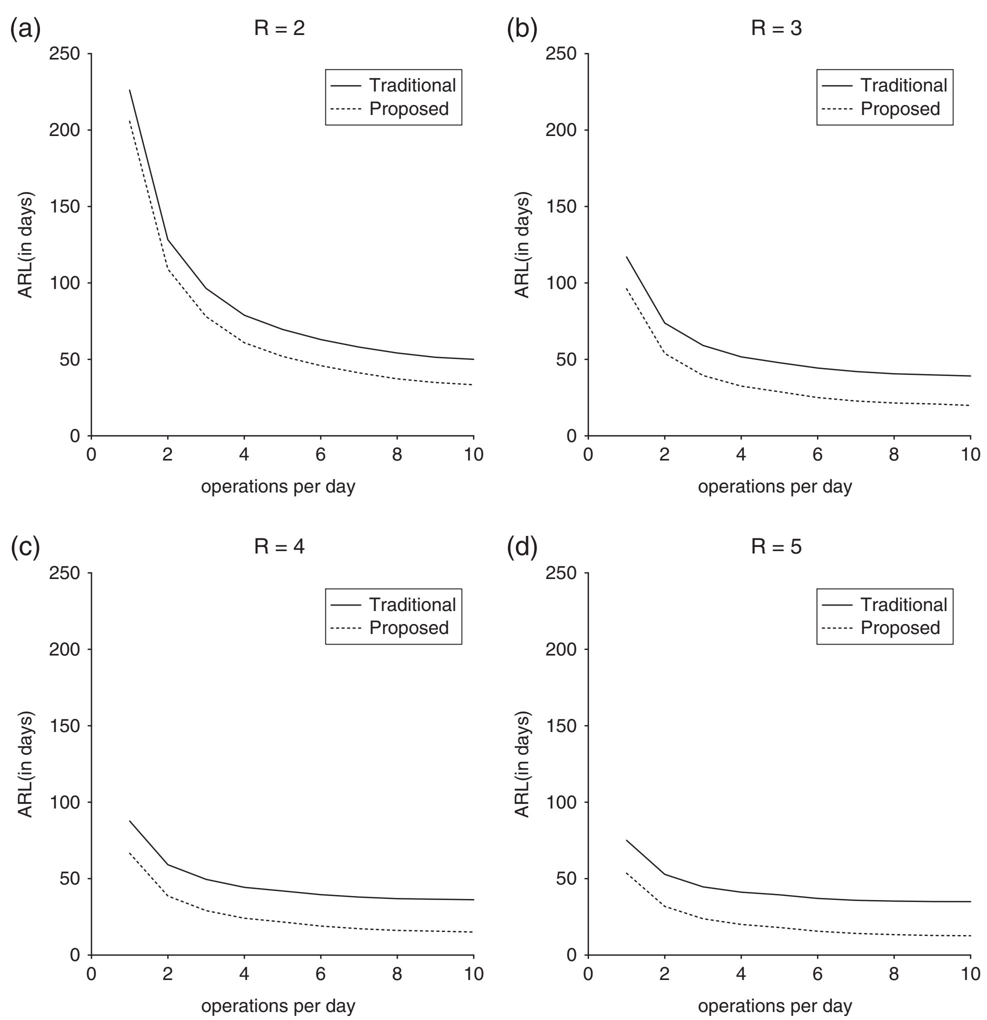

Figure 3 ARL (in days) by number of operations per day for (**a**) $R = 2$, (**b**) $R = 3$, (**c**) $R = 4$, and (**d**) $R = 5$ (based on 1000 simulated charts).

about some of the outcomes is available much sooner. We have proposed a more practical and appealing implementation scheme for the risk-adjusted Bernoulli CUSUM chart that incorporates outcome information as soon as it is available. We have illustrated that the proposed implementation significantly improves the time until detection of deterioration in the process, especially when most adverse outcomes occur toward the beginning of the waiting period.

Performance of our proposed monitoring scheme in terms of time until detection of process deterioration is limited by the effect of estimation error inherent in fitting the risk-adjustment model. Furthermore, this method is intended for binary outcomes that have a waiting period required before obtaining the outcome. If deaths occur closer to the end of the waiting period, then use of our proposed method would not be as advantageous, but would still perform as well as the traditional approach.

Other practical issues regarding the implementation of the risk-adjusted Bernoulli CUSUM chart have been addressed. For example, Tian *et al.* [7] discussed the impact of varying patient populations and its effect on chart performance. As a consequence, an appropriate monitoring scheme to account for varying patient populations is the use of dynamic probability control limits [11]. Additionally, situations arise in which observations happen concurrently and there is no way to determine the exact order [12]. Furthermore, Paynabar *et al.* [13] explored the importance of including other covariate information, such as surgeon information, into the risk-adjustment procedure. Jones and Steiner [14] studied the effect of Phase I estimation error on the performance of the risk-adjusted Bernoulli CUSUM chart. Also, Tang *et al.* [15] developed a risk-adjusted CUSUM chart for multi-responses, in cases where the response is not binary, but rather has several categories. This multi-response technique was further developed by Zhang *et al.* [16] to include dynamic probability control limits. Our proposed implementation could be applied directly to accommodate dynamic probability control limits, concurrent observations, covariate information, or multi-responses in order to improve performance.

Funding

The research of A.E. Elhabashy and W.H. Woodall was supported by National Science Foundation Grant CMMI-1436365.

References

1. Woodall WH, Fogel SL, Steiner SH. The monitoring and improvement of surgical-outcome quality. *J Qual Technol* 2015;**47**:383–99.
2. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics* 2000;**1**:441–52.
3. Steiner SH, Woodall WH. Debate: what is the best method to monitor surgical performance? *BMC Surg* 2016;**16**. DOI:10.1186/s12893-016-0131-8.
4. Jones-Farmer LA, Woodall WH, Steiner SH, Champ CW. An overview of phase I analysis for process improvement and monitoring. *J Qual Technol* 2014;**46**:265–82.
5. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. *Circulation* 1989;**79**:I3–I12.
6. Yashchin E. Design and implementation of systems for monitoring lifetime data. In: Lenz H-J, Schmid W, Wilrich P-T (eds). *Frontiers in Statistical Quality Control 10*. Heidelberg: Physica-Verlag, 2012: 171–95.
7. Tian W, Sun H, Zhang X, Woodall WH. The impact of varying patient populations on the in-control performance of the risk-adjusted CUSUM chart. *Int J Qual Health Care* 2015;**27**:31–6.
8. Biswas P, Kalbfleisch JD. A risk-adjusted CUSUM in continuous time based on the Cox model. *Stat Med* 2008;**27**:3382–3406.
9. Collett D, Sibanda N, Pioli S, Bradley JA, Rudge C. The UK scheme for mandatory continuous monitoring of early transplant outcome in all kidney transplant centers. *Transplantation* 2009;**88**:970–5.
10. Neuberger J, Madden S, Collett D. Review of methods for measuring and comparing center performance after organ transplantation. *Liver Transplant* 2010;**16**:1119–28.
11. Zhang X, Woodall WH. Dynamic probability control limits for risk-adjusted Bernoulli CUSUM control charts. *Stat Med* 2015;**34**:3336–48.
12. Keefe MJ, Franck CT, Woodall WH. Monitoring foreclosure rates with a spatially risk-adjusted Bernoulli CUSUM chart for concurrent observations. *J Appl Stat* 2017;**44**:325–41.
13. Paynabar K, Jin JH, Yeh AB. Phase I risk-adjusted control charts for monitoring surgical performance by considering categorical covariates. *J Qual Technol* 2012;**44**:39–53.
14. Jones MA, Steiner SH. Assessing the effect of estimation error on risk-adjusted CUSUM chart performance. *Int J Qual Health Care* 2012;**24**: 176–81.
15. Tang X, Gan FF, Zhang L. Risk-adjusted cumulative sum charting procedure based on multiresponses. *J Am Stat Assoc* 2015;**110**:16–26.
16. Zhang X, Loda JB, Woodall WH. Dynamic probability control limits for risk-adjusted CUSUM charts based on multi-responses. *Stat Med* 2017. (to appear).