应用概率统计　第 33 卷
第 5 期　2017 年 10 月

Chinese Journal of Applied Probability and Statistics
Oct., 2017, Vol. 33, No. 5, pp. 441-449
doi: 10.3969/j.issn.1001-4268.2017.05.001

Phase II Risk-Adjusted Geometric Control Charts for Monitoring Surgical Performance *

QI DeQuan

(*School of Biomedical Engineering, Jilin Medical University, Jilin, 132013*)

GENG Wei⋆

(*LPMC and Institute of Statistics, Nankai University, Tianjin, 300071*)

Abstract: In recent years, statistical process control (SPC) has been widely used to monitor the performance of clinical practitioners, such as surgeons and general practitioners. In this paper, two risk-adjusted geometric control charts namely cumulative sum (CUSUM) and weighted likelihood ratio test (WLRT) are proposed to monitor surgery performance in phase II. The performance of the proposed control charts is evaluated and compared by simulation experiments for different shift values in the parameters of a risk-adjusted logistic regression model in terms of the average run length (ARL) criterion. The results show that all methods work well in the sense that they can effectively detect shifts in the process parameters.

Keywords: change-point; risk-adjusted control chart; statistical process control; weighted likelihood ratio test

2010 Mathematics Subject Classification: 62P10

Citation: Qi D Q, Geng W. Phase II risk-adjusted geometric control charts for monitoring surgical performance [J]. Chinese J. Appl. Probab. Statist., 2017, 33(5): 441–449.

§1. Introduction

In the recent years, statistical monitoring the performance of clinical practitioners, such as surgeons and general practitioners, has increasingly attracted attention of researchers. Since Treasure et al. [1] and Waldie [2] brought the need of formally monitoring surgical outcomes to the forefront, most of the previous research consider patient's risks because a surgical outcome depends on not only surgical-operation performance but also patients' risk factors before surgery. A number of methods for surgical monitoring have been described, including the observed-expected plot [3], the risk-adjusted CUSUM (RA-CUSUM) chart [4], the resetting sequential probability ratio test [5], the risk-adjusted Shewhart p-chart [6], the risk-adjusted sets method [7], the risk-adjusted exponentially weighted

*The project was supported by National Natural Science Foundation of China (Grant Nos. 11371202; 11571191).

⋆Corresponding author, E-mail: gengwei@nankai.edu.cn.

Received September 14, 2015. Revised July 24, 2016.

moving average (RA-EWMA) chart [8] and the likelihood ratio test derived from a change-point model (LRTCP) [9]. In the related literature, the monitoring charts that took the risk factor of each patient into account are known as the "risk-adjusted" control charts. For excellent reviews about these charts, we refer to [9], [10] and [11].

Mohammadian et al. [12] proposed a phase I risk-adjusted geometric control chart via the LRTCP based on the risk-adjusted logistic regression. Nevertheless, the LRTCP method has to be computed by maximizing the likelihood ratio with respect to all possible change locations, which may lead to significant increasing of computational load in phase II. In phase I, the purpose is to check the quality of historical data and to obtain accurate estimates of the model parameters, while in phase II, one is interested in detecting shifts in the model parameters as quickly as possible. In this paper, two risk-adjusted geometric control charts namely CUSUM and WLRT are proposed to monitor surgery performance in phase II. The performance of the proposed control charts is evaluated and compared by simulation experiments for different shift values in the parameters of a risk-adjusted logistic regression model in terms of the ARL criterion. The results show that all methods work well in the sense that they can effectively detect shifts in the process parameters.

The remainder of this article is organized as follows. In Section 2, the statistical model and the proposed CUSUM and WLRT schemes are introduced. A real case study is given in Section 3. In Section 4, the performance of the proposed CUSUM and WLRT control charts are further compared through Monte Carlo simulations. Section 5 concludes this article and gives further discussion.

§2. The Statistical Model and Proposed Schemes

2.1 Statistical Model

Let $X_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^\top$ is a $p \times 1$ vector reflecting the risk factors of patient i, where patients were ordered according to the operation time. Following [4] and [12], the probability of unsuccessful surgery of the ith patient is π_i obtained using the logistic regression model

$$\pi_i = \frac{\exp(\beta^\top X_i)}{1 + \exp(\beta^\top X_i)}, \tag{1}$$

where $\beta = (\beta_1, \beta_2, \ldots, \beta_p)^\top$ is a vector of risk factors' coefficients.

We monitor the surgical performance every time when a new unsuccessful surgery emerge. Let I_j denotes the jth unsuccessful surgery occur at the I_jth patient, and Z_j be the number of patients between $(j-1)$th and jth unsuccessful surgery. It is easy to see that Z_j is a geometric random variable with the probability mass function shown in equation (2).

$$\mathsf{P}(Z_j = z_j) = \prod_{i=1}^{z_j-1} (1 - \pi_{I_{j-1}+i})\pi_{I_j}. \tag{2}$$

Following [12], we suppose that an assignable cause occurring at an unknown time τ causes the risk-adjustment's parameters β to changes from $\beta_{\rm IC}$ to another unknown value $\beta_{\rm OC}$, say,

$$\beta = \begin{cases} \beta_{\rm IC}, & j = 1, 2, \dots, \tau; \\ \beta_{\rm OC}, & j = \tau + 1, \tau + 2, \dots. \end{cases}$$

We will explore two risk-adjusted geometric control charts to detect the shifts in the model parameters as quickly as possible.

2.2 CUSUM Control Chart

The classic CUSUM chart methodology was initially developed by Page [13] for industrial problems. Hawkins and Olwell [14] give a detailed background of CUSUM charts and their applications. Steiner et al. [4] utilized a RA-CUSUM chart to monitor the binary survival status of any given patient during a 30-day period after surgery. In this subsection, we extend the RA-CUSUM chart proposed by Steiner et al. [4] to monitor surgical performance.

At any time when the jth unsuccessful surgery occur, according to equation (2), the log-likelihood function can be derived as

$$l_j(\beta) = \ln(\pi_{I_j}) + \sum_{i=1}^{z_j-1} \ln(1 - \pi_{I_{j-1}+i}). \tag{3}$$

Let β_d denotes the design parameter which represents the out of control (OC) parameter of most interest to detect. Then, we can express the loglikelihood CUSUM monitoring statistics as

$$C_j = \max\{0, C_{j-1} + \Delta_j\}, \tag{4}$$

where $C_0 = 0$ and $\Delta_j = l_j(\beta_d) - l_j(\beta_{\rm IC})$. When the CUSUM statistics C_j in (4) is larger than the control limit, we can declare the model parameter β has deviated from the nominal value, which means the process is OC.

Here, from equations (1), (2) and (4), we can see the proposed CUSUM chart is a risk-adjusted geometric control chart. Nevertheless, this control chart requires a preferred design parameter, which has significant influence on the control chart's performance. If we concern about both the increase and decrease of the $\beta^{\mathsf{T}} X_i$, we can combine two CUSUM charts (multiple CUSUM chart). If there lacks information about the OC scenarios to design effective CUSUM chart, a more robust procedure is to use the WLRT to construct control chart.

2.3 WLRT Control Chart

The control chart based on the WLRT was used by various authors, including Li et al. [15], Qi et al. [16], Shang et al. [17], Zhou et al. [18], etc. Li et al. [15] standardized the weights by dividing a sequence of constants to ensure that all of the weights sum up to one. Qi et al. [16] and Zhou et al. [18] included a pseudo "sample" in the weighted-log-likelihood function to confirm that the short-run false alarm of the WLRT chart is not too large. Shang et al. [17] expanded the WLRT statistics to asymptotically equivalent Wald-type charting statistics using standard Taylor's expansion arguments of likelihood functions. In this subsection, we extend the WLRT chart proposed by Qi et al. [16] and Zhou et al. [18] to monitor surgical performance because this scheme has desired in-control (IC) performance.

At any time when the jth unsuccessful surgery occur, we can first express the weighted-log-likelihood function as

$$wl_j(\beta) = \sum_{k=0}^{j} w_k l_k(\beta), \tag{5}$$

where the weights $w_0 = (1-\lambda)^j$, $w_k = \lambda(1-\lambda)^{j-k}$, $k = 1, 2, \ldots, j$, such that $\sum\limits_{k=0}^{j} w_k = 1$, and $\lambda \in (0, 1)$ is a smoothing parameter. Then, given the value of λ, we can obtain the maximum weighted likelihood estimate (MWLE) of β

$$\widehat{\beta}_j = \arg\max_{\beta} wl_j(\beta),$$

using the Newton-Raphson approximation. Finally, we can express the WLRT statistic as

$$W_j = 2[wl_j(\widehat{\beta}_j) - wl_j(\beta_{\mathrm{IC}})]. \tag{6}$$

When the WLRT statistic in (6) is larger than a prespecified upper control limit (UCL), the corresponding control chart generates OC signal.

Here, the initial value of $l_0(\beta)$ can be calculate by the IC dataset which can be obtained from Phase I or by simulation. In Section 3 and Section 4, we obtain the log-likelihood $l_0(\beta)$ by simulation, and the random seed is chosen as 174 109. To alleviate the computation burden, we denote q as a sufficiently large integer to make $\lambda(1-\lambda)^q$ close to 0. When $j \leqslant q$, we make use of all available samples up to the current time point j to estimate $\widehat{\beta}_j$ and calculate W_j. Otherwise, we only use the most recent q sets of sample observations to estimate $\widehat{\beta}_j$ and calculate W_j.

In Section 3 and Section 4, we use the bisection searching algorithms to find the control limits for the CUSUM and WLRT charts.

§3. A Real Data Example

In this section, we analyzed a cardiac surgery dataset which was used by various authors, including Steiner et al. [4], Paynabar et al. [9] and Mohammadian et al. [12], etc. Following [12], the probability of unsuccessful surgery given by (1) was appropriate:

$$\pi_i = \frac{\exp(-3.680 + 0.077x_i)}{1 + \exp(-3.680 + 0.077x_i)}, \tag{7}$$

where x_i is the Parsonnet score of ith patient. We present an application of our proposed methodology: monitoring the cardiac surgery dataset from 1994 to the end of 1995. There are 2 103 operation data with 164 deaths in 2 years (1994 − 1995). Hereafter, we use the notation h to denote the control limit coefficients, and adjust the control limits of different charts to make their ARL_0 as close as 200 by convention.

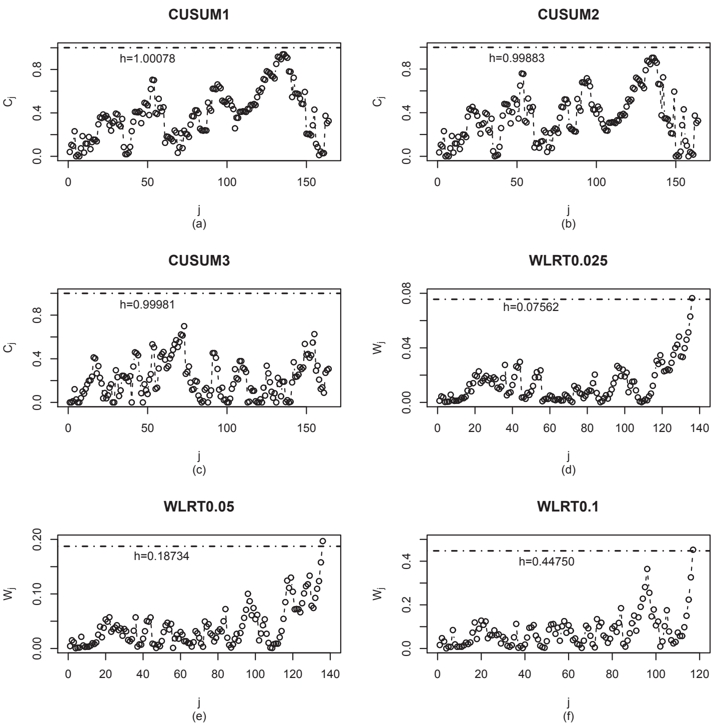

Figure 1 The CUSUM and WLRT control charts for the cardiac surgery dataset from 1994 to the end of 1995

It is important to point out that the CUSUM chart in Section 2.2 is a one-sided control chart, while the WLRT chart in Section 2.3 is a two-sided control chart. The proposed CUSUM and WLRT control charts are constructed in Figure 1. For a relatively fair comparison, the CUSUM1 chart in Figure 1 is a multiple CUSUM chart, which combines two CUSUM charts with different design parameters. We first adjust the control limits of these two CUSUM charts to make their ARL_0 as close as 400. Then we choose the control limit of the CUSUM1 chart as 1.0078 to make its ARL_0 as close as 200. The parameter of one CUSUM chart is chosen as $(-3.680, 0.087)$ for the increase of $\beta^{\mathsf{T}} X_i$. The parameter of the other CUSUM chart is chosen as $(-5.0, 0.077)$ for the decrease of $\beta^{\mathsf{T}} X_i$. Similarly, CUSUM2 with $(-3.680, 0.095)$ and $(-5.0, 0.077)$; CUSUM3 chart with $(-3.680, 0.177)$ and $(-4.0, 0.077)$. The smoothing parameter λ of the WLRT chart is chosen as 0.025, 0.05 and 0.1 respectively. The WLRT0.025, WLRT0.05 and WLRT0.1 charts generate OC signal when j equals 136, 136 and 117 respectively. However, all the CUSUM statistics fall below the control limits. In one sense, this shows that the CUSUM chart requires a preferred design parameter.

§4. Performance Comparisons

In this section, the performance of the proposed CUSUM and WLRT charts will be further compared through Monte Carlo simulations. We obtain all results in this section based on 5 000 replications. A Fortran program is also available from the authors upon request.

To generate the OC simulated data, a random Parsonnet score x_i is generated independently from the empirical distribution of Parsonnet scores obtained from the first 2 years of data (corresponding to 1992 and 1993). After that, the randomly generated x_i is substituted into equation (7) to calculate the probability of unsuccessful surgery π_i. Then, the obtained π_i is used to generate a binary outcome through a Bernoulli distribution. Finally, the geometric random variable is calculated by counting the number of binary variables with value of zero before a binary variable with value of one.

The comparisons of ARL are reported in Table 1. The exact values of ARL_0 are listed in the first row in Table 1, and the corresponding ARL_1 for different shifts are summarized in the rest of Table 1. From Table 1, we can observe that the performance of the CUSUM chart might deteriorate if the real OC parameter is far from the design parameter β_d. For instance, when $\beta_{\mathrm{OC}} = (-3.680, 0.087)$, the ARL_1 of the CUSUM1 chart is 81.7, while it is 169 for the CUSUM3 chart. As mentioned in Section 3, the design parameters of the CUSUM1 chart are $(-3.680, 0.087)$ and $(-5.0, 0.077)$, while the design parameters of the CUSUM3 chart are $(-3.680, 0.177)$ and $(-4.0, 0.077)$. Consequently, it is reasonable that the performance of the CUSUM1 chart is better than the CUSUM3 chart when $\beta_{\mathrm{OC}} = (-3.680, 0.087)$. In addition, the performance of the CUSUM3 chart is unsatisfactory when $\beta_{\mathrm{OC}} = (-3.55, 0.077)$. This also shows that the CUSUM chart requires a preferred design parameter. From Table 1, we can also find that the ARL_1 of the WLRT chart increases with the parameter λ above the line $\beta_{\mathrm{OC}} = (-2.96, 0.077)$, and decreases below

Table 1 Comparisons of ARL when $\tau = 0$

β_{OC}	CUSUM1 $h = 1.00078$	CUSUM2 $h = 0.99883$	CUSUM3 $h = 0.99981$	WLRT0.025 $h = 0.07562$	WLRT0.05 $h = 0.18734$	WLRT0.1 $h = 0.44750$
$(-3.68, 0.077)$	200 (184)	200 (192)	200 (190)	200 (186)	200 (191)	200 (193)
$(-3.55, 0.077)$	149 (116)	155 (134)	219 (215)	135 (112)	152 (137)	178 (167)
$(-3.68, 0.087)$	81.7 (52.1)	84.7 (62.1)	169 (164)	83.8 (63.6)	97.5 (82.1)	129 (116)
$(-3.40, 0.077)$	88.7 (56.5)	89.4 (66.3)	149 (142)	67.0 (42.5)	75.2 (56.3)	98.3 (84.6)
$(-3.68, 0.095)$	47.8 (21.1)	45.5 (25.4)	98.1 (93.6)	44.5 (25.1)	48.5 (33.2)	64.0 (52.4)
$(-2.96, 0.077)$	39.4 (13.7)	34.5 (15.0)	42.3 (37.3)	26.8 (8.04)	25.0 (9.53)	26.1 (14.0)
$(-3.68, 0.350)$	16.4 (1.79)	12.7 (1.55)	4.95 (0.88)	12.2 (0.91)	10.1 (0.82)	8.38 (0.76)
$(-3.68, 0.004)$	6.43 (4.98)	6.42 (4.98)	7.80 (4.61)	5.56 (2.87)	5.02 (2.75)	4.66 (2.77)
$(-5.00, 0.010)$	1.88 (1.11)	1.88 (1.11)	2.31 (1.30)	2.18 (1.19)	2.01 (1.10)	1.88 (1.04)
$(-6.00, 0.001)$	1.26 (0.53)	1.26 (0.53)	1.42 (0.68)	1.39 (0.65)	1.33 (0.60)	1.28 (0.55)

NOTE: Standard deviations are in parentheses.

the line $\beta_{\mathrm{OC}} = (-2.96, 0.077)$. This shows that the control chart WLRT0.025 is better than WLRT0.05, and WLRT0.05 is better than WLRT0.1 for detecting smaller shifts; while the control chart WLRT0.1 is better than WLRT0.05, and WLRT0.05 is better than WLRT0.025 for detecting larger shifts. In other words, the WLRT charts with smaller parameter λ perform better for detecting small shifts, while those with larger parameter λ perform better for detecting larger shifts. This result is consistent with the traditional EWMA control chart.

Table 2 Comparisons of CED when $\tau = 20$

β_{OC}	CUSUM1 $h = 1.00078$	CUSUM2 $h = 0.99883$	CUSUM3 $h = 0.99981$	WLRT0.025 $h = 0.07562$	WLRT0.05 $h = 0.18734$	WLRT0.1 $h = 0.44750$
$(-3.55, 0.077)$	137 (117)	144 (134)	213 (215)	127 (112)	146 (138)	174 (168)
$(-3.68, 0.087)$	71.7 (52.9)	76.7 (62.9)	164 (163)	80.7 (65.1)	94.5 (82.9)	125 (116)
$(-3.40, 0.077)$	78.2 (56.0)	79.9 (66.4)	146 (142)	63.1 (44.1)	70.7 (56.7)	95.6 (85.0)
$(-3.68, 0.095)$	40.7 (21.9)	39.2 (26.0)	95.6 (93.9)	44.0 (28.1)	47.9 (35.0)	63.4 (53.3)
$(-2.96, 0.077)$	32.8 (14.7)	28.6 (15.5)	41.1 (37.3)	25.2 (11.1)	23.4 (11.2)	24.3 (14.6)
$(-3.68, 0.350)$	13.2 (3.24)	10.2 (2.80)	4.48 (1.16)	12.0 (4.05)	9.95 (3.25)	8.21 (2.50)
$(-3.68, 0.004)$	6.21 (5.02)	6.21 (5.03)	6.99 (4.57)	5.93 (3.35)	5.50 (3.18)	5.12 (3.11)
$(-5.00, 0.010)$	1.81 (1.09)	1.81 (1.09)	2.15 (1.27)	2.24 (1.29)	2.08 (1.20)	1.94 (1.12)
$(-6.00, 0.001)$	1.26 (0.54)	1.25 (0.54)	1.36 (0.63)	1.41 (0.67)	1.35 (0.63)	1.31 (0.59)

NOTE: Standard deviations are in parentheses.

We also compare the conditional expected delay (CED) [19] in Table 2 due to the detection ability being dependent on the time point of the change [20]. For the illustration purpose, we only present the CED's results when $\tau = 20$, and discard any series in which a signal occurs before the $(\tau+1)^{\text{th}}$ observation. In general, the results show that all methods work well in the sense that they can effectively detect shifts in the process parameters.

§5. Conclusion Remarks

In this article, we propose two risk-adjusted geometric control charts, namely CUSUM and WLRT, derived from a change-point model based on the risk-adjusted logistic regression to monitor surgery performance in phase II. The performance of the proposed control charts is evaluated and compared by simulation experiments for different shift values in the parameters of a risk-adjusted logistic regression model in terms of the ARL criterion. The CUSUM control chart requires a preferred design parameter, which has significant influence on the control chart's performance. If there lacks information about the OC scenarios to design effective CUSUM chart, the more robust WLRT chart can be deployed.

In this article, we focused on each patient's binary survival status after surgery. In fact, the patient's survival status can also be divided into more than two outcomes, such as died, full recovery, survives an operation but remained bed-ridden for life. Tang et al.[21] developed a risk-adjusted CUSUM chart based on more than two outcomes. Thus, it makes sense to develop such a WLRT chart and study its performance.

Acknowledgements The authors would like to thank Kamran Paynabar for providing the surgical data set. The authors would also like to thank the Editor and two anonymous referees for their insightful comments, that improve this paper greatly.

References

[1] Treasure T, Taylor K, Black N. *Independent Review of Adult Cardiac Surgery – United Bristol* [R]. Bristol: Health Care Trust, 1997.

[2] Waldie P. Crisis in the cardiac unit [N] // *The Globe and Mail*, Canada's National Newspaper, Oct. 27, Section A:3, Column 1, 1998.

[3] Lovegrove J, Valencia O, Treasure T, et al. Monitoring the results of cardiac surgery by variable life-adjusted display [J]. *Lancet*, 1997, **350(9085)**: 1128–1130.

[4] Steiner S H, Cook R J, Farewell V T, et al. Monitoring surgical performance using risk-adjusted cumulative sum charts [J]. *Biostatistics*, 2000, **1(4)**: 441–452.

[5] Spiegelhalter D, Grigg O, Kinsman R, et al. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery [J]. *Int. J. Qual. Health Care*, 2003, **15(1)**: 7–13.

[6] Cook D A, Steiner S H, Cook R J, et al. Monitoring the evolutionary process of quality: risk-adjusted charting to track outcomes in intensive care [J]. *Crit. Care Med.*, 2003, **31(6)**: 1676–1682.

[7] Grigg O A, Farewell V T. A risk-adjusted Sets method for monitoring adverse medical outcomes [J]. *Stat. Med.*, 2004, **23(10)**: 1593–1602.
[8] Grigg O, Spiegelhalter D. A simple risk-adjusted exponentially weighted moving average [J]. *J. Amer. Statist. Assoc.*, 2007, **102(477)**: 140–152.
[9] Paynabar K, Jin J H, Yeh A B. Phase I risk-adjusted control charts for monitoring surgical performance by considering categorical covariates [J]. *J. Qual. Technol.*, 2012, **44(1)**: 39–53.
[10] Grigg O, Farewell V. An overview of risk-adjusted charts [J]. *J. Roy. Statist. Soc. Ser. A*, 2004, **167 (3)**: 523–539.
[11] Woodall W H. The use of control charts in health-care and public-health surveillance (with discussion) [J]. *J. Qual. Technol.*, 2006, **38(2)**: 89–104.
[12] Mohammadian F, Niaki S T A, Amiri A. Phase-I risk-adjusted geometric control charts to monitor health-care systems [J]. *Qual. Reliab. Eng. Int.*, 2016, **32(1)**: 19–28.
[13] Page E S. Continuous inspection schemes [J]. *Biometrika*, 1954, **41(1-2)**: 100–115.
[14] Hawkins D M, Olwell D H. *Cumulative Sum Charts and Charting for Quality Improvement* [M]. New York: Springer-Verlag, 1998.
[15] Li J, Tsung F, Zou C L. Multivariate binomial/multinomial control chart [J]. *IIE Trans.*, 2014, **46(5)**: 526–542.
[16] Qi D Q, Wang Z J, Zi X M, et al. Phase II monitoring of generalized linear profiles using weighted likelihood ratio charts [J]. *Comput. Ind. Eng.*, 2016, **94**: 178–187.
[17] Shang Y F, Tsung F, Zou C L. Profile monitoring with binary data and random predictors [J]. *J. Qual. Technol.*, 2011, **43(3)**: 196–208.
[18] Zhou Q, Zou C L, Wang Z J, et al. Likelihood-based EWMA charts for monitoring Poisson count data with time-varying sample sizes [J]. *J. Amer. Statist. Assoc.*, 2012, **107(499)**: 1049–1062.
[19] Kenett R S, Zacks S. *Modern Industrial Statistics: The Design and Control of Quality and Reliability* [M]. Pacific Grove, CA: Duxbury Press, 1998.
[20] Sonesson C, Bock D. A review and discussion of prospective statistical surveillance in public health [J]. *J. Roy. Statist. Soc. Ser. A*, 2003, **166(1)**: 5–21.
[21] Tang X, Gan F F, Zhang L Y. Risk-adjusted cumulative sum charting procedure based on multiresponses [J]. *J. Amer. Statist. Assoc.*, 2015, **110(509)**: 16–26.

阶段 II 监控手术表现的风险调整后的几何控制图

齐德全 耿 薇

(吉林医药学院生物医学工程学院, 吉林, 132013) (南开大学统计研究院, 天津, 300071)

摘 要: 近年来, 统计过程控制 (SPC) 已被广泛用于监测临床从业人员的表现. 本文建立了两个阶段 II 的风险调整的几何控制图, 即累积和 (CUSUM) 和加权似然比检验 (WLRT) 控制图, 用来监控手术的表现. 通过模拟实验, 以平均运行长度 (ARL) 为标准, 对所提出的控制图的性能进行了评估和比较. 结果表明, 所提出的控制图可以有效地检测到过程参数的变化.

关键词: 变点; 风险调整后的控制图; 统计过程控制; 加权似然比检验
中图分类号: O213.1