AUT Journal of Modeling and Simulation

AUT J. Model. Simul., 57(1) (2025) 17-28
DOI: 10.22060/miscj.2025.23841.5400

A Risk-Adjusted CUSUM Chart for Monitoring Surgical Performance with Ordinal Outcomes and Random Effects

Ramezan Khosravi[1] *, Hadi Gholinezhad[1], Hamid Esmaeeli[2]

[1] Department of Industrial Engineering, University of Gonabad, Gonabad, Iran.
[2] Department of Industrial Engineering, Islamic Azad University- North Tehran Branch, Tehran, Iran.

ABSTRACT: Monitoring healthcare processes poses unique challenges due to the substantial variability in patient risk profiles, which can significantly influence surgical outcomes. Traditional control charts often neglect these individual differences, leading to potentially biased and misleading performance assessments. To overcome these limitations, risk-adjusted control charts have been developed to incorporate patient-specific covariates for more equitable monitoring. This study extends previous approaches by proposing a risk-adjusted cumulative sum (RA-CUSUM) control chart that accommodates ordinal surgical outcomes and incorporates random effects to model unobserved heterogeneity among healthcare providers. The proposed RA-CUSUM chart employs dynamic probability control limits (DPCLs) to maintain a constant conditional false alarm rate, enabling consistent performance across heterogeneous patient populations. Through extensive simulation studies, we demonstrate its efficacy in detecting shifts in surgical performance stability, particularly in response to changes in location and scale. A real-world case study using cardiac surgery data demonstrates the practical applicability of the method. This work provides a more refined and fair framework for evaluating surgical quality and lays the groundwork for integrating adaptive techniques in future healthcare monitoring systems. In addition to healthcare monitoring, the method can be extended to other domains where ordinal outcomes and case heterogeneity are relevant, such as education and finance. This adaptability makes it a valuable decision-support tool for quality improvement programs and real-time risk management.

Review History:

Received: Jan. 19, 2025
Revised: Jul. 15, 2025
Accepted: Jul. 30, 2025
Available Online: Aug. 05, 2025

Keywords:

Risk-Adjusted Control Chart
Healthcare Process Monitoring
Categorical Covariates
Dynamic Probability Control Limits
Patient Heterogeneity

1- Introduction

Online monitoring of healthcare systems plays a pivotal role in safeguarding patient well-being and maintaining high standards of care. In particular, undetected declines in clinical performance can lead to severe consequences for both patients and healthcare institutions. At the hospital level, performance indicators such as infection rates, postoperative mortality, and recovery times are critical metrics that warrant continuous surveillance (Bersimis and Sachlas, 2022).

Statistical process monitoring (SPM) provides a framework for issuing timely alerts when performance deviates from expected baselines. However, applying SPM in healthcare presents unique challenges compared to industrial settings. Unlike manufactured products, patients vary significantly in their risk profiles due to pre-existing conditions, age, and other clinical factors. These risk factors are typically beyond the provider's control but must be accounted for to ensure fair performance evaluations (Grigg, 2019).

Control charts, a fundamental tool in SPM, have been extensively adopted in surgical monitoring. However, in such applications, patient heterogeneity can obscure true shifts in surgical quality. Risk-adjusted control charts (RA charts) address this issue by conditioning the monitoring process on patient-specific covariates. These models estimate the probability of adverse outcomes using logistic or survival regression, providing a more accurate and fair basis for comparing provider performance. For example, an adverse outcome for a low-risk patient is a stronger signal of poor care quality than the same outcome for a high-risk patient (Steiner et al., 2000).

RA-CUSUM charts have been developed for binary and continuous outcomes. The seminal work by Steiner et al. (2000) introduced a risk-adjusted CUSUM for binary outcomes (e.g., survival within 30 days), while Sego et al. (2009) extended the approach to continuous survival times. Later advancements incorporated both continuous and categorical covariates (Paynabar et al., 2012) and emphasized the importance of accounting for additional factors such as surgeon experience and procedure type.

Despite these advances, two critical limitations remain. First, the reliance on binary classifications ignores the ordinal nature of many clinical outcomes, such as partial versus full recovery. Researchers including Tang et al. (2015) and Khosravi et al. (2018), addressed this limitation by

*Corresponding author's email: khosravi.r@gonabad.ac.ir

Table 1. The contribution of the present study in the literature of RA control chars.

Study	Outcome Type	Random Effects	Control Limits	Contribution Summary
Steiner et al. (2000)	Binary	No	Fixed	Classical RA-CUSUM for binary outcomes
Sego et al. (2009)	Continuous	No	Fixed	Risk-adjusted CUSUM for survival outcomes
Tang et al. (2015)	Ordinal	No	Fixed	RA-CUSUM for ordinal outcomes using cumulative logit model
Aminnayeri and Sogandi (2016)	Binary	No	Dynamic	Self-starting RA-CUSUM with simulation-based dynamic thresholds
Li et al. (2023)	Binary	Yes	Dynamic	Novel performance evaluation by considering random effects
This study	Ordinal	Yes	Dynamic	First RA-CUSUM with ordinal outcomes, random effects, and DPCLs

developing RA-CUSUM charts for ordinal outcomes using proportional odds logistic models. Second, conventional control limits assume a homogeneous patient stream, yet real-world clinical settings often deal with fluctuating patient risks. Fixed control limits may lead to inconsistent false alarm rates across risk profiles. To resolve this, Zhang and Woodall (2015) proposed using dynamic probability control limits (DPCLs) to ensure a fixed conditional false alarm rate (CFAR) across varying patient populations. Aminnayeri and Sogandi (2016) extended this to a self-starting Bernoulli CUSUM chart with dynamic limits.

While previous studies have made important advances in risk-adjusted monitoring, they leave key challenges unaddressed in settings involving ordinal outcomes and heterogeneous patient populations. To address these limitations, this study is guided by three main research questions:

(1) How can we design a risk-adjusted control chart that captures multiple ordered levels of surgical outcomes, rather than binary success/failure metrics?

(2) How can we account for both systematic shifts in performance and variability across providers using a unified framework?

(3) How can we ensure fair signaling across heterogeneous patient populations through adaptive control limits?

To address these questions, this paper aims to develop and evaluate a novel (RA-CUSUM control chart for ordinal outcomes that incorporates a random effect term to monitor provider-specific variability and uses DPCLs to maintain a consistent false alarm rate across varying patient risk profiles. Building upon prior literature, the proposed method integrates ordinal outcome modeling with random effects to more accurately reflect clinical realities where both outcome severity and provider heterogeneity influence performance. The adoption of DPCLs further ensures equitable and robust monitoring in diverse patient populations.

To clarify the methodological contributions of this study, Table 1 summarizes and compares selected works on risk-adjusted control charts, focusing on key modeling components such as the type of outcome modeled, the incorporation of random effects, and the structure of the control limits.

As illustrated, prior studies have addressed specific aspects of the monitoring problem, but none have simultaneously incorporated all three dimensions. The proposed method is the first to combine ordinal outcome modeling, random effects for provider-specific variability, and DPCLs within a unified RA-CUSUM framework. This integration improves the relevance, fairness, and responsiveness of performance monitoring in clinical settings.

The remainder of the paper is structured as follows: Section 2 formulates the risk-adjusted ordinal regression model with random effects. Section 3 describes the CUSUM chart and its dynamic control limits. Section 4 evaluates performance using simulation studies. Section 5 applies the methodology to a real-world surgical dataset. Section 6 concludes with a discussion on implications and future directions.

2- Risk-adjusted ordinal modeling

Let Y_i denote the surgical outcome for patient *i*, which falls into one of J ordered categories. For instance, in a three-category case, Y_i=0 may indicate full recovery, Y_i=1 partial recovery, and Y_i=2 death. The ordering reflects the increasing severity of the outcome. This classification enables a more detailed evaluation of surgical performance compared to binary outcomes, which only distinguish between survival and death.

We assume that $Y_i \sim \text{Multinomial}(1; p_{i1},...,p_{iJ})$, where $p_{ij} = \Pr(Y_i = j)$ denotes the probability that patient *i*'s outcome falls into category j. These probabilities are modeled as functions of continuous risk factors $\mathbf{x}_i$ (e.g., Parsonnet

score) and categorical covariates $\mathbf{d}_i$ (e.g., type of surgery, surgeon identity).

To relate the covariates to the outcome probabilities, we use the proportional odds logistic regression model, a special case of the cumulative link model (Agresti, 2010). This model is well-suited to ordinal outcomes because it leverages the ordering information, improving estimation efficiency and interpretability. Cumulative link models are widely used for modeling ordinal response variables due to their ability to incorporate the ordered nature of responses and their flexible regression framework.

2- 1- Cumulative Link Model

Instead of modeling each category probability p_{ij} directly, the model estimates the cumulative probabilities:

$$\gamma_{ij} = \mathrm{P}\left(Y \le j\right) = \sum_{k=0}^{j} p_{ik}, \quad j = 0,1,...,J-1.$$

These are related to the covariates via the logit link function:

$$\mathrm{logit}\left(\gamma_{ij}\right) = \alpha_j - \mathbf{x}_i^T\boldsymbol{\beta}, \quad j = 1,...,J-1, \tag{1}$$

where α_j is the intercept (cutpoint) for category j and $\beta = [\beta_1,...,\beta_p]$ is a parameter vector for regression coefficients shared across categories. The negative sign ensures that higher values of $x_i\beta$ shift the response distribution toward worse outcomes (i.e., higher category indices), which aligns with clinical intuition when x_i encodes patient risk.

2- 2- Latent Variable Interpretation

This model can be interpreted in terms of a continuous latent health score S_i such that:

$$Y_i = j \quad \text{if} \quad \tau_{j-1} < S_i \le \tau_j, j = 0,1,...,J-1$$

where $\tau_0,\ldots,\tau_{J-1}$ are thresholds corresponding to the intercepts α_j, and $S_i = x_i\beta + \varepsilon_i$, with $\varepsilon_i \sim$ Logistic(0,1).

This framework connects the ordinal regression to an underlying severity continuum (Christensen, 2015).

2- 3- Categorical Covariates

The linear predictor is extended to incorporate categorical variables (e.g., procedure type, surgeon) using dummy coding.

$$\mathrm{logit}\left(\gamma_{ij}\right) = \alpha_j - \mathbf{x}_i^T\boldsymbol{\beta} - \mathbf{d}_i^T\boldsymbol{\gamma}, \tag{2}$$

where $\mathbf{d}_i$ is a vector of dummy variables for the categorical covariates and $\boldsymbol{\gamma}$ is the corresponding coefficient vector. If a covariate (e.g., surgery type) has three levels, two dummy variables are introduced, and one category is treated as the baseline.

3- The proposed RA-CUSUM control chart

Let us assume that surgical performance remains in control up to a certain point in the monitoring process and then undergoes a change at patient τ. This change may affect the central tendency or the variability of outcomes. We model this situation using the following change-point formulation:

$$\theta_i \sim \begin{cases} \theta_0 & \text{for } i \le t \\ \theta_1 & \text{for } i > t \end{cases} \tag{3}$$

Here, θ_i represents the parameter vector at time *i*, and the change may occur in the location or scale parameters (or both) of the ordinal logistic model.

3- 1- Incorporating Random Effects to Capture Variability

In real clinical settings, variability among surgeons, due to skill levels, techniques, or case complexity, can affect surgical outcomes. To model this unobserved heterogeneity, we introduce a random effect into the linear predictor:

$$\eta_i = \mathbf{x}_i^T\boldsymbol{\beta} + \mathbf{d}_i^T\boldsymbol{\gamma} + \delta_i, \tag{4}$$

where:

- $\delta_i \sim N\left(0,1\right)$ is a random effect associated with patient *i*'s surgery,
- θ is the variance of the random effect, used to monitor stability in performance.

When θ=0, the process is considered stable (in-control). A positive θ indicates excess variation, which may reflect a decline in surgical consistency.

To formally monitor performance, we test the following hypotheses:

$$\begin{aligned} &H_0 : \theta = 0 \ \text{(stable process)} \\ &\text{vs.} \\ &H_1 : \theta > 0 \ \text{(instablity present)} \end{aligned} \tag{5}$$

3- 2- Derivation of the Score Statistic

To evaluate evidence against the null hypothesis, we use a score statistic, denoted by W_i, which is derived from the first derivative of the log-likelihood function with respect to θ. Let y_i denote the observed outcome for patient *i*, $F_j(.)$ denote the cumulative distribution function corresponding to category j, and I(.) be the indicator function. The log-likelihood contribution for observation *i*, conditional on δ_i, is:

$$\ell_i = \log\left[F_j\left(\alpha_j - \eta_i\right) - F_{j-1}\left(\alpha_{j-1} - \eta_i\right)\right] \tag{6}$$

The score statistic W_i is then given by:

$$W_i = \frac{\partial \ell_i}{\partial \theta}\bigg|\theta = 0 \quad (7)$$

The intuition is that W_i captures how sensitive the likelihood is to variability at each patient's outcome. A larger W_i implies greater evidence of instability.

The details of the computation are provided in Appendix A.

3- 3- CUSUM Statistic and Monitoring Procedure

We now integrate the score statistics into a CUSUM chart, which accumulates small changes over time to detect sustained shifts. The one-sided CUSUM statistic is defined recursively as:

$$C_i^+ = \max\left(0, C_{i-1}^+ + W_i\right), \quad C_0^+ = 0 \quad (8)$$

where C_i^+ is cumulative sum of the score statistics up to patient *i*. The process signals an alarm if C_i^+ exceeds a control limit.

3- 4- Dynamic Probability Control Limits

Unlike industrial settings, patient covariates in healthcare vary significantly over time. Consequently, fixed control limits may produce inconsistent false alarm rates. To address this, we adopt dynamic probability control limits (DPCLs), which are individualized for each patient (Zhang, Loda, and Woodall, 2017; Zhang and Woodall, 2015).

The core idea is to maintain a fixed conditional false alarm rate (CFAR) α, such that (Aytaçoğlu, Driscoll, and Woodall, 2023):

$$\Pr\left(C_i^+ > h_i(\alpha) \mid C_{i-1}^+ \le h_{i-1}(\alpha), \ldots, C_1^+ \le h_1(\alpha)\right) = \alpha \quad (9)$$

where $h_i(\alpha)$ is the dynamic control limit for patient *i*, estimated via simulation.

These control limits are calculated using the following procedure:

1) Estimate in-control parameters: Use the historical patient data to fit the risk-adjusted ordinal logistic regression model. Obtain estimates of the intercepts α_j, coefficients $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ under the assumption that the process is stable.
2) Simulate in-control outcomes: Generate a large number of synthetic patient outcomes (e.g., N=100,000) under the in-control model.
   a) For each simulated patient, use the estimated regression model to compute category probabilities (e.g., probability of full recovery, partial recovery, or death).
   b) Randomly assign an outcome by drawing from a multinomial distribution with these probabilities.
   c) These simulated outcomes mimic how the process would behave if it remained stable, providing a reference distribution for comparison.
3) Compute simulated CUSUM values: For each simulated outcome, calculate the corresponding score statistic W_i and update the CUSUM statistic C_i^+ using Eq. (8).
4) Determine control limit h_i: Sort the N simulated CUSUM values in ascending order. Select the (1−α)-quantile of this empirical distribution as the control limit for patient *i*, denoted by h_i.
5) Decision rule: Compare the observed CUSUM statistic for patient *i* with the computed h_i. If $C_i^+ > h_i$, issue an out-of-control signal, otherwise, continue monitoring with the next observation.

This simulation-based procedure ensures that the control chart adjusts to individual patient characteristics while maintaining a consistent false alarm rate across varying risk levels. DPCLs are particularly beneficial during the early stages of monitoring or when patient populations are heterogeneous.

4- Performance evaluation

This section presents simulation studies designed to evaluate the performance of the proposed RA-CUSUM chart in detecting changes in both the location and scale parameters of surgical performance. The primary performance measures are:

- ARL_0 (in-control average run length): the expected number of observations until a false alarm occurs when the process is stable.
- SDRL: standard deviation of the run length under the in-control condition.
- ARL_0 (out-of-control average run length): the expected number of observations to signal after a shift has occurred.
- CFAR (conditional false alarm rate): used in DPCL design to maintain control over false signal probabilities.

We define a baseline RA-CUSUM model using the ordinal logistic regression structure discussed in Section 2. The model parameters for the in-control state are:

$$\alpha_1 = 0.5, \alpha_2 = 0.5, \beta = -0.1, \gamma_1 = -0.4, \gamma_1 = 0.6$$

$x_i \sim Exponential(\lambda)$: continuous patient risk score, d_{i1}, d_{i2}: dummy variables identifying the surgery type, drawn from a multinomial distribution over three categories.

Each patient's ordinal outcome is generated based on their covariate profile, with categories indicating different recovery levels (e.g., full recovery, partial recovery, death).

4- 1- In-Control Performance under Fixed vs. Dynamic Limits

First, we compare the in-control performance of the RA-CUSUM chart under fixed control limits and DPCLs. For the fixed-limit case, we select h=4.06 to target an ARL_0 of 400.

Table 2 shows that the achieved ARL_0 varies substantially across different risk distributions, indicating the sensitivity of fixed-limit charts to changes in patient risk profiles.

Table 2. In-control performance of the proposed control chart based on the fixed control limit .

Index	Risk distribution	ARL_0	SDRL
1	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	370.2	348.1
2	$x \sim \exp(0.2)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	430.5	339.6
3	$x \sim \exp(0.05)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	384.1	320.4
4	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.1,0.3)$	395.9	359.2
5	$x \sim \exp(0.1)$ $d \sim M(0.6,0.3,0.1)$	415.3	357.9
6	$x \sim \exp(0.1)$ $d \sim M(0.4,0.3,0.3)$	401.3	329.3

Table 3. In-control performance of the proposed control chart using DPCLs (α=0.0025).

		ARL_0	SDRL	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$	$\bar{\alpha}_i$	$\bar{\alpha}_i'$
1	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	399.4	384.6	45.1	117.7	278.0	553.5	916.4	2.4969E-03	2.5063E-03
2	$x \sim \exp(0.2)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	398.3	387.5	44.2	116.4	279.9	552.9	917.2	2.4996E-03	2.5011E-03
3	$x \sim \exp(0.05)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	398.5	383.5	42.4	117.2	279.4	553.7	919.0	2.4932E-03	2.5085E-03
4	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.1,0.3)$	401.0	387.6	42.6	116.2	277.6	554.4	916.4	2.4996E-03	2.5037E-03
5	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.1,0.3)$	400.5	383.4	43.7	115.3	279.5	554.2	916.9	2.4911E-03	2.5038E-03
6	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.4,0.3,0.3)$	398.8	386.8	44.3	115.5	277.5	552.0	917.1	2.4982E-03	2.5013E-03
	Geometric	**400.0**	**399.49**	**42**	**115**	**277**	**552**	**916**	-	-

In contrast, when using DPCLs with N=100,000 Monte Carlo replications and false alarm probabilities α=0.0025 and α=0.005, the chart consistently achieves the desired ARL_0 values across all distributions.

Tables 3 and 4 report the in-control ARL_0 and SDRL values, along with selected percentiles of the run length distribution (Q10–Q90) and the estimated CFAR. The results show that the DPCL-based chart consistently achieves the nominal in-control performance across varying patient risk profiles. The empirical run length distribution closely aligns with the theoretical geometric distribution, and CFAR values remain near the target levels. These findings confirm that the proposed method is both statistically reliable and well-suited to heterogeneous clinical settings.

To visualize how DPCLs adapt during the monitoring process, Fig. 1 compares the individualized control limits

Table 4. In-control performance of the proposed control chart using DPCLs (α=0.005).

		ARL_0	SDRL	$Q_{0.10}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.90}$	$\bar{\alpha}_i$	$\bar{\alpha}_i'$
1	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	200.7	193.2	23.0	58.9	139.1	276.8	458.3	4.994E-03	5.013E-03
2	$x \sim \exp(0.2)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	199.5	194.4	22.2	58.5	140.8	277.0	458.7	4.999E-03	5.002E-03
3	$x \sim \exp(0.05)$ $\mathbf{d} \sim M(0.6,0.2,0.2)$	199.4	191.8	21.8	59.6	140.6	277.0	460.3	4.986E-03	5.017E-03
4	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.1,0.3)$	200.6	194.2	21.7	58.5	139.5	277.9	458.6	4.999E-03	5.007E-03
5	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.6,0.1,0.3)$	200.8	191.7	22.0	58.1	140.3	277.5	458.8	4.982E-03	5.008E-03
6	$x \sim \exp(0.1)$ $\mathbf{d} \sim M(0.4,0.3,0.3)$	200.0	194.1	22.3	58.4	139.1	276.5	458.7	4.996E-03	5.003E-03
	Geometric	200	199.5	21	57	139	276	459	-	-

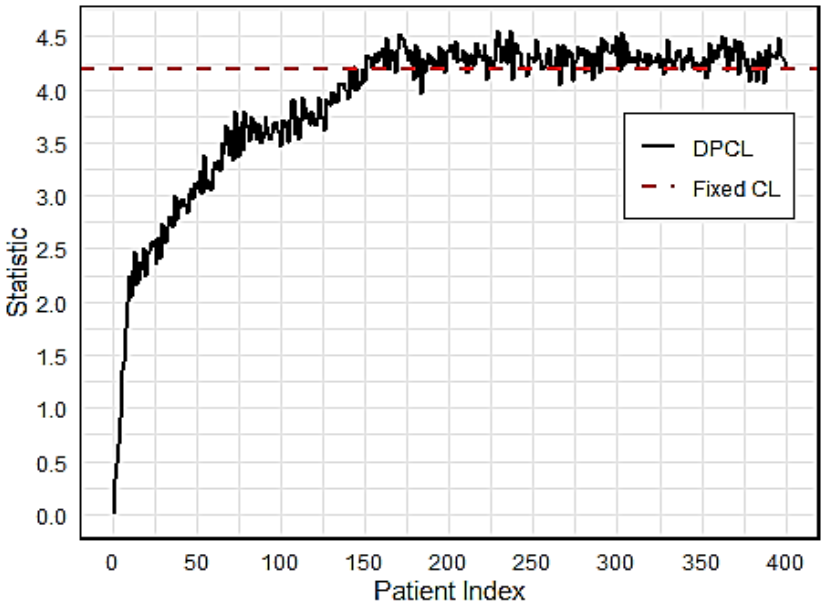


Fig. 1. The DPCLs for the first 400 patients

for the first 400 patients to the constant fixed control limit. The DPCLs start more conservatively in early stages (due to higher variability) but converge toward the fixed limit over time as more information accumulates. This adaptive behavior reduces early-stage false alarms and ensures fairness in the presence of diverse risk profiles.

4- 2- Out-of-Control Performance and Robustness

To evaluate the chart's detection capability, shifts were introduced in both the location and scale parameters of the baseline model at various change points Let ε_l and ε_s denote the magnitudes of changes in the location and scale parameters, respectively.

Table 5. ARL1 (SDRL) values of the proposed control charts with the shifts of different sizes and change points.

τ	ε_1 in φ						ε_s in θ					
	0.05	**0.1**	**0.2**	**0.5**	**1**	**2**	**0.5**	**0.7**	**0.9**	**1**	**2**	**3**
25	286.7 (40.5)	227.3 (35.4)	132.1 (25.1)	63.2 (13.1)	25.7 (5.6)	6.8 (1.3)	285.3 (40.6)	220.0 (36.6)	170.7 (26.3)	160.3 (22.1)	34.2 (6.8)	13.5 (1.9)
50	283.9 (40.4)	224.6 (36.3)	130.6 (24.8)	62.5 (12.8)	25.1 (5.1)	6.3 (1.2)	284.9 (42.8)	221.3 (36.2)	170.5 (26.6)	155.0 (21.9)	32.0 (6.3)	12.4 (1.6)
75	283.5 (39.1)	224.6 (36.1)	131.0 (24.3)	62.5 (12.5)	25.1 (5.1)	6.2 (1.1)	283.1 (42.1)	223.4 (34.9)	167.6 (25.5)	144.9 (21.9)	28.8 (6.5)	14.9 (1.8)
100	277.8 (39.9)	219.6 (36.2)	127.2 (23.8)	61.4 (12.4)	24.5 (5.3)	5.0 (1.1)	280.9 (40.7)	221.7 (35.6)	168.4 (24.9)	147.8 (22.8)	28.7 (6.2)	10.9 (1.7)

The results, summarized in Table 5, indicate that the proposed RA-CUSUM chart effectively detects both types of shifts, with faster detection occurring as the magnitude of change increases. The method demonstrates sensitivity not only to shifts in the average performance but also to increased variability, highlighting its strength in monitoring both systematic and unstable behavior. Moreover, the time at which the change occurs has a minimal effect on detection speed, suggesting that the chart maintains stable performance across different monitoring horizons. These findings confirm the method's utility for real-time surveillance in dynamic clinical environments.

5- Illustrative example

To demonstrate the practical application of the proposed RA-CUSUM control chart, we analyze a real-world dataset of surgical outcomes from cardiac procedures performed in UK hospitals between 1992 and 1998. The dataset includes information on 6,994 patients, along with demographic and preoperative clinical characteristics. These characteristics have been combined into the widely used Parsonnet score, which ranges from 0 to 71 and provides a composite measure of patient-specific surgical risk, with higher values indicating higher risk.

To define the ordinal response variable, we categorized surgical outcomes into three ordered levels based on postoperative survival duration: Y=0 was assigned to patients who survived more than 30 days, representing full recovery; Y=1 indicated partial recovery and was assigned to those who died between days 15 and 30; and Y=2 was used to denote mortality within the first 14 days after surgery. This ordering reflects increasing severity of outcome and allows the monitoring procedure to distinguish between varying degrees of recovery, rather than relying on a simple binary classification.

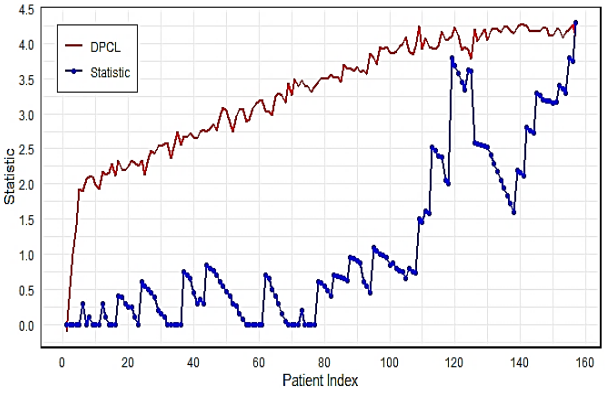
We used proportional odds logistic regression to estimate the cumulative category probabilities as functions of the Parsonnet score and type of surgery. The estimation was carried out using the *polr* function in the MASS package in R.

To evaluate the responsiveness of the proposed control chart to changes in surgical performance, we artificially introduced shifts in both the location and scale parameters of the model. For the location shift, a constant was added to the linear predictor, representing a deterioration in average performance. For the scale shift, a random effect term with increased variance was added, reflecting greater variability in performance stability (e.g., inconsistent surgical outcomes across patients). In both cases, the modified model was used to generate a sequence of observations, and the RA-CUSUM chart with DPCLs was applied for monitoring.

Figure 2 shows the resulting CUSUM trajectories under location and scale shifts, along with the corresponding dynamic control limits. The control chart successfully detected the changes in both scenarios. As expected, shifts in the location parameter resulted in earlier signals, whereas shifts in scale led to more gradual increases in the CUSUM statistic. These results are consistent with the simulation findings and confirm that the proposed method is capable of identifying both systematic and irregular changes in surgical outcomes.

From a managerial perspective, the ability to detect both average shifts and increased variability is of particular importance. Changes in the average outcome may signal systemic issues such as deteriorating surgical technique or process drift, while increased variability may reflect inconsistent performance among surgical teams or procedural complexity. Timely identification of such deviations enables hospital administrators to investigate root causes, initiate targeted interventions (e.g., training, standardization), and prevent further adverse events. Furthermore, the use of dynamic control limits ensures that performance evaluations are adjusted fairly based on each patient's risk profile, making the approach suitable for high-stakes clinical environments where equitable monitoring is essential.

(a) location parameter shift



(b) scale parameter shift

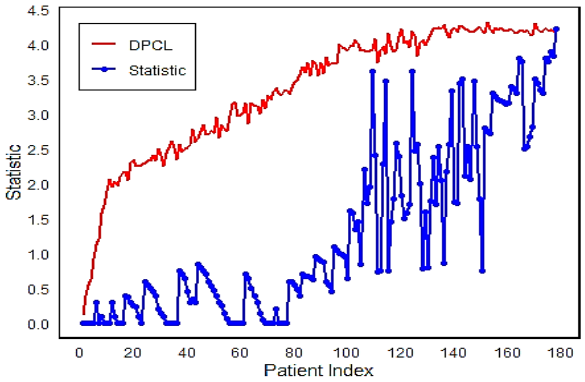


Fig. 2. The performance of the proposed CUSUM chart in detecting shifts in location and scale parameters.

In addition to surgical monitoring, the method can be extended to other healthcare domains where ordinal outcomes and patient heterogeneity are relevant, such as intensive care monitoring, readmission analysis, or rehabilitation assessment. Its adaptability makes it a valuable decision-support tool for quality improvement programs and real-time risk management.

6- Conclusion

Traditional control charts often overlook the variability introduced by patient-specific risk factors, which can lead to biased or inequitable assessments of healthcare performance. This paper proposed a novel RA-CUSUM control chart tailored to monitor surgical outcomes with ordinal structure, while accounting for unobserved heterogeneity through the inclusion of a random effect term. In contrast to binary-outcome models, our approach enables more nuanced performance evaluations by distinguishing varying levels of postoperative recovery. Furthermore, by employing DPCLs, the proposed chart adapts to patient-level risk profiles and ensures a consistent CFAR, even in heterogeneous patient populations.

Extensive simulation studies demonstrated the effectiveness of the method in detecting both location shifts and increased variability, confirming its utility for monitoring both systematic deterioration and unstable surgical performance. The real-world case study illustrated its applicability in clinical settings and emphasized its relevance for quality assurance and risk management in hospitals.

Looking ahead, the integration of machine learning (ML) techniques offers promising avenues for enhancing the flexibility and predictive power of risk-adjusted monitoring systems. For instance, data-driven models such as neural networks or gradient boosting machines could be trained to estimate ordinal outcome probabilities more flexibly than traditional logistic regression. These models may also identify latent interactions among risk factors and dynamically adjust to changes in patient demographics or treatment protocols. In addition, incorporating real-time learning mechanisms could allow control charts to evolve with incoming data, enabling early adaptation to new patterns of variation or risk. Such advancements would pave the way for more intelligent and adaptive monitoring frameworks in healthcare.

Future research may also explore the extension of the proposed method to other sectors where ordinal outcomes and contextual heterogeneity are prevalent, such as education, finance, and public health. Further investigation into joint monitoring of multiple indicators or integrating prior clinical knowledge into the model structure would also be valuable directions for advancing the practical impact of this work.

References

[1] Agresti, A. (2010). Analysis of ordinal categorical data: Wiley.

[2] Aminnayeri, M., & Sogandi, F. (2016). A risk-adjusted self-starting Bernoulli CUSUM control chart with dynamic probability control limits. AUT Journal of Modeling and Simulation, 48(2), 103-110.

[3] Aytaçoğlu, B., Driscoll, A. R., & Woodall, W. H. (2023). Design of adaptive EWMA control charts using the conditional false alarm rate. Quality and Reliability Engineering International, 39(6), 2206-2214.

[4] Bersimis, S., & Sachlas, A. (2022). Surveilling public health through statistical process monitoring: a literature review and a unified framework. Communications in Statistics: Case Studies, Data Analysis and Applications, 8(3), 515-543.

[5] Christensen, R. H. B. (2015). Analysis of ordinal data with cumulative link models—estimation with the R-package ordinal. R-package version, 28, 406.

[6] Grigg, O. A. (2019). The STRAND Chart: A survival

time control chart. Statistics in Medicine, 38(9), 1651-1661.

[7] Khosravi, R., Owlia, M. S., Fallahnezhad, M. S., & Amiri, A. (2018). Phase I risk-adjusted control charts for surgical data with ordinal outcomes. Communications in Statistics-Theory and Methods, 47(18), 4422-4432.

[8] Lai, X., Li, X., Liu, L., Tsung, F., Lai, P. B., Wang, J., . . . Liu, J. (2021). A risk-adjusted approach to monitoring surgery for survival outcomes based on a weighted score test. Computers & Industrial Engineering, 160, 107568.

[9] Li, L., Liu, Y., Shang, Y., & Liu, Z. (2023). A new phase II risk-adjusted CUSUM chart for monitoring surgical performance. Computers & Industrial Engineering, 186, 109738.

[10] Paynabar, K., Jin, J., & Yeh, A. B. (2012). Phase I risk-adjusted control charts for monitoring surgical performance by considering categorical covariates. Journal of Quality Technology, 44(1), 39-53.

[11] Sego, L. H., Reynolds Jr, M. R., & Woodall, W. H. (2009). Risk-adjusted monitoring of survival times. Statistics in Medicine, 28(9), 1386-1401.

[12] Steiner, S. H., Cook, R. J., Farewell, V. T., & Treasure, T. (2000). Monitoring surgical performance using risk-adjusted cumulative sum charts. Biostatistics, 1(4), 441-452.

[13] Tang, X., Gan, F. F., & Zhang, L. (2015). Risk-adjusted cumulative sum charting procedure based on multiresponses. Journal of the American Statistical Association, 110(509), 16-26.

[14] Zhang, X., Loda, J. B., & Woodall, W. H. (2017). Dynamic probability control limits for risk-adjusted CUSUM charts based on multiresponses. Statistics in medicine, 36(16), 2547-2558.

[15] Zhang, X., & Woodall, W. H. (2015). Dynamic probability control limits for risk-adjusted Bernoulli CUSUM charts. Statistics in Medicine, 34(25), 3336-3348.

Appendix A. Derivation of the Score Statistic

To monitor variability in surgical performance, we compute a score statistic that assesses whether the variance of the patient-specific random effect $\delta_i \sim N(0,\theta)$ is greater than zero.

The likelihood contribution of patient *i*, derived from the proportional odds model with random intercepts, for a single observation y_i is:

$$L_i(\theta) = \sum_{j=1}^{J} P(y_i = j)^{I(y_i=j)}.$$

Thus, the log-likelihood $\ell_i(\theta)$ is:

$$\ell_i(\theta) = \sum_{j=1}^{J} I(y_i = j) \log P(y_i = j)$$

where $P(y_i \le j) = \dfrac{\exp(\eta_j)}{1+\exp(\eta_j)}$ and $\eta_j = \alpha_j + \beta^T x_i + \gamma^T d_i + \delta$.

The derivative of $\log P(y_i = j)$ concerning to δ is:

$$\frac{\partial log\left(P(y_i = j)\right)}{\partial \delta} = \frac{\frac{\partial P(y_i = j)}{\partial \delta}}{P(y_i = j)}.$$

Derivative of $P(y_i = j)$:

$$\frac{\partial P(y_i = j)}{\partial \delta} = \frac{\partial P(y_i \le j)}{\partial \delta} - \frac{\partial P(y_i \le j\text{-}1)}{\partial \delta}.$$

For $P(y_i \le j)$:

$$\frac{\partial P(y_i \le j)}{\partial \delta} = P(y_i \le j)\left(1 - P(y_i \le j)\right).\frac{\partial \eta_j}{\partial \delta}.$$

Since $\frac{\partial \eta_j}{\partial \delta} = 1$,

$$\frac{\partial P(y_i \le j)}{\partial \delta} = P(y_i \le j)\left(1 - P(y_i \le j)\right) = s_{i,j}.$$

In the same way,

$$\frac{\partial P(y_i \le j)}{\partial \delta} = P(y_i \le j\text{-}1)\left(1 - P(y_i \le j\text{-}1)\right) = s_{i,j-1}.$$

Therefore, we have:

$$\frac{\partial P(y_i = j)}{\partial \delta} = P(y_i \le j)\left(1 - P(y_i \le j)\right) - P(y_i \le j\text{-}1)\left(1 - P(y_i \le j\text{-}1)\right) = s_{i,j} - s_{i,j-1},$$

and:

$$\frac{\partial \log P(y_i = j)}{\partial \delta} = \frac{P(y_i \le j)\left(1 - P(y_i \le j)\right) - P(y_i \le j\text{-}1)\left(1 - P(y_i \le j\text{-}1)\right)}{P(y_i = j)} = \frac{s_{i,j} - s_{i,j-1}}{p_{ij}}.$$

Second Derivative of the Log-likelihood

The second derivative of $\log P(y_i = j)$ with respect to δ is:

$$\frac{\partial^2 log P(y_i = j)}{\partial \delta^2} = \frac{\frac{\partial^2 P(y_i = j)}{\partial \delta^2}}{P(y_i = j)} - \left(\frac{\frac{\partial P(y_i = j)}{\partial \delta}}{P(y_i = j)}\right)^2.$$

Second Derivative of $P(y_i = j)$:

$$\frac{\partial^2 P(y_i = j)}{\partial \delta^2} = \frac{\partial}{\partial \delta}\left(\frac{\partial P(y_i = j)}{\partial \delta}\right).$$

Substituting the expression for the first derivative, we have:

$$\frac{\partial^2 P(y_i = j)}{\partial \delta^2} = \frac{\partial}{\partial \delta}\left[P(y_i \le j)\left(1 - P(y_i \le j)\right) - P(y_i \le j\text{-}1)\left(1 - P(y_i \le j\text{-}1)\right)\right].$$

For $P(y_i \le j)\left(1\text{-}P(y_i \le j)\right)$:

$$\frac{\partial}{\partial \delta}\left[P(y_i \le j)\left(1 - P(y_i \le j)\right)\right] = \frac{\partial P(y_i \le j)}{\partial \delta}\left(1 - 2P(y_i \le j)\right)$$
$$= P(y_i \le j)\left(1\text{-}P(y_i \le j)\right)\left(1 - 2P(y_i \le j)\right) = q_{i,j}$$

Similarly, for $P(y_i \le j-1)\left(1\text{-}P(y_i \le j-1)\right)$:

$$\frac{\partial}{\partial\delta}\Big[P\big(y_i \le j\big)\big(1-P\big(y_i \le j\big)\big)\Big] = P\big(y_i \le j\text{-}1\big)\big(1\text{-}P\big(y_i \le j\text{-}1\big)\big)\big(1-2P\big(y_i \le j\text{-}1\big)\big) = q_{i,j-1}$$

Thus, the second derivative is:

$$\frac{\partial^2 P\big(y_i = j\big)}{\partial\delta^2} = q_{i,j} - q_{i,j-1}$$

Accordingly, the second derivative of $\log P\big(y_i = j\big)$ is obtained as follows:

$$\frac{\partial^2 log P\big(y_i = j\big)}{\partial\delta^2} = \frac{q_{i,j}\text{-}q_{i,j\text{-}1}}{p_{ij}} - \left(\frac{s_{i,j}}{p_{ij}}\right)^2$$

Finally, the score statistic is given by:

$$W_i = \sum_{j=1}^{J} I\big(y_i = j\big)\left[\left(\frac{s_{i,j}\text{-}s_{i,j\text{-}1}}{p_{ij}}\right)^2 + \frac{q_{i,j}\text{-}q_{i,j\text{-}1}}{p_{ij}} - \left(\frac{s_{i,j}}{p_{ij}}\right)^2\right] = \sum_{j=1}^{J} I\big(y_i = j\big)\frac{\Big(s_{i,j\text{-}1}\big(s_{i,j\text{-}1}\text{-}2s_{i,j}\big)\Big)^2 + p_{ij}\big(q_{i,j}\text{-}q_{i,j\text{-}1}\big)}{p_{ij}^{\,2}}.$$

HOW TO CITE THIS ARTICLE
R. Khosravi, H. Gholinezhad, H. Esmaeeli, A Risk-Adjusted CUSUM Chart for Monitoring Surgical Performance with Ordinal Outcomes and Random Effects, AUT J. Model. Simul., 57(1) (2025) 17-28.

DOI: 10.22060/miscj.2025.23841.5400