

Graphical-model based O/E control chart for monitoring multiple outcomes from a multi-stage healthcare procedure

Nokuthaba Sibanda

School of Mathematics, Statistics and Operations Research, Victoria University of Wellington, P O Box 600, Wellington, New Zealand. Email: nokuthaba.sibanda@vuw.ac.nz

Summary. Most statistical process control (SPC) programmes in healthcare focus on surveillance of outcomes at the final stage of a procedure, such as mortality or failure rates. This approach ignores the multi-stage nature of such procedures, in which a patient progresses through several stages prior to the final stage. In this paper we introduce a novel approach to SPC programmes in healthcare. Our proposed approach uses regression adjustment, variance propagation and multi-stage control charts that have been in use in industrial applications for decades. A test statistic for the control charts is proposed. Simulations are performed to test the control charts and the results are summarised using a probability of true detection. An illustrative example using data from a maternity unit is included. Our results show that the multi-stage approach has two main advantages: (i) enables a closer study of all stages of a procedure and thus a clearer understanding of the potential sources of excess variation that lead to a shift in final stage outcome rates, (ii) changes in mid-stage outcomes rates can serve as a warning of pending changes in final stage outcomes rates if left unchecked.

1. Introduction

Healthcare procedures generally comprise multiple stages. For example, in a major surgical procedure the patient is prepared for the operation, anaesthetised and the surgery is then carried out. A poor outcome at an upstream stage is likely, through variance propagation, to result in poor outcomes at downstream stages. Most studies on healthcare performance monitoring focus on monitoring end-stage clinical outcomes^{1,2,3,4,5}, ignoring what occurs in the earlier stages of the procedure. In this paper we propose that healthcare performance monitoring should encompass all stages of a procedure. This approach has a three-fold advantage. First, the correlation and variance propagation across the various stages will be better understood and taken into account. Second, explicit monitoring of upstream stages may detect trends in practice that, if continued, will lead to more poor end-stage clinical outcomes. This allows these trends to be curtailed before they reach a level where they cause excessive poor end-stage outcomes. Third, when monitoring multiple stage procedures, understanding the impact of each stage is important for allocation of resources for quality improvement⁶.

The multi-stage nature of some healthcare procedures has parallels in manufacturing, where a production process comprises various stages. Examples are numerous and include vehicle assembly⁷ and semiconductor manufacturing^{8,9}. This paper presents a new approach, Multi-Stage O/E (MSOE) control charts, for monitoring outcomes of multi-stage healthcare procedures. The approach we propose is based on regression adjustment^{10,11} and variance transmission models¹² in that the outcome variable at each stage is regressed

on a subset of process and upstream outcome variables, but with two main differences. First, rather than charting residuals from the regression models, which we believe are hard to interpret, our approach charts a ratio of observed to expected outcomes. The expected outcomes are based on predictions from the regression model. Secondly, in healthcare, patients have different underlying risks for a given outcome. Therefore, in addition to the process and upstream quality variables, we regress the quality variable on patient-specific risk factors that influence the quality variable. To our knowledge this is the first time that a multi-stage process approach has been proposed for monitoring health outcomes. One main advantage of this approach is that it allows a whole system approach to the monitoring process, which in turn allows better understanding of patterns in outcomes seen in end-stage results.

The proposed approach is illustrated using outcomes from a maternity delivery unit. As a first step, we construct the joint conditional distribution for all outcome variables. A graphical model is used to represent the decomposition of this joint distribution into a product of conditional independence distributions, each of which determines the structure of the regression model for each outcome variable. In our application we use a combination of expert knowledge, literature and empirical analysis to determine the structure of the graphical model and no discussion is made of a statistical approach to estimation of the graph structure. We use simulations to examine the sensitivity of our proposed control charts to various types of shifts within the process.

The layout of the paper is as follows. Section 2 provides a review of multi-stage control charts in industrial applications. Section 3 describes the motivating example for the paper and Section 4 describes our proposed control charts and computation of control limits. Sections 5 and 6 give simulation and application results, respectively, and these are discussed in Section 7.

2. Multi-stage control charts in industrial applications

Mandel (1969) proposed a model-based approach to monitor a quality variable, man-hours used to process mail, that was highly correlated with a process input, mail volume¹³. The cause-selecting chart¹⁴ and the regression adjustment method^{10,11} extended Mandel's approach to multi-stage processes. In the cause-selecting approach, the quality variable at a given stage is regressed on a quality variable at a previous stage. The residuals from the regression model of each stage are then charted and monitored. The regression adjustment approach extended this idea by regressing the quality variable at a given stage on any subset of the other quality variables. The capability of these approaches in modelling interstage correlations makes them particularly suited to monitoring of multistage manufacturing process (MMP)¹⁵. If any quality or process variable undergoes a parameter shift, either in mean or variance, it may affect some or all of the quality variables in succeeding stages, but none in the preceding stages. The regression adjustment technique was also implemented by Rao *et al* (1996)¹⁶ in a Bayesian framework to monitor multistage semiconductor manufacturing processes.

Interstage correlations are the result of variance transmission from upstream to downstream stages in a MMP. The monitoring schemes can be extended further to estimate variance components due to the different stages. A variance transmission model was developed by Lawless *et al.* (1999)¹² for this purpose and reviewed by Agrawal *et al.* (1999)¹⁷. A first order autoregressive model given by

$$Y_1 \sim N(\mu_1, \sigma_1^2) \quad (1)$$

$$Y_i = \alpha_i + \beta_i Y_{i-1} + e_i, \quad i = 2, \dots, M \quad (2)$$

was used, where M is the number of stages. The model assumes that the output at stage i given all other preceding stages depends only on what is presented to it from stage $i - 1$. Taking the variance of both sides in equation (2), we can see that the variance of the output at stage i decomposes into variation transmitted to Y_i from stage $i - 1$ and variation added at stage i ^{12, 6}, present a further development of this model in which quality variable(s) at stage i are allowed to depend on *process variables*, \mathbf{X}_i , at stage i as well as quality variables at any of the preceding stages, \mathbf{Y}_k , $k \in D_i$, where $D_i = \{1, 2, \dots, i - 1\}$. The resulting model is written as a system of simultaneous equations given by

$$Y_{ij} = \sum_{l=1}^{p_i} \alpha_{ijl} X_{ijl} + \sum_{k \in D_i} \sum_{l=1}^{q_k} \beta_{ijkl} Y_{kl} + \epsilon_{ij}, \quad (3)$$

where at stage i , $i = 1, \dots, M$ and quality variable $j = 1, \dots, q_i$:

α_{ijl} = direct effect of l th process variable (PV) on ij th quality variable (QV),

D_i = set of preceding stages,

β_{ijkl} = direct effect of l th QV at stage k , $k \in D_i$, on ij th QV,

ϵ_{ij} = random error term with mean zero and variance σ_{ij}^2 .

Zantek (2002) then showed that through the model in (3) above, the variance of each quality variable can be decomposed into components attributable to each stage⁶.

The cause-selecting and regression adjustment charts are based on normal linear regression models, and therefore assume that each quality variable follows a Gaussian distribution and that there are linear relationships between the variables. The regression adjustment approach was further extended to cope with non-linear relationships between quality variables from different stages. Generalized linear models (GLM) were used to investigate cases with a count response¹⁸ and a gamma-distributed response¹⁹. Robust generalized linear models were used to address outliers in the historical data used to design the control charts⁸.

3. The multistage process in a maternity unit

Broadly speaking, the process of delivery can be divided into three stages: Dilation, Birth and Post-partum period²⁰. Each stage has at least one outcome measure and may or may not have associated process variables and patient-specific risk factors that affect each outcome variable.

The first stage, Dilation, consists of a latent and established first stage of labour. We have defined one outcome variable for this stage: length of Stage I labour. Pro-longed labour poses a risk to the infant and the mother. In practice the infant's vital signs are monitored during this stage and intervention is indicated when the risk is perceived to be too high. Labour is considered prolonged if it continues for longer than 18 hours in a primi-paras

4 Sibanda

(first time) mother or longer than 12 hours in a multi-paras mother²⁰. Note that in some cases, the first stage may be skipped due to an elective or emergency Ceasarean-section. In such cases, we set the length of the first stage to zero.

The second stage, Birth, consists of a passive and an active second stage of labour. We have defined two outcomes for this stage: length of Stage II labour and excessive maternal injury (third or fourth degree tear). Once labour is over the infant must be delivered quickly to minimise risk, and care must be taken not to cause excessive injury to the mother in the process. The second stage is pro-longed if it continues for more than two hours in a primi-paras mother or longer than one hour in a multi-paras mother²⁰.

The third stage, Post-partum period, is the time from the birth of the baby to the birth of the placenta. We have defined two outcomes for this stage: the infant status and amount of maternal blood loss. Infant status is commonly measured using the Apgar score^{21,22}, a composite measure on a scale from 0 (worst) to ten (best) used to assess the health of newborn infants. If, at five minutes after birth, a baby scores seven or more on the Apgar scale, they are considered to be in good health. A number of different definitions have been used for postpartum haemorrhage (PPH). The most common of these defines PPH as blood loss of more than 1000ml if delivery is through Caesarean section or more than 500ml without a Caesarean section²³.

3.1. Model specification for the delivery process

There are three major stages with a total of five outcome variables in the delivery process as stated above. We define the five binary outcome variables as follows:

$$\begin{aligned} Y_1 &= \begin{cases} 1 & \text{if } L_1 > 18 \text{ hours and Parity}=0 \text{ or if } L_1 > 12 \text{ hours and Parity} \geq 1 \\ 0 & \text{otherwise} \end{cases} \\ Y_2 &= \begin{cases} 1 & \text{if } L_2 > 2 \text{ hours and Parity}=0 \text{ or if } L_2 > 1 \text{ hour and Parity} \geq 1 \\ 0 & \text{otherwise} \end{cases} \\ Y_3 &= \begin{cases} 1 & \text{if a 3rd or 4th degree tear occurs} \\ 0 & \text{otherwise} \end{cases} \\ Y_4 &= \begin{cases} 1 & \text{if 5-minute Apgar score} < 7 \\ 0 & \text{otherwise} \end{cases} \\ Y_5 &= \begin{cases} 1 & \text{if maternal blood loss} > 500ml \text{ with no Caesarean, or } > 1000ml \text{ with a Caesarean} \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where L_1 is the length of the first stage and L_2 is the length of the second stage.

Each outcome variable is influenced by one or more process variables or risk factors, and may also depend on an upstream outcome variable. We consider the following process

variables

$$\begin{aligned} X_1 &= \begin{cases} 1 & \text{if labour induced} \\ 0 & \text{otherwise} \end{cases} \\ X_2 &= \begin{cases} 1 & \text{if mechanical instruments used during the second stage} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

and risk factors:

$$\begin{aligned} Z_1 &= \begin{cases} 1 & \text{if presentation is posterior or transverse} \\ 0 & \text{otherwise} \end{cases} \\ Z_2 &= \begin{cases} 1 & \text{if it is the mother's first birth} \\ 0 & \text{otherwise} \end{cases} \\ Z_3 &= \text{length of gestation (weeks)} \end{aligned}$$

Using a combination of literature^{24,23,25,26}, expert knowledge (Sibanda T, personal communication) and empirical evidence, we determined the model structure representing the relationships among the outcome, process and risk factor variables. The model structure we adopted is presented using a graphical model as shown in Figure 1.

The graphical model in Figure 1 can be viewed as a Bayesian network^{27,28}. A Bayesian network is a directed acyclic graph (DAG) with node set V representing random variables, $\mathbf{Y} = \{Y_{v \in V}\}$ having a joint probability distribution function that can be written as

$$P(\mathbf{Y}) = \prod_{v \in V} P(Y_v | Y_{\text{pa}(v)}). \quad (4)$$

The term $\text{pa}(v)$ represents the set of parent nodes of the node v . The power of a DAG representation is that once the structure is known, the joint probability distribution of \mathbf{Y} can be written in the form of 4 using the conditional independence axioms introduced by Dawid (1979)²⁹. In Equation (4), each node is conditionally independent of all non-descendants, given its parent nodes. Based on the DAG structure in Figure 1 we can write:

$$\begin{aligned} &P(Y_1, Y_2, Y_3, Y_4, Y_5 | \mathbf{X}, \mathbf{Z}) \\ &= P(Y_5, Y_4 | Y_1, Y_2, Y_3, \mathbf{X}, \mathbf{Z}) P(Y_2, Y_3 | Y_1, \mathbf{X}, \mathbf{Z}) P(Y_1 | \mathbf{X}, \mathbf{Z}) \\ &= P(Y_5 | Y_1, Y_2, \mathbf{X}, \mathbf{Z}) P(Y_4 | \mathbf{Z}) P(Y_3 | Y_2, \mathbf{X}, \mathbf{Z}) P(Y_2 | \mathbf{X}, \mathbf{Z}) P(Y_1 | \mathbf{X}, \mathbf{Z}) \end{aligned} \quad (5)$$

We can therefore fit a separate regression model for each outcome variable conditional on its parents, which are the relevant process variables, risk factors and upstream outcome variables. Each of the outcome variables Y_1, \dots, Y_5 is binary and an appropriate model is a logistic regression model. Using the formulation of generalised linear models, the model equation for outcome $Y_i, i = 1, \dots, 5$ is given by:

$$g(\pi_i) = \eta_i = \alpha_i + \sum_{l=1}^p \beta_{li} x_l + \sum_{l=1}^q \gamma_{li} y_l + \sum_{l=1}^r \delta_{li} z_l + \epsilon_i \quad (6)$$

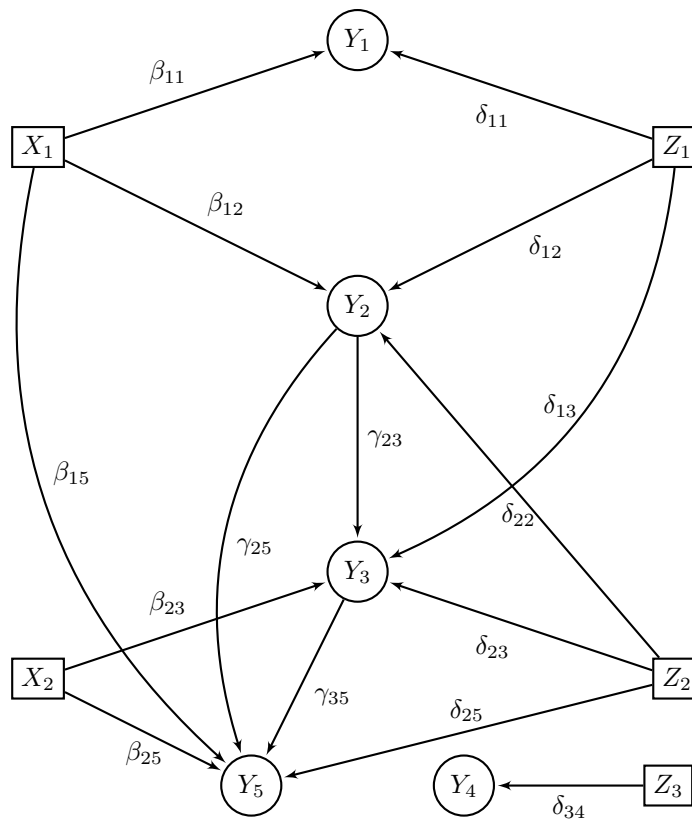


Fig. 1. Directed acyclic graph showing the model structure for a multi-stage process in a maternity unit. Edge directions indicate the direction of relationships and edge labels are the regression coefficients for the model in equation (6).

Table 1. Parameter estimates and standard errors for the logistic regression models represented in Figure 1 and equation (6).

Parameter	Estimate	Standard Error
β_{11}	-1.550	0.458
δ_{11}	0.688	0.281
β_{12}	0.287	0.105
δ_{12}	1.544	0.127
δ_{22}	1.383	0.086
β_{23}	0.678	0.162
γ_{23}	0.426	0.163
δ_{13}	0.414	0.209
δ_{23}	0.851	0.149
δ_{34}	-0.272	0.050
β_{15}	0.325	0.117
β_{25}	1.157	0.118
γ_{25}	0.491	0.117
γ_{35}	1.277	0.148
δ_{25}	0.341	0.103

where $\pi_i = E(Y_i)$ and

β_{li} = direct effect of l th process variable on Y_i

γ_{li} = direct effect of l th outcome variable on Y_i

δ_{li} = direct effect of l th risk factor on Y_i

p, q, r = number of process, outcome and risk factor variables respectively

ϵ_i = random error term with mean zero and variance σ_i^2

Assuming a logit link function and using matrix notation, we can re-write 6 as

$$\log(\pi_i/(1 - \pi_i)) = \eta_i = \alpha_i + \mathbf{X}'\boldsymbol{\beta}_i + \mathbf{Y}'\boldsymbol{\gamma}_i + \mathbf{Z}'\boldsymbol{\delta}_i$$

and

$$\pi_i = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)} = \frac{\exp(\alpha_i + \mathbf{X}'\boldsymbol{\beta}_i + \mathbf{Y}'\boldsymbol{\gamma}_i + \mathbf{Z}'\boldsymbol{\delta}_i)}{1 + \exp(\alpha_i + \mathbf{X}'\boldsymbol{\beta}_i + \mathbf{Y}'\boldsymbol{\gamma}_i + \mathbf{Z}'\boldsymbol{\delta}_i)}. \quad (7)$$

for $i = 1, \dots, 5$.

We used data from the Southmead maternity unit in Bristol, updated from that published in a previous paper⁴ to obtain model parameter estimates and these are shown in Table 1. We then used these model parameters to inform our simulations for testing the control charts discussed in the sections that follow.

4. Multi-stage O/E (MSOE) control chart

Previous research on control charts for multistage process control charts has been based largely on charting residuals^{19,9}. We propose a new multi-stage control chart based on the

ratio of the number, O , of observed counts to the number, E , of expected counts of the outcome variable of interest, for monitoring outcomes in multistage healthcare processes. The ratio O/E is an intuitive and easily interpretable statistic. It has been used in the past for cross-sectional provider profiling³⁰. The idea of comparing observed and expected counts is not new to control charts for longitudinal performance monitoring. Based on³¹, a cumulative sum (CUSUM) chart using an unadjusted³² and a risk-adjusted³³ difference between O and E was proposed for monitoring surgical outcomes. The control chart we propose can be used for surveillance of outcome rates over time within a single provider, where shifts from historical practice are the main targets for detection. It can also be used by a central governing body for surveillance of multiple providers individually over time where deviation from agreed national or regional standards is the main focus.

4.1. Test statistic

To detect deviation from historical or standard rates for an outcome variable Y_i , for patients in a given unit, we compare the number of observed outcomes to that predicted using a model developed from historical or standard practice. That is we test the hypotheses: $H_0 : R_i = O_i/E_i = 1$ vs $H_1 : R_i \neq 1$. As the model predictions are based on historical practice and outcomes, we expect a change in either to result in significant differences between observed and expected counts of a given outcome.

Suppose we have a sample of n patients whose outcomes we wish to monitor. These patients must be independent of those whose data is used for model development. We observe the i th binary outcome, y_{ij} , for patient $j, j = 1, \dots, n$, and the vectors $(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j)$. Denote the coefficient estimates obtained from historical data as $\hat{\Theta}_i = (\hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i, \hat{\delta}_i)$. The predicted value for patient j is

$$\hat{\pi}_{ij} = \frac{\exp(\hat{\eta}_{ij})}{1 + \exp(\hat{\eta}_{ij})}, \quad (8)$$

where $\hat{\eta}_{ij} = \alpha_i + \mathbf{x}'_j \hat{\beta}_i + \mathbf{y}'_j \hat{\gamma}_i + \mathbf{z}'_j \hat{\delta}_i$.

We propose that the test statistic, O_i/E_i , be calculated at regular time intervals, for example, on a monthly basis. For a given month we observe $o_i = \sum_{j=1}^n y_{ij}$ patients with $Y_{ij} = 1$. Under the assumption that the logistic model with coefficient estimates $\hat{\Theta}_i$ is correct and that outcome rates remain unchanged in the new set of n patients, E_i the expected value of O_i is estimated by $\hat{E}_i = \sum_{j=1}^n \hat{\pi}_{ij}$. Therefore at each time interval the statistic charted is given by:

$$\hat{R}_i = o_i / \hat{E}_i.$$

4.2. Control limits for MSOE control charts

Control charts consist of a centre line (CL) and upper (UL) and lower (LL) limits. The centre line is plotted at the value of the test statistic under the null hypothesis. The upper and lower limits are the confidence limits for a $100(1 - \alpha)\%$ confidence intervals. The observed values \hat{R}_i are plotted, and a point outside the limits is evidence against H_0 for that time interval.

For our control chart, if H_0 is true we expect $R_i = O_i/E_i = 1$. Therefore the centre line will be at drawn at 1 for all time points. To determine the control limits we consider

confidence interval for R_i . A number of approaches have been proposed in literature and we consider three of those approaches here: (1) treat \hat{E}_i as fixed due to unknown $Var(\hat{\Theta}_i)$ ³⁴; (2) assume an approximately normal distribution for R_i ³⁵ and use $Var(\hat{\Theta}_i)$ to estimate $Var(\hat{E}_i)$ and $Var(O_i/\hat{E}_i)$; (3) assume an approximately log-normal distribution for R_i ^{36,35} and use $Var(\hat{\Theta}_i)$ to estimate $Var(\hat{E}_i)$ and $Var(\ln(O_i/\hat{E}_i))$.

4.2.1. Estimating $Var(\hat{E}_i)$, $Var(O_i/\hat{E}_i)$ and $Var(\ln(O_i/\hat{E}_i))$

Let us assume that a development dataset, S_D with n_D patients, is used to fit logistic regression models for all binary outcome variables. We assume also that, as a minimum, we have available the estimated coefficients and an estimate of the covariance matrix of the estimated parameters for each fitted logistic model. For each successive month during the time period whose outcomes we are monitoring we have a monitoring sample, S_M with n_M patients.

Given the logistic regression model fitted using S_D , we denote the estimated parameters and estimated covariance matrix for the i th outcome variable as $\hat{\Theta}_i^D$ and $\hat{\mathbf{S}}(\hat{\Theta}_i^D)$ respectively. The estimated covariance matrix is given by

$$\hat{\mathbf{S}}(\hat{\Theta}_i^D) = (\mathbf{W}^{D'} \hat{\mathbf{V}}_i^D \mathbf{W}^D)^{-1}, \quad (9)$$

where $\mathbf{W}^D = (\mathbf{X}^D, \mathbf{Y}^D, \mathbf{Z}^D)$ is the matrix of covariates from S_D and $\hat{\mathbf{V}}_i^D$ is the $n_D \times n_D$ diagonal matrix that contains the model-based estimates of $Var(Y_{ij})$ for $j = 1, \dots, n_D$ with $\hat{\mathbf{V}}_i^D = \text{diag}\{\hat{\pi}_{ij}(1 - \hat{\pi}_{ij})\}_{n_D}$ ³⁴.

Given that $O_i = \sum_{j=1}^{n_M} Y_{ij}$ and assuming independence among patients in S_M we can write

$$Var(O_i) = \sum_{j=1}^{n_M} Var(Y_{ij} | \mathbf{w}_j^M) = \sum_{j=1}^{n_M} \pi_{ij}(1 - \pi_{ij}) \quad (10)$$

for which a consistent estimator is

$$\widehat{Var}(O_i) = \sum_{j=1}^{n_M} \hat{\pi}_{ij}(1 - \hat{\pi}_{ij}). \quad (11)$$

This can be expressed more concisely in matrix form as $\widehat{Var}(O_i) = \mathbf{1}' \hat{\mathbf{V}}_i^M \mathbf{1}$, where $\hat{\mathbf{V}}_i^M$ is the $n_M \times n_M$ diagonal matrix, $\hat{\mathbf{V}}_i^M = \text{diag}\{\hat{\pi}_{ij}(1 - \hat{\pi}_{ij})\}_{n_M}$ and $\mathbf{1} = \{1\}_{n_M}$ an n_M vector of ones.

\hat{E}_i is given by $\hat{E}_i = \sum_j^{n_M} \hat{\pi}_{ij}$, the sum of predicted probabilities, $\hat{\pi}_{ij}$, for individuals in S_M , where $\hat{\pi}_{ij}$ is given by (8). Since \hat{E}_i is based on predictions of new observations, to work out $Var(\hat{E}_i)$, we consider variability in $\hat{\pi}_{ij}$ and in $\hat{\Theta}_i^D$ ³⁴. The asymptotic ($n \rightarrow \infty$) variance of $\hat{\Theta}_i^D$ is given by equation (9).

An approximation of the variance matrix for $\hat{\boldsymbol{\pi}}_i = \{\hat{\pi}_{i1}, \hat{\pi}_{i2}, \dots, \hat{\pi}_{im}\}$ can be found using the delta method as follows:

$$Var(\hat{\boldsymbol{\pi}}_i)_{n_M \times n_M} = \left(\frac{\partial \boldsymbol{\pi}_i}{\partial \Theta_i^D} \right)_{n_M \times p} Var(\hat{\Theta}_i^D)_{p \times p} \left(\frac{\partial \boldsymbol{\pi}_i}{\partial \Theta_i^D} \right)_{p \times n_M}.$$

This can be written more concisely in matrix form as

$$Var(\hat{\boldsymbol{\pi}}_i) = \mathbf{V}_i^M \mathbf{W}^M (\mathbf{W}^{D'} \mathbf{V}_i^D \mathbf{W}^D)^{-1} \mathbf{W}^{M'} \mathbf{V}_i^M,$$

since $\frac{\partial \boldsymbol{\pi}_i}{\partial \boldsymbol{\Theta}_i^D} = \mathbf{V}_i^M \mathbf{W}^M$. We can therefore write

$$Var(\hat{E}_i) = Var\left(\sum_{j=1}^{n_M} \hat{\pi}_{ij}\right) = \mathbf{1}' \hat{\mathbf{V}}_i^M \mathbf{W}^M \hat{\mathbf{S}}(\hat{\boldsymbol{\Theta}}_i^D) \mathbf{W}^{M'} \mathbf{V}_i^M \mathbf{1} \quad (12)$$

We now consider various options for computing confidence intervals for the ratio $R_i = O_i/\hat{E}_i$.

4.2.2. Control limits with fixed \hat{E}_i

In this first approach we assume that only the model coefficient estimates $\hat{\boldsymbol{\Theta}}_i^D$ obtained using S_D are available and that \mathbf{W}_i^D is not available for the calculation of $\hat{\mathbf{S}}(\hat{\boldsymbol{\Theta}}_i^D)$. This may be the case where detailed historic data is not available and predictions are obtained from published risk stratification tools. In such a case we calculate exact or approximate confidence limits for O_i and divide these by \hat{E}_i .

Assuming O_i is asymptotically normally distributed, that is $O_i \overset{\text{approx}}{\sim} N(\mu_i, \sigma_i^2)$, with $\mu_i = \sum_j E(Y_{ij}|\mathbf{w}_j)$ and $\sigma_i^2 = \sum_j Var(Y_{ij}|\mathbf{w}_j)$. This assumption holds only when μ_i is sufficiently large, usually > 10 .

Based on the normal assumption and equation (11), approximate $100(1-\alpha)\%$ confidence limits for O_i can be calculated using

$$o_i \pm Z_{(1-\alpha/2)} \sqrt{\sum_{j=1}^{n_M} \hat{\pi}_{ij}(1 - \hat{\pi}_{ij})}.$$

The control chart with fixed \hat{E}_i for outcome variable Y_i can then be constructed using

$$\begin{aligned} CL &= 1 \\ LCL &= 1 - Z_{(1-\alpha/2)} \frac{\sqrt{\sum_{j=1}^{n_M} \hat{\pi}_{ij}(1 - \hat{\pi}_{ij})}}{\hat{E}_i} \\ UCL &= 1 + Z_{(1-\alpha/2)} \frac{\sqrt{\sum_{j=1}^{n_M} \hat{\pi}_{ij}(1 - \hat{\pi}_{ij})}}{\hat{E}_i}, \end{aligned} \quad (13)$$

4.2.3. Control limits for Normal R_i

We now consider the case where $\hat{\mathbf{S}}(\hat{\boldsymbol{\Theta}}_i^D)$ is available and R_i is assumed to be approximately normally distributed. The $100(1-\alpha)\%$ confidence interval for R_i in this case is given by

$$\frac{o_i}{e_i} \pm z_{(1-\alpha/2)} \sqrt{Var\left(\frac{O_i}{\hat{E}_i}\right)}.$$

Following³⁵ we can find an approximation of the variance of R_i using a first-degree Taylor expansion:

$$Var\left(\frac{O}{\hat{E}}\right) \approx \frac{1}{e^2}Var(O) + \frac{o^2}{e^4}Var(\hat{E}) \quad (14)$$

where $Var(O)$ and $Var(\hat{E})$ are given in equations (10) and (12), respectively. Using equation (14) we can therefore write

$$Var\left(\frac{O}{\hat{E}}\right) \approx \frac{1}{e^2} \sum_{j=1}^{n_M} \pi_{ij}(1 - \pi_{ij}) + \frac{o^2}{e^4} \mathbf{1}' \hat{\mathbf{V}}_i^M \mathbf{W}^M \hat{\mathbf{S}}(\hat{\Theta}^D) \mathbf{W}^{M'} \mathbf{V}_i^M \mathbf{1}. \quad (15)$$

Therefore, when we assume that R_i is approximately normally distributed, the control chart would have

$$\begin{aligned} CL &= 1 \\ LCL &= 1 - z_{(1-\alpha/2)} \sqrt{Var\left(\frac{O}{\hat{E}}\right)} \\ UCL &= 1 + z_{(1-\alpha/2)} \sqrt{Var\left(\frac{O}{\hat{E}}\right)}, \end{aligned} \quad (16)$$

where $Var(O/\hat{E})$ is given in equation (15).

4.2.4. Control limits with log-Normal R_i

We will now consider the case where $\hat{\mathbf{S}}(\hat{\Theta}_i^D)$ is available and R_i is considered to be log-normally distributed. An approximation of the variance of $\log R_i$ can be found using the delta method as follows^{34,35}:

$$Var\left[\ln\left(\frac{O}{\hat{E}}\right)\right] \approx \frac{1}{o^2}Var(O) + \frac{1}{e^2}Var(\hat{E}). \quad (17)$$

Therefore, when we assume that R_i is approximately log-normally distributed, the control chart will have

$$\begin{aligned} CL &= 1 \\ LCL &= \exp\left(-z_{(1-\alpha/2)} \sqrt{Var\left(\ln\frac{O}{\hat{E}}\right)}\right) \\ UCL &= \exp\left(z_{(1-\alpha/2)} \sqrt{Var\left(\ln\frac{O}{\hat{E}}\right)}\right), \end{aligned} \quad (18)$$

with $Var(\ln(O/\hat{E}))$ given by (17).

In equations (13), (16) and (18), the centre line is placed at 1 as that is the value of R_i when $H_0 : R_i = 1$ is true. A value of R_i that falls outside the control limits as described above indicates statistically significant differences between O_i and E_i . Assuming the model

used to determine \hat{E} is correct, differences between O_i and E_i could be due to a number of reasons. We use simulations to determine the probability of detecting departures from the null hypothesis that occur as a result of the changes listed below. For each scenario we introduce changes of various sizes in the appropriate parameter and then determine the proportion of times an out-of-control signal occurs in the control charts. To allow comparisons across the different types of changes, we focus on changes that affect Y_3 and Y_5 .

- (a) Change in \mathbf{X}_i or \mathbf{Z}_i . For example, there may be a change in $P(X_2 = 1)$, the probability of mechanical instrument use in a given month. Based on the dependencies shown in Figure 1 this will, in turn, result in a change in \hat{E}_3 and \hat{E}_5 .
- (b) Change in the effect, Θ_i of one or more of \mathbf{X} , \mathbf{Y} or \mathbf{Z} on Y_i . We expect this to be the most frequently detected change. \hat{E}_i will be based on predictions that assume the effects found in S_D still apply, when O_i will be drawn from a population where these effects no longer apply. We therefore expect that O_i will differ considerably from \hat{E}_i .

5. Simulation results

In this section we present results of simulations for the types of shifts listed at the end of section 4.2. For each type of shift, we simulate a reduction or increase in the appropriate parameters for data in S_M . We then generate 500 sets of monthly data with the simulated shifts applied. For each shift we estimate the probability of true detection, $P_D = \Pr(\text{true detection} \mid \text{shift})$ as

$$P_D = \begin{cases} \frac{n(\text{lower out-of-control signals} \mid \text{reduction})}{n(\text{monthly data sets})} \\ \frac{n(\text{upper out-of-control signals} \mid \text{increase})}{n(\text{monthly data sets})} \end{cases} \quad (19)$$

5.1. Change in process variable or risk factor

We investigate shifts in Y_i due to changes in $E(X_l)$ or $E(Z_l)$, with β_{li} or δ_{li} remaining unchanged. This would occur if there was a change in practice policy or a change in the underlying population. For example, a change in practice policy may result in fewer mechanical instruments being used (that is a reduction in $E(X_2)$ and this may, in turn, lead to a change in the associated outcome variables. Changes of various sizes are introduced to the proportion of instrumental deliveries ($E(X_2)$) in the period being monitored, with no change in the effect of X_2 on any of the outcome variables. The probability of true detection under various shifts is shown in Table 2 and an example set of charts is shown in Figure 2.

5.2. Change in effect of process variable or risk factor on outcome

We investigate shifts in Y_i due to changes in β_{li} , the effect of process variable X_l on Y_i , or δ_{li} , the effect of risk factor Z_l on Y_i , with $E(X_l)$ and $E(Z_l)$ remaining unchanged. Such changes would occur if, for example, new staff or equipment are introduced whose outcome rates differ from those in the past. For example, a new brand of mechanical instrument

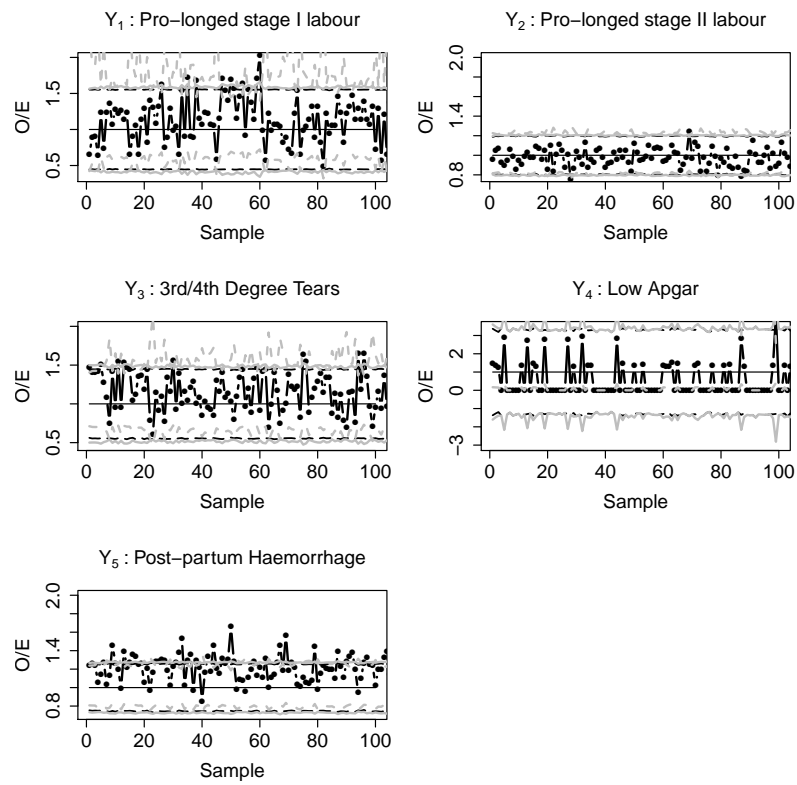


Fig. 2. MSOE control charts for a simulated change where $E(X_2^M) = 2E(X_2^D)$, that is a doubling of rate of deliveries using mechanical instruments. This results in increased rates of Y_3 (3rd/4th degree tears) and Y_5 (PPH) as shown by most points on the corresponding graphs lying above the centre line. The change is more marked for Y_5 than for Y_3 as the effect of X_2 is larger for Y_5 . Control limits are as follows: (- - - - Fixed \hat{E}_i), (— Normal R_i), (- . - . - log-Normal R_i)

Table 2. Proportion of points outside the lower (LL) and upper (UL) limits following change in $E(X_2)$, whereby $E(X_2^M) = cE(X_2^D)$ and the change factor, c , is as shown. The probability of true detection is shown in bold font where applicable.

Change factor, c	Variable	Fixed \hat{E}			Normal R_i			Log-normal R_i		
		$\% \leq LL$	$\% \geq UL$	Total	$\% \leq LL$	$\% \geq UL$	Total	$\% \leq LL$	$\% \geq UL$	Total
0.1	Y_1	0.020	0.024	0.044	0.020	0.016	0.036	0.000	0.036	0.036
	Y_2	0.022	0.028	0.050	0.020	0.020	0.040	0.010	0.030	0.040
	Y_3	0.014	0.040	0.054	0.008	0.026	0.034	0.000	0.054	0.054
	Y_4	0.000	0.014	0.014	0.000	0.004	0.004	0.506	0.040	0.546
	Y_5	0.116	0.006	0.122	0.108	0.006	0.114	0.046	0.006	0.052
0.5	Y_1	0.010	0.034	0.044	0.010	0.024	0.034	0.000	0.042	0.042
	Y_2	0.026	0.026	0.052	0.026	0.022	0.048	0.014	0.028	0.042
	Y_3	0.018	0.042	0.060	0.018	0.028	0.046	0.000	0.054	0.054
	Y_4	0.000	0.004	0.004	0.000	0.000	0.000	0.536	0.038	0.574
	Y_5	0.032	0.022	0.054	0.030	0.018	0.048	0.004	0.026	0.030
1.5	Y_1	0.008	0.032	0.040	0.008	0.026	0.034	0.000	0.040	0.040
	Y_2	0.020	0.022	0.042	0.018	0.016	0.034	0.008	0.024	0.032
	Y_3	0.008	0.056	0.064	0.008	0.034	0.042	0.000	0.062	0.062
	Y_4	0.000	0.002	0.002	0.000	0.000	0.000	0.528	0.030	0.558
	Y_5	0.000	0.316	0.316	0.000	0.276	0.276	0.000	0.338	0.338
2.0	Y_1	0.014	0.020	0.034	0.014	0.020	0.034	0.000	0.030	0.030
	Y_2	0.026	0.020	0.046	0.024	0.018	0.042	0.012	0.026	0.038
	Y_3	0.014	0.064	0.078	0.012	0.036	0.048	0.000	0.068	0.068
	Y_4	0.000	0.002	0.002	0.000	0.000	0.000	0.524	0.022	0.546
	Y_5	0.000	0.620	0.620	0.000	0.574	0.574	0.000	0.630	0.630
4.0	Y_1	0.016	0.036	0.052	0.016	0.026	0.042	0.000	0.048	0.048
	Y_2	0.028	0.020	0.048	0.024	0.014	0.038	0.008	0.026	0.034
	Y_3	0.002	0.094	0.096	0.002	0.028	0.030	0.000	0.070	0.070
	Y_4	0.000	0.002	0.002	0.000	0.000	0.000	0.476	0.024	0.500
	Y_5	0.000	0.998	0.998	0.000	0.996	0.996	0.000	0.998	0.998

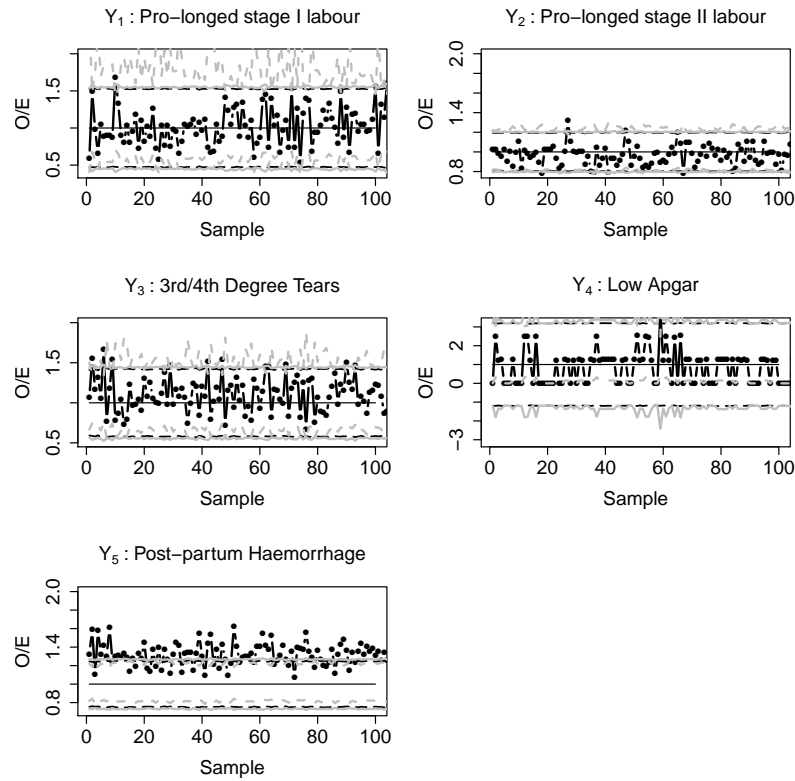


Fig. 3. MSOE control charts for a simulated change where $\beta_{2i}^M = 2\beta_{2i}^D$, for $i = 3, 5$. That is, we simulate a doubling of the effect of use of a mechanical instrument on the risk of 3rd or 4th degree tears and the risk of post-partum haemorrhage. This results in increased rates of Y_3 (3rd/4th degree tears) and Y_5 (PPH) as shown by most points on the corresponding graphs lying above the centre line. The change is more marked for Y_5 than for Y_3 as the effect of X_2 is larger for Y_5 . Control limits are as follows: (----- Fixed \hat{E}_i), (— Normal R_i), (- . - . - log-Normal R_i)

may be introduced that results in reduced post-partum haemorrhage rates. As illustrative examples, changes of various sizes are introduced to β_{23} and β_{25} , the regression coefficients for X_2 in the regression models for Y_3 and Y_5 , respectively. The probability of true detection under various shifts is shown in Table 3 and an example set of charts is shown in Figure 3.

In Table 3 and Figure 3 we simulated shifts in both β_{23} and β_{25} , the *direct* effect of X_2 on Y_3 and Y_5 , respectively. The graphical model presentation in Figure 1 shows that Y_3 is upstream of and has an effect on Y_5 . Therefore changes detected in Y_5 in the simulations above will be due to a combination of the direct effect of the X_2 on Y_5 and an additional *indirect* effect of X_2 through Y_3 . An important question is whether any change would be detected in Y_5 if there was no change in β_{25} , but there is a change in β_{23} . In other words, are indirect effects of process variables or risk factors transmitted through upstream outcomes detectable on downstream outcomes? As an example, we introduce a number of

Table 3. Proportion of points outside the lower (*LL*) and upper (*UL*) limits following change in β_{2i} , whereby $\beta_{2i}^M = c\beta_{2i}^D$, for $i = 3, 5$ and the change factor, c , as shown. The probability of true detection is shown in bold font where applicable.

Change factor, c	Variable	Fixed E			Normal			Log-normal		
		% $\leq LL$	% $\geq UL$	Total	% $\leq LL$	% $\geq UL$	Total	% $\leq LL$	% $\geq UL$	Total
0.1	Y_1	0.018	0.024	0.042	0.014	0.014	0.028	0.000	0.032	0.032
	Y_2	0.010	0.060	0.070	0.008	0.052	0.060	0.000	0.062	0.062
	Y_3	0.106	0.000	0.106	0.102	0.000	0.102	0.000	0.002	0.002
	Y_4	0.000	0.030	0.030	0.000	0.000	0.000	0.524	0.030	0.554
	Y_5	0.402	0.000	0.402	0.384	0.000	0.384	0.204	0.000	0.204
0.5	Y_1	0.030	0.020	0.050	0.026	0.010	0.036	0.000	0.034	0.034
	Y_2	0.020	0.064	0.084	0.016	0.056	0.072	0.004	0.066	0.070
	Y_3	0.076	0.008	0.084	0.076	0.008	0.084	0.000	0.012	0.012
	Y_4	0.000	0.030	0.030	0.000	0.000	0.000	0.516	0.030	0.546
	Y_5	0.200	0.000	0.200	0.190	0.000	0.190	0.098	0.000	0.098
1.5	Y_1	0.024	0.020	0.044	0.020	0.018	0.038	0.000	0.032	0.032
	Y_2	0.002	0.060	0.062	0.002	0.042	0.044	0.000	0.064	0.064
	Y_3	0.014	0.016	0.030	0.010	0.012	0.022	0.000	0.028	0.028
	Y_4	0.000	0.032	0.032	0.000	0.000	0.000	0.522	0.032	0.554
	Y_5	0.002	0.202	0.204	0.000	0.152	0.152	0.000	0.214	0.214
2.0	Y_1	0.024	0.012	0.036	0.018	0.006	0.024	0.000	0.022	0.022
	Y_2	0.018	0.074	0.092	0.016	0.062	0.078	0.010	0.078	0.088
	Y_3	0.002	0.074	0.076	0.002	0.050	0.052	0.000	0.096	0.096
	Y_4	0.000	0.036	0.036	0.000	0.000	0.000	0.542	0.036	0.578
	Y_5	0.000	0.704	0.704	0.000	0.654	0.654	0.000	0.720	0.720
4.0	Y_1	0.010	0.030	0.040	0.010	0.016	0.026	0.000	0.038	0.038
	Y_2	0.012	0.066	0.078	0.008	0.046	0.054	0.004	0.074	0.078
	Y_3	0.000	0.958	0.958	0.000	0.938	0.938	0.000	0.962	0.962
	Y_4	0.000	0.022	0.022	0.000	0.000	0.000	0.534	0.022	0.556
	Y_5	0.000	1.000	1.000	0.000	1.000	1.000	0.000	1.000	1.000

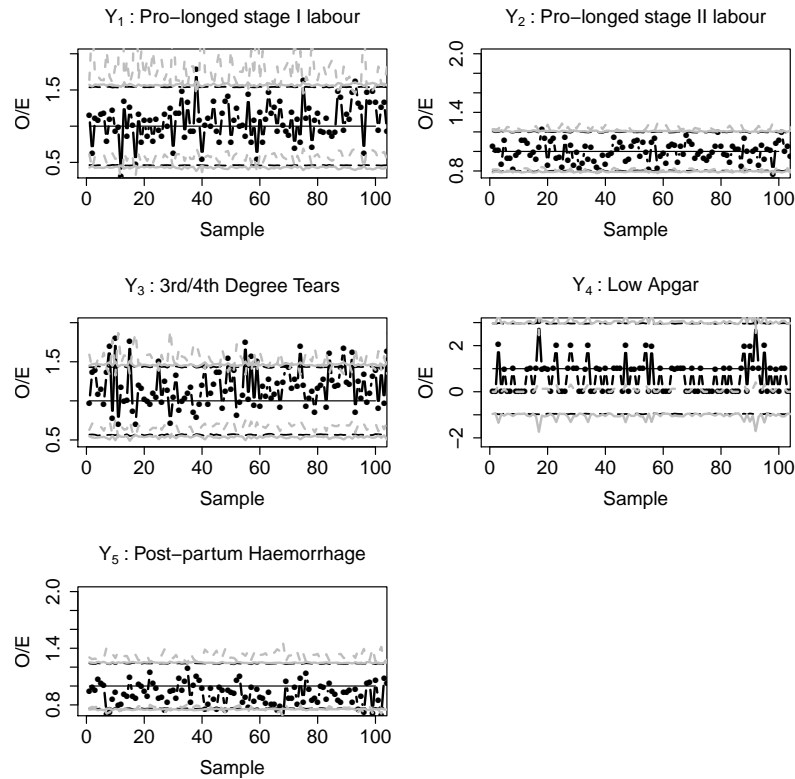


Fig. 4. MSOE control charts for a simulated change where $\beta_{23}^M = 2\beta_{23}^D$. That is, we simulate a doubling of the effect of use of a mechanical instrument on the risk of 3rd or 4th degree tears. There is an increased rate of Y_3 (3rd/4th degree tears) as shown by most points on the corresponding graphs lying above the centre line. There is no shift in Y_5 (PPH). Control limits are as follows: (- - - - Fixed \hat{E}_i), (— Normal R_i), (- . - . - log-Normal R_i)

changes of different sizes in β_{23} , but none in β_{25} and assess the probability of true detection in Y_3 and Y_5 . The results are shown in Table 4 and example charts are shown in Figure 4.

6. Application to Southmead Maternity Unit data

In this section we present results of an illustrative example in which multi-stage control charts were applied to data from a maternity unit. Data used were for deliveries carried out in 2008 at a large tertiary unit with about 5000 births a year on average. Deliveries with a breech presentation or those resulting in multiple births were excluded. Data for the first three months of the year (1204 births) were used to fit the model presented in Figure 1. The fitted model was then used to determine and to estimate expected counts for each of the outcome variables Y_1, \dots, Y_5 for the rest of the year. Observed counts were obtained and control charts constructed for the rest of the deliveries on a fortnightly frequency, with

Table 4. Proportion of points outside the lower (LL) and upper (UL) limits following change in β_{23} , whereby $\beta_{23}^M = c\beta_{23}^D$ and the change factor, c is as shown. The probability of true detection is shown in bold font where applicable.

Change factor, c	Variable	Fixed E			Normal			Log-normal		
		$\% \leq LL$	$\% \geq UL$	Total	$\% \leq LL$	$\% \geq UL$	Total	$\% \leq LL$	$\% \geq UL$	Total
0.1	Y_1	0.020	0.046	0.066	0.018	0.032	0.050	0.000	0.054	0.054
	Y_2	0.040	0.008	0.048	0.032	0.004	0.036	0.010	0.010	0.020
	Y_3	0.036	0.020	0.056	0.034	0.014	0.048	0.000	0.022	0.022
	Y_4	0.000	0.166	0.166	0.000	0.000	0.000	0.502	0.166	0.668
	Y_5	0.010	0.048	0.058	0.008	0.036	0.044	0.000	0.054	0.054
0.5	Y_1	0.020	0.034	0.054	0.020	0.024	0.044	0.000	0.044	0.044
	Y_2	0.028	0.018	0.046	0.024	0.012	0.036	0.012	0.020	0.032
	Y_3	0.030	0.014	0.044	0.024	0.006	0.030	0.000	0.018	0.018
	Y_4	0.000	0.118	0.118	0.000	0.000	0.000	0.548	0.118	0.666
	Y_5	0.006	0.028	0.034	0.006	0.016	0.022	0.002	0.040	0.042
1.5	Y_1	0.012	0.044	0.056	0.012	0.030	0.042	0.000	0.062	0.062
	Y_2	0.044	0.012	0.056	0.042	0.006	0.048	0.006	0.014	0.020
	Y_3	0.004	0.098	0.102	0.004	0.066	0.070	0.000	0.128	0.128
	Y_4	0.000	0.124	0.124	0.000	0.000	0.000	0.516	0.124	0.640
	Y_5	0.016	0.050	0.066	0.010	0.036	0.046	0.002	0.062	0.064
2.0	Y_1	0.016	0.062	0.078	0.014	0.028	0.042	0.000	0.094	0.094
	Y_2	0.022	0.014	0.036	0.018	0.010	0.028	0.006	0.016	0.022
	Y_3	0.000	0.228	0.228	0.000	0.192	0.192	0.000	0.270	0.270
	Y_4	0.000	0.112	0.112	0.000	0.000	0.000	0.536	0.112	0.648
	Y_5	0.018	0.042	0.060	0.014	0.018	0.032	0.000	0.052	0.052
4.0	Y_1	0.024	0.046	0.070	0.022	0.024	0.046	0.000	0.066	0.066
	Y_2	0.034	0.014	0.048	0.030	0.012	0.042	0.016	0.016	0.032
	Y_3	0.000	0.994	0.994	0.000	0.984	0.984	0.000	0.996	0.996
	Y_4	0.000	0.168	0.168	0.000	0.000	0.000	0.476	0.168	0.644
	Y_5	0.010	0.056	0.066	0.010	0.040	0.050	0.002	0.056	0.058

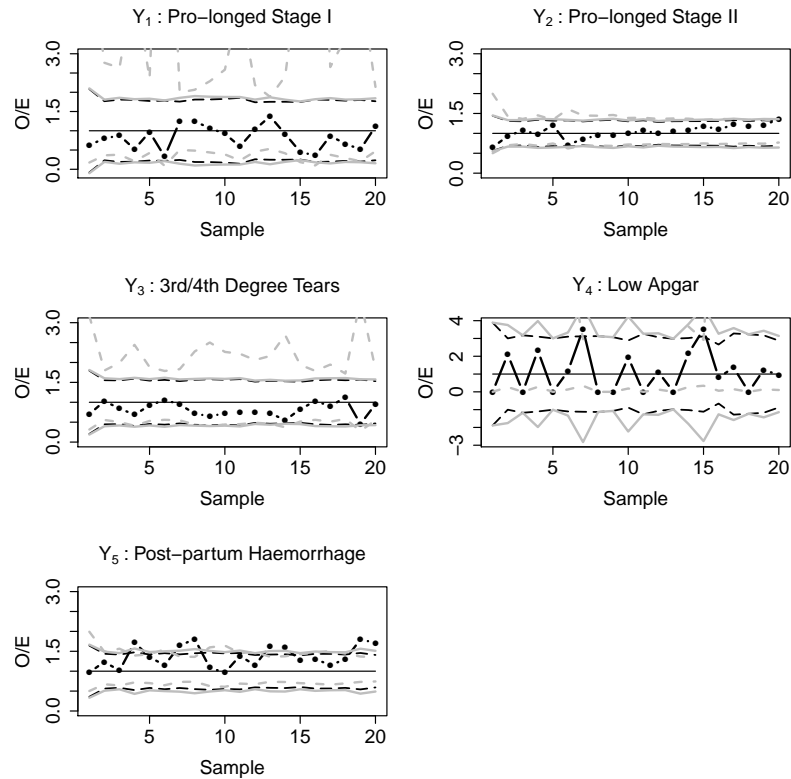


Fig. 5. Multi-stage O/E control charts for deliveries carried out in 2008 at a tertiary maternity unit. Control limits are as follows: (----- Fixed \hat{E}_i), (— Normal R_i), (- . - . - log-Normal R_i)

an average of 196 births a fortnight. The charts obtained are shown in Figure 5.

The MSOE charts in Figure 5 show a marked upward trend in rates of pro-longed stage II labour. Although all points are within the control limits for Y_3 , 80% of them are below the central line indicating lower than expected rates of 3rd or 4th degree tears. The chart for Y_5 indicates higher than expected rates for post-partum haemorrhage. There is no indication that rates for pro-longed stage I labour and low Apgar scores are different to those predicted by the model.

7. Discussion

In this paper, we proposed the use of multi-stage O/E (MSOE) control charts for monitoring outcomes of healthcare procedures. A multi-stage approach tracks outcomes at all stages of the procedure and takes into account the inherent correlation among these outcomes that exists due to variance propagation from upstream to downstream stages. We proposed use of O/E, the ratio of observed to expected outcomes, as the test statistic as this is easier to interpret than the residuals that are typically plotted in multi-stage control charts.

We investigated three options for constructing the control limits. Our simulation results

showed similarity in true detection rates between confidence limits based on a fixed \hat{E} and those based on a normal distribution assumption for the ratio O/E . The normal and log-normal approximations work well when O and \hat{E} are sufficiently large. When they are small (O, E less than about 10), then the use of exact methods is recommended.

We further investigated the true detection rate for various types of changes in outcome rates and found that the true detection rate is dependent on the size and type of change. Changes in the effect of a process variable or risk factor are detected quicker than changes in the population distribution of these variables. For example, a doubling in the effect of X_2 (use of mechanical instruments) on Y_5 was detected about 70% of the time for all three types of confidence limits. On the other hand, when the *number* of deliveries in which mechanical instruments were used doubled, the resulting increased post-partum haemorrhage rate was detected about 60% of the time.

The true detection rate was based on the number of points outside the control limits and are thus the minimum detection rates. Detection rates can be improved by the use of run rules. For example, in Figure 5 the chart for Y_2 shows a steady upward trend, but with none of the points falling outside the control limits. A run rule such as ‘five or more successive points above or below the central line’³⁷ would allow the change to be detected earlier than if the rules were based only on points outside the limits.

When comparison is to be made to historical practice, data from past patients is used to construct logistic regression models for each of the outcome variables. The models constructed at this stage are key to the entire procedure and must be well calibrated and have good predictive ability. In addition, the choice of candidate explanatory variables must be made carefully. As the goal of the monitoring exercise is to inform practice during the procedure, any information that is only available after completion of the procedure should be excluded. For example, an association between high birth weight and the risk of third or fourth degree lacerations has been reported in a number of articles^{38,26}. However, as birthweight is not available to practitioners until after birth, it cannot be included as a predictor in the regression model for the control charts. Ideally, the time period to use for obtaining S_D , the historical data, should be a time when practice and outcomes were at acceptable levels. For example, if a hospital has always had unacceptably high rates of one of the outcome variables, then S_D should ideally consist of births from other hospitals.

If there are no established models identifying relevant explanatory variables to each quality variables, then the model construction stage can be used to test hypotheses about these relationships. Note that even if there are published and accepted models, it is worth going through this testing exercise as the models determined from one institution may not hold true for another institution. For example, one hospital may have an established practice of delivering breech presentations by caesarean section. When assessing whether the type of presentation affects, say the risk of a prolonged labour in such a hospital, the model would indicate an association between breech presentation and *reduced* labour duration. This is counterintuitive and reflects the result of practice rather than an actual causal effect and it would therefore be prudent to exclude breech presentations from model construction. It is likely that a hospital that does not have this practice will see breech presentations associated with prolonged labour. Therefore model construction has to take into account local practice and guidelines.

Acknowledgements

This research is supported by the Victoria University Wellington University Research Fund (93710).

The authour would like to thank Southmead Hospital (Bristol, UK) for making the data available.

References

1. Lovegrove J, Sherlaw-Johnson C, Valencia O, Treasure T, Gallivan S. Monitoring the performance of cardiac surgeons. *Journal of the Operational Research Society*. 1999;50(7):684–689.
2. Steiner S, Cook R, Farewell V. Risk-adjusted monitoring of binary surgical outcomes. *Medical Decision Making*. 2001;21(3):163–169.
3. Grigg O, Farewell V, Spiegelhalter D. Use of risk-adjusted CUSUM and RSPRT charts for monitoring in medical contexts. *Statistical Methods in Medical Research*. 2003;12(2):147–170.
4. Sibanda T, Sibanda N. The CUSUM chart method as a tool for continuous monitoring of clinical outcomes using routinely collected data. *BMC Medical Research Methodology*. 2007;7.
5. 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(8):970–975.
6. Zantek P, Wright G, Plante R. Process and product improvement in manufacturing systems with correlated stages. *Management Science*. 2002;48(5):591–606.
7. Liu J. Variation reduction for multistage manufacturing processes: a comparison survey of statistical-process-control vs stream-of-variation methodologies. *Quality and Reliability Engineering International*. 2010;26(7, SI):645–661. INFORMS Annual Meeting on Recent Advancements in Quality and Reliability, San Diego, CA, 2009.
8. Jearkpaporn D, Montgomery D, Runger G, Borrer C. Model-based process monitoring using robust generalized linear models. *International Journal of Production Research*. 2005;43(7):1337–1354.
9. Li Y, Tsung F. False discovery rate-adjusted charting schemes for multistage process monitoring and fault identification. *Technometrics*. 2009;51(2):186–205.
10. Hawkins D. Multivariate quality-control based on regression-adjusted variables. *Technometrics*. 1991;33(1):61–75.
11. Hawkins D. Regression adjustment for variables in multivariate quality-control. *Journal of Quality Technology*. 1993;25(3):170–182.
12. Lawless J, Mackay R, Robinson J. Analysis of variation transmission in manufacturing processes - Part I. *Journal of Quality Technology*. 1999;31(2):131–142.

13. Mandel BJ. The regression control chart. *Journal of Quality Technology*. 1969;1:1–9.
14. Zhang GX. New type of quality control charts - cause-selecting control charts and a theory of diagnosis with control charts. In: *Proceedings of World Quality Congress Transactions*. Milwaukee, WI: American Society for Quality Control; 1984. p. 175–185.
15. Shu L, Tsung F, Kapur KC. Design of multiple cause-selecting charts for multistage processes with model uncertainty. *Quality Engineering*. 2004;16(3):437–450.
16. Rao S, Strojwas AJ, Lehoczky JP, Schervish MJ. Monitoring multistage integrated circuit fabrication processes. *IEEE Transactions on Semiconductor Manufacturing*. 1996;9(4):495–505.
17. Agrawal R, Lawless J, Mackay R. Analysis of variation transmission in manufacturing processes - Part II. *Journal of Quality Technology*. 1999;31(2):143–154.
18. Skinner K, Montgomery D, Runger G. Process monitoring for multiple count data using generalized linear model-based control charts. *International Journal of Production Research*. 2003;41(6):1167–1180.
19. Jearkpaporn D, Montgomery D, Runger G, Borrer C. Process monitoring for correlated gamma-distributed data using generalized-linear-model-based control charts. *Quality and Reliability Engineering International*. 2003;19(6):477–491.
20. National Collaborating Centre for Women’s & Children’s Health (UK). Intrapartum Care: Care of Healthy Women and Their Babies During Childbirth. London: RCOG Press; 2007. Accessed: 30/01/2013. <http://www.ncbi.nlm.nih.gov/books/NBK49388/>.
21. Apgar V. A proposal for a new method of evaluation of the newborn infant. *Current Researches in Anesthesia & Analgesia*. 1953;32:260–267.
22. Finster M, Wood M. The Apgar score has survived the test of time. *Anesthesiology*. 2005;102(4):855–857.
23. Cameron MJ, Robson SC. Vital statistics: An overview. In: B-Lynch C, Keith LG, Lalonde AB, Karoshi M, editors. *A Textbook of Postpartum Hemorrhage: A Comprehensive Guide to Evaluation, Management and Surgical Intervention*. Dumfriesshire, UK: Sapiens Publishing; 2006. p. 17–34. Available from: http://www.sapienspublishing.com/pph_pdf/PPH.pdf.
24. Thorngren-Jerneck K, Herbst A. Low 5-minute Apgar score: A population-based register study of 1 million term births. *Obstetrics and Gynecology*. 2001;98(1):65–70.
25. Rouse DJ, Weiner SJ, Bloom SL, et al. Second-stage labor duration in nulliparous women: relationship to maternal and perinatal outcomes. *American Journal of Obstetrics and Gynecology*. 2009;201(4):e.1–7.
26. Hirayama F, Koyanagi A, Mori R, Zhang J, Souza JP, Guelmezoglu AM. Prevalence and risk factors for third- and fourth-degree perineal lacerations during vaginal delivery: a multi-country study. *BJOG-An International Journal of Obstetrics and Gynaecology*. 2012;119(3):340–347.

27. Pearl J. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*. 1986;29(3):241–288.
28. Lauritzen S, Spiegelhalter D. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society Series B-Methodological*. 1988;50(2):157–224.
29. Dawid A. Conditional independence in statistical-theory. *Journal of the Royal Statistical Society Series B-Methodological*. 1979;41(1):1–31.
30. Glance LG, Dick A, Osler TM, Li Y, Mukamel DB. Impact of changing the statistical methodology on hospital and surgeon ranking: The case of the New York State cardiac surgery report card. *Medical Care*. 2006;44(4):311–319.
31. Altman DG, Royston JP. The hidden effect of time. *Statistics in Medicine*. 1988;7(6):629–637.
32. Williams SM, Parry BR, Schlup MMT. Quality control: An application of the cusum. *British Medical Journal*. 1992;304(6838):1359–1361.
33. 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(9085):1128–1130.
34. Hosmer D, Lemeshow S. Confidence-interval estimates of an index of quality performance-based on logistic-regression models. *Statistics in Medicine*. 1995;14(19):2161–2172.
35. Talbot D, Duchesne T, Brisson J, Vandale N. Variance estimation and confidence intervals for the standardized mortality ratio with application to the assessment of a cancer screening program. *Statistics in Medicine*. 2011;30(25):3024–3037.
36. Coldman A, Phillips N, Warren L, Kan L. Breast cancer mortality after screening mammography in British Columbia women. *International Journal of Cancer*. 2007;120(5):1076–1080.
37. Montgomery DC. *Introduction to statistical quality control, 6th ed.* Hoboken, N.J.: Wiley; 2009.
38. Landy HJ, Laughon SK, Bailit JL, et al. Characteristics Associated With Severe Perineal and Cervical Lacerations During Vaginal Delivery. *Obstetrics and Gynaecology*. 2011;117(3):627–635.