

# Bayesian Estimation of the Time of a Decrease in Risk-Adjusted Survival Time Control Charts

Hassan Assareh, Kerrie Mengersen

**Abstract**—Change point detection has been recognized as an essential effort of root cause analyses within quality control programs since enables clinical experts to search for potential causes of disturbance in hospital outcomes more effectively. In this paper, we consider estimation of the time when a drop has occurred in the mean survival time observed over patients undergone an in-control cardiac surgery with death and survive outcomes in the presence of variable patient mix. The data are right censored since the monitoring is conducted over a limited follow-up period. The effect of risk factors prior to the surgery is captured using a Weibull accelerated failure time regression model.

We apply Bayesian hierarchical models to formulate the change point. Markov Chain Monte Carlo is used to obtain posterior distributions of the change point parameters including location and magnitude of drops and also corresponding probabilistic intervals and inferences. The performance of the Bayesian estimator is investigated through simulations and the result shows that precise estimates can be obtained when they are used in conjunction with the risk-adjusted survival time CUSUM control charts for different magnitude scenarios. This advantage enhances when probability quantification, flexibility and generalizability of the Bayesian change point detection model are also considered.

**Index Terms**—Bayesian Hierarchical Model, Cardiac Surgery, Change Point, Markov Chain Monte Carlo, Risk-Adjusted Survival Time Control Charts.

## I. INTRODUCTION

A control chart monitors behavior of a process over time by taking into account the stability and dispersion of the process. The chart signals when a significant change has occurred. This signal can then be investigated to identify potential causes of the change and corrective or preventive actions can then be conducted. Following this cycle leads to variation reduction and process stabilization [1].

Risk adjustment has been considered in the development of control charts due to the impact of the human element in process outcomes. Steiner *et al.* [2] developed a Risk-adjusted type of Cumulative Sum control chart (CUSUM) to monitor surgical outcomes, death and survival, which are influenced by the state of a patient's health, age and other factors. This approach has been extended to Exponential Moving Average control charts (EWMA) [3], [4]. Both modified procedures have been intensively reviewed and are now well established for monitoring clinical outcomes where the observations are recorded as binary data [5], [6], [7].

Manuscript received August 20, 2011; revised October 31, 2011. This work was supported by Queensland University of Technology through an ARC Linkage Project.

H. Assareh is with the Discipline of Mathematical Sciences, Faculty of Science & Technology, Queensland University of Technology, Brisbane, Australia; Email: h.assareh@qut.edu.au.

K. Mengersen is with the Discipline of Mathematical sciences, Faculty of Science & Technology, Queensland University of Technology, Brisbane, Australia; Email: k.mengersen@qut.edu.au.

Monitoring patient survival time instead of binary outcomes of a process in presence of patient mix has recently been proposed in the healthcare context. In this setting a continuous time-to-event variable within a follow-up period is considered. The variable may be right censored due to a finite follow-up. Biswas and Kalbfleisch [8] developed a risk-adjusted CUSUM based on Cox model for failure time outcomes. Sego *et al.* [9] used accelerated failure time regression model to capture the heterogeneity among patients prior to the surgery and developed risk-adjusted survival time CUSUM (RAST CUSUM) scheme. Steiner and Jones [10] extended this approach by proposing an EWMA procedure based on the same survival time model discussed by Sego *et al.* [9] that can also be updated.

The need to know the time at which a process began to vary, the so-called change point, has recently been raised and discussed in the context of quality control. Accurate detection of the time of change can help in the search for a potential cause more efficiently as a tighter time-frame prior to the signal in the control charts is investigated. Assareh *et al.* [11] discussed the benefits of change point investigation in monitoring a cardiac surgery process and post-signal root causes analysis by providing precise estimates of the time of the change in the rates of use of blood products during surgery and adverse events in the follow-up period.

Samuel and Pignatiello [12] developed and applied a maximum likelihood estimator (MLE) for the change point in a process fraction nonconformity monitored by a p-chart, assuming that the change type is a step change. They showed how closely this new estimator detects the change point in comparison with the usual p-chart signal. Subsequently, Perry *et al.* [13] compared the performance of the derived MLE estimator with EWMA and CUSUM charts. These authors also constructed a confidence set based on the estimated change point which covers the true process change point with a given level of certainty using a likelihood function based on the method proposed by Box and Cox [14]. This approach was extended to other probability distributions and change type scenarios. In the case of a very low fraction non-conforming, Noorossana *et al.* [15] derived and analyzed the MLE estimator of a step change based on the geometric distribution control charts discussed by Xie *et al.* [16].

All MLE estimators described above were developed assuming that the underlying distribution is stable over time. This assumption cannot often be satisfied in monitoring clinical outcomes as the mean of the process being monitored is highly linked to individual characteristics of patients. Therefore it is required that the survival time model, which explains patient mix, be taken into consideration in detection of true change points in time-to-event control charts.

An interesting approach which has recently been considered in the SPC context is Bayesian hierarchical modelling

(BHM) using, where necessary, computational methods such as Markov Chain Monte Carlo (MCMC). Application of these theoretical and computational frameworks to change point estimation in a clinical context facilitates modelling the process and also provides a way of making a set of inferences based on posterior distributions for the time and the magnitude of a change [17]. This approach has been considered by Assareh *et al.* [11] in change point investigation of two clinical outcomes.

In recent studies Bayesian change point estimators have been developed in monitoring clinical outcomes where the mean of processes are highly linked to individual characteristics of patients. In monitoring outcomes of a surgery, Assareh *et al.* [18], captured the pre-operation risk of death using a logistic regression model. Assareh and Mengersen [19] also proposed this approach for monitoring survival time.

In this paper we model and detect the change point in a Bayesian framework. The change points are estimated assuming that the underlying change is a sudden drop in survival time which can be interpreted as an increase in odds of mortality following a surgical process. In this scenario, we model the step change in the mean of survival time of a clinical process. We analyze and discuss the performance of the Bayesian change point model through posterior estimates and probability based intervals. Risk-adjusted survival time CUSUM charts is reviewed in Section 2. The change point model is demonstrated and evaluated in Sections 3-5. We then summarize the study and obtained results in Section 6.

## II. RISK-ADJUSTED SURVIVAL TIME CONTROL CHARTS

Risk-adjusted control charts for time-to-event are monitoring procedures designed to detect changes in a process parameter of interest, such as survival time, where the process outcomes are affected by covariates, such as risk factors. In these procedures, regression models for time are used to adjust control charts in a way that the effects of covariates for each input, patient say, would be eliminated.

The RAST CUSUM proposed by Segó *et al.* [9] continuously evaluates a hypothesis of an unchanged and in-control survival time distribution,  $f(x_i, \theta_{i0})$ , against an alternative hypothesis of changed, out-of-control, distribution,  $f(x_i, \theta_{i1})$  for the  $i^{th}$  patient. In this setting the density function  $f(\cdot)$  explains the observed survival time,  $x_i$ , and are adjusted corresponds to the observed patient's covariates,  $u_i$ .

For the  $i^{th}$  patient,  $i = 1, 2, \dots$  corresponds to the order in which patient undergoes the surgery, we observe  $(t_i, \delta_i)$  where

$$t_i = \min(x_i, c) \quad \text{and} \quad \delta_i = \begin{cases} 1 & \text{if } x_i \leq c \\ 0 & \text{if } x_i > c. \end{cases} \quad (1)$$

$c$  is a fixed censoring time, equal to the follow-up period. We assume that the survival time,  $x_i$ , for the  $i^{th}$  patient and consequently  $(t_i, \delta_i)$ , are not updated after the follow-up period. This leads to a dataset of right censored time,  $t_i$ .

An accelerated failure time (AFT) regression model is used to predict survival time functions,  $f(\cdot)$ , for each patient in presence of covariates,  $u_i$ . However other models such as a

Cox model that allows to capture covariates can also be of interest.

In an AFT model the survival function for the  $i^{th}$  patient with covariates of  $u_i$ ,  $S(x_i, \theta_i | u_i)$ , is equivalent to the baseline survival function  $S_0(x_i \exp(\beta^T u_i))$ , where  $\beta$  is a vector of covariate coefficients.

Several distributions can be used to model survival time with an AFT, among those we here focus on the Weibull distribution and outline relevant RAST CUSUM statistics, see Klein & Moeschberger [20] for more details. For a Weibull distribution the baseline survival function is  $S_0(x) = \exp[-(x/\lambda)^\alpha]$  where  $\alpha > 0$  and  $\lambda > 0$  are shape and scale parameters, respectively. For the RAST CUSUM procedure, all parameters of the Weibull survival function,  $\beta, \alpha$  and  $\lambda$ , are estimated using training data, so-called phase I, assuming that the process is in-control.

It has been discussed that any shifts in the quality of the process of the interest can be interpreted in terms of shifts in the scale parameters,  $\lambda$ , see Segó *et al.* [9].

Hence the RAST CUSUM procedure can be constructed and calibrated to detect a drop in the average or median survival time (MST) since any shift in  $\lambda$  is equivalent to an identical shift in size in the average or median survival time. Thus the CUSUM score,  $W_i$ , is given by

$$W_i(t_i, \delta_i | u_i) = (1 - (\rho)^{-\alpha}) \left( \frac{t_i \exp(\beta^T u_i)}{\lambda_0} \right) - \delta_i \alpha \log \rho, \quad (2)$$

where it is designed to detect a decrease from  $\lambda_0$  to  $\lambda_1 = \rho \lambda_0$ . Upper CUSUM statistic is obtained through  $Z_i = \max\{0, Z_{i-1} + W_i\}$ , and then plotted over  $i$ . Often the CUSUM statistic initialized at 0.

Therefore a reduction in the the MST is detected when a plotted  $Z_i$  exceeds a specified decision threshold  $h$ . Although this interpretation of chart's signal is in contrast with the common expression used for standard risk-adjusted control charts for binary outcomes, it seems reasonable to take into account that any drop in the MST can be characterized as an increase in the odds of mortality. However in Weibull distribution scenario for a specific drop in the MST, the equivalent magnitude of the increase in odds is not obtainable, see Segó *et al.* [9] for more details.

The magnitude of the decision thresholds in RAST CUSUM,  $h$ , is determined in a way that the charts have a specified performance in terms of false alarm and detection of shifts in the MST. In this regard, Markov chain and simulation approaches can be applied; see Segó [21] for more details.

## III. CHANGE POINT MODEL

Statistical inferences for a quantity of interest in a Bayesian framework are described as the modification of the uncertainty about their value in the light of evidence, and Bayes' theorem precisely specifies how this modification should be made as below:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}, \quad (3)$$

where "Prior" is the state of knowledge about the quantity of interest in terms of a probability distribution before data are observed; "Likelihood" is a model underlying the

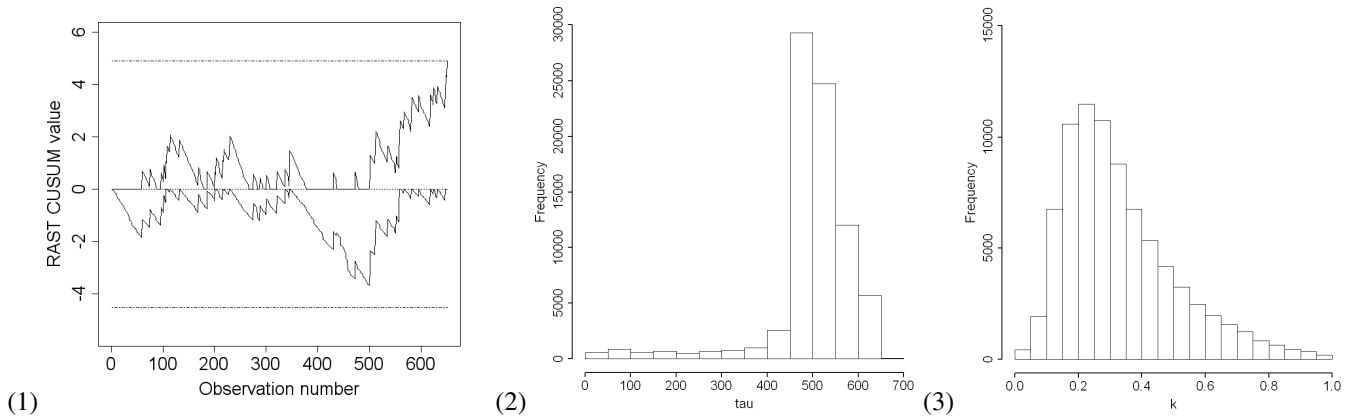


Fig. 1. (1) Risk-adjusted survival time CUSUM chart ( $h = 4.88$ ) and obtained posterior distributions of (2) time  $\tau$  and (3) magnitude  $k$  of a decrease of size  $k = 0.25$  in  $\lambda$  (mean survival time) where  $\lambda_0 = 42133.6$  and  $\tau = 500$ .

observations, and “Posterior” is the state of knowledge about the quantity after data are observed, which also is in the form of a probability distribution.

As discussed in Section I, in RAST CUSUM procedures, we let the survival function vary over patients and we control the observed survival time, which may be right censored, against the corresponding predicted survival function obtained through the survival time model. In this setting, a process is in the in-control state when observations can be statistically expressed by the underlying survival time model, taking into account their individual covariates. The RAST CUSUM signals when observations tend to violate the underlying model.

To model a change point in the presence of covariates, consider a process that results a survival time of  $t_i$ ,  $i = 1, \dots, T$ , that is initially in-control. The observations can be explained by a survival function  $S(t_i, u_i)$ , where the underlying distribution ( $f(\cdot)$ ) is a Weibull distribution with  $(\alpha_0, \lambda_0)$ , and  $u_i$  is a vectors of covariates. At an unknown point in time,  $\tau$ , the Weibull scale parameter changes from its in-control state of  $\lambda_0$  to  $\lambda_1$ ,  $\lambda_1 = k \times \lambda_0, 0 < k < 1$ . The right censored survival time step change model can thus be parameterized using survival function as follows:

$$S(t_i, u_i) = \begin{cases} \exp \left[ - \left( \frac{t_i \exp(\beta_0^T u_i)}{\lambda_0} \right)^{\alpha_0} \right] & \text{if } i = 1, 2, \dots, \tau \\ \exp \left[ - \left( \frac{t_i \exp(\beta_0^T u_i)}{\lambda_1} \right)^{\alpha_0} \right] & \text{if } i = \tau + 1, \dots, T \end{cases}$$

where  $\beta_0$  is the vector of covariate coefficients.

If desired, an overall estimation of change size in odds of mortality equivalent to a specific shift in the MST or  $\lambda$  can be obtained through simulation and averaging over different values of covariate,  $u_i$ .

Relating this to Equation (3), the likelihood that underlies the observations is obtained through  $f(\cdot)^\delta S(\cdot)^{1-\delta}$ ; see Sego *et al.* [9]. The time and the magnitude of a drop in the MST are the unknown parameters of interest; and the posterior distributions of these parameters will be investigated in the change point analysis.

Assume that the process  $t_i$  is monitored by a control chart that signals at time  $T$ . We assign a truncated normal distribution  $(\mu, \sigma)I(\cdot)$  for  $k$  as prior distribution where all parameters are set study-specific. For a decrease in  $k$  which is detected by the upper RAST CUSUM, exceeding the upper threshold  $h$ , we set  $N(\mu = 0.255, \sigma = 0.6)I(0.01, 0.99)$ .

This setting leads to relatively an informed prior for the magnitude of the fall.

Mean of the prior was set corresponds to the shift that the chart was calibrated to detect, see Section IV. The prior let to be sensitive in detection of low to nearly large falls in  $k$ . Note that other distributions such as uniform and Gamma might also be of interest for  $k$  since it is always a positive value; see Gelman *et al.* [17] for more details on selection of prior distributions.

We place a uniform distribution on the range of  $(1, T - 1)$  as prior for  $\tau$  where  $T$  is set to the time of the signal of control charts. See the Appendix for the step change model code in WinBUGS.

#### IV. EVALUATION

We used Monte Carlo simulation to study the performance of the constructed BHM in step change detection following a signal from a RAST CUSUM control chart when a fall in mean survival time is simulated to occur at  $\tau = 500$ . However, to extend to the results that would be obtained in practice, we considered the same cardiac surgery dataset that were used by Steiner *et al.* [2] and then Sego *et al.* [9] to construct risk-adjusted control charts for Bernoulli and time-to event variables, respectively. It was reported (4) that this dataset contains 6449 operations information that were performed between 1992-1998 at a single surgical center in U.K. The Parsonnet score [22] was recorded to quantify the patient’s risk prior to the cardiac surgery. A follow-up period of 30 days after the surgery was set as the censoring time. A Weibull AFT model with parameters of  $\hat{\alpha}_0 = 0.4909, \hat{\lambda}_0 = 42133.6$  and  $\hat{\beta}_0 = 0.1307$  was reported by Sego *et al.* [9] when the first two years of the data set were used as training data to fit the model and construct in-control state of the process and RAST CUSWUM. They also found that the recorded Parsonnet scores of the training data can be well approximated by an exponential distribution with a mean of 8.9.

To generate right censored survival time observations of a process in the in-control state  $t_i, i = 1, \dots, \tau$ , we first randomly generated the Parsonnet score,  $u_i, i = 1, \dots, \tau$ , from an exponential distribution with a mean of 8.9 and then drew associated survival time,  $x_i, i = 1, \dots, \tau$ , from the Weibull AFT model with

$\alpha_0 = 0.4909, \lambda_0 = 42133.6,$  and  $\beta_0 = 0.1307$ . Finally,  $t_i$  and  $\delta_i$  were obtained considering a censoring time of  $c = 30$  through Equation 1. Plotting the obtained observations when the associated covariates are considered results a RAST CUSUM chart that is in-control. To generate the drops in  $\lambda_0$ , or MST, we then induced changes of sizes  $k = \{0.05, 0.066, 0.1, 0.143, 0.20, 0.25, 0.33, 0.50, 0.66, 0.75\}$  and generated observations until the control charts signalled. These changes led to different change sizes in in-control estimated survival probability over days for a patient with  $u_i$  as well as survival curves between patients with different Parsonnet scores. Note that other distributions such as uniform distributions with proper parameters or even sampling randomly from the baseline Parsonnet scores can be applied to generate covariates directly.

To construct RAST CUSUM, we applied the procedures discussed in Section II. We calibrated RAST CUSUM to detect a decrease in the MST that correspond to a doubling of the odds ratio within the follow-up period and has an in-control average run length ( $ARL_0$ ) of approximately 10000 observations. As mentioned in Section II, for the Weibull AFT model the corresponding odds ratio formula, discussed by Sego *et al.* [9], is not reduced to a close form of  $\lambda_0$  and  $\rho$  since the covariate term is not simplified. in

$$OR = \frac{O_{i1}}{O_{i0}}, \quad \text{and} \quad O_i = \frac{1 - S(c | u_i)}{S(c | u_i)} \quad (5)$$

where  $S(c | u_i)$  is the probability of survival at the end of follow-up period,  $c$ .

Therefore we used Monte Carlo simulation to estimate corresponding  $\rho$ . To do so, we set  $\rho$  such over 100,000 replications of generating Parsonnet scores from the fitted exponential distribution with a mean of 8.9 and calculating the odds in Equation 5, the desired odds ratio of size  $OR = 2$  was obtained. A decrease of  $\hat{\rho} = 0.255$  in the MST was found correspond to the desired jump in odds ratio.

We also used Monte Carlo simulation to determine the decision interval,  $h$ . However other approaches may be of interest; see Steiner *et al.* [2] and Sego *et al.* [9]. This setting led to a decision interval of  $h = 4.88$ . The associated CUSUM scores were also obtained through Equation (2) considering the generated  $t_i, \delta_i$  and  $u_i$ .

The step change and control charts were simulated in the R package (<http://www.r-project.org>). To obtain posterior distributions of the time and the magnitude of the changes we used the R2WinBUGS interface [23] to generate 100,000 samples through MCMC iterations in WinBUGS [24] for all change point scenarios with the first 20000 samples ignored as burn-in. We then analyzed the results using the CODA package in R [25]. See the Appendix for the step change model code in WinBUGS.

V. PERFORMANCE ANALYSIS

To demonstrate the achievable results of Bayesian change point detection in risk-adjusted control charts, we induced a drop of size  $k = 0.25$  at time  $\tau = 500$  in an in-control process with an overall survival time of  $\lambda_0 = 42133.6$ . RAST CUSUM detected the drop and signalled at the 651<sup>st</sup> observation, corresponding to a delay of 151 observations as shown in Figure 1-1. The posterior distributions of time and magnitude of the change were then obtained using

TABLE I  
POSTERIOR ESTIMATES (MODE, SD.) OF STEP CHANGE POINT MODEL PARAMETERS ( $\tau$  AND  $k$ ) FOLLOWING SIGNALS (RL) FROM RAST CUSUM ( $h = 4.88$ ) WHERE  $\lambda_0 = 42133.6$  AND  $\tau = 500$ .

$k$	RL	$\hat{\tau}$	$\hat{\sigma}_{\hat{\tau}}$	$\hat{k}$	$\hat{\sigma}_{\hat{k}}$
0.25	651	499.8	96.0	0.226	0.18
0.33	722	494.8	160.6	0.27	0.19

MCMC discussed in Section IV. The distribution of the time of the change,  $\tau$ , concentrates on the 500<sup>th</sup> observation, approximately, as seen in Figure 1-2. The posterior for the magnitude of the change,  $k$ , also reasonably identified the exact change size as it highly concentrates on values of around 0.25 shown in Figure 1-3.

This investigation was replicated using a smaller shift of size  $k = 0.33$  in  $\lambda$ . Table I summarizes the posterior estimates for both change sizes. If the posterior was asymmetric and skewed, the mode of the posterior was used as an estimator for the change point model parameters ( $\tau$  and  $k$ ).

The RAST CUSUM signalled after 222 observations when the mean survival time became a third whereas the posterior distribution reported a drop at the 491<sup>st</sup> observation. This result implies that although the obtained posterior estimates underestimated the change point, they still performed significantly better than the RAST CUSUM charts.

Bayesian estimates of the magnitude of the change tend to be relatively accurate following signals of the control chart, see Figure 1-3 and Table I. The slight bias, here underestimation, observed in the figures must be considered in the context of their corresponding standard deviations.

Comparison of estimates obtained across change sizes reveals that although a shorter run of observations from the out-of control state of the process is used when a larger shift size occurred, less dispersed posteriors are obtained, particularly for posteriors of time.

Applying the Bayesian framework enables us to construct probability based intervals around estimated parameters. A credible interval (CI) is a posterior probability based interval which involves those values of highest probability in the posterior density of the parameter of interest. Table II presents 50% and 80% credible intervals for the estimated time and the magnitude of changes in  $\lambda_0$  for the RAST CUSUM chart. As expected, the CIs are affected by the dispersion and higher order behaviour of the posterior distributions. Under the same probability of 0.5, the CI for the time of the change of size  $k = 0.25$  covers 63 observations around the 500<sup>th</sup> observation whereas it increases and reaches to 161 observations for  $k = 0.33$  due to the larger standard deviation, see Table I. This investigation can be extended to other shift sizes for the time estimates. As shown in Table I and discussed above, the magnitude of the changes are

TABLE II  
CREDIBLE INTERVALS FOR STEP CHANGE POINT MODEL PARAMETERS ( $\tau$  AND  $k$ ) FOLLOWING SIGNALS (RL) FROM RAST CUSUM ( $h = 4.88$ ) WHERE  $\lambda_0 = 42133.6$  AND  $\tau = 500$ .

$k$	CI 50%		CI 80%	
	$\hat{\tau}$	$\hat{k}$	$\hat{\tau}$	$\hat{k}$
0.25	(488, 551)	(0.14, 0.33)	(453, 581)	(0.09, 0.48)
0.33	(487, 648)	(0.15, 0.40)	(359, 709)	(0.09, 0.57)

TABLE III  
 PROBABILITY OF THE OCCURRENCE OF THE CHANGE POINT IN THE LAST {25, 50, 100, 200, 300, 400, 500} OBSERVATIONS PRIOR TO SIGNALLING FOR RAST CUSUM ( $h = 4.88$ ) WHERE  $\lambda_0 = 42133.6$  AND  $\tau = 500$ .

$k$	25	50	100	200	300	400	500
0.25	0.03	0.07	0.20	0.89	0.94	0.96	0.97
0.33	0.01	0.05	0.20	0.59	0.77	0.82	0.90

also estimated reasonably well and Table II shows that in all cases the real sizes of changes are contained in the respective posterior 50% and 80% CIs.

Having a distribution for the time of the change enables us to make other probabilistic inferences. As an example, Table III shows the probability of the occurrence of the change point in the last {25, 50, 100, 200, 300, 400, 500} observations prior to signalling in the control charts. For a step change of size  $k = 0.33$  in the mean survival time, since the RAST CUSUM signals late (see Table I), it is unlikely that the change point occurred in the last 100 observations. A considerable growth in the probability is seen when the next 200 observations are included, reaching to 0.77, whereas for a larger drop of size  $k = 0.25$ , it is more certain that the change point has occurred in the last 200 observations with a probability of 0.89.

The above studies were based on a single sample drawn from the underlying distribution. To investigate the behavior of the Bayesian estimator over different sample datasets, for different reduction in  $\lambda_0$ , we replicated the simulation method explained in Section IV 100 times.

Table IV shows the average of the estimated parameters obtained from the replicated datasets where there exists a drop in  $\lambda_0$  of size  $k$ .

As seen, the RAST CUSUM control chart tends to detect larger shifts in the MST with less delays. For a large drop, a  $k$  of size 0.143 and less, the chart signals with a delay of at most 95 observations. This delay increases over moderate reductions in  $\lambda_0$ , reaching to 279 observations for  $k = 0.33$ . However, the chart is failed in detection of small drops since signals with a long delay of more than 639 observations obtained when the MST halved,  $k = 0.50$ .

For large drops in the MST, a  $k$  of size 0.143 and less, the average values of the modes,  $\hat{E}(\hat{\tau})$ , tends to underestimate the time of the change since it reports at best the 490<sup>th</sup> observation for  $k = 0.066$ . However, the Bayesian estimator still outperforms the chart signal with a less bias over large reductions. This superiority persists for moderate shifts in the MST, where a less bias is still associated with the Bayesian estimates of the time,  $\tau$ , at best three observations obtained for  $k = 0.20$ . Although the RAST CUSUM chart was designed to detect a moderate drop of 0.255 in the MST, it is outperformed by the posterior mode that detects the change point with a delay of 27 observations.

Table IV shows that the bias of the Bayesian estimator,  $\hat{E}(\hat{\tau})$ , did not exceed 55 observations over moderate reductions. This bias increased when the MST halved, reaching to 162 observations, yet significantly outperformed the chart's signal. For smaller reductions,  $k = (0.66, 0.75)$ , the posterior modes significantly overestimate the change point since the RAST CUSUM signals very late. The variation of the Bayesian estimates for time tends to reduce when the magnitude of shift in the MST increases. The mean of the standard deviation of the posterior estimates of time,  $\hat{E}(\sigma_{\hat{\tau}})$ , also decreases when shift sizes increase.

Table IV indicates that the average of the Bayesian estimator of the magnitude of the change,  $\hat{E}(\hat{\delta})$ , identifies change sizes with some biases. This estimator tends to overestimate and underestimate the sizes where there exist large drops and moderate to small drops, respectively. Having said that, Bayesian estimates of the magnitude of the change must be studied in conjunction with their corresponding standard deviations. In this manner, analysis of credible intervals is

TABLE IV  
 AVERAGE OF POSTERIOR ESTIMATES (MODE, SD.) OF STEP CHANGE POINT MODEL PARAMETERS ( $\tau$  AND  $k$ ) FOR A CHANGE IN THE MEAN SURVIVAL TIME FOLLOWING SIGNALS (RL) FROM RAST CUSUM ( $h = 4.88$ ) WHERE  $\lambda_0 = 42133.6$  AND  $\tau = 500$ . STANDARD DEVIATIONS ARE SHOWN IN PARENTHESES.

$k$	$\hat{E}(RL)$	Change point		Change size	
		$\hat{E}(\hat{\tau})$	$\hat{E}(\sigma_{\hat{\tau}})$	$\hat{E}(k)$	$\hat{E}(\sigma_{\hat{k}})$
0.05	542.4	486.0	91.2	0.077	0.173
	(16.2)	(57.3)	(34.7)	(0.086)	(0.022)
0.066	554.8	490.5	92.9	0.083	0.177
	(26.6)	(62.5)	(36.7)	(0.075)	(0.025)
0.10	568.3	485.7	99.4	0.127	0.183
	(39.7)	(70.9)	(33.9)	(0.094)	(0.017)
0.143	594.2	487.3	110.9	0.154	0.185
	(49.2)	(72.5)	(34.5)	(0.090)	(0.016)
0.20	624.7	503.7	119.5	0.182	0.183
	(71.3)	(87.1)	(36.6)	(0.103)	(0.018)
0.25	692.3	527.3	132.9	0.211	0.183
	(150.4)	(146.2)	(53.4)	(0.111)	(0.018)
0.33	779.6	554.3	153.9	0.25	0.176
	(187.7)	(162.3)	(58.9)	(0.118)	(0.023)
0.50	1139.0	661.8	258.9	0.43	0.178
	(605.0)	(287.7)	(173.0)	(0.16)	(0.028)
0.66	2469.4	1270.3	562.1	0.51	0.183
	(2169.8)	(783.2)	(456.6)	(0.22)	(0.047)
0.75	2773.4	1748.0	697.9	0.53	0.195
	(2195.4)	(1304.4)	(720.8)	(0.25)	(0.047)

effective.

## VI. CONCLUSION

Quality improvement programs and monitoring of medical process outcomes are now being widely implemented in the health context to achieve stability in outcomes through detection of shifts and investigation of potential causes. Obtaining accurate information about the time when a change occurred in the process has been recently considered within industrial and business quality control applications. Indeed, knowing the change point enhances efficiency of root cause analysis efforts by restricting the search to a tighter window of observations and related variables.

In this paper, using a Bayesian framework, we modeled change point detection in time-to-event data for a clinical process with dichotomous outcomes, death and survival, where patient mix was present. We considered a drop in the mean of survival time of an in-control process. We constructed Bayesian hierarchical models and derived posterior distributions for change point estimates using MCMC. The performance of the Bayesian estimators were investigated through simulation when they were used in conjunction with risk-adjusted survival time CUSUM control charts monitoring right censored survival time of patients who underwent cardiac surgery procedures within a follow-up period of 30 days where the severity of risk factors prior to the surgery was evaluated by the Parsonnet score. The results showed that the Bayesian estimates significantly outperform the RAST CUSUM control charts in change detection over different magnitude of drops in the mean survival time.

Apart from accuracy and precision criteria used for the comparison study, the posterior distributions for the time and the magnitude of a change enable us to construct probabilistic intervals around estimates and probabilistic inferences about the location of the change point. This is a significant advantage of the proposed Bayesian approach. Furthermore, flexibility of Bayesian hierarchical models, ease of extension to more complicated change scenarios such as linear and non-linear trends in survival time, relief of analytic calculation of likelihood function, particularly for non-tractable likelihood functions and ease of coding with available packages should be considered as additional benefits of the proposed Bayesian change point model for monitoring purposes.

The investigation conducted in this study was based on a specific in-control rate of mortality observed in the pilot hospital. Although it is expected that superiority of the proposed Bayesian estimator persists over other processes in which the in-control rate and the distribution of baseline risk may differ, the results obtained for estimators and control charts over various change scenarios motivates replication of the study using other patient mix profiles. Moreover modification of change point model elements such as replacing priors with more informative alternatives may be of interest.

The two-step approach to change-point identification described in this paper has the advantage of building on control charts that may be already in place in practice (as in the pilot hospital). An alternative may be to retain the two-step approach but to use a Bayesian framework in both stages. There is now a substantial literature on Bayesian formulation of control charts and extensions such as monitoring processes

with varying parameters [26], over-dispersed data [27], start-up and short runs [28], [29]. A further alternative is to consider a fully Bayesian, one-step approach, in which both the monitoring of the in-control process and the retrospective or prospective identification of changes is undertaken in the one analysis. This is the subject of further research.

## APPENDIX A

### CHANGE POINT MODEL CODE IN WINBUGS

```

model {
  for(i in 1 : RLcusum)
  {
    y[i] ~ dweib(alpha0, gamma[i])I(yc[i],)
    gamma[i] = pow(exp(beta0 * riskscore[i])/
      (lambda0+step(i-tau)
      * lambda0 * (k-1)), alpha)
  }
  RL=RLcusum-1
  k ~ dnorm(0.255, 2.77)I(0.01, 0.99)
  tau ~ dunif(1, RL) }
    
```

## REFERENCES

- [1] Montgomery D.C., *Introduction to Statistical Quality Control* (Sixth ed.): Wiley, 2008.
- [2] Steiner, S.H., Cook, R.J., Farewell, V.T., & Treasure, T., "Monitoring surgical performance using risk-adjusted cumulative sum charts", *Biostatistics*, V1, N4, pp. 441-452, 2000.
- [3] Cook, D., "The Development of Risk Adjusted Control Charts and Machine learning Models to monitor the Mortality of Intensive Care Unit Patients", *Ph.D. Thesis*; University of Queensland, 2004.
- [4] Grigg, O.V. & Spiegelhalter, D.J., "A Simple Risk-Adjusted Exponentially Weighted Moving Average", *Journal of the American Statistical Association*, V102, N477, pp. 140-152, 2007.
- [5] Grigg, O. V., & Farewell, V. T., "An overview of risk-adjusted charts", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, V167, N3, pp. 523-539, 2004.
- [6] Grigg, O. V., & Spiegelhalter D. J., "Discussion", *Journal of Quality Technology*, V38, N2, pp. 124-136, 2006.
- [7] Cook, D. A., Duke, G., Hart, G. K., Pilcher, D., Mullany, D., "Review of the application of risk-adjusted charts to analyse mortality outcomes in critical care", *Critical Care Resuscitation*, V10, N3, pp. 239-251, 2008.
- [8] Biswas, P. & Kalbfleisch, J.D., "A risk-adjusted CUSUM in continuous time based on the Cox model," *Statistics in Medicine*, V 27, N17, pp. 3382-3406, 2008.
- [9] Sego, L.H., Reynolds, Jr.M.R., Woodall, W.H., "Risk adjusted monitoring of survival times", *Statistics in Medicine*, V28, N9, pp. 1386-1401, 2009.
- [10] Steiner, S.H. & Jones, M., "Risk-adjusted survival time monitoring with an updating exponentially weighted moving average (EWMA) control chart", *Statistics in Medicine*, V29, N4, pp. 444-454, 2010.
- [11] Assareh, H., Smith, I., Mengersen, K., "Bayesian change point detection in monitoring cardiac surgery outcomes", *Quality Management in health Care*, V20, N3, pp. 207-222, 2011.
- [12] Samuel, T. & Pignatjello, J., "Identifying the time of a step change in the process fraction nonconforming", *Quality Engineering*, V13, N3, pp. 357-365, 2001.
- [13] Perry, M. & Pignatiello, J., "Estimating the change point of the process fraction nonconforming in SPC applications", *International Journal of Reliability, Quality and Safety Engineering*, V12, N2, pp. 95-110, 2005.
- [14] Box, G. & Cox, D., "An analysis of transformations", *Journal of the Royal Statistical Society. Series B (Methodological)*, V26, N2, pp. 211-252, 1964.
- [15] Noorossana, R., Saghafi, A., Paynabar, K. & Abdi, S., "Identifying the period of a step change in high-yield processes", *Quality and Reliability Engineering International*, V25, N7, pp. 875-883, 2009.
- [16] Xie, M., Goh, T. & Kuralmani, V., *Statistical Models and Control Charts for High-Quality Processes*: Kluwer Academic Publishers, 2002.
- [17] Gelman, A., Carlin, J., Stern, H. & Rubin, D., *Bayesian Data Analysis*: Chapman & Hall/CRC, 2004.

- [18] Assareh, H., Smith, I. & Mengersen, K., "Identifying the time of a linear trend disturbance in odds ratio of clinical outcomes", *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2011*, V1, WCE 2011, 6-8 July, London, U.K., pp. 365-370, 2011.
- [19] Assareh, H. & Mengersen, K., "Detection of the time of a step change in monitoring survival time", *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2011*, V1, WCE 2011, 6-8 July, London, U.K., pp. 314-319, 2011.
- [20] Klein, J.P. & Moeschberger, M.L., *Survival Analysis: Techniques for Censored and Truncated Data*. Springer: New York, 1997.
- [21] Sego, L.H., "Applications of control charts in medicine and epidemiology", *PhD Thesis*, United States-Virginia, Virginia Polytechnic Institute and State University, 2006.
- [22] Parsonnet, V., Dean, D., Bernstein, A.D., "A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease", *Circulation* V79, 6 Pt 2, pp. I3-12, 1989.
- [23] Sturtz, S., Ligges, U., & Gelman, A., "R2WinBUGS: A package for running WinBUGS from R", *Journal of Statistical Software*, V12, N3, pp. 1-16, 2005.
- [24] Spiegelhalter, D., Thomas, A. & Best, N., *WinBUGS version 1.4. Bayesian inference using Gibbs sampling*. Manual. MRC Biostatistics Unit. Institute for Public Health, Cambridge, United Kingdom, 2003.
- [25] Plummer, M., Best, N., Cowles, K. & Vines, K., Output analysis and diagnostics for MCMC. R package version 0.10-3, URL <http://cran.r-project.org>, 2005.
- [26] Feltz, C. J. & Shiau, J.-J. H., "Statistical process monitoring using an empirical Bayes multivariate process control chart", *Quality and Reliability Engineering International*, V17, N2, pp. 119-124, 2001.
- [27] Bayarri, M. J. & Garcia-Donato, G., "A Bayesian Sequential Look at u-Control Charts", *Technometrics*, V47, N2, pp. 142-151, 2005.
- [28] Tsiamyrtzis, P. & Hawkins, D. M., "A Bayesian Scheme to Detect Changes in the Mean of a Short-Run Process", *Technometrics*, V47, N4, 446-456, 2005.
- [29] Tsiamyrtzis, P. & Hawkins, D. M., "A Bayesian EWMA method to detect jumps at the start-up phase of a process", *Quality and Reliability Engineering International*, V24, N6, 721-735, 2008.