

**RESEARCH ARTICLE****The STRAND Chart: A survival time control chart**Olivia AJ Grigg MRC Biostatistics Unit, University of  
Cambridge, Cambridge, UK**Correspondence**Olivia AJ Grigg, MRC Biostatistics Unit,  
Institute of Public Health, Cambridge CB2  
0SR, UK.  
Email: mathsworkmusic@gmail.com

The STRAND Chart (Survival Time, Risk-Adjusted, N-Division Chart) is a new tool for online risk-adjusted (RA) monitoring of survival outcomes. The chart is drawn in continuous time, making it responsive to change in the process of interest, for example, performance over time of a surgical unit and the procedures that they employ. Though it is difficult to achieve with charts designed for the purpose described, we show that our suggested chart keeps patient ordering intact. We discuss the difficulties maintaining patient ordering poses, making reference to other charts in the literature. Our conclusion is that the best approach to preserving patient ordering on any chart of this nature involves compromising on the fullness of presentation of the recorded data. The chart is divided into  $N$  strands, each strand relating to a benchmark patient's survival information at  $t_n$  days following treatment,  $n = 1, 2, \dots, N$ . We present a simple version of the chart where the strands consist of Bernoulli RA exponentially weighted moving averages, recording RA failure rates at  $t_n$  days. It can detect recent process change *and* historical change. We illustrate the STRAND Chart using a well-known UK post cardiac surgery survival dataset, where the nature of a certain cluster in the data can be seen.

**KEYWORDS**

control chart, performance monitoring, quality control, risk adjustment, survival

**1 | INTRODUCTION**

Monitoring of patient outcomes over time is prolific in the field of medicine, most especially in the form of sequential clinical trials.<sup>1-3</sup> Outside of the clinical trials setting, there has been increasing interest in assessing quality of care provided, where there are potential underlying changes in performance and also planned changes in care that are not formally assessed by clinical trial (such as periodic retraining of surgeons).<sup>4-7</sup> While quality control and the use of control charts for process monitoring has its roots in the industrial setting where monitored items, for example, electrical components, are homogeneous in nature,<sup>8,9</sup> in the medical context, the “items” monitored (patients) are heterogeneous in nature.<sup>10-12</sup> Patients have different underlying characteristics and morbidities that, though they affect outcomes, are beyond the control of the medical practitioner or unit and hence need to be adjusted for. This aspect of monitoring in medical contexts has given rise to a growing body of work on risk-adjusted (RA) monitoring with control charts.<sup>10-20</sup>

In the particular area of RA monitoring of survival data (akin to reliability or lifetime data in the industrial setting),<sup>21,22</sup> development of methodology has moved from monitoring in discrete time to monitoring in near-continuous time, at finer timescales. Where more basic methods record outcomes after some fixed period of time (30 days say) has elapsed, methods

developed more recently react to survival events as and when they occur.<sup>10,14-16</sup> Though they are not strictly in continuous time, we will henceforth refer to near-continuous methods as continuous.

All of these methods fall broadly into two types: those that use sequential log-likelihood ratio (LLR) statistics and those that use smoothed estimators of some summary statistic of survival such as survival beyond 30 days, mean or median survival. The first type are cumulative sum (CUSUM) charts<sup>23</sup> that are cumulatively summed LLR statistics of a null process state versus an “out-of-control” process state, where an out-of-control process may have an elevated median survival for example. The second type are charts that smoothly estimate some summary statistic of survival, an example being exponentially weighted moving average (EWMA) charts.<sup>24,25</sup> These EWMA charts can be constructed from a Classical standpoint or from a Bayesian standpoint, where the former treats the chart measure as a statistic with sampling error and the latter treats it as a posterior estimate for a randomly wandering underlying process parameter.<sup>11,17</sup> Where Classical EWMA charts and CUSUM charts are set up using operating characteristics such as in- and out-of-control average run lengths to a signal of process change,<sup>26,27</sup> Bayesian EWMA charts would focus on quality of estimation and hence on width and placement of credible intervals.<sup>11,17,28-34</sup> The width of the credible intervals about the chart statistic determines how often the chart will signal and the placement determines the balance between signals of improvement and signals of deterioration in the process. As with Classical charts, Bayesian EWMA charts and similar Bayesian time series charts would be tuned on test data before being applied to online data.<sup>35</sup> Tuning the chart to the pilot data can be viewed as an exercise in analysis of sensitivity to changes in chart parameters.

The charts are based on varying degrees of parameterisation of the survival curve. While all methods use a fully parametric risk model, some methods specify the survival curve fully and others specify it only in part. We refer to the latter as semiparametric charts in this context.

In Section 2, we review existing charts in the literature designed to monitor RA survival outcomes. In Section 3, we outline the details of our chart, and in Section 3.1, we apply the chart to the now classic post cardiac surgery survival data from a UK hospital, 1992-1997.<sup>11,12,14,15,17,18</sup> In Section 4, we compare the performance of our chart to that of the best competing chart in the literature.

## 2 | REVIEW OF RISK-ADJUSTED SURVIVAL TIME CONTROL CHARTS

### 2.1 | The RA CUSUM and RAST CUSUM

These methods capture different aspects of the survival process being monitored. The RA Bernoulli CUSUM of Steiner et al accumulates information about the measure  $\Pr(X \leq d)$  number of days) where  $X$  is a dynamic censored survival process.<sup>12</sup> The information is gathered via a cumulatively summed (CUSUM) LLR statistic of a null risk model versus an alternative model where the failure rate before  $d$  number of days is assumed to be higher (or, alternatively, lower) across all patients. An LLR CUSUM can be written as

$$C_{t+1} = C_t + LLR_t, \quad t = 1, 2, \dots, \quad (1)$$

where  $t$  indexes arriving data.

The risk-adjusted survival time (RAST) CUSUM of Sego et al similarly accumulates an LLR CUSUM statistic of a null versus an alternative hypothesis,<sup>15</sup> but the likelihood used reflects the whole parametric survival model  $f_X(x)$  as opposed to a prespecified tail of the survival distribution  $\Pr(X \leq d)$ . Hence, the RAST CUSUM of Sego et al is fully parametric, whereas the RA Bernoulli CUSUM of Steiner et al is semiparametric.

The charts, with their differing specification of the likelihood, are presented as discrete time methods. They are updated after a fixed period of time,  $d$  days, has elapsed following treatment. They have been presented in application to the now classic UK cardiac surgery dataset of 1992-1997, where  $d = 30$ -day survival post cardiac surgery is of interest. An application of the STRAND Chart (Survival Time, Risk-Adjusted, N-Division Chart) to this same dataset is given in Section 3.1.

These charts, as discrete-time charts, are not as responsive to the continuous process as they might be, being slow to react to early failures. A move from discrete-time charts to continuous-time charts, then, would lead to an improvement in efficiency in detecting process change.

Methods developed subsequently have moved to continuous-time monitoring of survival data, where failures are recorded as and when they occur rather than being recorded  $d$  number of days after treatment. Since they update on a finer timescale, these charts are more reactive to deterioration in the survival process.

A challenge for continuous-time survival time control charts is the problem of maintaining patient ordering, so that time ordering according to the treatment timeline is preserved. The difficulty lies in the fact that patients do not fail in necessarily the same order that they arrive in. If patient ordering is maintained by a chart, clusters of unusual failures or successes in cohorts of patients on the treatment timeline can be identified by the chart effectively. Another challenge is presenting the data in a sensible and interpretable way. We believe that there is no perfect solution to creating a (risk-adjusted) survival time control chart that meets both of these criteria and that the best solutions given the first challenge is met will need to compromise on the fullness of presentation of data.

The reason survival time charts began as discrete as opposed to continuous is that, though recording in continuous time would on the surface appear quite feasible, updating in discrete time avoids corruption of patient ordering on the timeline. If failures were charted when they occurred instead of at the end of the censoring window, the ordering on the chart would then be by outcome day rather than surgery day. As a result, a significant and important localised change on the surgery timeline might be masked on the chart.

## 2.2 | RA CUSUM in continuous time based on the Cox model

This RA CUSUM in continuous time of Biswas and Kalbflesich records patient survival time data in continuous time as and when updates come in, by way of a Cox proportional hazards survival model.<sup>16,36</sup>

The chart is an LLR CUSUM, testing for a fixed change across patients in the instantaneous hazard, or intensity (the probability of imminent failure). The chart is semiparametric in nature since, though the risk model is fully parametric, the survival curve itself is left unspecified.

The continuous-time passage of patients through the censoring window is regularly recorded on the chart. Outside of these regular records, the chart is updated whenever a failure occurs. Since information is recorded as and when it arrives and is not reshuffled according to patient order, any clusters of failures or long ongoing survivals on the treatment timeline may be contaminated by noise and potentially not detected by the chart. A remedial step would be to adopt the Weighted CUSUM approach of Yashchin,<sup>21,37</sup> weighting data as it comes in, giving heavier weights to more recent patients. This would lead to a CUSUM method akin to the updating EWMA (uEWMA) Chart of Steiner and Jones discussed in the next section,<sup>14</sup> or the chart could be drawn reshuffled according to patient order, leading to a method akin to the Improved Bernoulli CUSUM of Keefe et al discussed in Section 2.4.<sup>10</sup> Such a chart would need to be regularly redrawn on new axes however. We elaborate on this in Section 2.4.

## 2.3 | uEWMA for survival time data and Weighted CUSUM

This method of Steiner and Jones, as with a Weighted Biswas and Kalbfleisch CUSUM, does respect patient ordering by weighting incoming survival data (failures or ongoing incrementing survival) according to how recently a patient was treated.<sup>14,16</sup> It exponentially weights the  $j$ th patient's current contribution  $S_{j,t}$  to the user's chosen survival summary statistic (the choice allowing for both fully parametric and semiparametric methods) according to where that patient is on the current list of patients treated  $j = [1, \dots, J_t]$ . Patients operated on longer ago are downweighted by a factor  $(1 - k)^r$  according to the distance on the list between them and the most recent patient. The uEWMA chart can be written as

$$S_t = kS_{J_t,t} + k(1 - k)S_{J_t-1,t} + k(1 - k)^2S_{J_t-2,t} + \dots + k(1 - k)^{J_t-1}S_{1,t} \quad t = 1, 2, \dots, \quad 0 < k < 1. \quad (2)$$

The chart effectively damps older cohorts of data down in the current cross-section of data at time  $t$ , lessening the effect of the disorder caused by patient failures not occurring in the same order as arrivals. In practice, especially for larger  $k$ , the weights decay quickly, so that the chart statistic  $S_t$  is little affected by only taking the summation over the most recent block of patients.

In damping the older data down, the chart's sensitivity to historical change is reduced. Its focus is drawn toward short-term behaviour of the process, rendering it more reactive to deterioration in surgical performance and less reactive, in particular, to improvements further back relating to long survivals of higher-risk patients. This is a drawback of a method such as this that respects patient ordering via the use of decaying weights. Valuable information in the data regarding older cohorts is downgraded.

Another method that weighs incoming data in this way is the Weighted CUSUM of Yashchin.<sup>21,37</sup> Yashchin remarks that their approach as applied to survival time (equivalently lifetime, or reliability) data is not "sufficient [alone] for efficient detection of change in the process level" and that "supplementary tests may be needed."<sup>22</sup> We suggest that rather than introducing such supplemental tests to make up for the limitations of the chart employed, the chart design should be reconsidered. We present our proposed method in Section 3. In the next section, we review a method that maintains patient ordering while accounting for ongoing survival in a conservative way.

## 2.4 | Improved Bernoulli CUSUM

This chart of Keefe et al is an improvement to the Bernoulli RA CUSUM of Steiner et al in that failures are recorded as and when they occur as opposed to when a fixed censoring time has elapsed.<sup>10,12</sup> The chart is redrawn at the occurrence of new failures in such a way that patient ordering is respected. Ongoing patient survival is conservatively treated in that patients yet to have, if at all, recorded failures are assumed up until that time to be cases that survive past the censoring time  $d$ .

This chart picks up process deterioration faster than the standard Bernoulli CUSUM because failures are recorded when they occur. However, a chart that did not employ such conservative treatment of patients yet to present as failures would identify process deterioration faster still. If full survival information were recorded for patients continuing to survive, rather than the conservative measure used, the chart would be more efficient, both for detecting process deterioration and improvement. This use of full survival information would make it akin to the Biswas and Kalbfleisch CUSUM,<sup>16</sup> but with reshuffling of points on the chart at regular redrawing of the chart, so that the chart adheres to patient ordering.

As we believe is necessary for a (risk-adjusted) survival time control chart, the chart does compromise on fullness of presentation of the data in that it requires being redrawn on a new set of axes as new failures occur. We consider the chart to be well suited to examining how clusters develop in a snapshot of time. However, we think it is not suited to viewing clusters and clusters of clusters in bigger frames of time, unless an additional low signal boundary were set and clusters of signals above that low boundary were recorded on a separate chart. The redrawing of the chart is not so much of a drawback when the censoring window is fairly short, as for the post cardiac surgery data (where  $d = 30$  days), since the chart is fixed once a lag the size of the censoring period has elapsed. However, it is more of a drawback when the censoring window is of the order of years, as for example, with the transplant data analysed by Biswas and Kalbfleisch<sup>16</sup>: one has to wait a year, say, before the plotted chart at least a year previous to that is no longer subject to change.

## 3 | METHOD INTRODUCED HERE: THE STRAND CHART

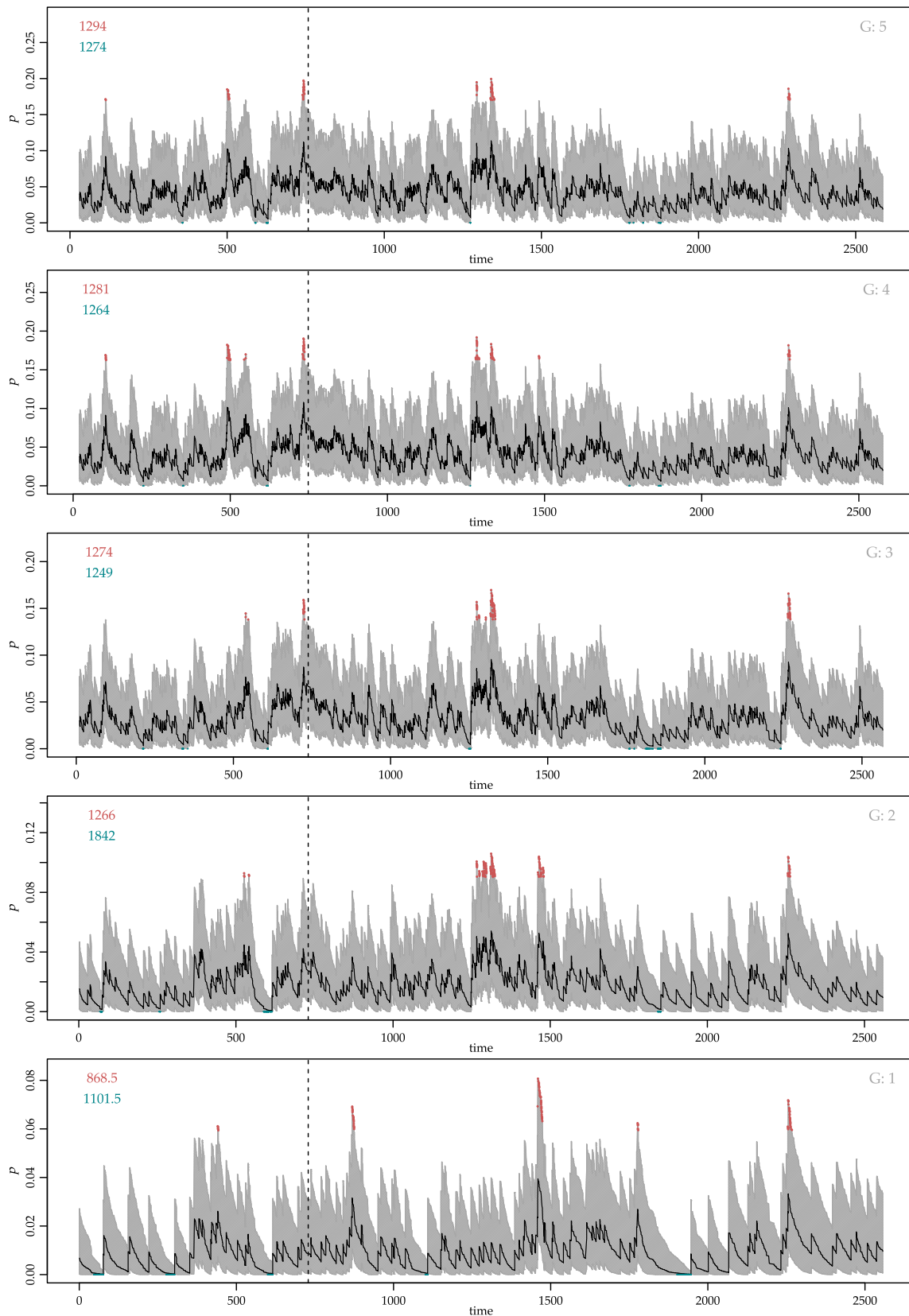
We propose a new method, the STRAND Chart, which is our solution to survival time monitoring where patient ordering is maintained and where balance between early failures and longer-term survivals within data cohorts is maintained. The chart can be a Bayesian or Classical chart. We present it as a Bayesian chart as we see this as the more natural approach. The theoretical origins of the chart are fairly complex but the resulting method is simple to implement and interpret.

If we consider a survival time distribution where patients survive a number of days following an intervention such as surgery, the STRAND Chart can estimate probabilities of failure at each of  $t_1, t_2, \dots, t_N$  days. The first strand estimates the failure rate at, say,  $t_1 = 1/2$  a day, the second the failure rate at, say,  $t_2 = 2$  days and so on. The last strand would, we suggest, relate to the failure rate at the censoring time of, say,  $t_N = 30$  days. We refer to survival past  $t_n$  days ( $n$  in  $1, \dots, N$ ) as survival past “gate”  $g_n$ . In the example application in Section 3.1, these failure rates are small, as one would expect from a high quality process. The chart signals when the credible interval about any strand excludes a target value, for example, the estimated failure rate from the pilot data.

The discretisation into  $N$  strands means that the chart is not fully continuous, though increasingly fine continuity can be achieved by adding more strands to the chart. Too many strands increase the complexity of the chart and so reduce human readability. Also, too many strands can result in a reduction in chart efficiency (see Section 4). However, with more strands, the passage of unusual cohorts of patients can be seen to track diagonally upward through the chart, by way of peaks and troughs, in a ripple effect. The nature of any clusters can be seen to play out (see Figure 1): a cluster of late failures in a cohort of patients would affect only upper strands of the chart; clusters of early failures might be balanced in a cohort by unusual longer-term survivals, so that an effect on lower strands would be nullified in upper strands. That the chart allows for the equal balance of failures and survivals in cohorts of data is an important feature.

As a Bayesian chart, the STRAND Chart can be formed of Bayesian time series estimators, updating estimates of the probability of failure at each gate  $g_n$ ,  $n = 1, 2, \dots, N$  in the light of the data. We present the chart here as a set of Bayesian Bernoulli risk-adjusted exponentially weighted moving average (RA EWMA).<sup>17</sup> The RA EWMA emanates from a model where the probability of failure at time  $i$ ,  $\mu_i$ , satisfies the following conditions:

$$\begin{aligned} E[\mu_i] &= E[\mu_{i-1}] \\ P[\mu_i] &= \kappa P[\mu_{i-1}], \end{aligned} \quad (3)$$



**FIGURE 1** STRAND Chart (Survival Time, Risk-Adjusted, N-Division Chart) showing dynamic risk-adjusted survival rate estimates for gates at (1/2, 2, 10, 20, 30) days, 1 Jan 1992 to 31 Dec 1997. First upper and lower signals respectively on each strand shown in upper left corners. Key: — rate; ■ 95% credible interval; ● signals of rise/fall in rate; - - pilot data cut-off [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

where  $E$  represents a priori expectation and  $P$  represents a priori precision, the latter being equal to one over the variance and a priori meaning before the data point  $y_i$  is seen. According to the model, the quantity  $\mu_i$  is said to be “mean steady” and is assumed to have decaying precision  $P$ , by a factor  $\kappa$ , from one time point to the next. It is assumed to change smoothly with time and this smooth process can be learnt about via the data  $y_i$ ,  $i = 1, 2, \dots$ .

From the mean steady model, the EWMA statistic

$$m_i = \kappa m_{i-1} + (1 - \kappa)y_i; \quad i = 1, 2, \dots; \quad 0 < \kappa < 1 \quad (4)$$

results as the posterior mean for  $\mu_i$  and the associated precision  $P[\mu_i]$  tends to  $1/(1 - \kappa)$  as  $i$  tends to infinity.<sup>17</sup>

As an additional step, to account for patient heterogeneity, risk adjustment of the EWMA can be made so that we have

$$m_i = \kappa m_{i-1} + (1 - \kappa)[y_{i,j} - \hat{\mu}_{i,j} + \hat{\mu}_{i,b}]; \quad i = 1, 2, \dots; \quad j = 0, 1, 2, \dots, b, \dots, J; \quad 0 < \kappa < 1, \quad (5)$$

where the  $i$ th outcome at a gate, relating to a patient of type  $j$ , is approximately adjusted to the metric of a baseline patient (one with a benchmark risk value) of type  $j = b$  via a risk model relating the expected rate for the arriving patient  $\mu_{i,j}$  to the current expected rate for a baseline patient  $\mu_{i,b}$ . This risk adjustment step is elaborated on in the application section, Section 3.1. Through the risk adjustment step,  $m_i$  becomes an estimate for  $\mu_{i,b}$ , which we refer to as simply  $\mu_i$ .

As discussed elsewhere,<sup>11,17</sup> these RA EWMA statistics can also be applied from within the frequentist framework. In the Bayesian approach, we note that there is formal admission that there is uncertainty about the first point on the chart,  $m_0$ .

The uncertainty about the RA EWMA statistic  $m_i$  is accounted for as follows. Bounds are placed about  $m_i$  to express uncertainty about the baseline mean  $\mu_i = \mu_{i,b}$ . Since the data are Bernoulli, we suggest for simplicity that a conjugate Beta distribution be chosen. We also suggest, again for simplicity, that the precision at time  $i$  be set to the steady-state value  $1/(1 - \kappa)$  so that the bounds about  $m_i$  are percentiles of a Beta distribution with mean  $m_i$  and precision  $1/(1 - \kappa)$ .

In summary, all that is required to apply the essential form of the STRAND Chart is a risk model relating patients to a baseline or benchmark patient, a choice of gates, a choice for  $\kappa$  (which we will elaborate on in the application section, Section 3.1), and outcomes  $y_{i,j}$  indicating whether the  $i$ th patient survives or fails at a gate along with risk scores  $z_{i,j}$  for those patients. Only the data are required, a choice of gates and an indication of the amount of data to be apportioned as pilot data, to implement the essential form of the STRAND Chart using the R package “strand” (available from the author).

It should be noted that there are existing Bayesian methods for monitoring survival data.<sup>38-40</sup> However, these focus on retrospective changepoint estimation and, by this, contrast with online monitoring methods. Where online methods are ideal for detecting process change forward in time, changepoint methods are ideal for locating with hindsight when that change occurred.

### 3.1 | Example: post cardiac surgery survival

There is an accompanying package in R, “strand,” see Supporting information, for users to implement the chart. We used the package to draw a chart of the now classic UK cardiac surgery data of 1992-1997, treating, as in other papers, the first two years’ data as pilot data.<sup>14,15</sup> We chose gates at (1/2, 2, 10, 2, 30) days with the aim of giving a comprehensive yet not too detailed cross-section of the overall survival distribution. Patient risk is summarised by the Parsonnet score,<sup>41</sup> and survival is censored at 30 days following surgery. The resulting chart can be seen in Figure 1.

The risk model, assumed to be a set of logistic models, one for each strand, is calibrated to pilot data. This means that the chart adopts an empirical Bayes approach where the initial prior for the variable of interest is calibrated on pilot data. For each strand, this provides a starting point  $m_0$  for the incrementing estimate,  $m_i$ , of the baseline failure rate  $\mu_i$  at a particular gate.

The STRAND Chart could alternatively be initialised by fitting an overarching fully parametric survival model as opposed to a separate model for each strand. However, fitting separate risk models for each strand allows the survival distribution to go unspecified, resulting in a semiparametric method that is more flexible.

The logistic risk model assumes that the odds of failure at a gate for a patient of type  $j$  is linked to the odds of failure for a baseline patient by a scaling factor, such that

$$\left( \frac{\hat{\mu}_j}{1 - \hat{\mu}_j} \right) = \left( \frac{\hat{\mu}_b}{1 - \hat{\mu}_b} \right) e^{\hat{\beta}(z_j - z_b)}, \quad (6)$$

where  $z_j$  is the patient risk score, in this case the Parsonnet score, and  $z_b$  is the baseline patient risk score. As per the paper of Steiner and Jones, we take the baseline Parsonnet score to be the median in the pilot data,  $z_b = 7$ . For these data, Equation (6) gives  $\hat{\beta} = (0.067, 0.071, 0.074, 0.074, 0.077)$  for the strands relating to the survival rates at (1/2, 2, 10, 2, 30) days.

When a patient of type  $j$  arrives at any of the gates at time  $i$ , their binary outcome of success is noted. The risk adjustment step can be carried out by assuming that the current predicted failure rate at that gate for a baseline patient,  $\hat{\mu}_{i,b}$ , is equal to the previous chart value  $m_{i-1}$ . Substituting this into (6) gives

$$\hat{\mu}_{i,j} = \frac{\left(\frac{m_{i-1}}{1-m_{i-1}}\right) e^{\hat{\beta}(z_j-z_b)}}{1 + \left(\frac{m_{i-1}}{1-m_{i-1}}\right) e^{\hat{\beta}(z_j-z_b)}}. \quad (7)$$

By using the logistic risk model, differing patient risks can be predicted by the STRAND Chart simply by scaling up or down on the odds scale by the factor  $e^{\hat{\beta}(z_j-z_b)}$ .

Given the vector of binary outcomes at time  $i$ ,  $y_{i,j}$ ,  $\hat{\mu}_{i,b} = m_{i-1}$  and  $\hat{\mu}_{i,j}$  as given by (7),  $y_{i,j} - \hat{\mu}_{i,j} + \hat{\mu}_{i,b}$  is the RA data value to be fed into the RA EWMA.

The (vector-valued) smoothing parameter  $\kappa$  can normally be chosen by minimising the mean squared error of prediction in the pilot data. However, since this curve tends to flatten out for  $\kappa \rightarrow 1$  for Bernoulli data, we recommend choosing a  $\kappa$  that is large enough for good prediction (low mean squared error) yet small enough to allow the chart to be reactive to change in the process. We recommend using the following formula as a guide:

$$\kappa = 0.993 - 0.192m_0, \quad (8)$$

where  $m_0$  is the starting value for a particular strand. This was developed from the example data, so is most useful for applications with similarly small failure rates  $\mu$ . For these data, Equation (8) gives  $\kappa = (0.992, 0.990, 0.987, 0.986, 0.985)$  for the strands relating to survival at (1/2, 2, 10, 20, 30) days, respectively.

The Bayesian version of the chart allows the chart statistic to be treated as the posterior mean of the rate of failure for a baseline patient at the survival gates,  $\mu_{i,b}$ . The steady-state Bayesian RA EWMA allows the precision of the posterior distribution for  $\mu_{i,b}$  to be fixed at  $1/(1 - \kappa)$ . Credible intervals around the RA EWMA estimate for  $\mu_{i,b}$ ,  $m_i$ , act as control limits for the chart. These can be set by adjusting the probability interval. The default interval in the strand package is a symmetric 95% credible interval.

In terms of placing the credible intervals, we do not recommend doing this by running simulations under null and alternative hypotheses where the true mean is fixed and considering run length properties. To tune to test criteria, we recommend setting the posterior credible bounds to the default and if necessary widening or narrowing one or both so that no or very few signals occur in the pilot data. Credible interval width and placement seems a more natural and simpler consideration to us than considering run length properties under basic null and alternative hypotheses. In the placement of the credible intervals, the user can make the chart as robust as they require using the pilot data as a guide. We chose the interval width and placement to give signals in the pilot data similar to those seen in other applications to these data in the literature,<sup>12,14,15,17</sup> that is, few signals of process improvement or deterioration.

Though we do not recommend the consideration of run length properties of the chart for setting chart parameters, it is nonetheless of interest to consider the performance of the STRAND Chart compared to competing charts, appealing to run length characteristics as standard in the literature.<sup>10,14</sup> In Section 4, we compare the chart to the best-performing chart in the literature, the uEWMA.<sup>14</sup>

To construct the posterior credible interval for  $\mu_{i,b}$ , we can assume a conjugate Beta( $m_i/(1 - \kappa)$ ,  $1/(1 - \kappa)$ ) distribution whose mean is  $m_i$  and whose precision is  $1/(1 - \kappa)$ . For 95% credible intervals, we take an interval of width 95% from this Beta distribution. For this application, we chose asymmetric bounds with the lower boundary at 5% and the upper boundary at 97.5%, increasing signals of deterioration compared to symmetric bounds of 2.5% and 97.5%. Choosing these bounds gives a similar signalling pattern to other charts in the literature applied to this particular dataset.<sup>12,14,15,17</sup>

In terms of results from the data, new information is available about these data when implementing the STRAND Chart. In particular, the cluster of unusual deaths around three and half years (1266 days) from the start of monitoring can be analysed in more detail: the nature of the cluster can be seen to play out in Figure 1. There are no signals at that particular period of time at the 1/2-day survival gate (the lowest strand). At the 2-day survival gate (the next strand up), a block of red indicating signals is evident, starting at 1266 days. As longer-term survivals are weighed in, however, at the 10- and 20-day gates (the next two strands up), the blocks of red reduce in size and the extremity of the overall signal of a cluster is

reduced. This equal balancing of failures and survivals in cohorts of data that the STRAND Chart allows is an important feature.

## 4 | CHART PERFORMANCE

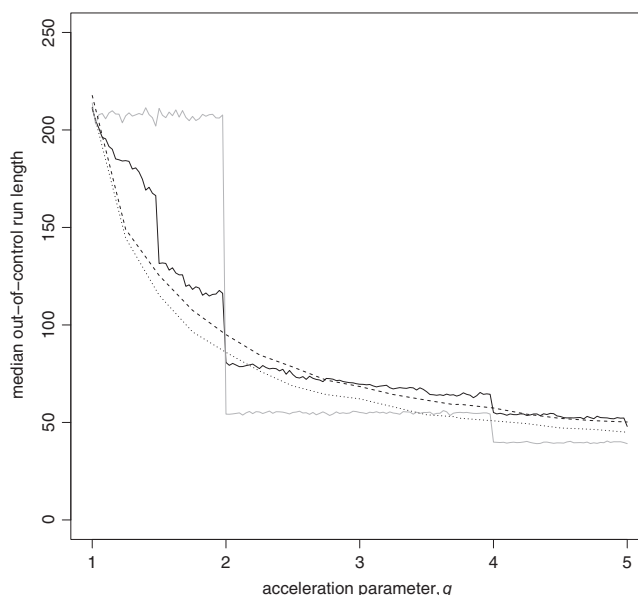
The best-performing chart in the literature appears to be the uEWMA with LLR scores.<sup>14</sup> We compare this chart, through consideration of run length properties under basic null and alternative hypotheses, to a STRAND Chart with the same parameters as that in Figure 1. We also introduce into the comparison the uEWMA with scores relating to 30-day survival as well as a STRAND Chart with a single gate estimating 1/2-day survival.

The simulations upon which the performance is assessed were carried out via empirical bootstrapping from the example data, that is, sampling without replacement from the example data pairs of survival times and Parsonnet scores. For the “in-control” scenario the data were sampled without any alteration. For the “out-of-control” scenario, the bootstrapped survival times were accelerated by a factor  $q$ , as per the paper of Steiner and Jones.<sup>14</sup> Figure 2 shows the results. All simulated values were based on 20 000 chart runs. As a performance measure, median run lengths were estimated.

It can be seen from Figure 2 that, for survival times shortened by a factor  $q$ , the uEWMA with LLR scores appears to be more efficient than the STRAND Chart in the format of that in Figure 1 for most values of  $q$ . The uEWMA with scores relating to 30-day survival, however, is sometimes more and sometimes less efficient than that STRAND Chart. Contrastingly, the STRAND Chart with a single gate at 1/2 a day is sometimes more and sometimes less efficient than the uEWMA with LLR scores. It seems that a STRAND Chart with a single early gate may be more efficient than one with multiple gates, though, due to the step-function nature of the median run length function, a STRAND Chart with a single early gate may not be sensitive to small changes in the survival distribution compared to other charts.

Although having a single early STRAND might improve efficiency for moderate to large values of  $q$ , less can be deciphered about the monitored process by just having one early strand. In reality, process changes are likely to be more complex than a simple acceleration of survival times such that the survival curve is affected in different ways. Hence, we recommend having more strands to capture more information about the process as it changes with time and to capture more information about any clusters of failures. The potential sacrifice of efficiency we believe is worth the gain in information.

Note that, for the STRAND Chart with gates at (1/2, 2, 10, 20, 30) days,  $\kappa$  was selected by the formula given in (8) and the credible bounds were set at 5% and 97.5% giving an “in-control” median run length of 210. For the STRAND Chart



**FIGURE 2** Efficiency of STRAND Chart (Survival Time, Risk-Adjusted, N-Division Chart) versus uEWMA (updating exponentially weighted moving average) for the example data as measured by median “out-of-control” run length for a given survival acceleration  $q$ . Key: — STRAND Chart as per Figure 1; — STRAND Chart with 1 strand at 1/2 day; ··· uEWMA with log-likelihood ratio weights; - - - uEWMA with 30-day survival weights



with a gate at  $1/2$  day,  $\kappa$  was again selected by the formula given in (8) and the credible bounds were set at 13.6% and 97.5%. For the uEWMA with LLR scores, all parameters were set at the values specified in the paper of Steiner and Jones,<sup>14</sup> except the upper signalling boundary that was set at  $h_u = 0.009$ . For the uEWMA with scores relating to 30-day survival,  $\kappa$  was set at 0.01, the lower reflecting boundary was set at  $h_l = 0.01$ , the upper boundary at  $h_u = 0.048$ , and the chart was initialised at  $S_0 = 0.025$ .

It makes intuitive sense that the STRAND Chart is more comparable in terms of performance to the uEWMA with scores relating to 30-day survival, it being based on  $n$ -day survival rates, than to the uEWMA with LLR scores. It is possible to run the STRAND Chart as a set of EWMA or CUSUMs with LLR scores and it is likely that, in that case, the performance would be more alike to the uEWMA with LLR scores, though this is a hypothesis we have not formally tested. We recommend using RA EWMA for the strands, however, as this provides a direct way of estimating features of interest of the process. Where LLR scores are more optimal, scores relating to summary measures of the survival distribution, such as tail areas, are more interpretable. LLR scores also require a fully specified survival model for the data and specification of the alternative hypothesis to be tested for.

## DISCUSSION

The STRAND Chart for online monitoring of survival outcomes in continuous time keeps patient ordering intact, as the uEWMA,<sup>14</sup> Weighted CUSUM,<sup>21</sup> and Improved Bernoulli CUSUM<sup>10</sup> do. The Biswas and Kalbfleisch continuous-time Cox proportional hazards CUSUM does not keep patient ordering intact in the same way,<sup>16</sup> as we have argued, but does update as and when information comes in and so is more sensitive to recent change in the process being monitored.

As in the papers of Steiner and Jones, Yashchin, and Keefe et al,<sup>10,14,21</sup> patient ordering according to real-time ordering of surgery is maintained, but here, short survival times do not affect all strands until the censoring period is up and all survival times for that cohort of patients who are operated on the same day (or on neighbouring days) have been collected. Thus, the short survival times do not get larger weights than the longer survival times in the same cohort simply because they are more contemporary. Also, the longer survival times are not lost in the chart, since the chart in effect waits for them to come through and to be weighted fairly against the short times. The layered nature of the chart and the natural real-time delay mean that short-term survival patterns affect the lower strands of the chart, where these relate to short-term survival rates, earlier in real monitoring time than they affect the upper strands of the chart, where these relate to longer-term survival rates (within the censoring window).

The STRAND chart can be applied within a Bayesian framework, so is akin to Bayesian changepoint methods for monitoring survival outcomes,<sup>39</sup> but, in contrast to those methods, is an online as opposed to retrospective method. The methods we review in this paper are, like the STRAND chart we present, also online methods.<sup>14-16,21</sup> Where online methods can detect process change quickly, retrospective methods are ideally used as tools for locating the time of change (or changepoint).

We do not try to discuss in full the relative merits of the Classical and Bayesian approach to applying the EWMA here, as it is beyond the scope of this paper. For more insight, there is a background of available literature.<sup>11,17,28-33,42</sup>

The basic form of the STRAND Chart that we present is semiparametric in nature, much like the Cox proportional hazards CUSUM of Biswas and Kalbfleisch<sup>16</sup> in that the survival distribution is not fully specified. However, any statistic could be used on each of the strands, in particular, the more efficient LLR CUSUMs or log hazard ratio CUSUMs. We present the chart as a collection of RA EWMA as this provides semiparametric smoothed estimates of percentiles of the survival distribution that are optimal time series smoothers.<sup>43</sup> Fully parametric methods have less freedom, and CUSUMs based on the likelihood, although optimal in terms of efficiency, do not immediately give smoothed estimates of features of the survival distribution.

In the example application to the now classic UK post cardiac surgery dataset of 1992-1997, we show that the STRAND Chart can describe the nature of clusters of failures as they ripple through the chart, and we highlight where unusually long survivals of patients are balancing out the unusual failures. Previously, undetected features of the data are highlighted by the chart. Through bootstrapping of this dataset, we show that, given the format of estimating  $n$ -day survival, the performance of the STRAND Chart is closer to that of the uEWMA with scores relating to 30-day survival than to the more optimal uEWMA with LLR scores. We also show that STRAND Charts with a single early strand might be more efficient for detecting process change but we note that such a chart would perhaps not describe the survival process in enough detail. We expect that a STRAND Chart based on CUSUMs would have efficiency more alike to that of the uEWMA with LLR scores. This is potential area for future investigation.

## ORCID

Olivia AJ Grigg  <https://orcid.org/0000-0002-2967-0863>

## REFERENCES

1. Bartroff J, Lai TL, Shih M. *Sequential Experimentation in Clinical Trials: Design and Analysis*. New York, NY: Springer Science & Business Media; 2012.
2. Whitehead J. *The Design and Analysis of Sequential Clinical Trials*. Chichester, UK: John Wiley & Sons; 1997.
3. Whitehead J, Jones D. The analysis of sequential clinical trials. *Biometrika*. 1979;66(3):443-452.
4. Spiegelhalter D, Sherlaw-Johnson C, Bardsley M, Blunt I, Wood C, Grigg O. Statistical methods for healthcare regulation: rating, screening and surveillance. *J R Statist Soc A*. 2012;175(1):1-47.
5. Woodall WH, Grigg OA, Burkom HS. Research issues and ideas on health related surveillance. In: *Frontiers in Statistical Quality Control*. Vol-9. Berlin, Germany: Physica-Verlag; 2010:145-155.
6. Woodall WH, Mohammed MA, Lucas JM, Watkins R, et al. The use of control charts in health-care and public-health surveillance. Discussion/Rejoinder. *J Qual Technol*. 2006;38(2):89-134.
7. Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care*. 2003;15(1):7-13.
8. Montgomery DC. *Statistical Quality Control*. New York, NY: Wiley New York; 2009.
9. Shewhart WA. *Economic Control of Quality of Manufactured Product*. Milwaukee, WI: ASQ Quality Press; 1931.
10. Keefe MJ, Loda JB, Elhabashy AE, Woodall WH. Improved implementation of the risk-adjusted Bernoulli CUSUM chart to monitor surgical outcome quality. *Int J Qual Health Care*. 2017;29(3):343-348.
11. Grigg OA. *Risk-Adjusted Monitoring and Smoothing in Medical Contexts* [PhD thesis]. Cambridge, UK: University of Cambridge; 2004.
12. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics*. 2000;1(4):441-452.
13. Zhang X, Woodall WH. Dynamic probability control limits for risk-adjusted Bernoulli CUSUM charts. *Statist Med*. 2015;34(25):3336-3348.
14. Steiner SH, Jones M. Risk-adjusted survival time monitoring with an updating exponentially weighted moving average (EWMA) control chart. *Statist Med*. 2010;29(4):444-454.
15. Sego LH, Reynolds Jr MR, Woodall WH. Risk adjusted monitoring of survival times. *Statist Med*. 2009;28(9):1386-1401.
16. Biswas P, Kalbfleisch JD. A risk-adjusted CUSUM in continuous time based on the Cox model. *Statist Med*. 2008;27(17):3382-3406.
17. Grigg OA, Spiegelhalter DJ. A simple risk-adjusted exponentially weighted moving average. *J Am Statist Ass*. 2007;102(477):140-152.
18. Grigg OA, Farewell VT. An overview of risk-adjusted charts. *J R Statist Soc Series A*. 2004;167(3):523-539.
19. Lovegrove J, Sherlaw-Johnson C, Valencia O, Treasure T, Gallivan S. Monitoring the performance of cardiac surgeons. *J Oper Res Soc*. 1999;50(7):684-689.
20. Poloniecki J, Valencia O, Littlejohns P. Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. *Br Med J*. 1998;316:1697-1700.
21. Yashchin E. Design and implementation of systems for monitoring lifetime data. In: *Frontiers in Statistical Quality Control*. Vol 10. Berlin, Germany: Physica-Verlag; 2012:171-195.
22. Yashchin E. Computational and Monte-Carlo aspects of systems for monitoring reliability data. In: *Proceedings of COMPSTAT'2010*. Berlin, Germany: Physica-Verlag; 2010:253-262.
23. Page ES. Continuous inspection schemes. *Biometrika*. 1954;41:100-115.
24. Hunter JS. The exponentially weighted moving average. *J Qual Technol*. 1986;18(4):203-210.
25. Roberts SW. Control chart tests based on geometric moving averages. *Technometrics*. 1959;1(3):239-250.
26. Sonesson C. Evaluations of some exponentially weighted moving average methods. *J Appl Stat*. 2003;30(10):1115-1133.
27. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*. 1990;32(1):1-12.
28. Vidoni P. Exponential family state space models based on a conjugate latent process. *J R Statist Soc B*. 1999;61(1):213-221.
29. Grunwald GK, Hamza K, Hyndman RJ. Some properties and generalizations of non-negative Bayesian time series models. *J R Statist Soc Some B*. 1997;59(3):615-626.
30. West M, Harrison PJ. *Bayesian Forecasting and Dynamic Models*. 2nd ed. New York, NY: Springer-Verlag; 1997.
31. Pole A, West M, Harrison PJ. *Applied Bayesian Forecasting and Time Series Analysis*. London, UK: Chapman & Hall/CRC; 1994.
32. West M, Harrison PJ, Migon HS. Dynamic generalized linear models and Bayesian forecasting. *J Am Statist Ass*. 1985;80(389):73-83.
33. Smith JQ. A generalization of the Bayesian steady forecasting model. *J R Statist Soc B*. 1979;41(3):375-387.
34. Harrison PJ, Stevens CF. Bayesian forecasting (with discussion). *J R Statist Soc B*. 1976;38:205-247.
35. Colosimo BM. Bayesian control charts. In: *Encyclopedia of Statistics in Quality and Reliability*. Hoboken, NJ: John Wiley & Sons, Inc; 2008.
36. Cox DR. Regression models life tables. *J R Statist Soc Series B*. 1972;34(2):187-220.
37. Yashchin E. Weighted cumulative sum technique. *Technometrics*. 1989;31(3):321-338.

38. Assareh H, Mengersen KL. Estimation of the time of a linear trend in monitoring survival time. *Health Serv Outcomes Res Methodol*. 2014;14(1):15-33.
39. Assareh H, Mengersen KL. Change point estimation in monitoring survival time. *PloS One*. 2012;7(3):e33630.
40. Assareh H, Mengersen KL. Bayesian estimation of the time of a decrease in risk-adjusted survival time control charts. *IAENG Int J Appl Math*. 2011;41(4):360-366.
41. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risks for evaluating the results of surgery in acquired adult heart disease. *Circulation*. 1989;779(1):1-12.
42. Key PB, Godolphin EJ. On the Bayesian steady forecasting model. *J R Statist Soc B*. 1981;43(1):92-96.
43. Muth JF. Optimal properties of exponentially weighted forecasts. *J Am Statist Ass*. 1960;55(2):299-306.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Grigg OAJ. The STRAND Chart: A survival time control chart. *Statistics in Medicine*. 2019;38:1651–1661. <https://doi.org/10.1002/sim.8065>