Research Article

Quality and Reliability Engineering International

(www.interscience.wiley.com) DOI: 10.1002/qre.1056 Published online 27 July 2009 in Wiley InterScience

A Comparison of CUSUM, EWMA, and Temporal Scan Statistics for Detection of Increases in Poisson Rates

Sung Won Han,[a] Kwok-Leung Tsui,[a] Bancha Ariyajunya[b] and Seoung Bum Kim[b]*†

Various control chart methods have been used in healthcare and public health surveillance to detect increases in the rates of diseases or their symptoms. Although the observations in many health surveillance applications are often discrete, few efforts have been made to explore the behavior of detection methods in discrete distributions. Joner *et al.* (*Statist. Med.* 2008; 27:2555–2575) investigated and compared the performance of the scan statistic methods with the cumulative sum (CUSUM) charts under a Bernoulli distribution. In this paper we compare the performance of three detection methods: temporal scan statistic, CUSUM, and exponential weighted moving average (EWMA) when the observations follow the Poisson distribution. A simulation study showed that the Poisson CUSUM and EWMA charts generally outperformed the Poisson scan statistic methods. In comparisons between CUSUM and EWMA, the CUSUM charts were superior in dealing with a large shift with a later change in time. However, the EWMA charts outperformed the CUSUM charts in situations with a small shift and an early change in time. The methods were also compared with thyroid cancer using a real data set. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: health surveillance; scan statistic; CUSUM; EWMA; online monitoring; Poisson distribution; temporal surveillance; conditional expected delay

1. Introduction

The timely detection of increases in the rate of unusual events is an important objective in public health and healthcare surveillance[1–4]. The objective of health surveillance, especially syndromic surveillance, is to detect a change in the incidence of natural outbreaks or bioterrorism and to issue an emergency alarm as soon as possible[5–7]. Detection of such changes is based on count data, such as a count of the respiratory diagnoses from civilian office visits, or measurement data such as emergency department (ED) visits, sales of over-the-counter remedies, and the number of visits to military clinics[2]. These count data, observed sequentially, are often assumed to follow a certain discrete distribution such as the Poisson distribution. If an undesirable event occurs, the rate of the Poisson distribution will change. The popular methods for detecting any such change include the scan statistic methods in biosurveillance, and the cumulative sum (CUSUM) and exponential weighted moving average (EWMA) charts in engineering statistical process control (SPC).

Scan statistic methods have been widely used for the detection of change in finite time domains and for detection of offline changes[8–11]. When used for online monitoring, scan statistic methods should be modified for the detection of a rate change[12]. A comprehensive review of the use of scan statistic methods for online monitoring can be found in Woodall *et al.*[13]. When used in online monitoring, scan statistic methods that use a fixed window size are indeed moving average control charts[12].

The CUSUM and EWMA charts are very popular methods in SPC application of manufacturing[14]. Both types of charts were devised to increase the capability to detect small process shifts. The theoretical properties of a CUSUM chart were investigated in Page[15], Shiryayev[16], Lorden[17], Pollak[18], and Lai[19]. In health surveillance, Hill *et al.*[20] and Weatherall and Haskey[21] adapted CUSUM charts mainly for the surveillance of congenital malformations. Brook and Evan[22] discussed the CUSUM chart for the detection of a mean change in the Poisson distribution. Lucas[23] studied the run length of the Poisson CUSUM charts. White and Keats[24] used a Markov chain approximation algorithm to find the threshold to use in obtaining the target of in-control average

[a]*H. Milton Stewart School of Industrial Systems and Engineering, Georgia Institute of Technology, Atlanta, GA, U.S.A.*
[b]*Industrial Systems and Information Engineering, Korea University, Seoul, Korea*
**Correspondence to: Seoung Bum Kim, Industrial Systems and Information Engineering, Korea University, Seoul, Korea.*
†*E-mail: seoungbum.kim@gmail.com*

Contract/grant sponsor: DHHS/PHS/NIOSH/CDC/Center for Disease Control Award; contract/grant number: 200-2008-M-28123

run length (i.e. ARL_0) that can be defined as the average number of observations needed for the method to detect a change in the normal state.

The EWMA charts were originally developed for two-sided tests[25–29]. Robinson and Ho[30] examined a one-sided control limit for a conventional EWMA chart. A discussion and comparison of the existing one-sided EWMA charts can be found in Shu *et al.*[31]. Gan[32] introduced a modified version of the EWMA chart to monitor the mean shift of a process away from the Poisson distribution. Borror *et al.*[33, 34] used the Markov chain approach and simulation to obtain the *ARL* of the Poisson EWMA chart. Joner *et al.*[35] used multivariate EWMA (MEWMA) for health surveillance under Poisson count data by proposing one-sided MEWMA.

Classical SPC methods such as CUSUM and EWMA have been applied for prospective health surveillance and studied their performance in recent periods. Cowling *et al.*[36] proposed an upper CUSUM chart by using a seven-week buffer interval and compared his proposed method with various time series methods. Jackson *et al.*[37] compared the performance of EMWA, Shewhart chart, and a general linear model method based on the data set having day-of-the-week effects. Because of the spatiotemporal characteristics of biosurveillance data, modifications of the multivariate SPC techniques such as MCUSUM and MEWMA have been suggested by Joner *et al.*[35] and Fricker *et al.*[38] Comprehensive review of the applications of SPC methods to health surveillance can be found in Woodall[1] and Tsui *et al.* [4].

A number of studies compared the performance of the aforementioned detection methods under continuous baseline distributions, typically, a normal distribution. Lucas and Saccucci[29] and Yashchin[39] showed that if a shift size is equal to a standard deviation, CUSUM performs slightly better than EWMA. Srivastava and Wu[40] suggested that EWMA is less efficient than CUSUM under stationary conditions when $ARL_0 \to \infty$. The other study revealed that the performance of EWMA is as good as, or slightly better than that of CUSUM in a two-sided test[41]. Recently, Joner and Woodall[12] compared the scan statistic methods with CUSUM for Bernoulli observations and concluded that based on a steady-state *ARL*, Bernoulli-based CUSUM charts performed better than Bernoulli scan statistic methods.

From a theoretical point of view, Lorden[17] proved that when Lorden's worst average detection delay is used, CUSUM charts are asymptotically optimal for the one-sided test in all baseline distributions. Moustakides[42] proved that the CUSUM method is exactly optimal when Lorden's performance measures are used. However, Lorden's criteria rate the performance of detection methods in extreme situations that rarely happen. More realistically, consideration of general and reasonable performance such as the conditional expected delay should be considered. This paper compares the conditional expected delay among the scan statistics method, CUSUM, and EWMA by evaluating their performance to detect increases in rates under the discrete Poisson distribution.

The outline of this paper is as follows: Section 2 describes the problem. Section 3 briefly reviews the Poisson version of the scan statistic, CUSUM, and EWMA methods. In Section 4, we compare the performance of these three detection methods under various simulation scenarios. Section 5 presents a real application study for the detection of male thyroid cancer data in New Mexico in the United States. Section 6 presents the concluding remarks and future research directions.

2. Problem formulation

In Poisson count data, we are interested in monitoring and detecting a shift in the rate of occurrences. In some health surveillances, the shift pattern is transient and has a finite duration, but in this problem formulation, we assume that the pattern is the persistent jump change in the mean (or rate), which is the classic SPC problem and the theoretical change point detection problem.

Assume that we observe a sequence of independent Poisson random variables over time, $\{Y_1, Y_2, \ldots\}$, in which the true mean of Y_i is μ_i. Under the normal state, $\mu_i = \lambda_0$. After an undesired event occurs at an *unknown* time v, the values of the μ_i's change from λ_0 to λ_1. In other words, for some $v \geq 1$, $Y_1, \ldots, Y_{v-1}$ are Poisson random variables with a mean of λ_0, whereas $Y_v, Y_{v+1}, \ldots$ are Poisson random variables with a mean of λ_1. The main goal here is to detect an increase in the rate (λ) as soon as possible after an undesirable event occurs. This problem can be formulated based on the following hypothesis testing problem:

$$H: \mu_i = \lambda_0 \quad \text{for all } i \geq 1 \quad \text{(i.e. no change)}$$

against the composite alternative hypothesis

$$K: \mu_i = \begin{cases} \lambda_0 & \text{if } 1 \leq i \leq v-1 \\ \lambda_1 & \text{if } i \geq v \end{cases} \qquad \text{for some } unknown\ v \geq 1 \quad \text{(i.e. a change)}$$

This hypothesis test (*H* versus *K*) is conducted each time based on a sequence of independent Poisson random variables. The null hypothesis *H* states that there is no change, which indicates $\mu_i = \lambda_0$ until the most recent time point in monitoring. The alternative hypothesis *K* states that a change occurred at an unknown time v and the values of the μ_i's changed from λ_0 to λ_1 at that time.

In general, the performance of a detection method *u* is evaluated by the following two criteria: the *false alarm rate* during the time of the in-control state and the *detection delay* in the out-of-control state. In the in-control state, we want the alarm time T^u to occupy as much time as possible so that the false alarm rate, measured by $1/E[T^u | v = \infty]$, is minimized. Here, v is the time of the onset of change and hence, $v = \infty$ indicates that a change never occurs. Note that a necessary condition for $ARL_0^u (= E[T^u | v = \infty])$ to be finite is that $P[T < \infty | v = \infty] = 1$. This implies that a false alarm is triggered with a probability of 1 even without any change.

In the out-of-control state, the detection delay of method u is $T^u - v$ for $(T^u \geq v)$ in which the conditional constraint guarantees that T^u generates a true alarm after a change occurs at time v. Because the stopping time T^u is a random variable, we may consider the *conditional expected delay*, defined as $CED^u(v) = E[T^u - v | T^u \geq v]$. In the present paper we use $CED^u(v, \lambda_1)$ to compare the detection methods in Poisson data with different shift sizes of change (i.e. λ_1) and different time points of changes (v).

It should be noted that $CED^u(v, \lambda_1)$ cannot be used for a direct comparison because the exact time that a change in rate occurs is generally unknown. Several approaches have addressed this issue and proposed various ways to make $CED^u(v, \lambda_1)$ applicable for comparison. The first approach is to consider $CED^u(v=1, \lambda_1)$, as the detection delay when a change occurs at time $v=1$[14]. The second approach, proposed by Shiryayev[16] and Pollak[18], uses $\sup_{1 \leq v < \infty} CED^u(v, \lambda_1)$. In many cases, $\sup_{1 \leq v < \infty} CED^u(v, \lambda_1)$ often equals to $CED^u(v=1, \lambda_1)$. The third approach is to consider $\lim_{v \longrightarrow \infty} CED^u(v, \lambda_1)$, which is called the steady-state *ARL* in the out-of-control state[40].

For comparison of the detection methods in the present study, we primarily use $CED(v, \lambda_1)$ for each method. That is, *CED*s are compared by graphical justification at each point in time that a change v occurs. Further, we considered different points of time for change v in $CED^u(v, \lambda_1)$. We focused on comparing three widely used detection methods that are based on scan statistic, CUSUM, and EWMA, by investigating their behavior on $CED^u(v, \lambda_1)$ for different v's and a certain range of shift sizes λ_1, subject to

$$ARL_0^u \geq \gamma \tag{1}$$

in which γ is the lower boundary of the target.

3. Detection methods

3.1. Scan statistics

A scan statistic generates an alarm at time T^S, which is the first time $n(\geq 1)$ such that the scan statistic (S_n) exceeds the threshold $k(\geq 1)$. Consequently, S_n is defined by

$$S_n = \max_{1 \leq i \leq n} \left\{ \sum_{j=i-m+1}^{i} Y_j \right\} \tag{2}$$

where Y_j is the observation at time j, n is the current time point, and m is the fixed window size. k is a prespecified threshold chosen to satisfy the constraint (1). Conventionally, we assume that $Y_0 = Y_{-1} = \cdots = Y_{-m} = 0$.

The scan statistic S_n has been used extensively in many areas of offline decision problems (for example, Glaz *et al.*[8, 9]). However, the direct use of scan statistics in (2) for online monitoring is inefficient. To efficiently implement this methodology, scan statistics can be modified as follows:

$$S_n = \max(S_{n-1}, M_n) \tag{3}$$

with

$$M_n = \sum_{j=n-m+1}^{n} Y_j \tag{4}$$

Instead of using S_n, we only need to monitor M_n. This modified scan statistic produces an alarm at time T^M, which is the first time n such that $M_n = \sum_{j=n-m+1}^{n} Y_j \geq k$.

To show that T^M is equivalent to T^S for k, first note that $T^S \leq T^M$ because the scan statistic S_n includes the statistic M_n. We can also show that $T^M \leq T^S$ as follows: If T^S issues an alarm at time n, S_{n-1} should be less than k, and S_n should be greater than or equal to k. Because $S_n = \max(S_{n-1}, M_n)$, M_n is greater than or equal to k, and T^M sets off an alarm at time n or earlier, implying that $T^M \leq T^S$. Combining the above statements yields an alarm time of $T^S = \inf\{n \geq 1 : S_n \geq k\}$, which is equivalent to $T^M = \inf\{n \geq 1 : M_n \geq k\}$.

Note that the method $M_n = \sum_{j=n-m+1}^{n} Y_j$ is simply referred to as the scan statistic method in Joner and Woodall[12] for Bernoulli random variables. However, in the SPC literature, the method is known as an unweighted moving average[14].

3.2. CUSUM charts

The Poisson CUSUM chart[22] triggers an alarm at time T^C when the CUSUM statistic C_n exceeds h. $C_n (n \geq 1)$ can be recursively calculated by

$$C_n = \max\{0, C_{n-1} + Y_n - r\} \quad \text{with } r = \frac{\lambda_1^* - \lambda_0^*}{\ln \frac{\lambda_1^*}{\lambda_0^*}} \tag{5}$$

where λ_0^* and λ_1^* are target in-control and out-of-control parameters, and $C_0=0$. Lorden[17] showed that the CUSUM chart is *asymptotically* optimal in detecting a one-sided mean shift according to Lorden's criteria, which is defined as

$$\bar{E}_1(T) = \sup_{1\leq\nu\leq\infty} \text{ess} \sup_{(X_1,\ldots,X_{\nu-1})} E_T((T-\nu+1)^+ | X_1,\ldots,X_{\nu-1}, \nu, T\geq\nu)$$

Furthermore, Moustakides[42] showed that the CUSUM chart is *exactly* optimal according to Lorden's criteria. However, when criteria based on *CED* (i.e $\sup_{1\leq\nu\leq\infty} E[T-\nu|T\geq\nu]$) is used, it is no longer clear if CUSUM continues to be optimal.

CUSUM can be considered as a scan statistic method with a variable window size for the following reasons. First, the CUSUM chart can be interpreted from an offline hypothesis testing viewpoint as generalized likelihood ratios in terms of changes in points of time[43]. More precisely, in detecting a rate change (from λ_0 to λ_1) from the distribution f_λ for the first n observations $Y_1,\ldots,Y_n$, we encounter the problem of testing H: *no change occurs* versus K: *a change occurred and* $1\leq\nu\leq n$. Based on this, the statistic W_n can be obtained from the following maximum of likelihood ratio with variable time windows:

$$W_n = \max_{1\leq\nu\leq n} \left\{ \sum_{j=\nu}^{n} \ln \frac{f_{\lambda_1}(Y_j)}{f_{\lambda_0}(Y_j)} \right\} \quad (6)$$

As illustrated in Moustakides[42], for $h\geq 0$, the alarm time $T^W = \inf\{n\geq 1 : W_n \geq h\}$ is equivalent to the alarm time $T^{W'} = \inf\{n\geq 1 : W_n' \geq h\}$ with $W_n' = \max(0, W_n)$. It is easy to see that

$$W_n' = \max\left(0, W_{n-1}' + \ln\frac{f_{\lambda_1}(Y_n)}{f_{\lambda_0}(Y_n)}\right)$$

and W_n' is equivalent to C_n in (5) for the Poisson count data.

Next, the maximum of likelihood ratio in (6) is equivalent to the scan statistic method derived from the likelihood ratio within variable time windows. We may define scan statistics with variable window size as

$$V_n = \max_{1\leq\nu\leq i\leq n} \left\{ \sum_{j=\nu}^{i} \ln \frac{f_{\lambda_1}(Y_j)}{f_{\lambda_0}(Y_j)} \right\} \quad (7)$$

which becomes scan statistics with a fixed window size if $\nu = i-m+1$. It is easy to see that V_n satisfies

$$V_n = \max(V_{n-1}, W_n) \quad (8)$$

For $h\geq 0$, the alarm time $T^V = \inf\{n\geq 1 : V_n \geq h\}$ is equivalent to the corresponding alarm times based on W_n or W_n'. Hence, the CUSUM chart can also be thought of as the scan statistic method with a variable window size for online monitoring[44].

3.3. *EWMA charts*

The EWMA chart for Poisson data, proposed by Borror *et al.*[33], causes an alarm at time $T^E =$ first n such that $E_n \geq b$. The EWMA statistic E_n can be recursively calculated by

$$E_n = \alpha Y_n + (1-\alpha) E_{n-1} \quad (9)$$

where $0<\alpha\leq 1$ and $E_0 = E[Y]$.

For early detection of a change, Montgomery[14] suggests an exact control limit, but we used a constant threshold in this study because the other two methods also use constant thresholds. We also applied the EWMA to log-likelihood ratios defined by

$$E_n = \alpha \frac{f_{\lambda_1}(Y_n)}{f_{\lambda_0}(Y_n)} + (1-\alpha) E_{n-1} \quad (10)$$

Note that the EWMA statistic of (10) can be simplified to the EWMA of the actual observation if the distribution $f(\cdot)$ belongs to the exponential family[43].

4. Simulation study

A simulation study was conducted to explore the detection ability of the three methods (scan statistics, CUSUM, and EWMA) for Poisson count data and to compare their performance under various scenarios. Our simulation is motivated by male thyroid cancer data in New Mexico[45].

In our simulation, we set ARL_0 as close to 1500 as possible without going below it. In other words, for all the detection methods, the corresponding parameters were chosen so that the methods generate a false alarm not less than once every 1500 time periods under the baseline incidence rate.

The rate $\lambda_0 = 1.4$ is the same baseline rate used for the spatiotemporal application in Sonesson[44]. Next, we decided to target shift sizes $\lambda_1^* = 1.75$, 2.1, 2.45, and 2.8, respectively, 25, 50, 75, and 100% larger than λ_0. We searched the parameters and thresholds for the targeted ARL_0 based on 1 600 000 replications. Further, we simulated $CED(v, \lambda_1)$ for different changes in points of time and shift sizes based on 50 000 replicates.

4.1. Parameter selection for target ARL_0

To test the scan statistic methods, we obtained a set of parameters (m, k) that yield ARL_0 as close to 1500 as possible without going below it. In the in-control state, if m decreased, ARL_0 increased because H_0 is hard to reject. The performance of the scan statistic methods depends on both m and k. Different m and k should be chosen for different shift sizes λ_1. Here is how we chose m and k for various specified shift sizes. We first tried to find all the possible combinations of m and k that result in the target ARL_0. Next, for each specified shift size (λ_1), we selected the combinations of m and k values that minimize ARL_1. Table I displays a set of the parameters for changes in target size λ_1^* and the corresponding ARL_0. The values in parentheses next to ARL_0 represent standard errors of run length. Note that the values of m in Table I are the optimal choice of m in terms of ARL_1 for a given shift size λ_1.

For the CUSUM charts, if λ_1^* in the CUSUM formula in (5) is equal to λ_1, then it yields a small ARL_0. First, we found the smallest threshold h under $\lambda_1^* = \lambda_1$ for the target ARL_0. However, the CUSUM statistic is discrete for Poisson observations. As a consequence, ARL_0 is a step function with respect to h (Appendix A). Table II shows a set of the CUSUM parameters for different true shift sizes λ_1 $(=\lambda_1^*)$, h, and the corresponding ARL_0.

The parameters for the EWMA charts are the weighting coefficient (α) and the threshold (b). The performance of the EWMA charts depends on both α and b. The way to select the parameters α and b is similar to the procedure used in the scan statistic methods. We first tried to find all the possible combinations of α and b that result in ARL_0's, which are close to ARL_0's in CUSUM. Next, for each specified shift size (λ_1), we selected a set of α and b values that minimize ARL_1. Table III shows a set of the EWMA parameters for change in target size λ_1^* and the corresponding ARL_0's.

4.2. Comparison of $CED(v, \lambda_1)$s at different points of time for changes v at fixed shift size λ_1

We first compared the three detection methods by considering different points of time at which change v occurred. To present the simulation results efficiently, we considered the following finite number of time points: $v = 1, 2, 3, \ldots, 48, 49, 50$. Note that most of $CED(v, \lambda_1)$s converge as early as 50. Figure 1 shows the resulting CED values of scan statistics, CUSUM, and EWMA for changes

Table I. A set of the parameters of the scan statistic method with different shift sizes given the targeted $ARL_0 = 1500$

λ_1 $(v=1)$	m	k	ARL_0 (s.e.)
1.75	37	72	1535.61 (1.19)
2.1	21	46	1500.57 (1.17)
2.45	14	34	1504.16 (1.18)
2.8	8	23	1519.24 (1.20)
3.15	5	17	1579.68 (1.25)

Table II. A set of the parameters of the CUSUM charts with different shift sizes given the targeted $ARL_0 = 1500$

λ_1 $(v=1)$	λ_1^*	h	ARL_0 (s.e.)
1.75	1.75	17.15	1547.35 (1.19)
2.10	2.10	11.68	1550.26 (1.21)
2.45	2.45	9.12	1533.51 (1.20)
2.80	2.80	7.8419	1576.68 (1.24)
3.15	3.15	6.7	1538.29 (1.21)

Table III. A set of the parameters in the EWMA charts with different shift sizes given the targeted $ARL_0 = 1500$

λ_1 $(v=1)$	α	b	ARL_0 (s.e.)
1.75	0.02	1.7038	1547.81 (1.20)
2.10	0.05	1.9651	1551.12 (1.22)
2.45	0.08	2.1720	1533.37 (1.20)
2.80	0.12	2.4155	1576.19 (1.24)
3.15	0.17	2.6792	1538.09 (1.21)

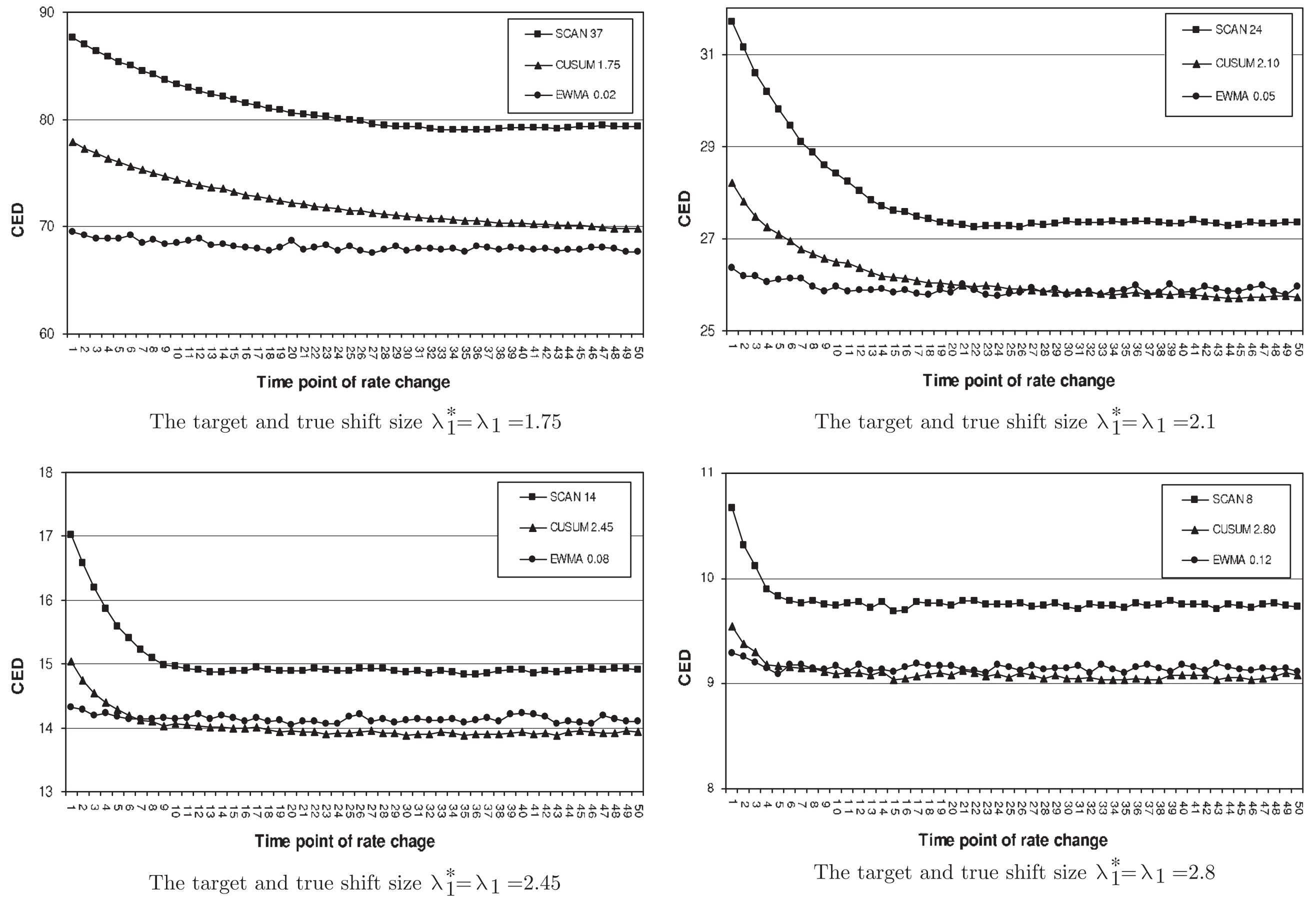


Figure 1. Comparison of the three detection methods with optimal parameters for given shift size λ_1 across different points of time for change v. The *x*-axis represents time point of rate change (vs) and the *y*-axis represents $CED^u(v, \lambda_1)$s

at different points in time (v), given four different shift sizes (λ_1). It can be seen that the values of $CED(v, \lambda_1)$ for CUSUM and EWMA are consistently smaller than those for the scan statistic methods, demonstrating that CUSUM and EWMA are more efficient than the scan statistic methods at detecting the points at which changes occur.

Furthermore, we generally observed that CUSUM performed slightly better than EWMA for a large shift in size and changes at later times. However, EWMA performed slightly better than CUSUM for smaller shifts and changes at early points in time.

Each detection method produced the maximum *CED* when $v=1$. In scan statistic methods, the $CED(v, \lambda_1)$ tends to rapidly decrease when change occurs at early time points and converges after $v \geq m$ because the scan statistics use a fixed window size *m*. The values of *CED* in CUSUM and EWMA do not significantly change over the time change points, implying that they are more robust than the scan statistic method with respect to the time change points. In addition, EWMA is more robust than CUSUM with respect to the point time that change occurs.

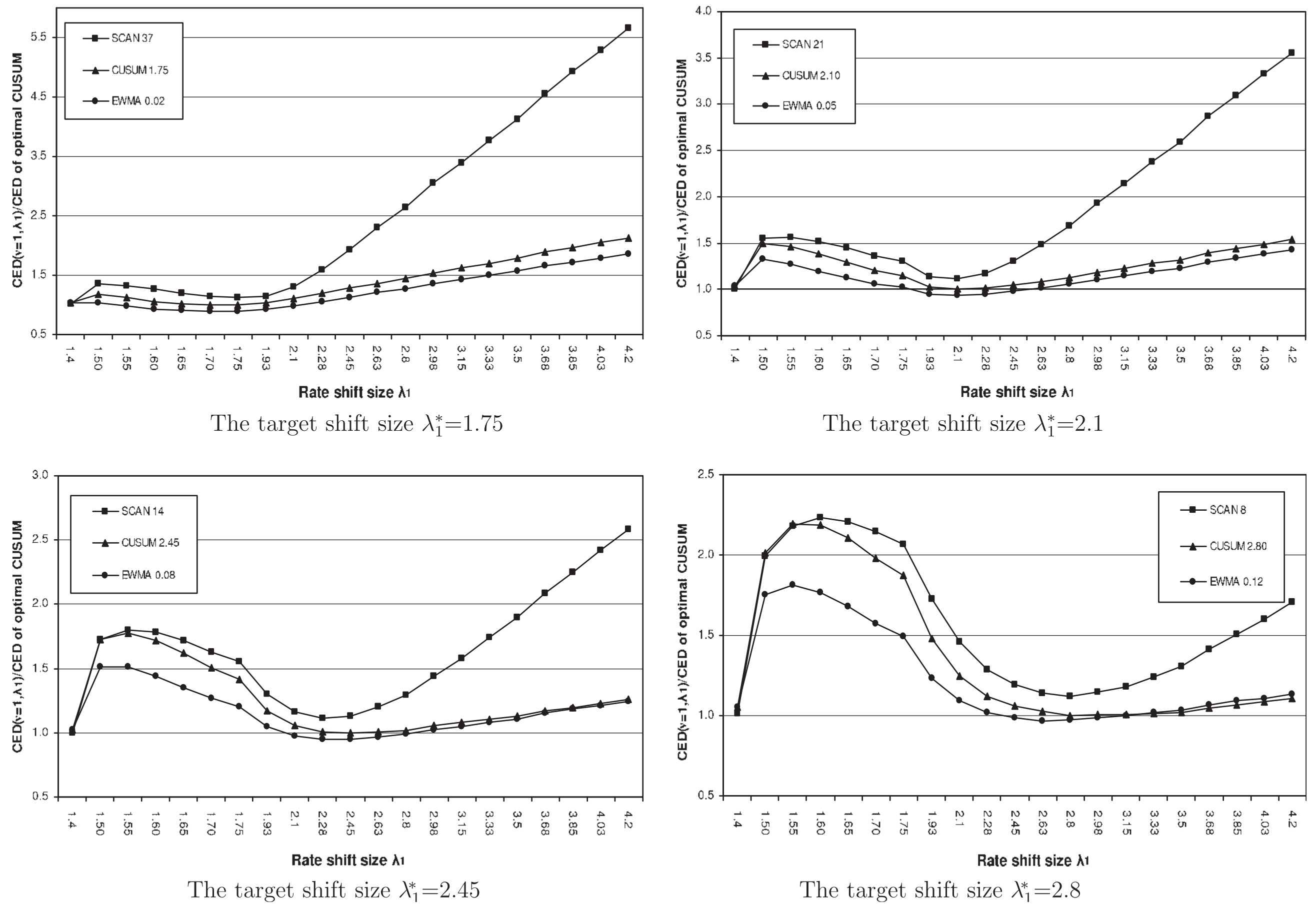
4.3. *Comparison of $CED(v, \lambda_1)$s under different shift sizes*

In the previous section, we investigated the performance of the three detection methods when λ_1 is known. This section contains the results of our investigation of their performance when λ_1 is unknown. In practice, if λ_1 is unknown, one might choose a set of parameters to detect the targeted shift size λ_1^* and investigate the pattern of *CED*s under a different true shift size λ_1 at a fixed change in time $v=v^*$, where v^* is some constant time point.

When we compare *CED*s between the different methods, there may be a scaling issue. In order to address this scaling problem, we used a scaled version of the *CED* value that can be computed by dividing the $CED^u(v=v^*, \lambda_1)$ in each detection method (*u*) by the $CED(v=v^*, \lambda_1)$ of the optimal CUSUM

$$\left(i.e. \frac{CED^u(v=v^*, \lambda_1)}{CED^{optimal.CUSUM}(v=v^*, \lambda_1)} \right)$$

Figures 2 and 3 display the scaled *CED* values of the scan statistic methods, CUSUM, and EWMA, over different values of λ_1 given $v^*=1$ and $v^*=1500$, respectively. Both figures showed that CUSUM and EWMA uniformly produced smaller scaled *CED* values than the scan statistic methods, implying that CUSUM and EWMA are superior to the scan statistic method for rapid detection of shifts of various sizes at both early (i.e. $v=1$) and later (i.e. $v=1500$) points in time that change occurs.



The target shift size λ_1^*=1.75 The target shift size λ_1^*=2.1

The target shift size λ_1^*=2.45 The target shift size λ_1^*=2.8

Figure 2. Comparison of the three detection methods with optimal parameters for a target shift size λ_1^* ($v=1$). Sets of parameters for the scan statistics, CUSUM, and EWMA for a target size λ_1^* and change times $v=1$ were used. The *x*-axis represents a true shift size (λ_1) and the *y*-axis represents $\frac{CED(v=1,\lambda_1)}{CED^{optimal.CUSUM}(v=1,\lambda_1)}$

In comparisons between EWMA and CUSUM, Figures 2 and 3 indicated that EWMA tended to perform better than CUSUM for a small shift with an early change in time, whereas CUSUM tended to perform better than EWMA for a large shift with a late change in time.

5. Example: the detection of increased rate in male thyroid cancer data

In addition to the simulation results, we considered how to apply these detection methods in order to assess their practical applications. Our data set contains the incidence of male thyroid cancer in New Mexico, 1973–2005, which is available through data from the Surveillance, Epidemiology, and End Results (SEER) Program at the National Cancer Institute (www.seer.cancer.gov/data/). The SEER program collects cancer incidence and mortality from the cancer registries in the United States. Figure 4 plots the annual incidence of thyroid cancer per 100 000 men.

The main goal of this application is to detect the change in rates as early as possible. It can be seen from Figure 4 that the rate increases after 1989 or so, and therefore, we assume that there are no shifts between 1973 and 1988. We used this steady-state period to estimate the baseline rate λ_0, which is $\hat{\lambda}_0 \approx 2$. We tried to detect a 25% increase of λ_0 to λ_1, which is the equivalent of the targeted shift size $\lambda_1^* = 2.5$.

We determined the parameters of each detection method as we did in the simulation (see Section 4.1). The target ARL_0 was set to 1000. The parameters were $\lambda_0^* = 2$ and $\lambda_1^* = 2.5$ for CUSUM, $m = 37$ for scan statistics, and $\alpha = 0.02$ for EWMA. Consequently, the thresholds of the three methods for target ARL_0 are 97 (scan statistics), 16.6 (CUSUM), and 2.33 (EWMA).

Figure 5 shows the statistics over time from the three detection methods. In order to ensure the comparability of the different methods and use the same threshold, we adjusted the values of the statistics of the scan statistic method and CUSUM by dividing them by 41.63 and 7.124, respectively.

It can be observed that the scan statistic, EWMA, and CUSUM methods trigger an alarm in 2004, 1999, and 2000, respectively. Assuming that early detection is desirable, EWMA and CUSUM triggered an alarm faster than the scan statistic method. This is consistent with our simulation results. Nevertheless, we cannot make any concrete conclusion about which methods perform best

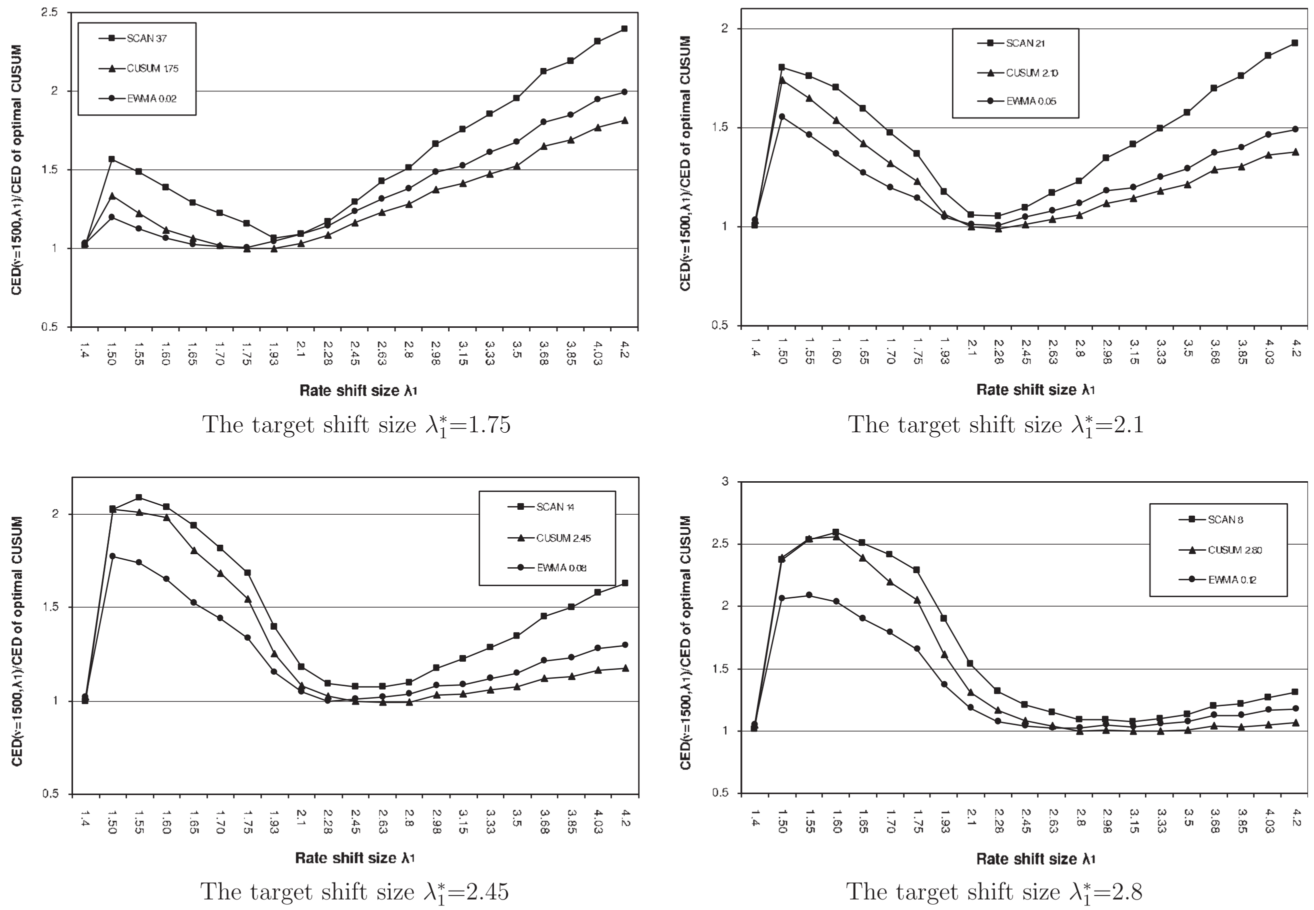


Figure 3. Comparison of the three detection methods with optimal parameters for a target shift size λ_1^* (ν=1500). Sets of parameters for the scan statistics, CUSUM, and EWMA for a target size λ_1^* and change times ν=1500 were used. The *x*-axis represents a true shift size (λ_1) and the *y*-axis corresponds to $\frac{CED(\nu=1500,\lambda_1)}{CED^{optimal.CUSUM}(\nu=1500,\lambda_1)}$

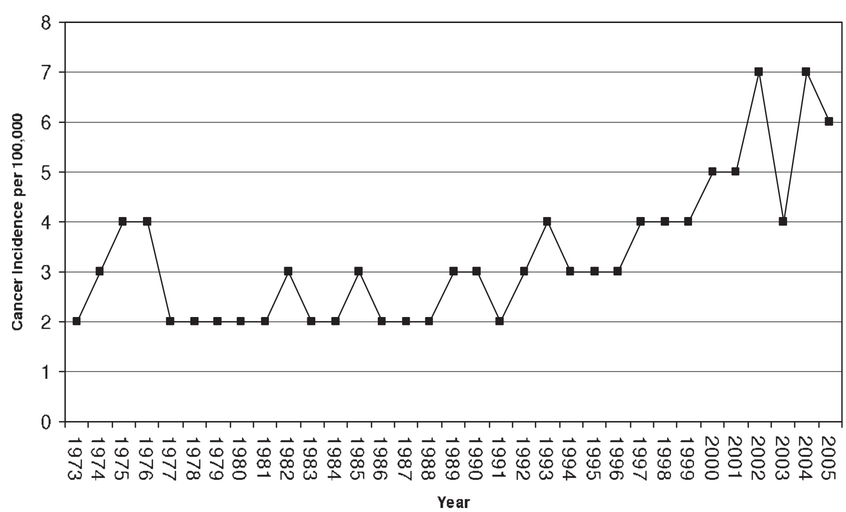


Figure 4. The trend of male thyroid cancer incidence between 1973 and 2005

because this real data set lacks information on when the shift actually occurred. The intention of this case study is to show that the three detection methods discussed here produce different results, but the overall result is consistent with the comparative performances determined in the simulation.

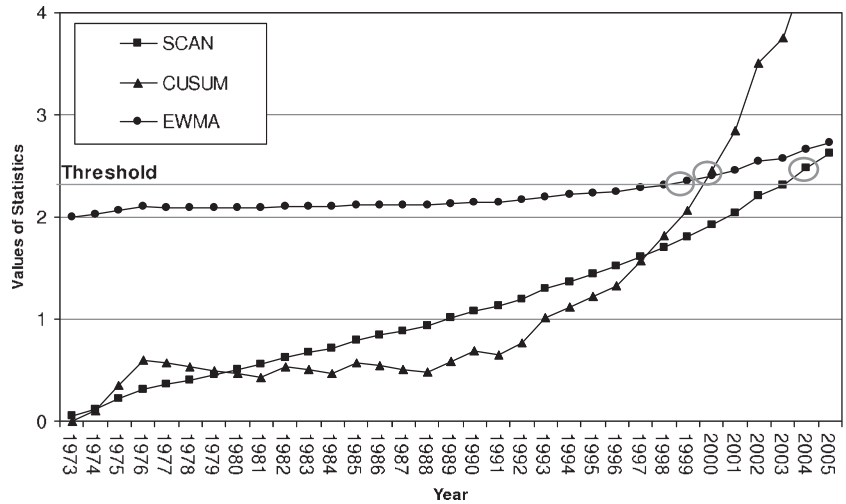


Figure 5. Plots of statistics of the scan statistic method with $m=37$, the CUSUM chart with $\lambda_0^*=2$ and $\lambda_1^*=2.5$, the EWMA chart with $\alpha=0.02$. The circle indicates the first time point when each method triggers an alarm

6. Conclusions

This paper investigates the properties of the scan statistic methods and CUSUM and EWMA charts when their observations follow the Poisson distribution and compares the performance of the three methods through simulation and a case study.

The results showed that the CUSUM and EMWA charts outperformed the scan statistic methods in Poisson cases. The simulation study revealed that the CUSUM and EWMA charts were better than the scan statistic method across all possible points of time for change and across a composite range of shift sizes λ_1. Given targeted shift sizes $\lambda_1^*(=\lambda_1)$, the CUSUM and EWMA charts outperformed the scan statistic methods across all the points of time at which they occurred. In addition, given the best methods for a targeted shift size λ_1^*, the CUSUM and EWMA charts were uniformly better than the scan statistic methods over different true shift sizes λ_1.

We also compared the CUSUM charts with the EWMA charts and obtained some interesting results from the simulation. The EWMA charts were slightly better than the CUSUM charts for a small shift size with changes of early points in time, and the CUSUM charts were better than the EWMA charts in the reverse situation.

There are interesting directions for the future research. The present study was performed with an assumption that the baseline parameter remains unchanged over time. In reality, however, the parameter under an in-control process often shifts and drifts over time because of changes in population or seasonal effects. For example, patient emergency room visits by persons with gastrointestinal symptoms or calls to nurse advice hotline from persons with respiratory symptoms may have seasonal patterns[23]. In other cases, the in-control rate depends on factors such as patient characteristics and methodological improvements. These cases require the application of a risk adjustment method[12, 46, 47]. Another important issue in health surveillance problems is that the quality of the baseline data can be readily contaminated by unexplained noise spikes. These noises may adversely affect the performance of the detection methods. Thus, efficient detection methods that can handle these noises should be developed.

In addition to baseline deviation, we need to study other types of outbreak patterns, especially transient outbreaks. Conventional detection methods have focused on detection of step shifts in mean. In health surveillance, many types of change can occur. Tsung and Tsui[48] showed that typical detection methods may be inefficient for detecting the special mean shift pattern such as a shift with finite duration. Shu *et al.*[49] proposed a weighted CUSUM chart in detecting shifts with patterned mean and found that it is efficient when the mean shift of time series data varies over time.

Acknowledgements

We thank the editor and the referees for their constructive comments and suggestions, which greatly improved the quality of the paper. This work was supported from DHHS/PHS/NIOSH/CDC/Center for Disease Control Award Number 200-2008-M-28123.

References

1. Woodall WH. The use of control charts in health-care and public health surveillance. *Journal of Quality Technology* 2006; **38**:89–104.
2. Rolka H, Burkom H, Cooper GF, Kulldorff M, Madigan D, Wong W-K. Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: research needs. *Statistics in Medicine* 2007; **26**:1834–1856.
3. Shmueli G, Burkom HS. Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics* (*Special Issue on Anomaly Detection*); in press.
4. Tsui K-L, Chiu W, Gierlich P, Goldsman D, Liu X, Maschek T. A review of healthcare, public health, and syndromic surveillance. *Quality Engineering* 2008; **20**:435–450.

5. Hutwagner L, Thompson GM, Seeman GM, Treadwell T. The bioterrorism preparedness and response early aberration reporting system (EARS). *Journal of Urban Health* 2003; **80**:i89–96.
6. Fricker RD Jr, Rolka HR. Protecting against biological terrorism: statistical issues in electronic biosurveillance. *Chance* 2006; **19**:4–13.
7. Fricker RD Jr. Syndromic surveillance. In *Encyclopedia of Quantitative Risk Assessment*, Melnick E, Everitt B (eds). Wiley: New York, 2008; 1743–1752.
8. Glaz J, Balakrishnan N (eds). *Scan Statistics and Applications*. Birkhäuser: Boston, MA, 1999.
9. Glaz J, Naus J, Wallenstein S. *Scan Statistics*. Springer: New York, 2001.
10. Balikrishina N, Koutras MV. *Runs and Scans with Applications*. Wiley: New York, 2002.
11. Naus J, Wallenstein S. Temporal surveillance using scan statistics. *Statistics in Medicine* 2006; **25**:311–324.
12. Joner MD, Woodall WH, Reynolds MR. Detecting a rate increase using a Bernoulli scan statistic. *Statistics in Medicine* 2008; **27**:2555–2575.
13. Woodall WH, Marshall JB, Joner MD, Fraker SE, Abdel-Salam G. On the use and evaluation of prospective scan methods for health-related surveillance. *Journal of the Royal Statistical Society, Series A* 2008; **171**:223–237.
14. Montgomery DC. *Statistical Quality Control* (5th edn). Wiley: New York, 2005.
15. Page ES. Continuous inspection schemes. *Biometrika* 1954; **41**:100–115.
16. Shiryayev AN. On optimum methods in quickest detection problems. *Theory of Probability and its Applications* 1963; **8**:22–46.
17. Lorden G. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics* 1971; **42**:1897–1908.
18. Pollak M. Optimal detection of a change in distribution. *Annals of Statistics* 1985; **13**:206–227.
19. Lai TL. Sequential analysis: some classical problems and new challenges. *Statistica Sinica* 2001; **11**:303–408.
20. Hill GB, Spicer CC, Weatherall JAC. The computer surveillance of congenital malformations. *British Medical Journal* 1968; **24**:215–218.
21. Weatherall JAC, Haskey JC. Surveillance of malformations. *British Medical Journal* 1976; **32**:39–44.
22. Brook D, Evan DA. An approach to the probability distribution of CUSUM run length. *Biometrika* 1972; **59**:539–549.
23. Lucas JM. Counted data CUSUMs. *Technometrics* 1985; **27**:129–144.
24. White CH, Keats JB. ARLs and higher order run length moments for Poisson CUSUM. *Journal of Quality Technology* 1996; **28**:363–369.
25. Hunter JS. The exponentially weighted moving average. *Journal of Quality Technology* 1986; **18**:203–210.
26. Crowder SV. A simple method for studying run length distributions of exponentially weighted moving average control charts. *Technometrics* 1987; **29**:401–407.
27. Crowder SV. Design of exponentially weighted moving average schemes. *Journal of Quality Technology* 1989; **21**:155–162.
28. Ng CH, Case KE. Development and evaluation of control charts using exponentially weighted moving averages. *Journal of Quality Technology* 1989; **21**:242–250.
29. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 1990; **32**:1–12.
30. Robinson PB, Ho TY. Average run lengths of geometric moving average charts by numerical methods. *Technometrics* 1978; **20**:85–93.
31. Shu L, Jiang W, Wu S. A one-sided EWMA control chart for monitoring process means. *Communications in Statistics—Simulation and Computation* 2007; **36**:901–920.
32. Gan FF. Monitoring Poisson observations using modified exponentially weighted moving average control charts. *Communications in Statistics—Simulation and Computation* 1990; **19**:103–124.
33. Borror CM, Rigdon SE, Champ CW. An exponentially weighted moving average control chart for Poisson data. *Proceedings of the 23rd Annual Modeling and Simulation Conference*, Pittsburgh, PA, 1992; 1775–1782.
34. Borror CM, Champ CW, Rigdon SE. Poisson EWMA control chart. *Journal of Quality Technology* 1998; **30**:352–361.
35. Joner MD, Woodall WH, Reynolds MR, Fricker RD. A one-sided MEWMA chart for health surveillance. *Quality and Reliability Engineering International* 2008; **24**:503–519.
36. Cowling BJ, Wong IOL, Ho LM, Riley S, Leung GM. Methods for monitoring influenza surveillance data. *International Journal of Epidemiology* 2006; **35**:1314–1321.
37. Jackson ML, Baer A, Painter I, Duchin J. A simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Medical Informatics and Decision Making* 2007; **7**:6.
38. Friker RD, Knitt MC, Hu CX. Comparing directionally sensitive MCUSUM and MEWMA procedures with application to biosurveillance. *Quality Engineering* 2008; **20**:478–494.
39. Yashchin E. Statistical control schemes: methods, applications and generalizations. *International Statistical Review* 1993; **61**:41–66.
40. Srivastava MS, Wu YH. Comparison of EWMA, CUSUM and Shiryayev–Roberts procedures for detecting a shift in the mean. *Annals of Statistics* 1993; **21**:645–670.
41. Srivastava MS, Wu YH. Evaluation of optimum weights and average run lengths in EWMA control schemes. *Communications in Statistics—Theory and Methods* 1997; **26**:1253–1267.
42. Moustakides GV. Optimal stopping times for detecting changes in distribution. *Annals of Statistics* 1986; **14**:1379–1387.
43. Basseville M, Nikiforov IV. *Detection of Abrupt Changes—Theory and Application*. Prentice-Hall: Englewoood Cliffs, NJ, 1993.
44. Sonesson C. A CUSUM framework for detection of space–time disease clustering using scan statistics. *Statistics in Medicine* 2007; **26**:4770–4789.
45. Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society, Series A* 2001; **164**:61–72.
46. Steiner SH, Cook RJ, Farewell VT, Treasure T. Monitoring surgical performance using risk-adjusted cumulative sum charts. *Biostatistics* 2000; **1**:441–452.
47. Grigg O, Farewell V. An overview of risk-adjusted charts. *Journal of the Royal Statistical Society, Series A* 2004; **167**:523–539.
48. Tsung FG, Tsui K-L. A mean-shift pattern study on integration of SPC and APC for process monitoring. *IIE Transactions* 2003; **35**:231–242.
49. Shu LJ, Jiang W, Tsui K-L. A weighted CUSUM chart for detecting patterned mean shifts. *Journal of Quality Technology* 2008; **40**:194–213.

Appendix A: obtaining target ARL_0 from discrete statistics

We explain why $ARL_0 = E[T|\nu=\infty]$ in CUSUM charts is a step function with respect to h. The condition sufficient for the statement is that if S_i is discrete and $S_{(j)}$ is the order statistics of $\{S_i, i=1,2,\ldots\}$, all $T(h)$ are identical for $h \in (S_{(j-1)}, S_{(j)})$, where $T(h) = \inf_n\{n : S_n \geq h\}$.

Suppose that $S_n \geq S_{(j)}$. Consequently, $S_n \geq h$, because $S_{(j-1)} < h < S_{(j)}$. Thus,

$$T(S_{(j)}) = \inf_n\{n : S_n \geq S_{(j)}\} \geq T(h) = \inf_n\{n : S_n \geq h\} \quad \text{for } h \in (S_{(j-1)}, S_{(j)}) \tag{A1}$$

If $S_n \geq h$, then $S_n > S_{(j-1)}$. S_n is discrete, and $S_{(j-1)}$ and $S_{(j)}$ are adjacent-order statistics. This results in $S_n \geq S_{(j)}$. Thus,

$$T(S_{(j)}) = \inf_n\{n : S_n \geq S_{(j)}\} \leq T(h) = \inf_n\{n : S_n \geq h\} \quad \text{for } h \in (S_{(j-1)}, S_{(j)}) \tag{A2}$$

By (A1) and (A2), $T(h) = T(S_{(j)})$ for all $h \in (S_{(j-1)}, S_{(j)})$. Therefore, $T(h)$ is identical for all $h \in (S_{(j-1)}, S_{(j)})$.

Authors' biographies

Sung Won Han is a PhD candidate in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He received an MS in Operations Research in 2006 and an MS in Statistics in 2007 from the Georgia Institute of Technology. His research interest is the sequential change-point detection problem in bio and health care application regarding temporal and spatiotemporal surveillance.

Kwok-Leung Tsui is professor in Industrial and Systems Engineering at Georgia Tech. He received his PhD in Statistics from the University of Wisconsin. He was a recipient of the NSF Young Investigator Award in 1992. He is a fellow of the American Statistical Association (ASA) and a US representative in the ISO Technical Committee. He was President of the ASA Atlanta Chapter in 1992; Chair of the INFORMS Section in Quality, Statistics, and Reliability in 2000; and Chair of the Data Mining Section in 2004. His research interests include data mining, health surveillance, computer model validation, bioinformatics, process control, and Taguchi method.

Bancha Ariyajunya is a PhD student in Industrial and Manufacturing Systems Engineering at the University of Texas at Arlington. He was a lecturer in Industrial Engineering at Burapha University, Thailand. He received a BS from Burapha University, Thailand in 1999 and an MS in Manufacturing Systems Engineering from the Asian Institute of Technology, Thailand in 2002. His current research interests include statistical process control and data mining algorithms.

Seoung Bum Kim is an assistant professor of Industrial Systems and Information Engineering at the Korea University, Seoul, Korea. From 2005–2009, he was an assistant professor in Industrial and Manufacturing Systems Engineering at the University of Texas at Arlington. He received an MS in Industrial and Systems Engineering in 2001, an MS in Statistics in 2004, and a PhD in Industrial and Systems Engineering in 2005 from the Georgia Institute of Technology. He was awarded the Jack Youden Prize as the best expository paper in Technometrics. Dr Kim's research interests include statistical process control, data mining algorithms, functional data analysis, and bioinformatics.