Some Zero Inflated Poisson-Based Combined Exponentially Weighted Moving Average Control Charts for Disease Surveillance

Robert Neil F. Leong
Mathematics Department, De La Salle University

Frumencio F. Co
Mathematics Department, De La Salle University

Daniel Stanley Y. Tan
Software Technology Department, De La Salle University

One of the main areas of public health surveillance is infectious disease surveillance. With infectious disease backgrounds usually being more complex, appropriate surveillance schemes must be in order. One such procedure is through the use of control charts. However, with most background processes following a zero-inflated Poisson (ZIP) distribution as brought about by the extra variability due to excess zeros, the control charting procedures must be properly developed to address this issue. Hence in this paper, drawing inspiration from the development of combined control charting procedures for simultaneously monitoring each ZIP parameter individually in the context of statistical process control (SPC), several combined exponentially weighted moving average (EWMA) control charting procedures were proposed (Bernoulli-ZIP and CRL-ZTP EWMA charts). Through an extensive simulation study involving multiple parameter settings and outbreak model considerations (i.e., different shapes, magnitude, and duration), some key results were observed. These include the applicability of performing combined control charting procedures for disease surveillance with a ZIP background using EWMA techniques. For demonstration purposes, application with an actual data, using confirmed measles cases in the National Capital Region (NCR) from January 1, 2010 to January 14 2015, revealed the comparability of the Bernoulli-ZIP EWMA scheme to historical limits method currently in use.

Keywords: EWMA control charts, disease surveillance, ZIP distribution, measles

1. Introduction

Control charts have long been in use in healthcare and public health surveillance. One specific application is on the prospective surveillance of diseases with the aim of early outbreak detections (Unkel et al., 2012; Woodall, 2006). In particular, the earliest temporal outbreak detection algorithms are variants of several control charting procedures. Also, while the health surveillance literature has now advanced to include new methodologies as motivated by regression and time series techniques, control charts are still widely used in most general settings because of their simplicity and ease-of-implementation (Fricker, 2013; Unkel et al., 2012).

However, despite the long history of applications of control charting procedures in public health surveillance, only a few were developed for infectious disease surveillance (Woodall, 2006). The complexity of developing surveillance procedures for infectious diseases is because of the problems associated with on-line outbreak detection: reporting delays, background incidence characterization, disease dynamics, etc. among others (Farrington and Andrews, 2004; Shmueli and Burkom, 2010). Nevertheless, the development of prospective surveillance procedures directly inspired by control charting procedures have garnered interest because of the parallelism between the problems involved – detecting outbreaks temporally is similar to that of detecting out-of-control shifts in industrial manufacturing processes. Particularly, one of the most common variants of statistical process control (SPC) procedures studied for disease surveillance are the Exponentially Weighted Moving Average (EWMA) control charts. EWMA control charts, first studied by Dr. S. W. Roberts in 1959, have gained substantial interest due to their simplicity and their ability to balance the optimal properties and the shortcomings of both Shewhart and Cumulative Sum (CUSUM) charts.

Applications of EWMA charts in prospective disease surveillance have first been loosely used in the development of the numerical derivative-based method by Nobre and Stroup (1994) for monitoring measles in the United States. Extensive studies by Burkom et al. (2007) and Elbert and Burkom (2009) proposed the use of the Holt-Winters technique for generalized exponential smoothing to account for trends and seasonal features. Moreover, Dong et al. (2008) proposed three types of EWMA charts that do not require an assumption of identical distributions of counts to detect positive shifts in incidence rates.

Empirical evaluations of EWMA charts have revealed their comparability with CUSUM variants in terms of sensitivity in detecting small shifts in process mean (Wong and Moore, 2006). Furthermore, Han et al. (2010) showed that for appropriately selected parameters, EWMA charts generally perform well for detecting shifts in Bernoulli observations. Also, Szarka et al. (2011) concluded that EWMA methods using adaptive thresholds generally performed better than most procedures in the Early Aberration Reporting System (EARS) of the Center for Disease Control (CDC). In the context of real-time monitoring of hospital

acquired infections, Morton et al. (2001) concluded that EWMA charts are ideal for monitoring bacteraemia and multi-resistant organism rates.

In spite of a long list of literature in the applications of SPC procedures in on-line disease surveillance, none in the aforementioned list has considered EWMA variants based on a ZIP distribution despite the knowledge that disease counts may contain excess zeros particularly for rare diseases (Ngatchou-Wandji and Paris, 2011). The lack of development of ZIP-based monitoring schemes in infectious disease surveillance may be due to the later developments of ZIP-based control charting procedures. A more recent development in the utilization of ZIP-based chart in disease surveillance is the development of an EWMA chart by Fatahi et al. (2012) to monitor unintentional needle-stick occurrences in Tehran. However, a perceived drawback of such is that it only monitors shifts in the mean process level, as opposed to the knowledge that a ZIP process is characterized by two parameters (π: vulnerability level, and λ: intensity level). Hence, a combination of control charting procedures must be ideal to simultaneously monitor the parameters of a ZIP process individually. The rationale behind simultaneous monitoring is to better identify the nature of the shift in the mean disease incidence level when it occurs. While this problem has never been tackled in any disease surveillance literature, several studies in the context of SPC have considered this before. In particular, as motivated by the Generalized ZIP model of Chen et al. (2008), He et al. (2012) developed the p–λ CUSUM chart, based on the combination of individual CUSUM charts. He et al. (2014) likewise developed the CRL-ZTP CUSUM, which is a combination of two CUSUM charts based on conforming run length to detect shifts in the occurrence of random shocks and on a zero-truncated Poisson distribution to monitor the magnitude of the occurrences of the random shocks.

Inspired by these developments of combined control charting procedures, this paper aims to showcase the usage of combination of control charting procedures to simultaneously detect increases in both parameters of a ZIP process in the context of disease surveillance. In particular, the paper aims to develop a combination of EWMA control charting procedures based on existing individual EWMA charts for a Bernoulli parameter and a Poisson parameter for simultaneously monitoring parameters of a ZIP process. It is desired that the results of this study will contribute to the literature of using control charts in performing disease surveillance, with an aim of developing improved early warning systems for outbreaks.

The next section provides a short discussion of the ZIP distribution followed by the development of two combined EWMA control charting procedures. A simulation study follows in the next section to provide empirical performance evaluations of the proposed charts in the context of disease surveillance. The fourth section then presents a case application of the proposed procedures with actual measles data in comparison with the historical limits procedure which is

currently in use. Lastly, conclusions and recommendations for future works are provided.

2. Conceptual Framework

2.1. Zero-inflated Poisson distribution

A random variable Y is defined to have a zero-inflated Poisson distribution if its probability mass function is given by

$$f(y)=\begin{cases}1-\pi+\pi e^{-\lambda} & , y=0\\ \pi\dfrac{e^{-\lambda}\lambda^{y}}{y!} & , y=1,2,3,...'\end{cases} \tag{1}$$

where $0 \leq \pi \leq 1$ is the probability that a random shock is observed (where in the context of disease surveillance will be known as the vulnerability level) and $\lambda > 0$ is the Poisson intensity of the shocks as they occur (where in the same context will be known as the intensity level). Clearly, $1-\pi$ is the excess proportion of zeros causing the extra poisson variation (i.e., overdispersion). Some of its moments are as follows:

$$E(Y) = \lambda\pi \tag{2}$$

$$Var(Y) = \lambda\pi\,(\lambda + 1 - \lambda\pi). \tag{3}$$

It can be shown that the maximum likelihood estimator of π is given by

$$\hat{\pi}_{mle}=\frac{\sum_{i=1}^{n}\left(1-I_0\left(y_i\right)\right)\Big/n}{1-e^{-\hat{\lambda}_{mle}}}, \tag{4}$$

where $I_0\,(y_i)$ is the indicator function for $y_i = 0$. Likewise, it can be shown that the maximum likelihood estimator of λ is the solution to the nonlinear equation

$$\frac{\hat{\lambda}_{mle}}{1-e^{-\hat{\lambda}_{mle}}}=\frac{\sum_{i-1}^{n}y_i\left(1-I_0(y)\right)}{\sum_{i-1}^{n}\left(1-I_0(y_i)\right)}. \tag{5}$$

Note that $\sum_{i=1}^{n}y_i\left(1-I_0(y_i)\right)$ is the sum of all nonzero observations and $\sum_{i=1}^{n}\left(1-I_0(y_i)\right)$ is the number of nonzero observations in the sample.

2.2 Bernoulli-ZIP EWMA chart

An initial attempt in using EWMA control charting procedure for ZIP-distributed data was developed by Fatahi et al. (2012). The control chart was called ZIP-EWMA control chart, specifically developed to monitor rare health-related events. The chart monitors the statistic

$$E_t = \max[0, (1-\kappa)\, E_{t-1} + \kappa Y_t], \quad t = 1,2,3,\ldots, \tag{6}$$

where Y_t is the observed number of cases at time t, $0 \leq \kappa \leq 1$, being the weighting parameter, and $E_0 = E(Y) = \lambda\pi$ as an initialization procedure. The ZIP-EWMA chart signals whenever $E_t > h_{ZIP\text{-}EWMA}$, where $h_{ZIP\text{-}EWMA}$ is the steady-state threshold taken as

$$h_{ZIP-EWMA} = \lambda\pi + L_{ZIP-EWMA}\sqrt{\left(\frac{\kappa}{2-\kappa}\right)\left[\lambda\pi(\lambda+1-\lambda\pi)\right]}, \tag{7}$$

and $L_{ZIP\text{-}EWMA}$ is appropriately chosen to achieve the desired average false signal rate.

However, the ZIP-EWMA control charting procedure was developed only to monitor shifts in λ. Thus, to augment its ability to detect shifts in π, it is proposed to be utilized simultaneously with an EWMA chart for Bernoulli data. In this case, the EWMA control chart to be used for Bernoulli data will be called Bernoulli-EWMA control chart, wherein the Bernoulli data are generated through the indicator function $I_t(y \geq 1)$ which takes on the value 1 whenever the observed number of cases y_t is at least one. Note that some of the moments of $I_t(y \geq 1)$ are as follows:

$$E(I_t(y \geq 1)) = \pi(1-e^{-\lambda}), \tag{8}$$

$$Var(I_t(y \geq 1)) = \pi(1-e^{-\lambda})[1-\pi(1-e^{-\lambda})]. \tag{9}$$

Thus, the Bernoulli-EWMA chart is designed to monitor the statistic

$$F_t = \max\,[0, (1-\kappa)\, F_{t-1} + \kappa I_t(y \geq 1)], \; t = 1,2,3,\ldots \tag{10}$$

whereas an initialization procedure, $F_0 = E(I_t(y \geq 1)) = \pi(1-e^{-\lambda})$, and the Bernoulli-EWMA chart signals whenever $F_t > h_{Ber\text{-}EWMA}$, where $h_{Ber\text{-}EWMA}$ is the steady-state threshold taken as

$$h_{Ber-EWMA} = \pi(1-e^{-\lambda}) + L_{Ber-EWMA}\sqrt{\left(\frac{\kappa}{2-\kappa}\right)\left\{\pi\left(1-e^{-\lambda}\right)\left[1-\pi\left(1-e^{-\lambda}\right)\right]\right\}}, \tag{11}$$

and $h_{Ber\text{-}EWMA}$ is appropriately chosen to achieve the desired average false signal rate.

In this paper, a combined EWMA control charting scheme is proposed to simultaneously monitor shifts in λ (based on the ZIP-EWMA chart) and in π (based on the Bernoulli-EWMA chart). This control charting procedure will be called as the Bernoulli-ZIP EWMA chart, and is made to generate a signal whenever one of the individual EWMA chart signals.

2.3 CRL-ZTP EWMA Chart

As an alternative EWMA procedure to monitor upward shifts in λ, a zero-truncated Poisson (ZTP)-based procedure is proposed following that of He et al. (2014) which viewed the ZIP process as a combination of Bernoulli and ZTP processes. That is, for the proposed ZTP-EWMA control chart, the procedure monitors the statistic

$$W_t = \max[0,(1-\kappa)W_{t-1} + \kappa Y_t], \;\; t = 1,2,3,\ldots, \tag{12}$$

where Y_t is the observed number of cases at time t assumed to follow a ZTP distribution with parameter λ, and $W_0 = E(Y) = \dfrac{\lambda}{1-e^{-\lambda}}$ as an initialization procedure. Here, it should be noted that W_t updates only whenever nonzero cases are observed in time t. The ZTP-EWMA chart signals whenever $W_t > h_{ZTP\text{-}EWMA}$, where $h_{ZTP\text{-}EWMA}$ is the steady-state threshold taken as

$$h_{ZTP-EWMA} = \frac{\lambda}{1-e^{-\lambda}} + L_{ZTP-EWMA}\sqrt{\left(\frac{\kappa}{2-\kappa}\right)\left[\frac{\lambda}{1-e^{-\lambda}}\left(1-\frac{\lambda e^{-\lambda}}{1-e^{-\lambda}}\right)\right]}, \tag{13}$$

and $L_{ZTP-EWMA}$ is appropriately chosen to achieve the desired average false signal rate.

To be able to detect shifts in π, the CRL-EWMA chart is proposed to be simultaneously utilized. The CRL-EWMA chart is a Bernoulli-type EWMA chart based on geometric conforming run lengths (CRL) as developed by Mavroudis and Nicolas (2013). The CRL-EWMA chart is an extension of the EWMA chart of Yeh et al. (2008) with increased detection sensitivity. That is, only the accumulation of positive deviations from the target are considered and is based on the CRL, which is assumed to follow a geometric distribution with mean $\mu_{CRL} = 1/p_0$, between nonzero cases. The control charting procedure transforms each CRL count into

$$CRL^{-} = \min\,[\mu_{CRL}, CRL] = \mu_{CRL} - \max[0, \mu_{CRL} - CRL\,]. \tag{14}$$

Mavroudis and Nicolas (2013) have shown that some of the moments of CRL^{-} are as follows:

$$E(CRL^{-}) = \frac{1 - 2p_0 - (1-p_0)^{\frac{1+p_0}{p_0}}}{p_0(1-p_0)}, \tag{15}$$

$$E\left(\left[CRL^{-}\right]^2\right) = \frac{p_0^2 - 4p_0 + 2 + (1-p_0)^{\frac{1}{p_0}}(-p_0^2 + 5p_0 - 4)}{p_0^2(1-p_0)}, \tag{16}$$

$$Var(CRL^{-}) = E\left(\left[CRL^{-}\right]^2\right) - \left[E(CRL^{-})\right]^2. \tag{17}$$

Hence, the chart monitors the statistic

$$G_t = \max\left[0, (1-\kappa)G_{t-1} + \kappa CRL_t^{-}\right], \quad t = 1, 2, 3, \ldots \tag{18}$$

where CRL_t^{-} is the transformed conforming run length at time t assuming nonzero cases were observed at time t and $G_0 = E(CRL^{-})$ as an initialization procedure. Likewise, note that G_t only updates whenever nonzero cases are observed in time t. The CRL-EWMA chart signals whenever $G_t < h_{CRL-EWMA}$, where $h_{CRL-EWMA}$ is the steady-state threshold taken as

$$h_{CRL-EWMA} = E(CRL^{-}) - L_{ZTP-EWMA}\sqrt{\left(\frac{\kappa}{2-\kappa}\right)Var(CRL^{-})}, \tag{19}$$

and $L_{CRL-EWMA}$ is appropriately chosen to achieve the desired average false signal rate.

Similar to the other combined charting procedure, this control charting procedure, to be called the *CRL-ZTP EWMA chart* henceforth, will signal whenever one of the individual EWMA chart signals.

3. Simulation Study

3.1. Simulation Framework

To empirically compare the proposed combined EWMA charting procedures in the context of disease surveillance, a simulation study was performed using MATLAB routines. The algorithms are available from the authors upon request. The study was done as follows:

1. From a ZIP distribution as characterized by each π–λ parameter combination (i.e., the disease background) from the sets π:{0.30, 0.60} (i.e., low vulnerability, high vulnerability) and λ:{1.0, 4.0} (i.e., low intensity, high intensity), 1,000 data sets, each having a length of 1,000 time points, were simulated. The choice of these parameter sets are based on the recommendations by Sim and Lim (2008), which are typically found in those of biological applications. Each simulated data set was subjected to a $\chi 2$ goodness-of-fit test to ensure that the particular data followed the desired distribution.

2. The first 250 observations of each series were used to determine the parameter settings of each of the control charting procedure, i.e., the Phase I procedure. That is, for the EWMA charts, in-control π_0 was estimated by $\hat{\pi}_{mle}$ (see Equation (4)), while in-control λ_0 was estimated by $\hat{\lambda}_{mle}$ (see Equation (5)). Hence, for the Bernoulli-EWMA chart,

$$E\left(I_t(\widehat{y \geq 1})\right) = \hat{\pi}_{mle}\left(1 - e^{-\hat{\lambda}_{mle}}\right), \tag{20}$$

which is equivalent to the sample proportion of nonzero cases. For the CRL-EWMA control chart, CRLs were computed based on the notion of time between *signal events*, defined by Fraker et al. (2008) as two consecutive time periods wherein the charting procedure generated signals. Thus, p_0 was estimated as

$$\widehat{p_0} = \left[\hat{\pi}_{mle}(1 - e^{-\hat{\lambda}_{mle}})\right]^2, \tag{21}$$

i.e., the square of the sample proportion of nonzero cases. For the weighting parameter κ, two settings were considered, namely $\kappa = 0.25$ (i.e., relatively small weight on most recent value) and $\kappa = 0.45$ (i.e., relatively large weight on most recent value). The former is based on the usual setting done in the application of EWMA charts in infection control (Fatahi et al., 2012), while the latter to only serve as a comparison when κ is large.

3. To ensure comparability of the proposed schemes with regard to their abilities to detect parameter shifts, the average time between false signals (ATFS) was preset at 90 for each combined charts. That is, appropriate thresholds for each chart were determined to achieve the desired ATFS values via simulation. For each data set, the charting schemes were applied for the remaining 750 observations as a Phase II procedure. Following Fricker (2013), the estimated ATFS for each data set is computed as

$$\widehat{ATFS} = \frac{750}{\text{number of signals generated}} \tag{22}$$

For such ATFS formula to be applicable for charting schemes with memory, the statistics for such charts were designed to reset whenever a signal was generated. Based on the desired ATFS level, thresholds for each combined charting procedure was adjusted accordingly. The calibrated threshold parameters which achieved such ATFS are summarized in Tables 1 and 2 for the Bernoulli-ZIP EWMA and the CRL-ZTP EWMA charts, respectively. It is worth noting that for a combined mean ATFS of 90, individual mean ATFS values for both π– and λ– designed charts ranged from 200 to 250.

Table 1. Threshold Parameters for Bernoulli-ZIP EWMA Charts for ATFS = 90

ZIP Parameters		Threshold Parameter Values			
		κ = 0.25		κ = 0.45	
π_0	λ_0	L_π	L_λ	L_π	L_λ
0.30	1.0	2.8113	3.1366	2.9037	3.7901
0.30	4.0	2.5483	2.7960	2.4317	3.1301
0.60	1.0	2.3548	2.7885	2.1300	3.2568
0.60	4.0	1.8928	2.4516	1.4616	2.6546

Table 2. Threshold Parameters for CRL-ZTP EWMA Charts for ATFS = 90

ZIP Parameters		Threshold Parameter Values			
		κ = 0.25		κ = 0.45	
π_0	λ_0	L_π	L_λ	L_π	L_λ
0.30	1.0	0.4261	6.2749	0.7875	5.4816
0.30	4.0	0.8403	5.0894	1.1953	4.3238
0.60	1.0	0.6801	4.6344	1.0388	4.1749
0.60	4.0	0.0001	3.1460	0.0001	2.8630

4. Outbreaks, as determined by shifts in either one of the ZIP parameters or both, were injected in three phases: phase one (time periods 251-500), phase two (time periods 501-750), and phase three (time periods 751-1,000). These outbreaks were determined by shifting the parameter magnitudes upward. Three shapes were considered, namely spike, triangular, and ramp. Two intensity levels (M_θ) were determined for each parameter: small (M_λ = +3; M_π = +25%) and large (M_λ = +6; M_π = +50%). Two outbreak durations (D) were considered, namely short (D = 15) and long (D = 30). For each intensity-duration (M–D) combination, random starts for the injected outbreaks for different phases were determined. For each shape-intensity-

length combination, three scenarios were considered, namely (Scenario 1) only λ increased, (Scenario 2) only π increased, and (Scenario 3) both λ and π increased. It is emphasized that the outbreaks are assumed to be stochastic and additive on the mean background disease process. Moreover, the outbreak equations added to the simulated disease background data followed those which were discussed in Fricker (2013).

5.) The charting procedures were empirically evaluated using several performance metrics discussed by Sonesson and Bock (2003) and Fricker (2013), with emphasis on their ability to properly detect the injected outbreaks. In particular, charts with lower conditional expected delay (CED) value and higher probability of successful detection (PSD), proportion of outbreaks detected (POD), and probability of true detection (PTD) values are determined to perform better. For each outbreak phase, the metrics were computed as follows:

$$\widehat{PSD} = \begin{cases} 1, \text{if the chart signaled at least once during the outbreak} \\ 0. \text{ otherwise} \end{cases}, \tag{23}$$

$$\widehat{CED} = \text{time of first true signal} - \text{time of start of outbreak}, \tag{24}$$

$$\widehat{POD} = \frac{\text{total number of signals generated within the outbreak}}{\text{total outbreak duration}}, \tag{25}$$

$$\widehat{PTD} = \frac{\text{total number of signals generated within the outbreak}}{\text{total number of signals generated during the phase}}, \tag{26}$$

As numerical summaries, means were computed. Also, for comparison purposes, 95% confidence intervals for the mean values were estimated using bias-corrected bootstrapping procedure with 1,000 replications.

3.2. Simulation results and discussion

In this section, results of the simulation study are presented and discussed. To simplify the presentation, chart legends are used which are given in Table 3.

Table 3. Chart Legends

Legend	Chart Description
A	Bernoulli-ZIP EWMA Chart ($\kappa = 0.25$)
B	Bernoulli-ZIP EWMA Chart($\kappa = 0.45$)
C	CRL-ZTP EWMA Chart($\kappa = 0.25$)
D	CRL-ZTP EWMA Chart($\kappa = 0.45$)

Table 4 provides the means, as well as the 95% confidence intervals, of performance metrics of the four charts considered for small short spike outbreaks with different ZIP backgrounds. As a summary, all charts were able to soundly identify outbreaks defined by any shift scenarios, except the cases wherein outbreaks were only defined by shifts in π for disease backgrounds with smaller π_0. Noticeably, mean PSDs were almost 100.00% (i.e., almost perfect detection success) when disease backgrounds had larger π_0. Meanwhile, CEDs were typically longer (hence, slower detections) for outbreaks defined by shifts in π. Lastly, PODs and PTDs were generally higher when outbreaks were defined by shifts in both λ and π, but were observably lowest when outbreaks were only defined by shifts in π.

Table 4. Performance Metrics for Small Short Spike Outbreaks

Shifts	Chart	PSD		CED		POD		PTD	
		mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
π_0 = 0.30 and λ_0 = 1.0									
λ only	A	96.87%	(96.27%, 97.40%)	4.4580	(4.3318, 4.5759)	26.49%	(25.85%, 27.06%)	52.58%	(51.74%, 53.46%)
	B	97.03%	(96.43%, 97.60%)	4.2943	(4.1889, 4.4188)	24.65%	(24.16%, 25.14%)	52.54%	(51.78%, 53.40%)
	C	96.80%	(96.17%, 97.37%)	4.8533	(4.7328, 4.9961)	20.12%	(19.69%, 20.60%)	46.41%	(45.54%, 47.20%)
	D	97.03%	(96.40%, 97.67%)	4.4367	(4.3166, 4.5704)	22.43%	(21.95%, 22.92%)	50.38%	(49.62%, 51.26%)
π only	A	66.50%	(64.88%, 68.23%)	7.3037	(7.1317, 7.4963)	11.85%	(11.36%, 12.36%)	31.97%	(30.94%, 32.98%)
	B	65.03%	(63.23%, 66.87%)	7.1910	(7.0112, 7.3968)	9.12%	(8.74%, 9.46%)	28.18%	(27.23%, 29.14%)
	C	63.20%	(61.46%, 64.83%)	7.4262	(7.2208, 7.6172)	9.24%	(8.86%, 9.59%)	27.68%	(26.83%, 28.72%)
	D	63.07%	(61.35%, 64.70%)	7.2894	(7.1042, 7.4930)	9.12%	(8.77%, 9.50%)	28.34%	(27.35%, 29.31%)
both λ and π	A	98.93%	(98.53%, 99.30%)	3.9068	(3.7908, 4.0187)	33.99%	(33.34%, 34.69%)	58.14%	(57.39%, 58.90%)
	B	99.03%	(98.70%, 99.40%)	3.7800	(3.6741, 3.8879)	30.70%	(30.18%, 31.26%)	57.84%	(57.07%, 58.52%)
	C	98.73%	(98.22%, 99.10%)	4.2482	(4.1471, 4.3546)	25.67%	(25.19%, 26.18%)	51.97%	(51.24%, 52.71%)
	D	98.97%	(98.53%, 99.30%)	3.8703	(3.7592, 3.9717)	28.39%	(27.87%, 28.96%)	55.89%	(55.18%, 56.63%)
π_0 = 0.30 and λ_0 = 4.0									
λ only	A	72.67%	(70.97%, 74.32%)	6.9577	(6.7816, 7.1342)	12.98%	(12.52%, 13.44%)	34.20%	(33.19%, 35.17%)
	B	72.43%	(70.97%, 73.99%)	6.7061	(6.5306, 6.8928)	11.78%	(11.36%, 12.16%)	32.03%	(31.11%, 33.14%)
	C	69.77%	(68.07%, 71.30%)	7.1824	(7.0002, 7.3531)	9.59%	(9.23%, 9.97%)	28.82%	(27.87%, 29.70%)
	D	72.07%	(70.33%, 73.57%)	6.8442	(6.6612, 7.0257)	9.06%	(8.77%, 9.36%)	29.27%	(28.47%, 30.17%)

π only	A	45.87%	(44.13%, 47.62%)	7.6677	(7.4206, 7.9091)	7.22%	(6.83%, 7.61%)	21.58%	(20.58%, 22.58%)
	B	47.23%	(45.27%, 49.07%)	7.6135	(7.3888, 7.8452)	6.40%	(6.08%, 6.74%)	19.60%	(18.66%, 20.54%)
	C	45.33%	(43.50%, 47.00%)	7.8395	(7.5971, 8.0943)	5.44%	(5.15%, 5.74%)	18.55%	(17.73%, 19.47%)
	D	46.53%	(44.87%, 48.27%)	7.7007	(7.4783, 7.9499)	4.75%	(4.50%, 4.98%)	17.72%	(16.87%, 18.47%)
both λ and π	A	82.23%	(80.90%, 83.60%)	6.6640	(6.5062, 6.8327)	17.70%	(17.12%, 18.26%)	40.90%	(39.88%, 42.08%)
	B	81.87%	(80.24%, 83.17%)	6.4268	(6.2598, 6.6207)	15.78%	(15.23%, 16.30%)	38.30%	(37.36%, 39.27%)
	C	79.83%	(78.37%, 81.23%)	6.9173	(6.7478, 7.0756)	12.74%	(12.33%, 13.19%)	34.68%	(33.83%, 35.57%)
	D	81.03%	(79.59%, 82.39%)	6.7195	(6.5516, 6.8820)	11.72%	(11.41%, 12.12%)	34.52%	(33.70%, 35.29%)

π_0 = 0.60 and λ_0=1.0									
λ only	A	100%	(44.13%, 47.62%)	6.9577	(7.4206, 7.9091)	12.98%	(6.83%, 7.61%)	34.20%	(20.58%, 22.58%)
	B	100.%	(45.27%, 49.07%)	6.7061	(7.3888, 7.8452)	11.78%	(6.08%, 6.74%)	32.03%	(18.66%, 20.54%)
	C	100%	(43.50%, 47.00%)	7.1824	(7.5971, 8.0943)	9.59%	(5.15%, 5.74%)	28.82%	(17.73%, 19.47%)
	D	100%	(44.87%, 48.27%)	6.8442	(7.4783, 7.9499)	9.06%	(4.50%, 4.98%)	29.27%	(16.87%, 18.47%)
π only	A	100%	(100%, 100%)	7.6677	(2.1875, 2.2824)	7.22%	(62.53%, 63.74%)	21.58%	(68.03%, 69.36%)
	B	100%	(100%, 100%)	7.6135	(2.1100, 2.2123)	6.40%	(60.75%, 62.07%)	19.60%	(67.47%, 68.96%)
	C	100%	(100%, 100%)	7.8395	(2.2630, 2.3738)	5.44%	(46.89%, 48.17%)	18.55%	(64.21%, 65.47%)
	D	100%	(100%, 100%)	7.7007	(2.1133, 2.2137)	4.75%	(49.42%, 50.42%)	17.72%	(67.08%, 68.30%)
both λ and π	A	100%	(100%, 100%)	6.6640	(1.4669, 1.5267)	17.70%	(74.85%, 75.90%)	40.90%	(71.65%, 72.92%)
	B	100%	(100%, 100%)	6.4268	(1.4083, 1.4653)	15.78%	(74.62%, 75.62%)	38.30%	(71.47%, 72.88%)
	C	100%	(100%, 100%)	6.9173	(1.4672, 1.5270)	12.74%	(62.25%, 63.38%)	34.68%	(70.24%, 71.52%)
	D	100%	(100%, 100%)	6.7195	(1.4007, 1.4554)	11.72%	(67.79%, 68.69%)	34.52%	(73.61%, 74.65%)

π_0 = 0.60 and λ_0 = 4.0									
λ only	A	99.77%	(99.52%, 99.87%)	3.5340	(3.4441, 3.6303)	37.97%	(37.26%, 38.72%)	52.41%	(51.49%, 53.42%)
	B	99.77%	(99.57%, 99.90%)	3.1713	(3.0940, 3.2645)	43.96%	(43.17%, 44.93%)	44.98%	(43.54%, 46.25%)
	C	99.90%	(99.73%, 99.97%)	3.1875	(3.1086, 3.2716)	23.84%	(23.58%, 24.12%)	53.38%	(52.70%, 54.07%)
	D	99.90%	(99.73%, 99.97%)	2.9485	(2.8665, 3.0330)	29.68%	(29.33%, 30.06%)	58.47%	(57.77%, 59.18%)

π only	A	98.87%	(98.50%, 99.23%)	4.8485	(4.7394, 4.9671)	38.36%	(37.48%, 39.20%)	50.90%	(50.05%, 51.81%)
	B	98.70%	(98.17%, 99.07%)	4.2597	(4.1503, 4.3768)	43.70%	(42.46%, 44.75%)	43.51%	(42.35%, 44.57%)
	C	97.27%	(96.63%, 97.80%)	4.9360	(4.8163, 5.0598)	14.43%	(14.20%, 14.68%)	41.01%	(40.32%, 41.70%)
	D	96.77%	(96.10%, 97.37%)	4.6513	(4.5334, 4.7797)	17.34%	(17.02%, 17.69%)	45.15%	(44.35%, 45.94%)
Both λand π	A	99.97%	(99.80%, 100.00%)	2.8160	(2.7550, 2.8768)	50.99%	(50.22%, 51.77%)	58.27%	(57.35%, 59.18%)
	B	99.97%	(99.83%, 100%)	2.5277	(2.4661, 2.5966)	56.73%	(55.80%, 57.62%)	50.16%	(48.95%, 51.38%)
	C	100%	(100%, 100%)	2.5783	(2.5157, 2.6381)	29.34%	(29.04%, 29.61%)	58.39%	(57.74%, 59.08%)
	D	100%	(100%, 100%)	2.3840	(2.3191, 2.4477)	36.80%	(36.41%, 37.20%)	63.48%	(62.82%, 64.08%)

Table 5 provides the means, as well as the 95% confidence intervals, of performance metrics of the four charts considered for small short triangular outbreaks with different ZIP backgrounds. As a summary, all charts were able to soundly identify outbreaks defined by any shift scenarios, particularly when outbreaks were defined by either shifts in λ only or in both λ and π when disease background has lower π_0, and when outbreaks were defined by shifts in π only when disease background has higher π_0. Meanwhile, CEDs were mostly similar in any scenario considered. Lastly, PODs and PTDs were generally higher when outbreaks were defined by shifts in both λ and π, but were observably lowest when outbreaks were only defined by shifts in π whenever the disease background has lower π_0. However, the converse was true whenever the disease background has higher π_0.

Table 5. Performance Metrics for Small Short Triangular Outbreaks

Shifts	Chart	PSD		CED		POD		PTD	
		mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
$\pi_0 = 0.30$ and $\lambda_0 = 1.0$									
λ only	A	69.67%	(68.00%, 71.26%)	6.6021	(6.3819, 6.8145)	10.78%	(10.34%, 11.18%)	30.44%	(29.48%, 31.40%)
	B	69.97%	(68.40%, 71.67%)	6.5695	(6.3409, 6.8074)	9.56%	(9.25%, 9.89%)	29.42%	(28.47%, 30.25%)
	C	68.07%	(66.32%, 69.63%)	7.0038	(6.7759, 7.2398)	10.05%	(9.70%, 10.42%)	29.36%	(28.53%, 30.27%)
	D	67.93%	(66.23%, 69.49%)	6.7933	(6.5645, 7.0363)	9.39%	(9.09%, 9.74%)	29.01%	(28.10%, 29.82%)

π only	A	62.87%	(60.97%, 64.63%)	7.2641	(7.0825, 7.4493)	10.34%	(9.91%, 10.78%)	29.26%	(28.32%, 30.26%)
	B	61.20%	(59.58%, 62.92%)	7.0940	(6.9008, 7.2895)	8.24%	(7.91%, 8.57%)	26.32%	(25.27%, 27.21%)
	C	60.07%	(58.17%, 61.70%)	7.4921	(7.2938, 7.6907)	8.03%	(7.68%, 8.37%)	25.11%	(24.24%, 26.02%)
	D	61.00%	(59.26%, 62.76%)	7.4678	(7.2686, 7.6780)	8.10%	(7.80%, 8.44%)	26.22%	(25.29%, 27.19%)
both λ and π	A	73.87%	(72.37%, 75.46%)	6.4251	(6.2336, 6.6403)	12.35%	(11.90%, 12.78%)	33.35%	(32.41%, 34.35%)
	B	73.97%	(72.43%, 75.53%)	6.3209	(6.1169, 6.5359)	11.00%	(10.66%, 11.35%)	32.41%	(31.52%, 33.33%)
	C	72.73%	(71.10%, 74.10%)	6.7644	(6.5646, 6.9758)	11.55%	(11.17%, 11.96%)	32.18%	(31.28%, 33.16%)
	D	72.53%	(70.93%, 74.10%)	6.5334	(6.3271, 6.7352)	10.84%	(10.46%, 11.22%)	32.11%	(31.00%, 32.92%)

$\pi_0 = 0.30$ *and* $\lambda_0 = 4.0$									
λ only	A	86.37%	(85.17%, 87.63%)	5.8687	(5.6822, 6.0477)	16.58%	(16.06%, 17.09%)	39.60%	(38.66%, 40.48%)
	B	87.67%	(86.47%, 88.90%)	5.5153	(5.3527, 5.7203)	16.50%	(16.06%, 16.95%)	39.58%	(38.71%, 40.43%)
	C	85.03%	(83.70%, 86.27%)	6.2239	(6.0366, 6.4057)	13.58%	(13.16%, 13.96%)	36.11%	(35.32%, 37.00%)
	D	87.37%	(86.20%, 88.57%)	5.8422	(5.6620, 6.0172)	12.65%	(12.36%, 12.95%)	36.47%	(35.68%, 37.26%)
π only	A	73.10%	(71.53%, 74.60%)	7.0012	(6.8188, 7.1730)	13.17%	(12.69%, 13.67%)	34.22%	(33.22%, 35.23%)
	B	73.47%	(71.77%, 74.90%)	6.6621	(6.4830, 6.8560)	12.03%	(11.60%, 12.46%)	32.38%	(31.41%, 33.38%)
	C	71.63%	(69.90%, 73.43%)	7.2344	(7.0519, 7.4315)	9.83%	(9.49%, 10.21%)	29.51%	(28.57%, 30.41%)
	D	73.60%	(72.00%, 75.20%)	6.9666	(6.7879, 7.1624)	9.27%	(8.97%, 9.60%)	29.65%	(28.84%, 30.52%)
both λ and π	A	87.17%	(85.93%, 88.40%)	5.6874	(5.5208, 5.8572)	16.97%	(16.53%, 17.49%)	40.31%	(39.48%, 41.20%)
	B	89.17%	(88.00%, 90.20%)	5.5603	(5.3790, 5.7279)	16.71%	(16.28%, 17.13%)	40.06%	(39.24%, 41.03%)
	C	87.03%	(85.70%, 88.17%)	6.3059	(6.1136, 6.4904)	13.84%	(13.47%, 14.26%)	36.70%	(35.91%, 37.46%)
	D	88.57%	(87.33%, 89.63%)	5.9071	(5.7367, 6.1098)	12.85%	(12.53%, 13.14%)	37.10%	(36.29%, 37.88%)

$\pi_0 = 0.60$ *and* $\lambda_0 = 1.0$									
λ only	A	25.80%	(24.40%, 27.47%)	6.6615	(6.2849, 7.0497)	2.77%	(2.55%, 2.99%)	9.58%	(8.78%, 10.39%)
	B	27.87%	(26.26%, 29.50%)	6.5482	(6.1783, 6.9538)	2.93%	(2.71%, 3.16%)	9.22%	(8.61%, 10.07%)
	C	27.37%	(25.73%, 29.00%)	6.8589	(6.4840, 7.2494)	2.47%	(2.29%, 2.65%)	8.72%	(8.12%, 9.40%)
	D	28.37%	(26.67%, 30.00%)	6.8241	(6.4691, 7.1906)	2.39%	(2.22%, 2.55%)	9.26%	(8.62%, 9.95%)

π only	A	83.70%	(82.27%, 84.83%)	6.8014	(6.6449, 6.9492)	18.37%	(17.75%, 18.96%)	40.14%	(39.11%, 41.17%)
	B	80.13%	(78.63%, 81.53%)	6.6598	(6.4893, 6.8152)	16.55%	(15.92%, 17.16%)	36.85%	(35.85%, 38.03%)
	C	81.40%	(80.07%, 82.80%)	6.8789	(6.6959, 7.0307)	13.26%	(12.77%, 13.74%)	34.12%	(33.25%, 34.98%)
	D	80.63%	(79.23%, 82.11%)	6.7995	(6.6210, 6.9606)	12.14%	(11.75%, 12.52%)	33.97%	(33.11%, 34.81%)
both λ and π	A	25.97%	(24.46%, 27.51%)	6.5317	(6.1599, 6.9128)	2.81%	(2.61%, 3.03%)	9.65%	(8.94%, 10.47%)
	B	28.17%	(26.53%, 29.80%)	6.8700	(6.4717, 7.2602)	2.98%	(2.77%, 3.20%)	9.61%	(8.90%, 10.42%)
	C	29.20%	(27.57%, 30.90%)	7.0600	(6.6986, 7.4108)	2.57%	(2.40%, 2.76%)	9.11%	(8.44%, 9.76%)
	D	29.90%	(28.23%, 31.60%)	7.0976	(6.7435, 7.4850)	2.49%	(2.35%, 2.66%)	9.84%	(9.25%, 10.54%)
$\pi_0 = 0.60$ *and* $\lambda_0 = 4.0$									
λ only	A	35.60%	(33.83%, 37.40%)	6.3201	(5.9810, 6.6761)	4.39%	(4.10%, 4.73%)	11.66%	(10.88%, 12.67%)
	B	48.77%	(46.55%, 50.70%)	6.1421	(5.8308, 6.4630)	7.42%	(7.00%, 7.91%)	11.34%	(10.48%, 12.27%)
	C	40.20%	(38.50%, 42.30%)	7.0474	(6.6965, 7.4023)	3.00%	(2.86%, 3.15%)	12.00%	(11.39%, 12.63%)
	D	43.23%	(41.51%, 45.03%)	7.0572	(6.6952, 7.3958)	3.40%	(3.24%, 3.56%)	13.79%	(13.14%, 14.41%)
π only	A	95.47%	(94.67%, 96.17%)	5.5207	(5.3954, 5.6443)	27.76%	(26.92%, 28.55%)	44.19%	(43.11%, 45.13%)
	B	95.70%	(94.83%, 96.33%)	4.8818	(4.7510, 5.0069)	33.11%	(32.18%, 34.17%)	37.67%	(36.62%, 39.00%)
	C	94.27%	(93.43%, 95.13%)	5.3640	(5.2241, 5.4854)	12.42%	(12.20%, 12.65%)	37.49%	(36.71%, 38.26%)
	D	94.10%	(93.20%, 94.93%)	4.9977	(4.8694, 5.1323)	14.82%	(14.51%, 15.12%)	41.50%	(40.77%, 42.33%)
both λ and π	A	35.53%	(33.77%, 37.30%)	6.4587	(6.0988, 6.8245)	4.37%	(4.06%, 4.70%)	11.59%	(10.77%, 12.46%)
	B	49.17%	(47.30%, 51.17%)	6.0970	(5.7786, 6.3881)	7.41%	(6.98%, 7.88%)	11.33%	(10.52%, 12.33%)
	C	41.17%	(39.41%, 43.02%)	7.0133	(6.7023, 7.3658)	3.11%	(2.95%, 3.27%)	12.55%	(11.94%, 13.28%)
	D	44.53%	(42.77%, 46.43%)	6.8932	(6.5029, 7.2135)	3.55%	(3.41%, 3.74%)	14.28%	(13.65%, 14.99%)

Table 6 provides the means, as well as the 95% confidence intervals, of performance metrics of the four charts considered for small short ramp outbreaks with different ZIP backgrounds. As a summary, observed patterns were largely similar to those of the cases considered for small short triangular outbreaks. The only exception was that CEDs where shortest when outbreaks were defined by shifts only in π whenever the disease background has higher π_0.

Table 6. Performance Metrics for Small Short Ramp Outbreaks

Shifts	Chart	PSD		CED		POD		PTD	
		mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
$\pi_0 = 0.30$ and $\lambda_0 = 1.0$									
λ only	A	86.27%	(85.07%, 87.47%)	6.4167	(6.2644, 6.5695)	20.64%	(20.05%, 21.29%)	45.09%	(44.15%, 45.96%)
	B	83.53%	(82.33%, 84.95%)	6.3819	(6.2251, 6.5587)	15.83%	(15.39%, 16.23%)	40.94%	(40.04%, 41.88%)
	C	82.47%	(81.07%, 83.83%)	6.7249	(6.5638, 6.8994)	16.69%	(16.24%, 17.16%)	40.45%	(39.49%, 41.35%)
	D	82.90%	(81.50%, 84.37%)	6.5679	(6.4001, 6.7014)	16.42%	(15.91%, 16.86%)	41.50%	(40.61%, 42.41%)
π only	A	63.70%	(61.90%, 65.47%)	7.2701	(7.0565, 7.4680)	10.52%	(10.09%, 10.99%)	29.49%	(28.42%, 30.53%)
	B	61.90%	(60.07%, 63.67%)	7.0936	(6.9018, 7.3155)	8.36%	(8.03%, 8.71%)	26.48%	(25.55%, 27.37%)
	C	59.87%	(58.07%, 61.70%)	7.3736	(7.1606, 7.5847)	8.19%	(7.86%, 8.55%)	25.46%	(24.52%, 26.38%)
	D	60.13%	(58.23%, 61.83%)	7.2368	(7.0373, 7.4475)	8.14%	(7.84%, 8.49%)	26.23%	(25.37%, 27.15%)
both λ and π	A	89.97%	(88.85%, 91.02%)	6.1556	(6.0112, 6.2959)	23.39%	(22.69%, 24.09%)	48.25%	(47.34%, 49.17%)
	B	87.80%	(86.67%, 89.05%)	6.0199	(5.8789, 6.1849)	17.97%	(17.50%, 18.48%)	44.04%	(43.22%, 44.96%)
	C	86.93%	(85.77%, 88.17%)	6.4415	(6.2913, 6.5830)	18.94%	(18.45%, 19.49%)	43.67%	(42.79%, 44.43%)
	D	87.50%	(86.33%, 88.70%)	6.3385	(6.1827, 6.4848)	18.78%	(18.29%, 19.24%)	44.92%	(43.98%, 45.77%)
$\pi_0 = 0.30$ and $\lambda_0 = 4.0$									
λ only	A	97.83%	(97.29%, 98.30%)	5.3138	(5.1915, 5.4239)	33.25%	(32.59%, 33.87%)	56.11%	(55.32%, 56.93%)
	B	97.23%	(96.57%, 97.77%)	5.0823	(4.9707, 5.1910)	29.59%	(28.97%, 30.16%)	54.18%	(53.36%, 54.98%)
	C	97.60%	(96.97%, 98.10%)	5.7182	(5.5929, 5.8340)	24.59%	(24.08%, 25.12%)	50.60%	(49.89%, 51.31%)
	D	97.50%	(96.90%, 98.07%)	5.3493	(5.2381, 5.4738)	21.19%	(20.81%, 21.58%)	48.96%	(48.24%, 49.68%)
π only	A	73.00%	(71.36%, 74.57%)	6.9622	(6.7905, 7.1622)	13.28%	(12.78%, 13.80%)	34.40%	(33.32%, 35.47%)
	B	73.63%	(72.03%, 75.16%)	6.7343	(6.5583, 6.9135)	12.02%	(11.60%, 12.47%)	32.16%	(31.23%, 33.14%)
	C	70.57%	(68.93%, 72.23%)	7.0836	(6.8921, 7.2601)	9.86%	(9.47%, 10.23%)	29.30%	(28.37%, 30.17%)
	D	73.77%	(72.13%, 75.40%)	6.8600	(6.6730, 7.0615)	9.53%	(9.23%, 9.85%)	30.11%	(29.29%, 31.05%)
both λ and π	A	97.97%	(97.43%, 98.43%)	5.2302	(5.1136, 5.3421)	33.75%	(32.98%, 34.46%)	56.38%	(55.59%, 57.14%)
	B	97.33%	(96.63%, 97.83%)	5.0488	(4.9322, 5.1609)	30.00%	(29.40%, 30.65%)	54.44%	(53.54%, 55.24%)
	C	97.77%	(97.23%, 98.27%)	5.6067	(5.4931, 5.7285)	24.92%	(24.34%, 25.49%)	50.82%	(50.15%, 51.53%)
	D	97.53%	(96.93%, 98.03%)	5.3120	(5.2057, 5.4387)	21.61%	(21.25%, 22.00%)	49.56%	(48.84%, 50.31%)

$\pi_0 = 0.60$ *and* $\lambda_0 = 1.0$									
λ only	A	26.73%	(25.10%, 28.27%)	6.3671	(6.0084, 6.7365)	2.91%	(2.69%, 3.15%)	10.07%	(9.31%, 10.91%)
	B	28.67%	(26.95%, 30.37%)	6.4879	(6.1465, 6.8835)	3.04%	(2.82%, 3.30%)	9.82%	(9.08%, 10.62%)
	C	29.57%	(28.00%, 31.43%)	6.8969	(6.5725, 7.2832)	2.64%	(2.45%, 2.82%)	9.70%	(9.06%, 10.45%)
	D	30.23%	(28.63%, 31.83%)	6.8179	(6.4013, 7.2016)	2.56%	(2.40%, 2.71%)	10.11%	(9.43%, 10.86%)
π only	A	83.83%	(82.50%, 85.10%)	6.6583	(6.5067, 6.8053)	18.68%	(18.03%, 19.27%)	40.69%	(39.70%, 41.60%)
	B	81.07%	(79.57%, 82.53%)	6.5051	(6.3494, 6.6638)	16.93%	(16.34%, 17.53%)	37.22%	(36.21%, 38.31%)
	C	81.30%	(79.81%, 82.67%)	6.8588	(6.7014, 7.0255)	13.38%	(12.90%, 13.86%)	34.30%	(33.49%, 35.16%)
	D	81.40%	(80.00%, 82.80%)	6.6858	(6.5438, 6.8456)	12.31%	(11.98%, 12.69%)	34.27%	(33.49%, 35.25%)
both λ and π	A	28.20%	(26.63%, 29.90%)	6.6360	(6.2451, 6.9932)	3.01%	(2.79%, 3.23%)	10.29%	(9.59%, 11.18%)
	B	30.40%	(28.53%, 32.07%)	6.6260	(6.2221, 7.0071)	3.18%	(2.96%, 3.41%)	10.07%	(9.37%, 10.79%)
	C	30.83%	(29.12%, 32.53%)	7.0906	(6.7383, 7.4784)	2.72%	(2.55%, 2.92%)	9.66%	(9.02%, 10.31%)
	D	32.20%	(30.56%, 33.93%)	7.0815	(6.7327, 7.4898)	2.71%	(2.55%, 2.89%)	10.42%	(9.78%, 11.11%)

$\pi_0 = 0.60$ *and* $\lambda_0 = 4.0$									
λ only	A	37.23%	(35.53%, 39.18%)	6.5174	(6.1689, 6.9199)	4.62%	(4.32%, 4.98%)	12.30%	(11.42%, 13.22%)
	B	50.57%	(48.53%, 52.42%)	6.2088	(5.9072, 6.5339)	7.79%	(7.31%, 8.29%)	12.18%	(11.33%, 13.14%)
	C	43.23%	(41.63%, 45.20%)	7.1837	(6.8084, 7.5188)	3.26%	(3.11%, 3.40%)	13.21%	(12.57%, 13.84%)
	D	46.30%	(44.50%, 48.03%)	6.9779	(6.6418, 7.3625)	3.72%	(3.57%, 3.89%)	14.96%	(14.32%, 15.65%)
π only	A	95.77%	(95.05%, 96.50%)	5.5022	(5.3707, 5.6328)	27.46%	(26.64%, 28.29%)	43.94%	(42.82%, 44.86%)
	B	96.40%	(95.60%, 97.00%)	4.9072	(4.7759, 5.0304)	33.23%	(32.21%, 34.15%)	37.83%	(36.60%, 38.98%)
	C	95.47%	(94.67%, 96.20%)	5.4215	(5.2895, 5.5522)	12.61%	(12.37%, 12.83%)	37.90%	(37.20%, 38.55%)
	D	95.57%	(94.77%, 96.37%)	5.1413	(5.0031, 5.2890)	15.00%	(14.70%, 15.28%)	41.69%	(40.90%, 42.50%)
both λ and π	A	37.37%	(35.49%, 39.19%)	6.4357	(6.1012, 6.7668)	4.56%	(4.25%, 4.90%)	12.18%	(11.35%, 13.04%)
	B	50.63%	(48.77%, 52.68%)	6.1160	(5.8223, 6.4872)	7.72%	(7.27%, 8.18%)	11.77%	(10.98%, 12.77%)
	C	43.67%	(41.87%, 45.57%)	7.1542	(6.8193, 7.5104)	3.29%	(3.14%, 3.45%)	13.39%	(12.69%, 14.03%)
	D	46.77%	(44.94%, 48.53%)	6.9035	(6.5870, 7.2489)	3.79%	(3.62%, 3.95%)	15.24%	(14.53%, 15.90%)

Mostly similar results and patterns were observed for any other simulation scenarios considered (i.e., different outbreak magnitudes and lengths). However, they were not presented anymore due to space limitations. Nevertheless, all results and associated tables are available from the authors upon request.

In summary, most outbreaks were detected by both EWMA charts, especially when the outbreaks were determined by shifts in λ for scenarios with small disease backgrounds, and when the outbreaks were determined by shifts in π for scenarios with large disease backgrounds. In particular, with PSDs, both proposed EWMA charts performed considerably well, with detection rates nearly comparable with one another. In particular, Bernoulli-ZIP EWMA charts mostly had higher PSD rates when π_0 is smaller, while the same statement is true for CRL-ZTP EWMA charts when π_0 is higher. Regarding CEDs, several situations likewise saw both EWMA charts having comparable values, wherein detection lags were mostly shorter for shifts in λ only or in both λ and π when π_0 is smaller, and for shifts in π only when π_0 is larger. Concerning PODs, the proposed Bernoulli-ZIP EWMA charts were found to perform best in almost all situations considered. In particular, it was observed that a smaller weighting parameter κ must be desired if the disease background is smaller, and vice versa. Likewise with PTDs, Bernoulli-ZIP EWMA charts mostly had better performances than the CRL-ZTP EWMA charts, except the situations when the disease background had higher π_0 and λ_0. Also, quite noticeable was that regardless of the disease background or the outbreak situation, smaller κ must be desired for Bernoulli-ZIP EWMA charts, while larger κ must be desired for CRL-ZTP EWMA charts.

Concerning types of outbreaks, all charts performed best when spike outbreaks were considered; worst when triangular outbreaks were considered. In particular, PTDs were typically higher and CEDs were generally shorter by almost two to three days with spike outbreaks as against similar situations but with triangular or ramp shapes. These findings can be explained by that spike outbreaks had sudden shifts; hence, were easier to detect. On the other hand, triangular outbreaks gradually shift and only peak once; hence, were harder to detect particularly by those charts with memory. With ramp outbreaks, while shifts were also gradual, the peaks lasted for a considerable amount of time; hence, were detected better than triangular outbreaks. Nevertheless, all charts were able to make the detection mostly during the first half of the outbreak period, which presented an opportunity to possibly introduce an intervention to control the infection spread before the outbreak is at its peak.

4. Case Application

4.1 The measles dataset

To demonstrate how the proposed combined EWMA schemes may be used with real data, all four previously considered charts were applied with actual

measles data consisting of both laboratory confirmed and epidemiologically-linked measles daily cases in the National Capital Region (NCR) spanning the dates January 1, 2010 to January 14, 2015, obtained from the Philippine Department of Health - National Epidemiology Center, NCR.

A plot of the data is provided in Figure 1, albeit truncated at 100. It should be noted that recorded cases from December 17, 2013 up to February 12, 2014 exceeded 100, hence are not shown in the graph. Three significant shifts can be observed; which upon confirmation from news reports, they were identified as outbreaks (Aragones, 2013; Agreda, 2013; RSJ, GMA News, 2013; Santos, 2013). The identified outbreak periods are the following:

Outbreak 1: January 1, 2010 to May 31, 2010 (151 days)

Outbreak 2: January 1, 2011 to May 31, 2011 (151 days)

Outbreak 3: September 1, 2013 to April 15, 2014 (227 days)

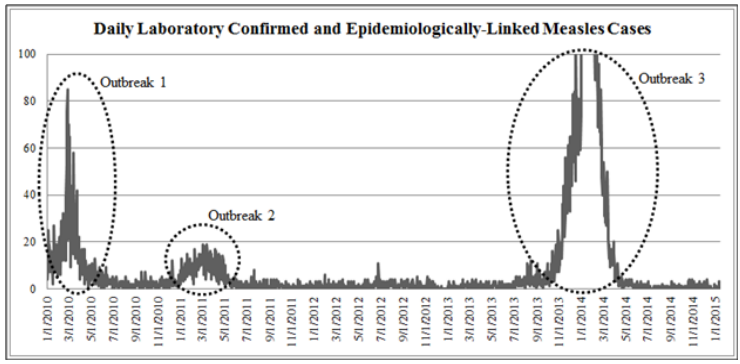

Figure 1. Time-series plot of daily laboratory confirmed and epidemiologically-linked measles cases in NCR from January 1, 2010 to January 14, 2015.

4.2. Phase I procedure

For the Phase I procedure, daily measles counts from January 1, 2010 up to December 31, 2012 were used to identify the disease background characteristics. In particular, the ZIP parameters π and λ were estimated. Prior to the estimation, pre-processing of the data was done. In which case, this procedure involved the deletion of the identified outbreak periods (January 1, 2010 to May 31, 2010 and January 1, 2011 to May 31, 2011) in the background. The resulting data consisted of 794 observations.

It was observed that the mean daily count was 1.3438, while the variance was 2.3116; accounting for a variance-to-mean dispersion index of 72.02%. Furthermore, out of the 794 observations, 35.26% are zeros. These findings may have provided hints that the underlying disease background follows that of the ZIP distribution. Hence, based on the assumption of a ZIP background process,

estimates of the parameters were provided by their corresponding MLEs, in which case were computed to be $\hat{\pi}_{mle} = 0.7930$ and $\hat{\lambda}_{mle} = 1.6946$. To formally determine whether the background indeed followed a ZIP process, the ZIP-Score Test by Xie et al. (2001) was employed. In which case, the test statistic was computed to be

$$S = \frac{(n_0 - np_0)^2}{np_0(1-p_0) - n\bar{y}p_0^2} = 216.73, \tag{27}$$

where $n = 794$ is the total number of observations, $n_0 = 280$ is the number of zeros, $\bar{y} = 1.3438$ is the sample mean of the observations, and $p_0 = e^{-\hat{\lambda}_{mle}} = e^{-1.6946}$. Asymptotically, $S \sim \chi^2_{(1)}$; hence, the observed test statistic $S = 216.73$ had a *p–value* < 0.0001. Clearly, the p-value provided overwhelming evidence against H_0, in which case the decision to reject H_0 meant that the overdispersion index was significant. Thus, the underlying distribution was indeed that of the ZIP distribution.

For the chart thresholds, since the estimated ZIP parameters were closest to the simulation situation wherein $\pi = 0.60$ and $\lambda = 1.0$, then the threshold parameters for all the combined control charting procedures were correspondingly set to be equivalent to those of the aforementioned situation. In addition to the simultaneous charting procedures, the historical limits procedure was also applied in order to demonstrate the comparability of the proposed charting schemes with the current system in use (Catindig and Niñal, 2008). It can be recalled that the historical limits procedure is simply a variant of the standard upper Shewhart chart utilizing a 2σ-limit. Thus, the historical limits threshold was computed as follows:

$$h_{HL} = \hat{\mu} + 2\hat{\sigma} = 4.3846, \tag{28}$$

i.e., the chart was made to generate a signal whenever the observed count for the particular day is at least five. A summary of the computed thresholds for each chart is provided in Table 7.

Table 7. Chart Thresholds Used as Applied in Measles Data

Chart	h_λ	h_π
Historical Limits	4.3846	*
Bernoulli-ZIP EWMA Chart (κ=0.25)	2.7638	0.9860
Bernoulli-ZIP EWMA Chart (κ=0.45)	3.7081	1.0840
CRL-ZTP EWMA Chart (κ=0.25)	4.0610	2.0537
CRL-ZTP EWMA Chart (κ=0.45)	4.0646	1.4280

**The historical limits procedure only used one threshold (i.e., a single-charting scheme), which simply monitors upward shifts in λ.*

4.3. Phase II procedure

A summary of all the charts considered regarding their performances in Phase II are provided in Table 8. The charts were assessed similarly with the framework provided in the simulations, wherein their abilities to detect the September 1, 2013 – April 15, 2014 outbreak (lasting for 227 days) were given emphasis. In addition, ATFS measures were also provided to assess their regularity of generating false signals during non-outbreak periods (January 1, 2013 – August 31, 2013 and April 16, 2014 – January 14, 2015; lasting for 517 days). Visualizations of these charts are provided in the corresponding figures identified in the second column of Table 8.

Regarding ATFS, it can be observed that the least specific were the CRL-ZTP EWMA variants, which generate false signals almost at least once in a span of less than two weeks. The Bernoulli-ZIP EWMA charts, on the other hand, had widely varying ATFSs as evidenced by a value of 32.31 for $\kappa = 0.25$ and a value of 129.25 for $\kappa = 0.45$. Lastly, the historical limits procedure was seen to generate false signals almost once a month, with ATFS value of 34.47. In which case, only the Bernoulli-ZIP EWMA chart with $\kappa = 0.25$ was found to have ATFS performances which were comparable to the current system in use.

Table 8. Performance Assessment of the Different Charts as Applied in Measles Data

Chart	Figure	ATFS	PSD*	CED**	POD	PTD
Historical Limits	Figure 2	34.47	✓	Sept. 2, 2013	84.58%	92.75%
Bernoulli-ZIP EWMA Chart ($\kappa = 0.25$)	Figure 3	32.31	✓	Sept. 2, 2013	82.38%	92.12%
Bernoulli-ZIP EWMA Chart ($\kappa = 0.45$)	Figure 4	129.25	✓	Sept. 2, 2013	79.30%	97.83%
CRL-ZTP EWMA Chart ($\kappa = 0.25$)	Figure 5	9.40	✓	Sept. 2, 2013	84.58%	77.73%
CRL-ZTP EWMA Chart ($\kappa = 0.45$)	Figure 6	5.07	✓	Sept. 1, 2013	85.02%	65.42%

**The PSD in this table is a pseudo-measurement, in which case only identified if the Sept. 1, 2013 - Apr. 15, 2014 outbreak was detected.*

***The CED in this table is a pseudo-measurement, in which case is the first day that the chart signaled during the identified outbreak period.*

Since it is impossible to measure PSD using only a single dataset, then a pseudo-measurement is provided instead. To that end, the PSD was taken as simply either the outbreak was detected or not (i.e., the control charting procedure signaled at least once during the entire outbreak duration). Clearly, all charts were able to detect the outbreak. In fact, all were able to detect the outbreak as soon as the second day (September 2, 2013), except for the CRL-ZTP EWMA chart with

$\kappa = 0.45$, wherein it was able to detect sooner (September 1, 2013). The reason for the detection right on the second day was quite trivial because 11 confirmed cases were recorded on that particular day. Regarding the earlier detection of CRL-ZTP EWMA chart with $\kappa = 0.45$, the reason was identified to be because it was the least specific (which also translates to being the most sensitive); hence, an observed value of three cases on September 1, 2013 was already enough to generate a signal. The first day that these charts signaled during the entire outbreak duration (i.e., the time of outbreak detection) was used as a pseudo-measurement of CED due to the same reasons cited above.

Further discussing their abilities to soundly detect the outbreak, POD and PTD metrics were also computed. Likewise, seen in the table was that CRL-ZTP EWMA charts, in general, were found to be the most sensitive (i.e., with typically higher PODs). The historical limits procedure, on the other hand, was observed to have a POD of 84.58%. Lastly, concerning PTDs, the CRL-ZTP EWMA charts were observed to generate the least true positives (as evidenced by typically lower PTDs). With the Bernoulli-ZIP EWMA charts applied in the given dataset, higher κ was observed to have higher PTD (97.83%), and vice versa (92.12%). Regarding the performance of the historical limits procedure, a PTD of 92.75% was observed. In which case, for both POD and PTD, only the Bernoulli-ZIP EWMA charts were found to have comparable performances with the current system in use.

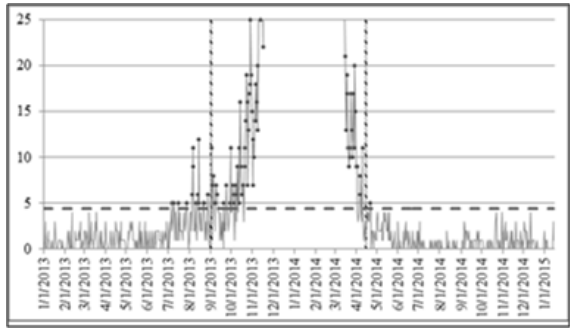

Figure 2. Visualization of the historical limits control charting procedure. The threshold is given by the broken horizontal line. Bold markers denote the instances the chart signaled. The Sept. 1, 2013 to Apr. 15, 2014 outbreak period is enclosed within the black dashed vertical lines.

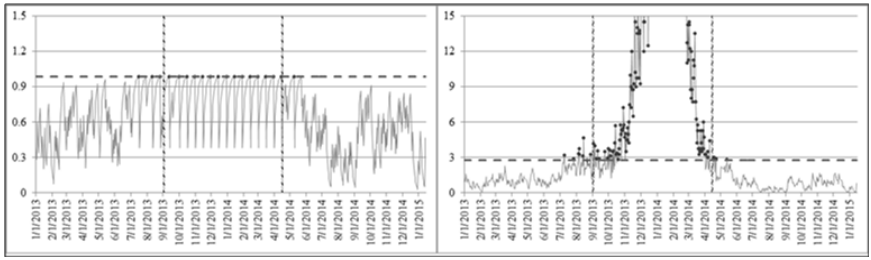

Figure 3. Visualization of the Bernoulli-ZIP EWMA charting procedures ($\kappa = 0.25$). [Left: Bernoulli-EWMA Chart, Right: ZIP-EWMA Chart] The thresholds are given by the broken horizontal lines. Bold markers denote the instances the chart signaled. The Sept. 1, 2013 to Apr. 15, 2014 outbreak period is enclosed within the black dashed vertical lines.

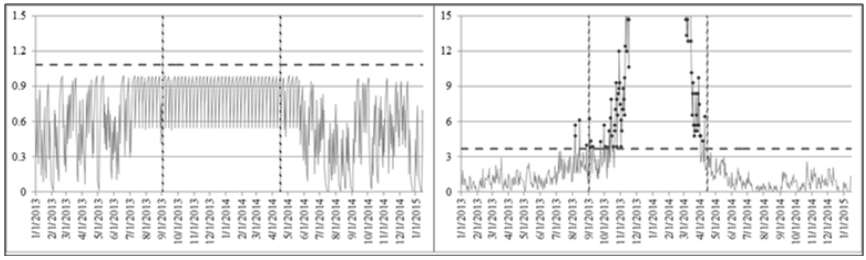

Figure 4. Visualization of the Bernoulli-ZIP EWMA charting procedures ($\kappa = 0.45$). [Left: Bernoulli-EWMA Chart, Right: ZIP-EWMA Chart] The thresholds are given by the broken horizontal lines. Bold markers denote the instances the chart signaled. The Sept. 1, 2013 to Apr. 15, 2014 outbreak period is enclosed within the black dashed vertical lines.

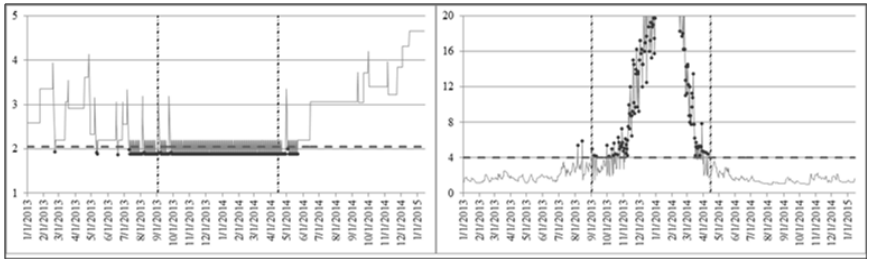

Figure 5. Visualization of the CRL-ZTP EWMA charting procedures ($\kappa = 0.25$).[Left: CRL-EWMA Chart, Right: ZTP-EWMA Chart] The thresholds are given bythe broken horizontal lines. Bold markers denote the instances the chart signaled. The Sept. 1, 2013 to Apr. 15, 2014 outbreak period is enclosed within the black dashed vertical lines.

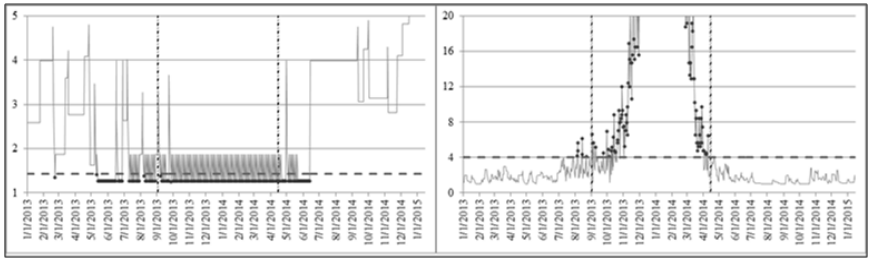

Figure 6. Visualization of the CRL-ZTP EWMA charting procedures (κ = 0.45). [Left: CRL-EWMA Chart, Right: ZTP-EWMA Chart] The thresholds are given by the broken horizontal lines. Bold markers denote the instances the chart signaled. The Sept. 1, 2013 to Apr. 15, 2014 outbreak period is enclosed within the black dashed vertical lines.

5. Conclusion

This paper has discussed and demonstrated the possibility of using a combined charting procedure for disease surveillance, wherein disease background follows a ZIP distribution. The rationale behind a combined charting procedure is that the ZIP process is characterized by two parameters. Hence, simultaneous individual charting procedure for each parameter is proposed in an effort to improve shift detection (i.e., outbreak detection in the context of disease surveillance). While this idea has already been explored in the context of SPC, two EWMA charts were developed, namely Bernoulli-ZIP EWMA and CRL-ZTP EWMA charts.

Based on an extensive simulation study, the proposed Bernoulli-ZIP EWMA charting procedure mostly had better performances than the CRL-ZTP EWMA chart regardless of outbreak scenario considered. Furthermore, it was empirically shown that a higher weighting parameter must be desired if the disease background has large parameters, and vice versa. These results were mostly consistent with the demonstration on actual measles data. For demonstration purposes, the current system in place (the historical limits method) was set as the reference chart. In which case, it was identified that the Bernoulli-ZIP EWMA chart had mostly equivalent performances to those of the historical limits procedure.

6. Recommendations

While this paper has offered some perspectives on the application of combined charting procedures for disease surveillance, several research directions can still be explored to enhance the proposed surveillance schemes. First, a fast initial response (FIR) feature may be added to improve sensitivity without compromising specificity. In this case, FIR is defined to be an initialization enhancement wherein the control chart starts at some value closer to the threshold (which, in most cases, is taken to be halfway between the threshold and the original initialization value). Furthermore, when the disease background is dynamic, an adaptive threshold scheme may be used to account for the gradual (or even seasonal) changes in

the background process. The most common procedure is a moving average and a moving standard deviation with a pre-identified window. Lastly, additional considerations in the simulation framework may be in order to further improve the evaluation of the proposed control charting procedures. These may include consideration of a more complex outbreak model, consideration of a few more ZIP parameters, and an extensive comparison with the modified Shewhart, CUSUM, and historical limits procedures.

References

AGREDA, J.M., 2013, May 24, Measles cases soar high by 157%, Retrieved from Sun Star:http://www.sunstar.com.ph/baguio/local-news/2013/05/24/measles-cases-soar-157-283866

ARAGONES, S., 2011, April 5, Over 2000 measles cases in Jan-March, Retrieved from ABS-CBN News: http://www.abs-cbnnews.com/lifestyle/04/05/11/over-2000-measles-cases-jan-march

BURKOM, H.S., MURPHY, S.P., and SHMUELI, G., 2007, Automated time series forecasting for biosurveillance. *Statistics in Medicine* 26 (22): 4202-4218.

CATINDIG, N. T. and NIÑAL, M. O., 2008, Manual of Procedures for the Philippine Integrated Disease Surveillance and Response. National Epidemiology Center of the Department of Health, Philippines.

CHEN, N., ZHOU, S., CHANG, T.S., and HUANG, H., 2008, Attribute control charts using generalized zero-inflated Poisson distribution, *Quality and Reliability Engineering International* 24 (7): 793-806.

DONG, Y., HEDAYAT, A.S., and SINHA, B.K., 2008, Surveillance strategies for detecting changepoint in incidence rate based on exponentially weighted moving average methods, *Journal of the American Statistical Association* 103 (482): 843-853.

ELBERT, Y. and BURKOM, H.S., 2009, Development and evaluation of a data-adaptive alerting algorithm for univariate temporal biosurveillance data, *Statistics in Medicine* 28(26): 3226-3248.

FARRINGTON, P. and ANDREWS, N., 2004, Outbreak detection: Application to infectious disease surveillance. In Brookmeyer, R. and Stroup, D.F., (editors) *Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance*, New York: Oxford University Press, Inc, Pp. 203-231.

FATAHI, A.A., NOOROSSANA, R., DOKOUHAKI, P., and MOGHADDAM, B.F., 2012, Zero-inflated Poisson EWMA control chart for monitoring rare health-related events, *Journal of Mechanics in Medicine and Biology* 12(4): 881-895.

FRAKER, S.E., WOODALL, W.H., and MOUSAVI, S., 2008, Performance metrics for surveillance schemes, *Quality Engineering* 20(4): 451-464.

FRICKER, R.D. 2013, *Introduction to Statistical Methods for Biosurveillance: With an Emphasis on Syndromic Surveillance*, New York: Cambridge University Press.

HAN, S.W., TSUI, K.L., ARIYAJUNYA, B., and KIM, S.B., 2010, A comparison of CUSUM, EWMA, and temporal scan statistics for detection of increases in Poisson rates, *Quality and Reliability Engineering International* 26 (3): 279-289.

HE, S., HUANG, W., and WOODALL, W.H., 2012, CUSUM charts for monitoring a zero-inflated Poisson process. *Quality and Reliability Engineering International* 28 (2): 181-192.

HE, S., LI, S., and HE, Z., 2014, A combination of CUSUM charts for monitoring a zero-inflated Poisson process, *Communications in Statistics – Simulation and Computation* 43 (10): 2482-2497.

MAVROUDIS, E. and NICOLAS, F., 2013, EWMA control charts for monitoring high yield processes, *Communications in Statistics – Theory and Methods* 42 (20): 3639-3654.

MORTON, A. P., WHITBY, M., MCLAWS, M.L., DOBSON, A., MCELWAIN, S., LOOKE, D.,STACKELROTH, J., and SARTOR, A., 2001, The application of statistical process control charts to the detection and monitoring of hospital-acquired infections, *Journal of Quality in Clinical Practice* 21 (4): 112-117.

NGATCHOU-WANDJI, J. and PARIS, C., 2011, On the zero-inflated count models with application to modelling annual trends in incidences of some occupational allergic diseases in France, *Journal of Data Science* 9 (4): 639-659.

NOBRE, F.F. and STROUP, D.F., 1994, A monitoring system to detect changes in public health surveillance data, *International Journal of Epidemiology* 23 (2): 408-418.

ROBERTS, S.W., 1959, Control chart tests based on geometric moving averages, *Technometrics* 42 (1): 97-101.

RSJ, GMA NEWS, 2013, DOH issues measles alert in NCR as cases rise. Retrieved from GMA Network News: http://www.gmanetwork.com/news/story /339346/news/metromanila/doh-issues-measles-alert-in-ncr-as-cases-rise, December 11.

SANTOS, T. G., 2013, Measles cases up in Metro, says DOH. Retrieved from *The Philippine Daily Inquirer*: http://newsinfo.inquirer.net/544665/measles-cases-up-in-metro-says-doh, December 11.

SHMUELI, G. and BURKOM, H.S., 2010, Statistical challenges facing early outbreak detection in biosurveillance, *Technometrics* 52 (1): 39-51.

SIM, C.H. and LIM, M.H., 2008, Attribute charts for zero-inflated Poisson processes. *Journal of Communications in Statistics – Simulation and Computation* 37 (7): 1440-1452.

SONESSON, C. and BOCK, D., 2003, A review and discussion of prospective statistical surveillance in public health, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 166 (1): 5-21.

SZARKA, J. L., GAN, L., and WOODALL, W., 2011, Comparison of the early aberration reporting system (EARS) W2 methods to an adaptive threshold method, *Statistics in Medicine* 30 (5): 489-504.

UNKEL, S., FARRINGTON, P.C., GARTHWAITE, P.H., ROBERTSON, C., and ANDREWS, N., 2012, Statistical methods for the prospective detection of infectious disease outbreaks: A review, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175 (1): 49-82.

WONG, W.K. and MOORE, A.W. ,2006, Classical time-series method for biosurveillance. In M.M. Wagner, A.W. Moore, and R.M. Aryel (Eds.), *Handbook of Biosurveillance*, Burlington: Elsevier Academic Press, Pp.217-234.

WOODALL, W.H., 2006, The use of control charts in health-care and public-health surveillance, *Journal of Quality Technology* 38 (2): 89-104.

XIE, M., HE, B., and GOH, T.N. (2001). Zero-inflated Poisson model in statistical process control, *Computational Statistics and Data Analysis* 38 (2): 191-201.

YEH, A. B., MCGRATH, R.N., SEMBOWER, M.A., and SHEN, Q., 2008, EWMA control charts for monitoring high-yield processes based on nontransformed observation. *International Journal of Production Research* 46 (20): 5679-5699.