Adaptive Threshold Method for Monitoring Rates in Public Health Surveillance

Linmin Gan

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Statistics

William H. Woodall, Chair
Marion R. Reynolds, Jr.
Dong-Yun Kim
Scotland Leman

April.30, 2010
Blacksburg, Virginia

Keywords: Biosurveillance; Exponentially weighted moving average chart; Negative binomial distribution; Outbreak detection; Recurrence interval.

Adaptive Threshold Method for Monitoring Rates in Public Health Surveillance

Linmin Gan

ABSTRACT

We examine some of the methodologies implemented by the Centers for Disease Control and Prevention's (CDC) BioSense program. The program uses data from hospitals and public health departments to detect outbreaks using the Early Aberration Reporting System (EARS). The EARS method W2 allows one to monitor syndrome counts (W2count) from each source and the proportion of counts of a particular syndrome relative to the total number of visits (W2rate). We investigate the performance of the W2r method designed using an empiric recurrence interval (RI) in this dissertation research. An adaptive threshold monitoring method is introduced based on fitting sample data to the underlying distributions, then converting the current value to a Z-score through a p-value. We compare the upper thresholds on the Z-scores required to obtain given values of the recurrence interval for different sets of parameter values. We then simulate one-week outbreaks in our data and calculate the proportion of times these methods correctly signal an outbreak using Shewhart and exponentially weighted moving average (EWMA) charts. Our results indicate the adaptive threshold method gives more consistent statistical performance across different parameter sets and amounts of baseline historical data used for computing the statistics. For the power analysis, the EWMA chart is superior to its Shewhart counterpart in nearly all cases, and the adaptive threshold method tends to outperform the W2 rate method. Two modified W2r methods proposed in the dissertation also tend to outperform the W2r method in terms of the RI threshold functions and in the power analysis.

Acknowledgement

I would like to thank my dissertation advisor, Dr. William H. Woodall, for his time and great advice over the years. I would like to also thank my dissertation committee members, Dr. Marion R. Reynolds, Jr., Dr. Dong-Yun Kim and Dr. Scotland Leman, for their valuable advice and comments throughout the dissertation process. Moreover, I would like to thank John L. Szarka III for his helpful input in this work and Merck Research Lab for their funding for my dissertation research.

I would like to thank my mom and Jianmin for their selfless help and support during my process of pursuing academic achievement. I would tell my dad in heaven, "Hey, I finally made it."

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

The Centers for Disease Control and Prevention (CDC) established the BioSense program with the intent of providing real-time biosurveillance for early disease outbreak detection [1]. The primary purpose of Early Aberration Reporting System (EARS) within BioSense is to provide national, state, and local health departments with several alternative aberration detection methods that have been developed for syndromic surveillance by CDC and non-CDC epidemiologists [2]. Currently, hundreds of hospitals and public health departments across the United States provide data to BioSense where the EARS methods are used for determining whether or not syndromic outbreaks have occurred [3].

There are two methodologies EARS uses for detecting these types of outbreaks. The W2 count (W2c) method focuses on the number of cases of a particular syndrome on a given day. The W2rate (W2r) method is based on the proportion of visits corresponding to a particular syndrome which accounts for the total number of visits to a health facility on a given day. The W2 statistics are based on 7-day moving windows. The short baseline is intended to accumulate recent information on a given syndrome. A 2-day lag is also incorporated in the calculation of the statistics, meaning the previous two days are not included in the baselines. If the current day's value is large relative to the baseline data, this will result in a large W2 statistic. If a W2 value exceeds a specified threshold, an alarm is given.

The W2 statistics are calculated separately for weekdays and weekends. This is done because many health care facilities have fewer visits during weekends. We first examine the simplified case where weekday and weekend counts follow the same distribution. We also examine the weekend effects where the average count is significantly lower on weekends for Poisson inputs. The number of cases of a syndrome relative to the total number of daily visits for the W2r method follows a conditional binomial distribution for Poisson inputs and follows what we refer to as a "conditional negative binomial distribution" for negative binomial inputs. Two modified W2r methods are proposed in this dissertation.

An adaptive threshold method proposed by Lambert and Liu [4] for computer network monitoring is also considered in our study. Using the baseline data, the parameters of a distribution are fit using maximum likelihood (ML) or method of moments (MOM) estimators. The current day's count or rate has an upper-tail p-value then calculated from the estimated distribution. A Z-score is computed by taking the inverse standard normal cumulative distribution function (CDF) of one minus the p-value, giving an approximately standard normal statistic when there is no outbreak. The successive Z-scores are used for process monitoring.

The W2 method relies on the use of an empiric recurrence interval (RI). Kleinman et al. [5] explained that if monitoring of a process continues without interruption after any alarm, the RI is the fixed number of time periods for which the expected number of false alarms is one. Table 3 of the CDC's Hospital User Guide [1] gives the W2r thresholds associated with a range of RI values from 10 to 2000 when $n = 7$, where n is the length of the baseline. Using our simulations, we computed our own empiric RI values and compared these to the results from BioSense. We also compared the RI threshold functions of the adaptive threshold, the W2r and the modified W2r methods across different parameter sets and baseline lengths. Since a single upper threshold value is used once a specified RI value is selected, it is important that the non-outbreak performance of the method not depend too much on the characteristics of the input data.

We evaluated the various methods using baselines of $n = 7$, 14, and 28 days. These baseline lengths were used in Tokars et al. [6], but with no more than 56 days of historical data being used. Therefore for weekends, only eight weeks of data were available, leading to only 16 days of data in their baseline. The current baseline of n=7 used by BioSense is a short baseline that in many instances is insufficient for estimation. However, a baseline that is too long will mitigate the ability of the statistic to adjust to seasonal variation. This can lead to a decreased chance in signaling an outbreak. Traditional approaches for detecting false alarms focus on the current day's statistic exceeding a particular threshold. However, we can also use a statistic that accumulates information over time. In our study we considered use of the exponentially weighted moving average (EWMA) statistic with both the W2r and adaptive threshold approaches.

A separate simulation study analyzes the ability of the W2r, modified W2r, and adaptive threshold methods to detect outbreaks, i.e., a power analysis (or sensitivity analysis). This is done by generating samples from a reference distribution for several weeks, then systematically injecting a specified increase in the average number of syndrome counts. The outbreaks are assumed to last for 7 days. It is of interest to determine how frequently the various methods signal, given different magnitudes of shifts and various baseline window sizes. We considered use of both the Shewhart and EWMA approaches for detecting outbreaks.

The W2 methods are reviewed in Chapter 2. In Chapter 3, we introduce the adaptive threshold methods for both the conditional binomial distribution and conditional negative binomial distribution. The performance evaluation with Poisson inputs is presented and discussed in Chapter 4. The performance evaluation with negative binomial inputs is presented and discussed in Chapter 5. Conclusions and planned research on both the W2r and the adaptive threshold methods for prospective public health surveillance are outlined in Chapter 6.

Chapter 2 The Early Aberration Reporting System (EARS) W2 Methods

The Early Aberration Reporting System (EARS) of the Centers for Disease Control and Prevention (CDC) has been implemented throughout the United States in a number of state and local health departments and in health departments in several other countries. The EARS has also been used for syndromic surveillance at several large public events in the United States, including the 2000 Democratic National Convention, the 2001 Super Bowl, and the 2001 World Series [2]. The EARS uses the W2 methods currently implemented in version 2.11 of the BioSense application for early outbreak detection and health situational awareness by all levels of public health and the health care community. The W2 methods are undergoing continued evaluation and may be modified in future releases.

John L. Szarka III has investigated the W2c method, while I report on the performance of the W2r method in my dissertation. Both W2c and W2r methods are based on centered and scaled statistics, using expected values and standard deviations estimated using past data. A minimum value of one is set for the estimated standard deviation. We consider a baseline of n days, where n=7, 14, and 28 in our study. We use a two-day lag when partitioning by weekday and weekend. For a given week k, Table 2-1 shows all of the previous days used in the windows when n=7. There are four distinct baseline groups formed for each week, i.e., Monday to Wednesday, Thursday, Friday, and Saturday to Sunday.

Table 2-1: 7-Day Baseline Days Used for Week *k*

Day in Week *k*	Baseline Data
Monday(k)	Thu-Fri(k-2), Mon-Fri(k-1)
Tuesday(k)	Thu-Fri(k-2), Mon-Fri(k-1)
Wednesday(k)	Thu-Fri(k-2), Mon-Fri(k-1)
Thursday(k)	Fri(k-2), Mon-Fri(k-1), Mon(k)
Friday(k)	Mon-Fri(k-1), Mon-Tue(k)
Saturday(k)	Sun(k-4),…,Sat-Sun(k-1)
Sunday(k)	Sun(k-4),…,Sat-Sun(k-1)

The W2c and W2r methods will signal whenever the corresponding statistic exceeds a given threshold. These thresholds will be determined from the RI threshold functions obtained from our simulation study, which is illustrated in Chapters 4 and 5.

2.1 The W2c Method

Let X_t be the count of a specific syndrome for day t. The baseline data for day t is dependent on its day of the week, as shown in Table 2-1. The W2c value for day t is

$$W2c(t) = \frac{x_t - \bar{x}_t}{s_t} \tag{2.1}$$

where $\overline{x_t}$ and s_t are the sample mean and standard deviation from the baseline period. These values are expressed as

$$\overline{x_t} = \frac{1}{n}\sum_{i=1}^{n} y_{it}, \qquad s_t = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_{it} - \overline{x_t})^2}\ , \tag{2.2}$$

where y_{it}, i=1,2,…,n, correspond to the eligible baseline data for day t. If s_t is less than one, it is reassigned a value of one.

The W2c method is similar to the C2 method formerly used with BioSense. The previous

CDC methods for monitoring counts include methods C1, C2, and C3 and can be found in Table 5 of the CDC's User Guide [7]. These methods do not partition the data by weekday and weekend. The reader is referred to Fricker et al. [8], Hutwagner et al. [2, 9, 10], Zhu et al. [11], and Watkins et al. [12] for analyses of these methods. See also Szarka, Gan, and Woodall [13], where some of the work in this dissertation is summarized.

2.2 The W2r Method

Tokars et al. [6] designed four algorithm modifications to address shortcomings in the C2 algorithm. Those modifications included stratifying the baseline days into weekdays versus weekends, lengthening the baseline period, adjustment for total daily visits (refer to this adjustment as the W2 rate algorithm), and increasing the minimum value for the estimated standard deviation. We study in detail the W2 rate method in this dissertation.

For day t, let X_{1t} represent the syndrome count, X_{2t} be the non-syndrome count, $D_t = X_{1t} + X_{2t}$ be the total number of visits to a facility, t=1,2,…, and d_t be the observed value of D_t. The corresponding counts and numbers of visits for the baseline days are Y_{it} and D_{it}, i = 1, 2, …, n. We let $\mathrm{BLS_t}$ and $\mathrm{BLV_t}$ be the total number of syndromic counts and facility visits over the baseline period. Thus, the average rate of syndrome counts over this period equals $\frac{\mathrm{BLS_t}}{\mathrm{BLV_t}}$. The W2r value for day t is

$$W2r(t) = \frac{x_{1t}-\hat{\mu}_t}{MAR_t}, \tag{2.3}$$

where the expected value for day t is a function of the average rate, and the estimated standard deviation is based on the mean absolute residual (MAR), i.e.,

$$\hat{\mu}_t = d_t \frac{\mathrm{BLS_t}}{\mathrm{BLV_t}} \quad \text{and} \quad MAR_t = \frac{1}{n}\sum_{i=1}^{n} |y_{it} - \hat{\mu}_{it}|, \tag{2.4}$$

where $\hat{\mu}_{it}$ refers to the estimated mean count for day i in the baseline period. Similar to s_t in Eq.

(2.2), if MAR_t is less than one, it is assigned a value of one.

We propose two modified W2r methods in this dissertation, primarily for two reasons. First, the definition of the MAR_t does not reflect the total number of counts or visits at time t. Second, other estimators for the standard deviation may perform better than the mean absolute residual, also called the mean absolute deviation.

Tokars et al. [6] reported that use of the W2 rate method produces a more accurate expected count value and lower residuals than with use of the W2 count method. They used real daily syndrome counts from two sources as baseline data and assessed the ability of the rate algorithm to improve the performance of the EARS approach in terms of sensitivity, i.e. the power values.

We consider an adaptive threshold method originally proposed by Lambert and Liu [4] in Chapter 3. Performance evaluations of the W2 rate, the modified W2 rate, and the adaptive threshold methods are further explored in Chapters 4 and 5.

Chapter 3 Adaptive Threshold Method

An adaptive threshold method used by Lambert and Liu [4] for computer network monitoring leads to an alternative to the W2 rate method. It is interesting to note that Lambert and Liu [4] mentioned that a referee said their method could be modified for use in public health surveillance.

Using the same baseline information as the W2r method, we can estimate the parameters of an assumed underlying parametric distribution. We use the conditional binomial distribution and conditional negative binomial distribution as the reference distributions with the adaptive threshold method.

3.1 Conditional Binomial Distribution

We consider two independent Poisson distributions for modeling count data for the W2 rate method. For day t, we let X_{1t} be the syndrome count, X_{2t} be the non-syndrome count, $D_t = X_{1t} + X_{2t}$ be the total number of visits for that day, and d_t be the observed value of D_t. The probability mass function (pmf) for the count X_{1t} or X_{2t} is

$$f(x_{it}|\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{x_{it}}}{x_{it}!}; x_{it} = 0,1,\dots; \lambda_i > 0;\ i = 1,2, \tag{3.1}$$

where λ_1 is the Poisson parameter for syndrome counts and λ_2 is the Poisson parameter for non-syndrome counts. Conditional on the total number of visits for day t, the syndrome count X_{1t} is distributed as a binomial random variable with parameters d_t and $\pi = \frac{\lambda_1}{\lambda_1+\lambda_2}$. See Przyborowski and Wilenski [14] for this conditional binomial result related to the two Poisson variables. The probability mass function for the count X_{1t} conditioned on d_t is

$$f(x_{1t}|d_t,\pi) = \binom{d_t}{x_{1t}} \pi^{x_{1t}}(1-\pi)^{(d_t-x_{1t})};\ x_{1t} = 0,1,\dots,d_t. \qquad (3.2)$$

For the maximum likelihood (ML) estimators, we have

$$\hat{\pi}_t = \frac{BLS_t}{BLV_t} \qquad (3.3)$$

where $\mathrm{BLS_t}$ and $\mathrm{BLV_t}$ are the total syndrome counts and total visits over the baseline period, respectively.

3.2 Conditional Negative Binomial Distribution

We also consider two independent negative binomial distributions for modeling count data for the W2 rate method. For day t, again let X_{1t} be the syndrome count, X_{2t} be the non-syndrome count, and $D_t = X_{1t} + X_{2t}$ be the total number of visits for that day, and d_t be the observed value of D_t. The probability mass function (pmf) for the count X_{1t} or X_{2t} is

$$f(x_{it}|\lambda_i) = \binom{\mathrm{r_i}+x_{it}-1}{x_{it}} \mathrm{p_i^{r_i}}(1-\mathrm{p_i})^{x_{it}}; x_{it} = 0,1,\dots; r_i > 0; 0 < \mathrm{p_i} < 1; i = 1,2, \quad (3.4)$$

where r_1 and p_1 are the negative binomial parameters for syndrome counts, and r_2 and p_2 are the negative binomial parameters for non-syndrome counts. The mean and variance of the negative binomial distributions are $\mu_i = \frac{r_i(1-p_i)}{p_i}$ and $\sigma_i^2 = \frac{r_i(1-p_i)}{p_i^2}$, i=1,2, respectively. Conditional on the total number of visits for day t, the syndrome count X_{1t} is distributed as a conditional negative binomial random variable. The probability mass function (pmf) for the count X_{1t} conditioned on d_t is

$$f(x_{1t}|d_t,r_1,p_1,r_2,p_2) = \frac{\binom{r_2+d_t-x_{1t}-1}{d_t-x_{1t}}\binom{r_1+x_{1t}-1}{x_{1t}}\left(\frac{1-p_1}{1-p_2}\right)^{x_{1t}}}{\sum_{x=0}^{d_t}\binom{r_2+d_t-x_{1t}-1}{d_t-x_{1t}}\binom{r_1+x_{1t}-1}{x_{1t}}\left(\frac{1-p_1}{1-p_2}\right)^{x_{1t}}} I_{x_{1t}\in(0,d_t)} \qquad (3.5)$$

The pmf given by Eq. (3.5) is derived below in the proof of Theorem 3-1.

Theorem 3-1: Let $X\sim$negative binomial(r_1, p_1), $Y\sim$negative binomial(r_2, p_2), X and Y are independent. Let V=X+Y. The pmf of X given V is given by

$$f(x|v, r_1, p_1, r_2, p_2) = \frac{\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{1-p_1}{1-p_2}\right)^x}{\sum_{x=0}^{v}\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{1-p_1}{1-p_2}\right)^x} I_{x\in(0,v)},$$

Proof:

The sample space for X is $\mathcal{X} = \{x: 0,1, \ldots\}$, and the sample space for Y is $\mathcal{Y} = \{y: 0,1, \ldots\}$. Therefore the sample space for V is $\mathcal{V} = \{v: v = g(y), y \in \mathcal{Y}, x \in \mathcal{X}\} = \{x, x+1, \ldots\}$. For any $v \in \mathcal{V}$, $y + x = g(y) = v$ if and only if $y = v - x$. So $g^{-1}(v)$ is the single point $y = v - x$ when x is given. Let $q_1 = 1 - p_1$, and $q_2 = 1 - p_2$. We have

$$\begin{aligned}
f_V(v|x) &= \sum_{y\in g^{-1}(v)} f_Y(y|x) = \sum_{y\in g^{-1}(v)} f_Y(y) \\
&= f_Y(v-x) \\
&= \binom{r_2+v-x-1}{v-x} p_2^{r_2} {q_2}^{(v-x)}
\end{aligned}$$

because X and Y are independent. It follows that

$$\begin{aligned}
f_V(v) &= \sum_{all\ x} f(x,v) = \sum_{x=0}^{\infty} f(v|x)\, f_X(x) \\
&= \sum_{x=0}^{\infty} \binom{r_2+v-x-1}{v-x} p_2^{r_2} {q_2}^{(v-x)} \binom{r_1+x-1}{x} p_1^{r_1} {q_1}^{x}.
\end{aligned}$$

For any fixed nonnegative integer v, $f(x,v)>0$ only for x=0, 1,..., v. Since,

$$\begin{aligned}
f_V(v) &= \sum_{x=0}^{v} \binom{r_2+v-x-1}{v-x} p_2^{r_2} {q_2}^{(v-x)} \binom{r_1+x-1}{x} p_1^{r_1} {q_1}^{x} \\
&= p_2^{r_2} p_1^{r_1} {q_2}^{v} \sum_{x=0}^{v} \binom{r_2+v-x-1}{v-x} \binom{r_1+x-1}{x} \left(\frac{q_1}{q_2}\right)^x,
\end{aligned}$$

we have

$$f(x|v) = \frac{f(x,v)}{f_V(v)} = \frac{f(v|x)f_X(x)}{f_V(v)} = \frac{\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{q_1}{q_2}\right)^x}{\sum_{x=0}^{v}\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{q_1}{q_2}\right)^x}$$

$$= \frac{\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{1-p_1}{1-p_2}\right)^x}{\sum_{x=0}^{v}\binom{r_2+v-x-1}{v-x}\binom{r_1+x-1}{x}\left(\frac{1-p_1}{1-p_2}\right)^x}.$$

Note: If $p_1 = p_2$, then $X|X+Y$ follows a negative hypergeometric distribution. See Jain and Consul [15].

A Monte-Carlo simulation was carried out to illustrate and provide a check on the proof of Theorem 3-1. We assume the parameter set is given as (r_1, p_1, r_2, p_2) = (80, 0.1, 50, 0.3). In Figure 3-1 the red dotted line with legend Analytical refers to the probability mass function for the conditional negative binomial distribution of X given v. The blue solid line with legend Monte-Carlo refers to the probability mass function of X given v estimated using Monte-Carlo simulation with 1,000,000 replications. The two curves are very close to each other given v = 800, 850, 900, and 950, respectively, as shown in Figure 3-1.

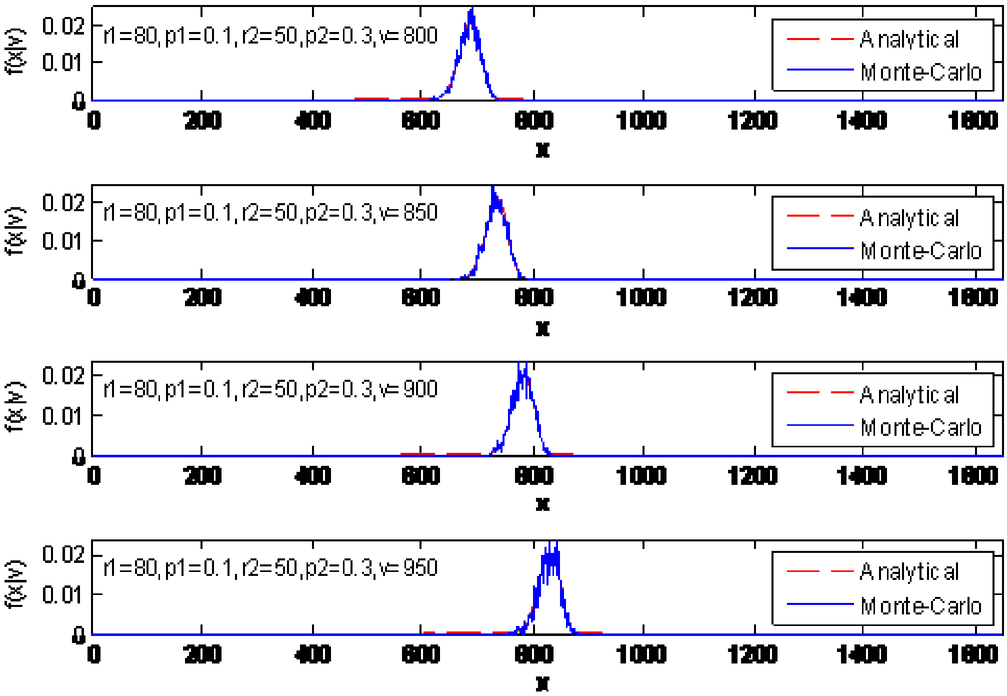

Figure 3-1: Example of Monte-Carlo Simulation on the Conditional Negative Binomial Distribution Given by Theorem 3-1.

For the method of moments (MOM) estimators of the negative binomial parameters for syndrome counts and nonsyndrome counts, we have

$$\hat{p}_{it} = \frac{\bar{x}_{it}}{\hat{\sigma}_{it}^2} \text{ and } \hat{r}_{it} = \frac{\bar{x}_{it}^2}{\hat{\sigma}_{it}^2 - \bar{x}_{it}} \text{ ; } i = 1,2, \tag{3.6}$$

where $\bar{x}_{it} = \frac{1}{n}\sum_{j=1}^{n} y_{ijt}$, $\hat{\sigma}_{it}^2 = \frac{1}{n}\sum_{j=1}^{n}(y_{ijt} - \bar{x}_{it})^2$, and y_{ijt}, i=1,2, j=1,2,…,n, correspond to the eligible baseline syndrome data (i=1) and nonsyndrome data (i=2) for day t . Clearly the domain of these parameters is violated when $\bar{x}_{it} > \hat{\sigma}_{it}^2, i = 1,2$. The MOM estimation problem will be discussed in more detail in Chapter 5.

3.3 Z-Scores Approach

For day t, an upper-tail p-value, P_t, can be computed based on the conditional binomial distribution with the estimated parameter $\hat{\pi}_t$. Then P_t is approximately distributed uniformly over [0,1] when the underlying distribution is in-control and there is no outbreak. Using the inverse normal CDF, we can obtain a standard normal Z-score, Z_t, with an approximate mean zero and variance of one. The equations for our conditional binomial variable X_{1t} with observed value x_{1t} are

$$P_t = P(X_{1t} \geq x_{1t} | \hat{\pi}_t), \tag{3.7}$$

and

$$Z_t = \Phi^{-1}(1 - P_t). \tag{3.8}$$

For a Shewhart approach, a signal is given when $Z_t \geq h_{Z\_AT}$, where $h_{Z\_AT} > 0$ is a specified threshold value. We also study the properties of an EWMA chart based on the Z-score values. We will refer to this approach as either the Z-Score method or the adaptive threshold method throughout the dissertation.

3.4 One-Sided EWMA Method

The exponentially weighted moving average (or EWMA) control chart has been widely used in traditional quality control applications since it was first proposed by Roberts [16]. See Crowder [17,18] and Lucas and Saccucci [19] for good discussions of the EWMA method.

While the Shewhart decision rule relies on using one observation at a time, the EWMA statistic incorporates information using past observations with observations closer to the current time point given larger weights than those further back in time. For standardized variables, say v_t, $t = 1, 2, \ldots$, the EWMA statistics E_t are

$$E_t = \alpha v_t + (1-\alpha)E_{t-1}, t = 1,2,\dots, \qquad (3.9)$$

where α is the weight given to the current observation and $E_0 = 0$. When $\alpha = 1$, the EWMA method reduces to a Shewhart chart. Montgomery [20] recommended using weights between .05 and .25 for EWMA charts. Smaller values of α are recommended for detecting smaller shifts quickly, and larger values are recommended for larger shifts.

In most industrial applications, a two-sided EWMA chart is used, signaling for abnormally low or large values of the EWMA statistic. However, we are only concerned with outbreaks in our applications, so a one-sided chart is used. The one-sided EWMA statistics are expressed as

$$E_t = max\,[0, \alpha v_t + (1-\alpha)E_{t-1}], t = 1,2,\dots, \qquad (3.10)$$

A signal is given if $E_t \geq h_{ET}$, where h_{ET} >0 is a specified threshold. The reflecting barrier at zero is used so that the statistic does not become very small. If this is not done and an outbreak occurred when the statistic is very small, it would be more difficult to signal. Lambert and Liu [4] recommended using a one-sided EWMA chart, but did not use the reflecting barrier at zero shown in Eq. (3.10) that we recommend and use in our RI threshold function and power analyses. Failure to use a reflecting barrier in a one-sided EWMA chart can lead to serious inertial problems, a topic discussed by Woodall and Mahmoud [21]. For more on a one-sided EWMA method, see Crowder and Hamilton [22] and Champ, Woodall, and Mohsen [23].

In traditional quality control applications, the EWMA statistic is reset to zero after a signal. This happens as a result of stopping a process, taking a corrective action, and then resuming the process. However, the EWMA statistic will not be reset after a signal in our applications because the monitoring statistics are not usually reset following a signal in public health surveillance.

To further motivate use of the EWMA in our dissertation, consider Figure 3-2. Figure 3-2 (above) represents 100 simulated values of Z-scores using Eq. (3.1), Eq. (3.2) and Eq. (3.8) given a window size n=7 and $\lambda_1 = \lambda_2$=50, with a 20% increase in λ_1 for seven days beginning at observation 29. In Figure 3-2 (below), the same observations are transformed and smoothed using Eq. (3.10) with α =0.2. Clearly, it is easier to observe the increase in the mean using the EWMA of the normal scores.

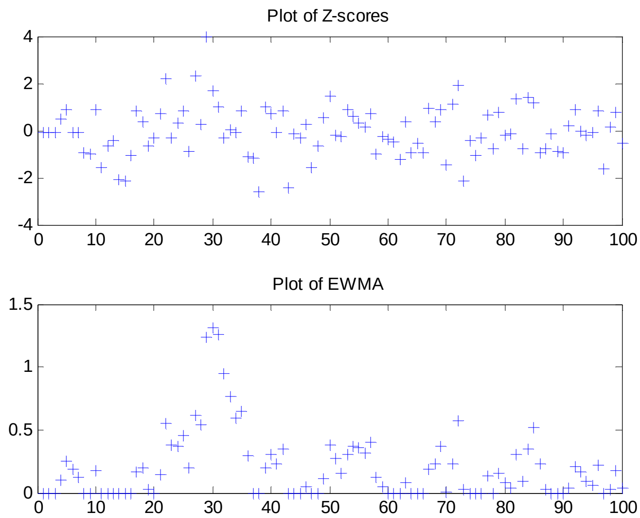

Figure 3-2: Example of the Statistic Values of Z-score and the EWMA Methods

In summary, for an incoming observed count x_{1t} at time *t*, the proposed adaptive threshold method consists of three basic steps:

1. Obtain the estimated parameters for the reference conditional binomial distribution or conditional negative binomial distribution at time *t* based on the baseline results.
2. Compute the tail probability p-value, P_t, and the normal score $Z_t = \Phi^{-1}(1 - P_t)$ for each incoming count x_{1t} under its reference distribution. Signal an outbreak when $Z_t \geq h_{Z\_AT}$, where $h_{Z\_AT} > 0$ is a specified threshold value, for the Shewhart-type approach.
3. Update the EWMA of the normal scores, $E_t = max\,[0, \alpha Z_t + (1 - \alpha)E_{t-1}]$, and signal an outbreak when $E_t \geq h_{ET}$, where h_{ET} >0 is a specified threshold value, for the EWMA approach.

As Lambert and Liu [4] reported, the *p* values for the counts provide a natural way to monitor the performance of the approach. These *p* values are approximately uniformly distributed when there is no outbreak; if not, a different reference distribution may be required. Lambert and Liu [4] pointed out that the way they define an EWMA of the Z-scores and then threshold it against a constant limit gives a Q-chart in the terminology of statistical quality control [24], although with the Q-chart approach all of the past data are used as the baseline, not the limited baseline of the adaptive threshold method.

We studied the effect of using an EWMA approach versus the traditional Shewhart approach for the adaptive threshold, W2r, and modified W2r methods for the RI threshold function and power analyses in Chapters 4 and 5.

Chapter 4 Performance Evaluation of Adaptive Threshold and W2r Methods with Poisson Inputs

In this chapter, we report the results of a simulation study for the conditional binomial distribution with two independent Poisson inputs. We explore the RI threshold function analysis and the power analysis for the adaptive threshold method, W2r method, and a modified W2r method. We examine the performance of both the Shewhart and the one-sided EWMA approaches for these methods. An analysis of the weekend effects follows in Section 4.5.

4.1 Simulation Plan

4.1.1 In-control Data

We assumed weekday and weekend counts each follow independent Poisson distributions where there is no outbreak. More precisely, we assumed the syndrome counts in weekdays follows a Poisson distribution with the parameter λ_1, the non-syndrome counts in weekdays follows a Poisson distribution with the parameter λ_2, the syndrome counts in weekends follows a Poisson distribution with the parameter λ_3, and the non-syndrome counts in weekends follows a Poisson distribution with the parameter λ_4. For simplicity, we first assume $\lambda_1 = \lambda_3$ and $\lambda_2 = \lambda_4$. We used the parameter combinations as listed in Table 4-1 for the conditional binomial study. The ratio of λ_1 and λ_2 was varied from 0.1 to 10. Correspondingly, the conditional binomial proportion π, which is defined as $\pi = \frac{\lambda_1}{\lambda_1+\lambda_2}$, took values from .0909 to .9091. We used the simulated in-control data to check how closely the uniform(0,1) distribution fits the in-control p-values for the adaptive threshold methods in Section 4.2.1, and then used the data to do the Recurrence Interval (RI) threshold function analysis described in Section 4.3.

Table 4-1: Poisson Parameters Used in the Conditional Binomial Study

λ_2 ($\lambda_1 = 10$)	λ_2 ($\lambda_1 = 50$)	λ_2 ($\lambda_1 = 100$)	λ_2 ($\lambda_1 = 200$)	λ_2 ($\lambda_1 = 500$)	$\lambda_1 : \lambda_2$	π
100	500	1000	2000	5000	0.1	.0909
50	250	500	1000	2500	0.2	.1667
20	100	200	400	1000	0.5	.3333
10	50	100	200	500	1	.5000
5	25	50	100	250	2	.6667
2	10	20	40	100	5	.8333
1	5	10	20	50	10	.9091

4.1.2 Outbreak Data

In Section 4.1.1 we discussed the simulation of in-control data over time, where the distribution parameters stay constant. In this section we examine syndromic outbreaks. We are only interested in an increase in syndrome counts and rates, so one-sided methods are used. Baseline data were first simulated from the in-control distributions for ten weeks, and then an outbreak lasting seven days was injected. This process was repeated 100,000 times for each parameter combination considered. For each of these transient shifts, we determined the proportion of times the various methods signaled during the outbreak. We used the simulated outbreak data to do the power analysis with results reported in Section 4.4.

4.2 Methods

4.2.1 Adaptive Threshold Methods

As shown in Section 3.1, if X_{1t} and X_{2t} are two independent Poisson variables with $X_{1t} \sim \text{Poisson}(\lambda_1), X_{2t} \sim \text{Poisson}(\lambda_2)$, $D_t = X_{1t} + X_{2t}$, and d_t as the observation value of D_t, then $X_{1t} | d_t \sim \text{Bin}(d_t, \pi)$, where $\pi = \frac{\lambda_1}{\lambda_1 + \lambda_2}$. We let BLS_t and BLV_t be the total number of syndromic counts and facility visits over the baseline period. The MLE for π is $\widehat{\pi}_t = \frac{\text{BLS}_\text{t}}{\text{BLV}_\text{t}}$, and $\hat{\mu}_t$ is the MLE estimator for $\text{E}(x_{1t} | d_t)$, i. e., $\hat{\mu}_t = d_t \frac{\text{BLS}_\text{t}}{\text{BLV}_\text{t}}$. We consider the adaptive threshold method using MLE estimators and the adaptive threshold method assuming the baseline parameters are known in the following simulation study.

The adaptive threshold method works best if the in-control p-values are approximately distributed uniformly over [0,1]. Figures 4-1 to 4-5 show how the in-control p-values of the adaptive threshold methods, assuming known parameters (left) or using MLE estimators (right), are distributed with n=7 given λ_1=10, 50 100, 200 and 500, respectively. The Q-Q plots show some tails deviated from the reference line for the adaptive threshold method when λ_1 is as small as 10 or 50. There are very good matches when λ_1 =100, 200, and 500 in the Q-Q plots for the adaptive threshold method assuming known parameters. There are only slight deviations from the uniform(0,1) distribution for the adaptive threshold method using MLE when λ_1=100, 200 and 500. Overall it can be seen that in-control p-values are approximately uniformly distributed over (0,1) for the cases considered here.

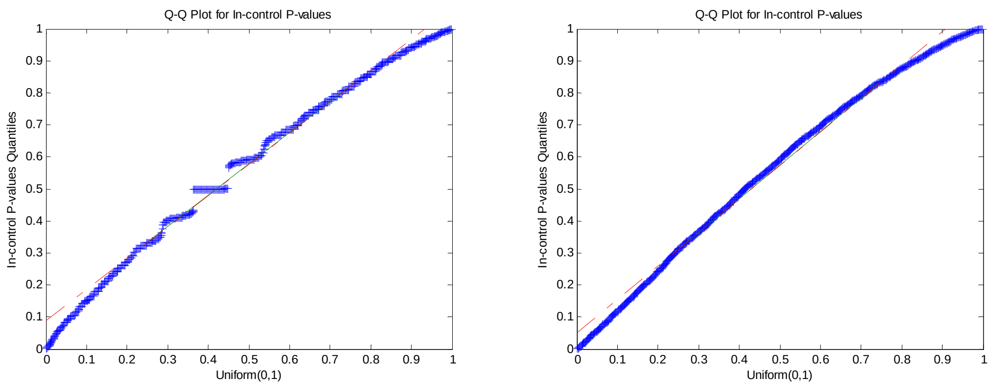

Figure 4-1: Q-Q Plots of In-control P-values for Adaptive Threshold Method Using Known Parameters (left) and MLE (right)-Conditional Binomial Distribution with Poisson Inputs- λ_1=10, n=7

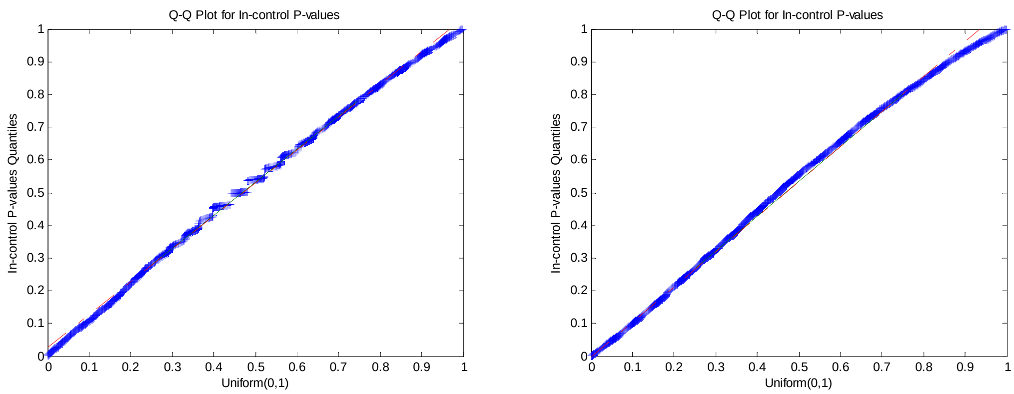

Figure 4-2: Q-Q Plots of In-control P-values for Adaptive Threshold Method Using Known Parameters (left) and MLE (right)-Conditional Binomial Distribution with Poisson Inputs- λ_1=50, n=7

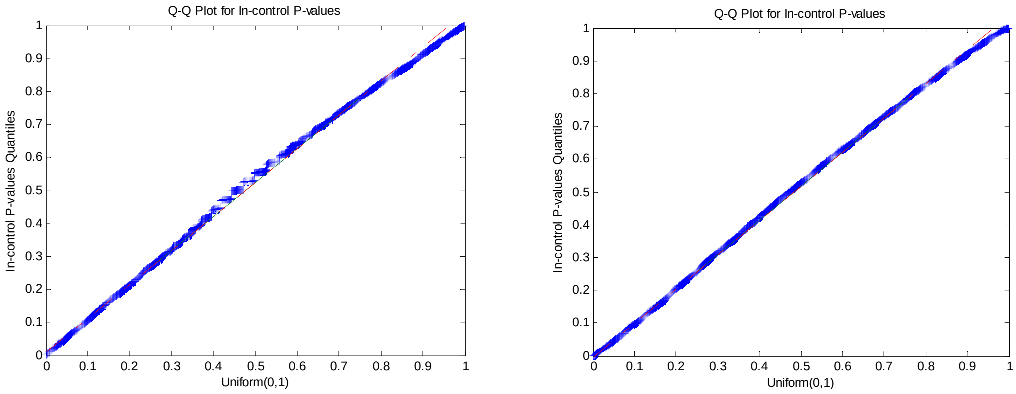

Figure 4-3: Q-Q Plots of In-control P-values for Adaptive Threshold Method Using Known Parameters (left) and MLE (right)-Conditional Binomial Distribution with Poisson Inputs- λ_1=100, n=7

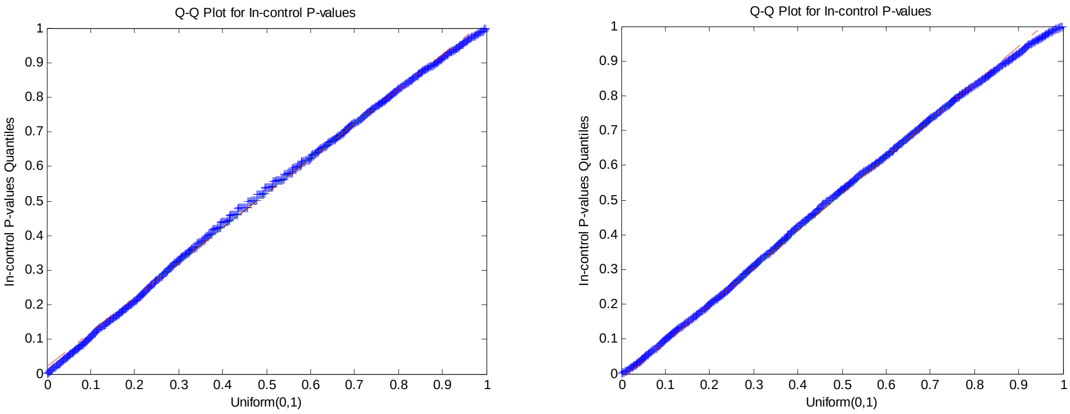

Figure 4-4: Q-Q Plots of In-control P-values for Adaptive Threshold Method Using Known Parameters (left) and MLE (right)-Conditional Binomial Distribution with Poisson Inputs- λ_1=200, n=7

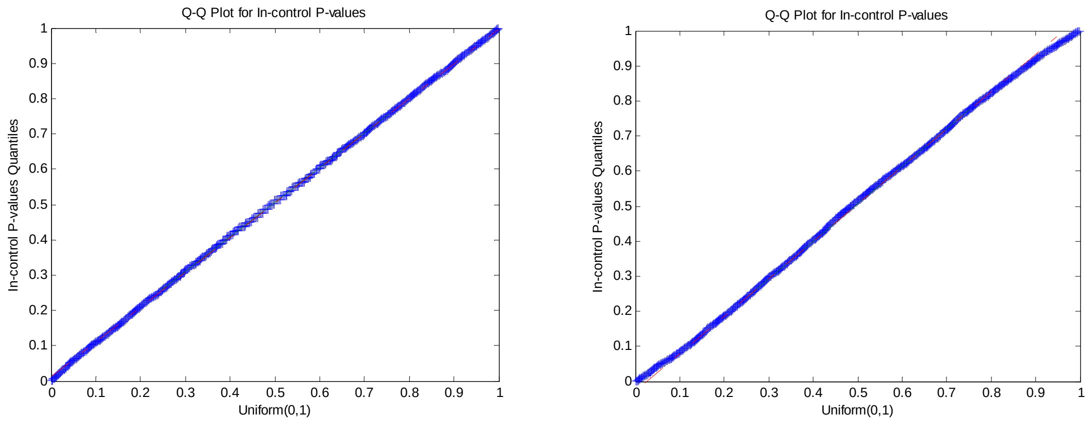

Figure 4-5: Q-Q Plots of In-control P-values for Adaptive Threshold Method Using Known Parameters (left) and MLE (right)-Conditional Binomial Distribution with Poisson Inputs- λ_1=500, *n*=7

4.2.2 W2r and Modified W2r Methods

We propose a modified W2r method in this section, W2r_1. The surveillance statistics of the W2r_1 method are defined as

$$W2r\_1(\mathrm{t}) = \frac{x_{1t} - \hat{\mu}_t}{max\,(1,\sqrt{d_t\,\hat{\pi}_t(1-\hat{\pi}_t)})}, \qquad (4.1)$$

where $\hat{\mu}_t = d_t \frac{\mathrm{BLS_t}}{\mathrm{BLV_t}}$, $\hat{\pi}_t = \frac{BLS_t}{BLV_t}$, and $\sqrt{d_t\,\hat{\pi}_t(1-\hat{\pi}_t)}$ is the estimated standard deviation based on the conditional binomial distribution with Poisson inputs. Similar to MAR_t in Eq. (2.4), if the standard deviation is less than one, it is assigned a value of one. We stated in Section 2.2 that the definition of the mean absolute residual (MAR_t) for the W2r statistic did not reflect the total number of counts or visits at time *t*. The standard deviation defined in Eq. (4.1) solves this problem.

4.3 RI Threshold Function Analysis

We used an empiric recurrence interval (RI) as one of the performance measures, which is defined as the fixed number of time periods for which the expected number of false alarms is one. Using a large number of values of a statistic, empiric recurrence intervals (RI's) can be calculated. The CDC did not specify the data input to calculate their RI's for the W2r method. They only mentioned that "the frequency distributions used to calculate the RI's will be updated periodically"[1]. For any frequency distribution, we can find the percentage of days with values beyond a given threshold. Taking the reciprocal of this percentage gives the corresponding RI value. For instance, if five percent of values exceed 3.4 in a frequency distribution, this threshold has an estimated RI value of 20.

According to BioSense [1], parametric methods for developing recurrence intervals are not valid in the public health application because real health care data are typically non-normal [1]. Several discrete distributions may be of use, however, to better understand the W2 methods. If different types of simulated data were to yield W2 RI thresholds similar to those found in the CDC handbook, the W2 method would be considered robust to differences in the characteristics of the different input data streams. Tokars et al. [6] also noted that BioSense uses past data that may contain outbreaks [25] to obtain empiric RIs. Note that we assume there are no outbreaks in the simulations for the RI threshold analysis in this dissertation.

In this section the RI threshold functions are compared for the adaptive threshold method, W2r and modified W2r methods, and the BioSense method. The BioSense threshold values were obtained from Table 3 of the CDC's Hospital Data User Guide [1] when n=7. The first methods compared were Shewhart-based methods, where if a daily statistic, i.e., W2r or Z-scores, exceeds a particular threshold, a signal is given. The second set of methods compared were the EWMA-based methods with a smoothing parameter of α=0.2. The *RI-Shewhart* algorithm was used to estimate the RI values for the adaptive threshold and the W2r Shewhart approach as described below:

1. Simulate 1,000,000 days as the in-control data and calculate the statistic values of W2r, Z-score, or W2r_1 methods.

2. Compare the statistic values calculated in step 1 with a predetermined threshold (initially 0.2). Count the number of values signaling, i.e., statistics exceeding the threshold value.
3. Calculate the empirical RI as the reciprocal of the percentage of days signaling from the simulated 1,000,000 days.
4. Change the threshold from 0.2 to 9 with a step size of 0.1, and for each threshold value, repeat steps 2-3 to estimate the empirical RI.

The algorithm we used to derive the RI values for EWMA approach was similar to the *RI-Shewhart* algorithm. The only difference is that in step 1, we calculated the EWMA statistic values based on values of W2r, W2r_1, or Z-score with a smoothing constant of α=0.2.

4.3.1 Comparison of Adaptive Threshold and W2r Methods

For comparisons to the BioSense W2r RI functions, we first used $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_0$, where λ_0 =10, 50, 100, 200 and 500 for the Shewhart approach. When $\lambda_1 = \lambda_2$, we expect half of the daily visits to a health facility to be for a particular syndrome. In general this ratio is $\pi = \frac{\lambda_1}{\lambda_1+\lambda_2}$. Figure 4-6 shows the RI threshold functions for the W2r method, compared to the BioSense threshold functions. With systems such as the CDC EARS, it is important for the threshold functions not to vary very much since typically the same thresholds are used for all data input streams.

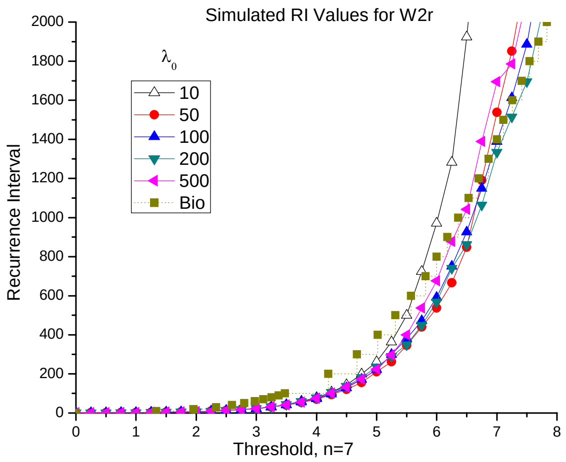

Figure 4-6: Threshold Curves Based on RIs: W2r Method Compared to BioSense W2r

The threshold functions for W2r under our model and the BioSense function are quite close to one another except for the case λ_0 =10 as shown in Figure 4-6. More specifically, when the threshold values range from approximately 2 to 5.5, the thresholds for BioSense are somewhat higher than all those for the W2r. The BioSense thresholds are very close to the W2r threshold function with λ_0 =500 when the thresholds range approximately from 6 to 7, and are very close to the W2r threshold functions with λ_0 =200 when the thresholds range approximately from 7 to 8. The BioSense data used in constructing these empiric RIs include outbreak data. This may explain why the thresholds differ somewhat in the Figure 4-6. Overall, however, it is surprising that the RI threshold functions based on the assumption of Poisson input data streams so closely match the RI threshold function used in BioSense that was based on actual syndromic data. This result did not hold for the W2c method as reported by Szarka, Gan, and Woodall [13].

Figures 4-7 to 4-11 show the RI threshold functions for the W2r method and adaptive threshold method using MLE estimators across different baselines, i.e., 7-day, 14-day, and 28-day, for the Shewhart approach. We used parameters with the ratio λ_1/λ_2 changing from 0.1 to 10 given λ_1=10, 50, 100, 200, and 500 in our study of the robustness of the threshold functions to parameter changes. Figure 4-7 shows there is considerable variability in the RI threshold curves for small counts with λ_1=10 for both the adaptive threshold and the W2r methods. Figures 4-8 and 4-9 show less variation in the RI threshold curves as $\lambda_1$increases. Figures 4-10 and 4-11 show that generally both methods work well for large counts with λ_1=200 and 500 since there is less variation in the corresponding RI threshold curves. Generally we observed that for any level of expected counts used in our simulation study, the Z-score threshold functions are more robust to parameter changes than the W2r threshold functions when n=7. As n increases, the Z-score threshold functions show almost the same robustness to the parameter changes, while the W2r threshold functions become more robust to the parameter changes. We also observed that the Z-score threshold values stay consistent for the different window sizes, while the W2r threshold values tend to decrease. Robustness to changes in parameter values, however, is much more important than robustness to different baseline lengths since different tables of thresholds can be given easily for different values of the baseline window lengths.

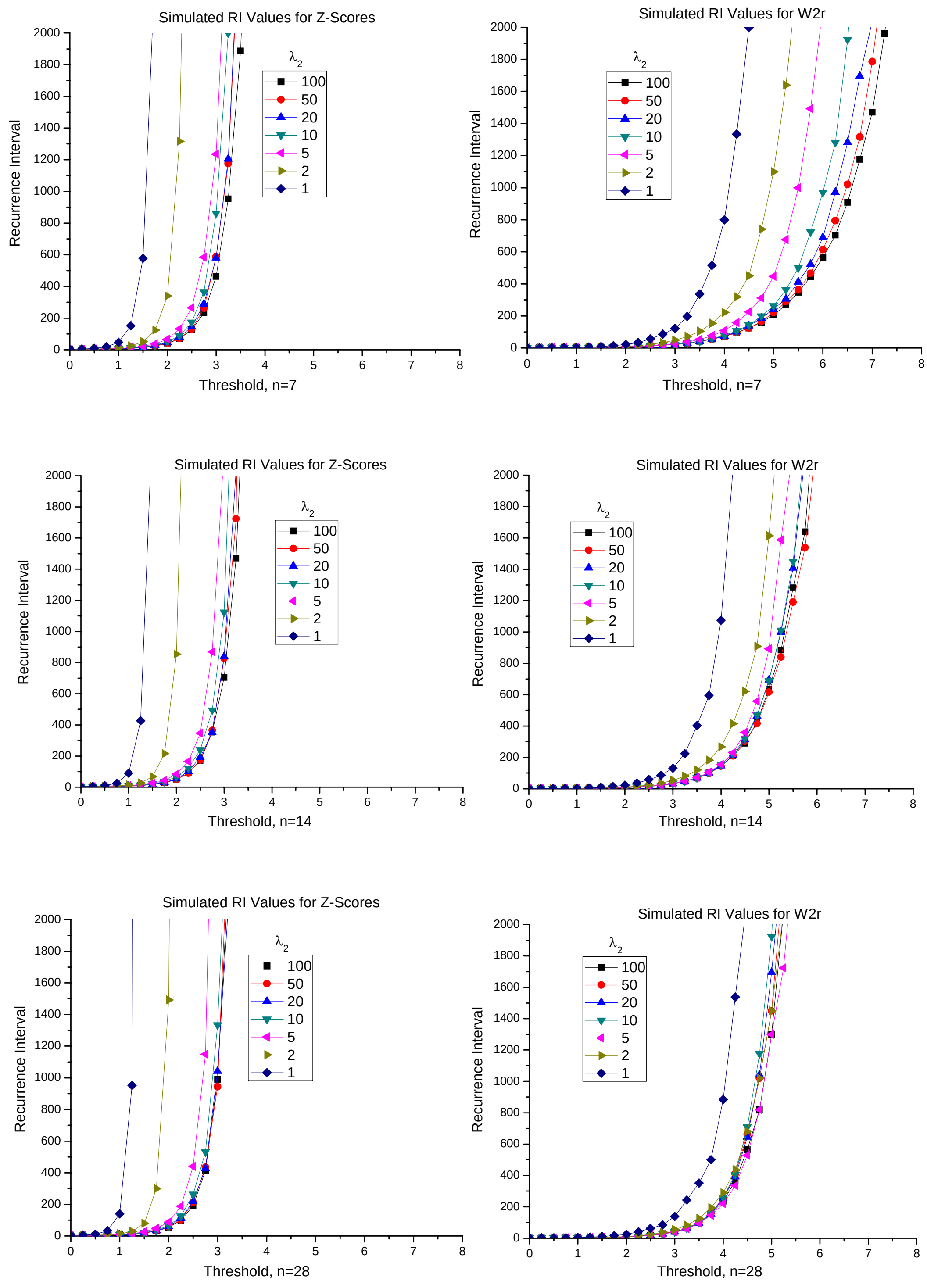

Figure 4-7: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=10-Shewhart

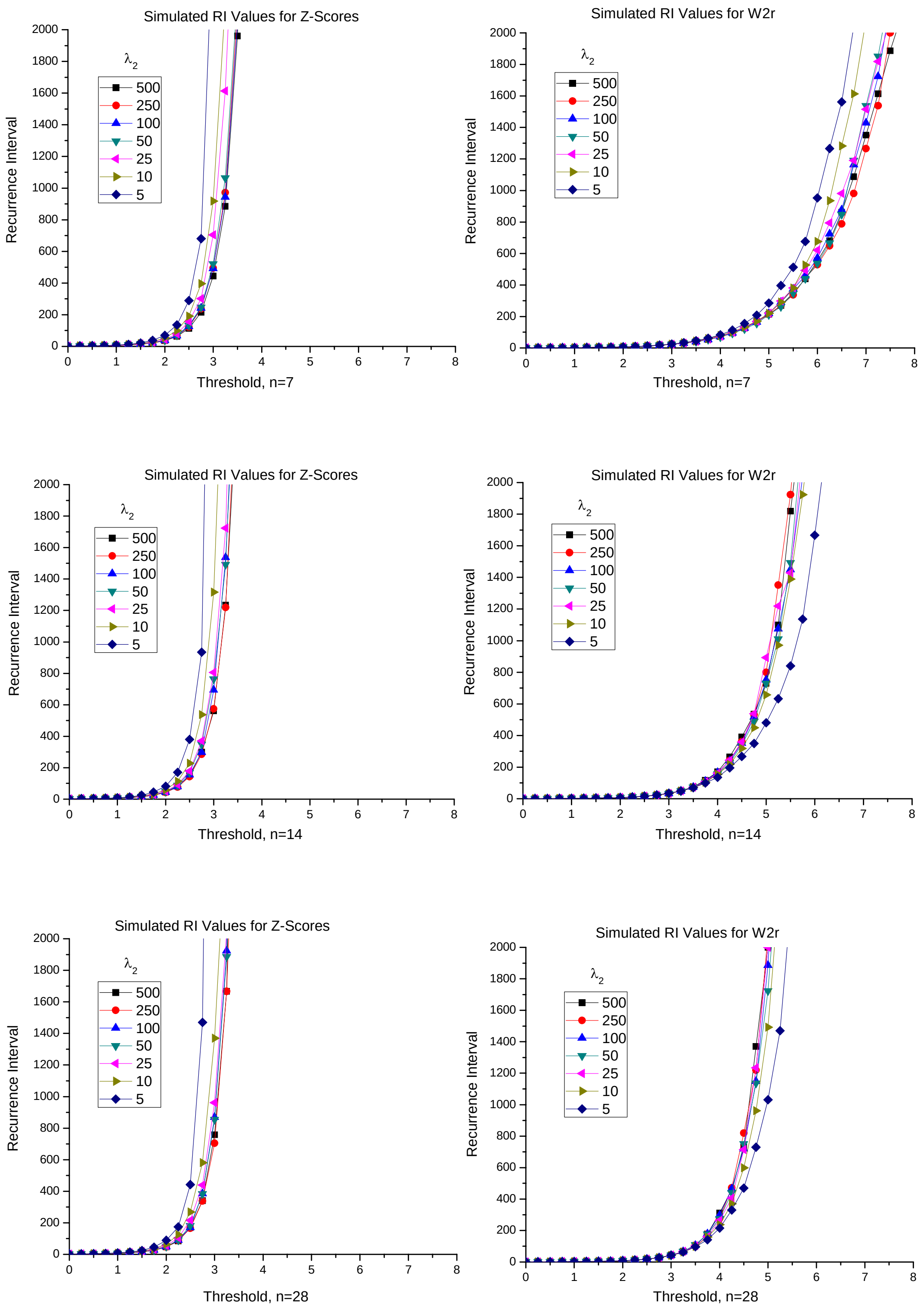

Figure 4-8: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=50-Shewhart

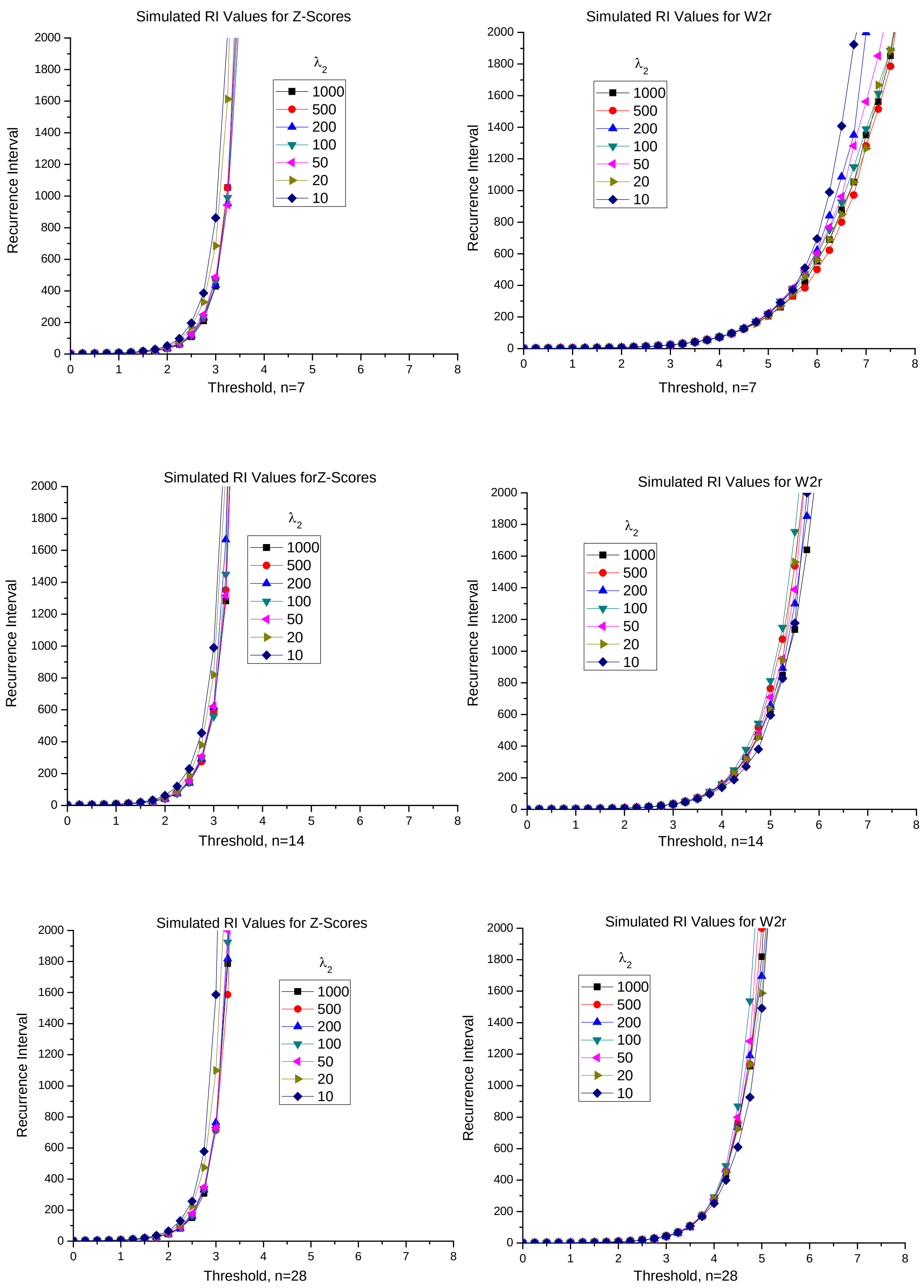

Figure 4-9: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=100-Shewhart

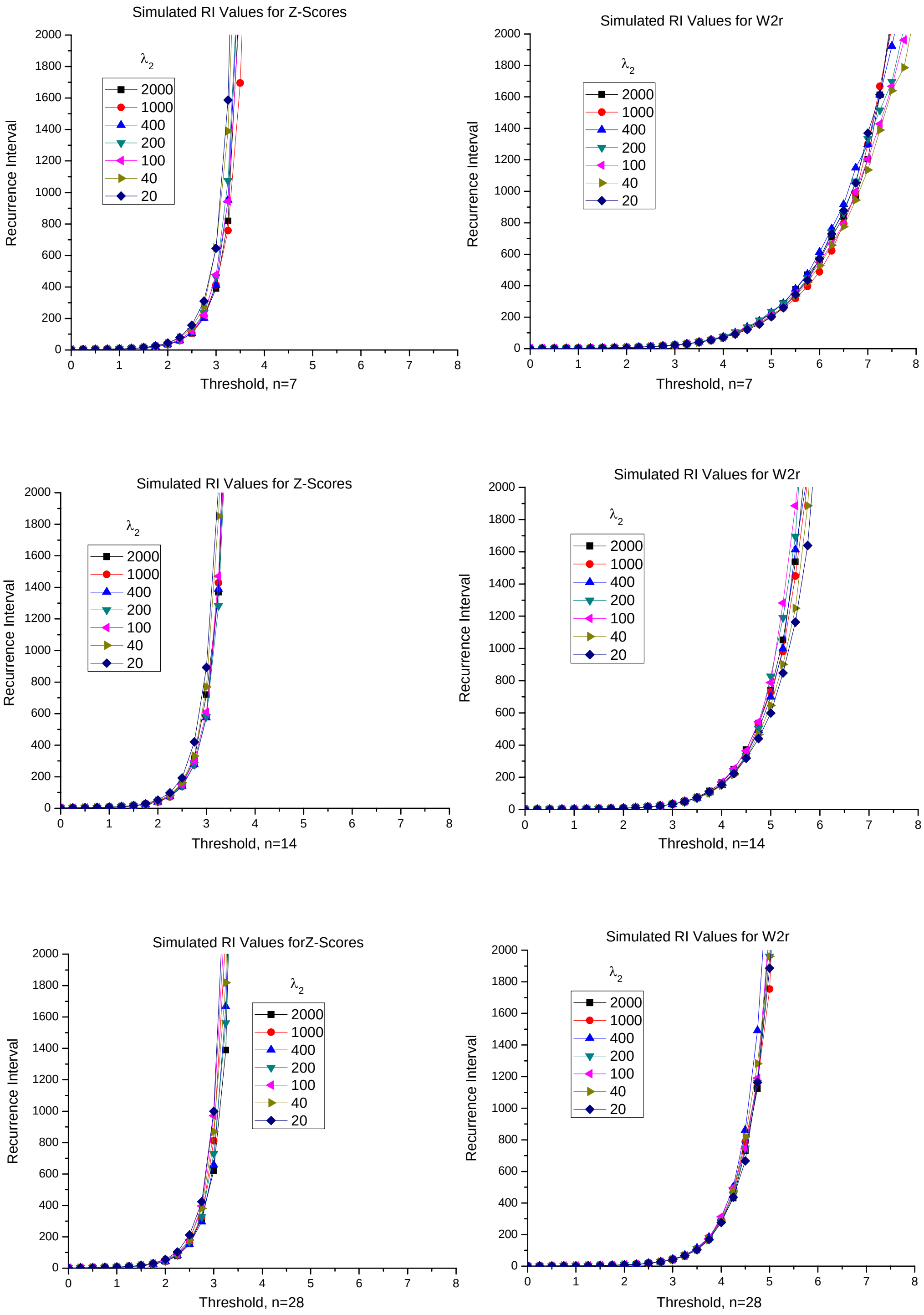

Figure 4-10: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=200-Shewhart

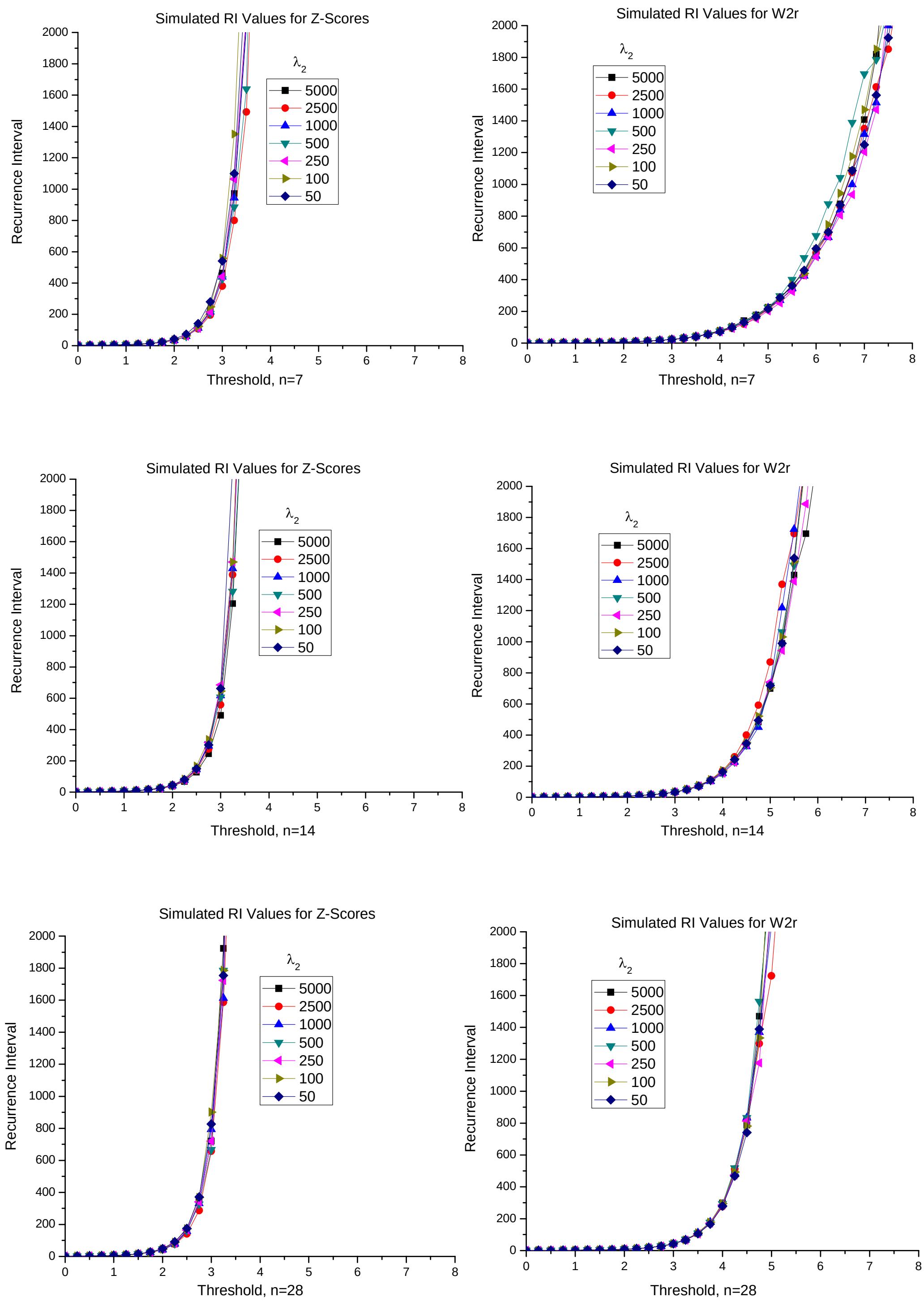

Figure 4-11: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=500-Shewhart

Figures 4-12 to 4-16 show, for the EWMA approach, the RI threshold functions for the W2r method and the MLE-based adaptive threshold method across different baselines with λ_1=10, 50, 100, 200, and 500. Figure 4-12 shows there is considerable variability in the RI threshold curves for small counts with the parameter λ_1=10 for both the adaptive threshold and W2r methods. Figures 4-13 to 4-16 show that the Z-score threshold functions are more robust to parameter changes than the W2r threshold functions when n=7. As n increases, the Z-scores threshold functions show almost the same robustness to the parameter changes, while the W2r threshold functions become more robust to the parameter changes.

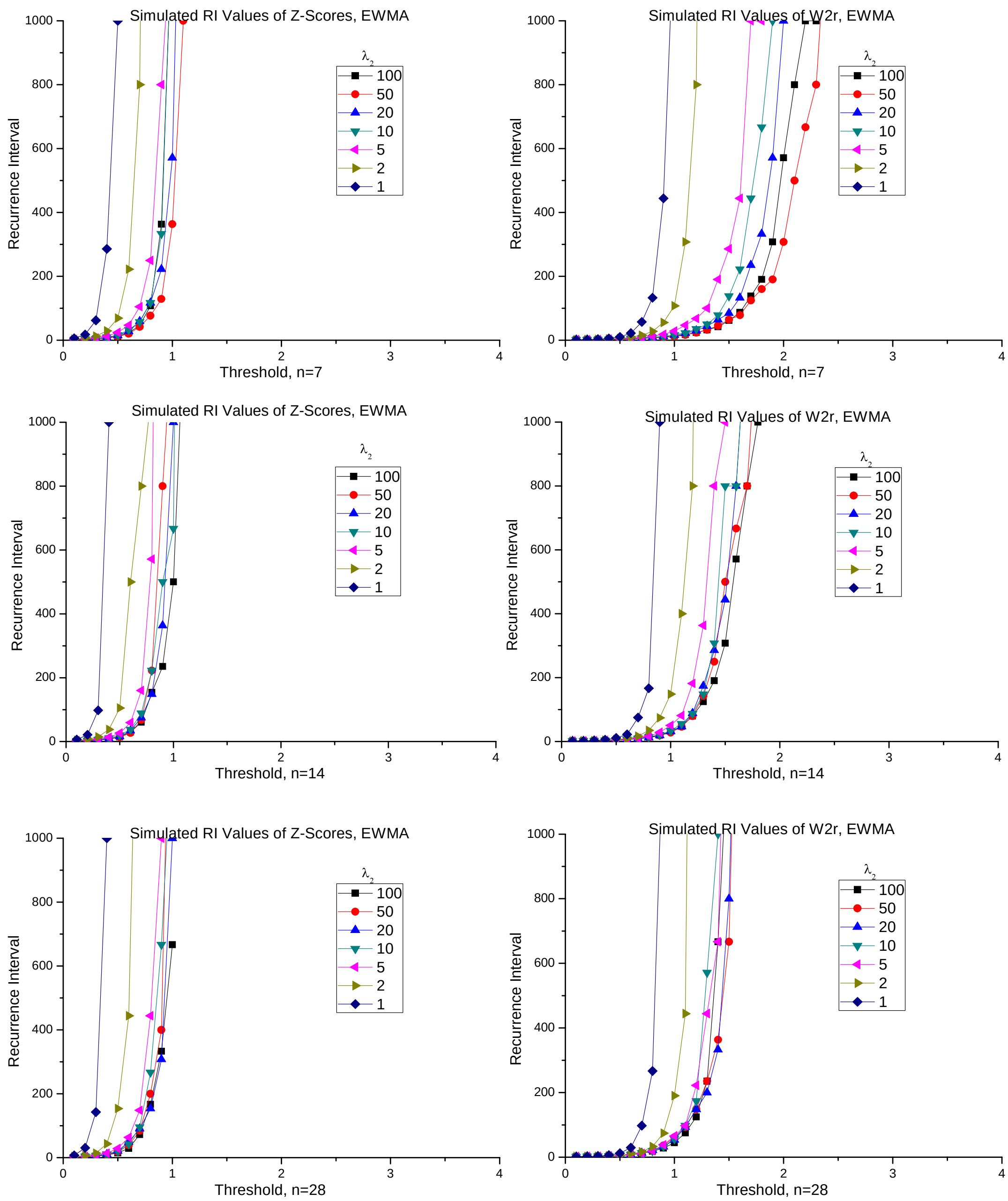

Figure 4-12: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=10-EWMA

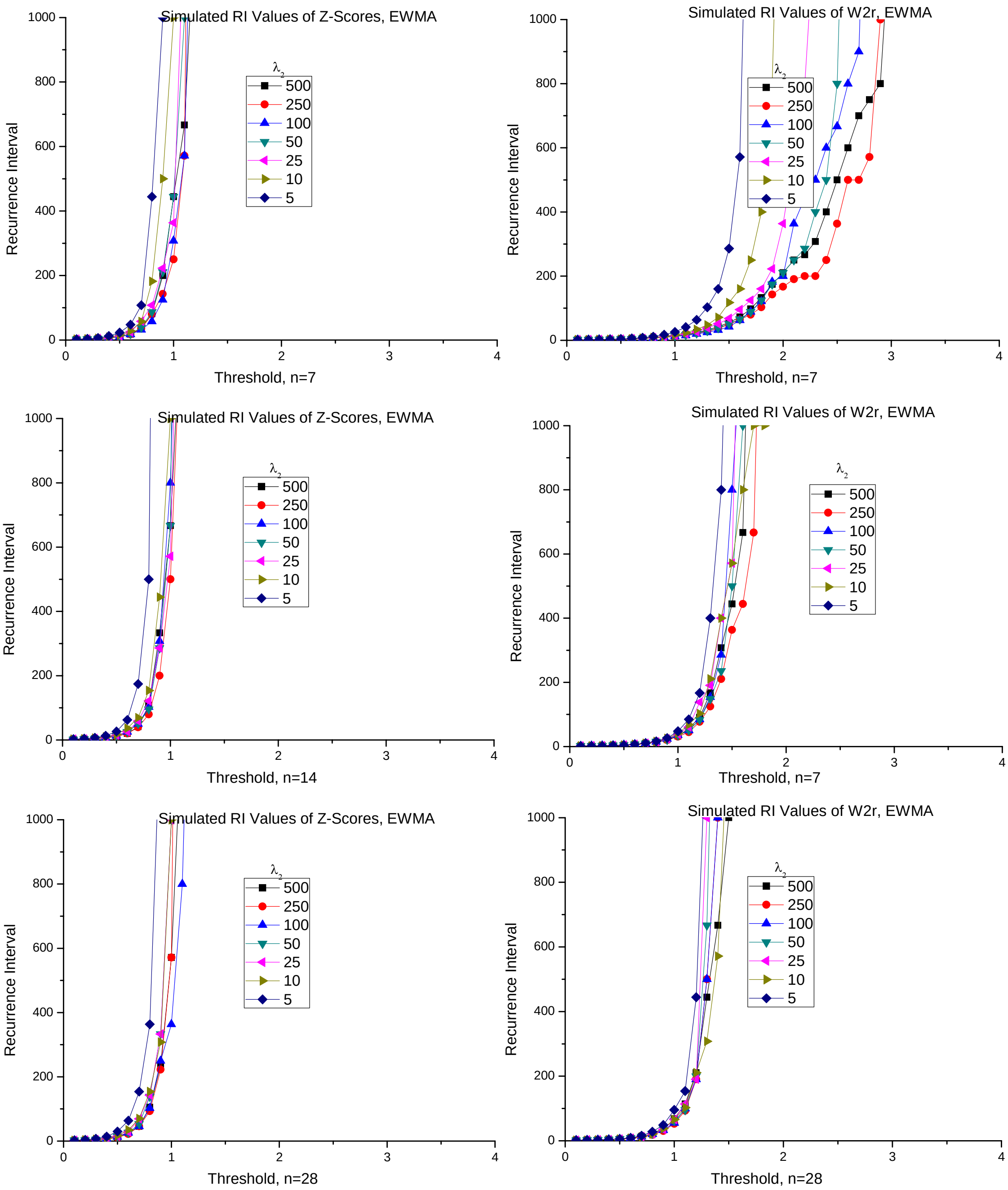

Figure 4-13: RI Thresholds for Adaptive Threshold Method (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=50-EWMA

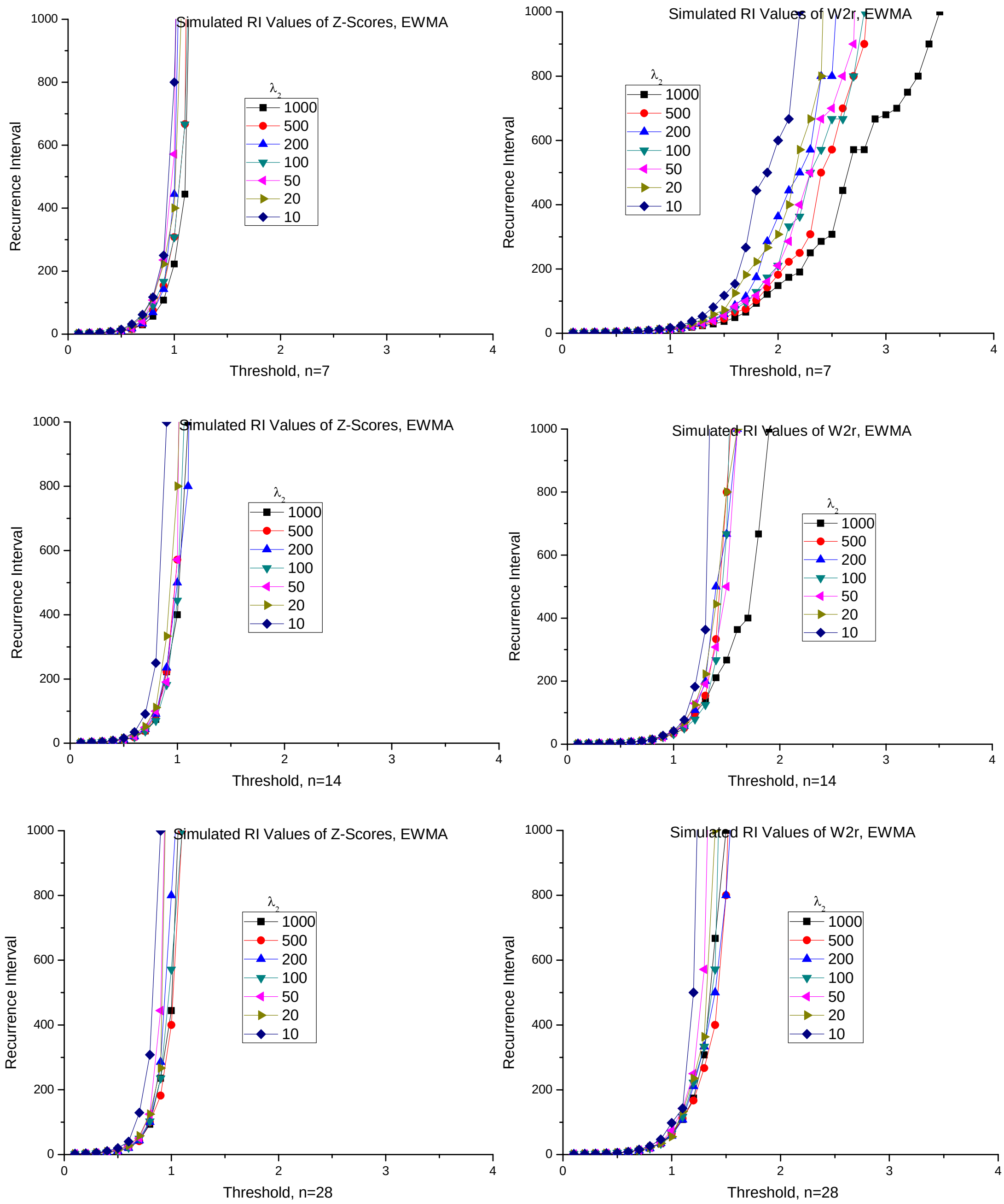

Figure 4-14: RI Thresholds for Adaptive Threshold Method Using MLE (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=100-EWMA

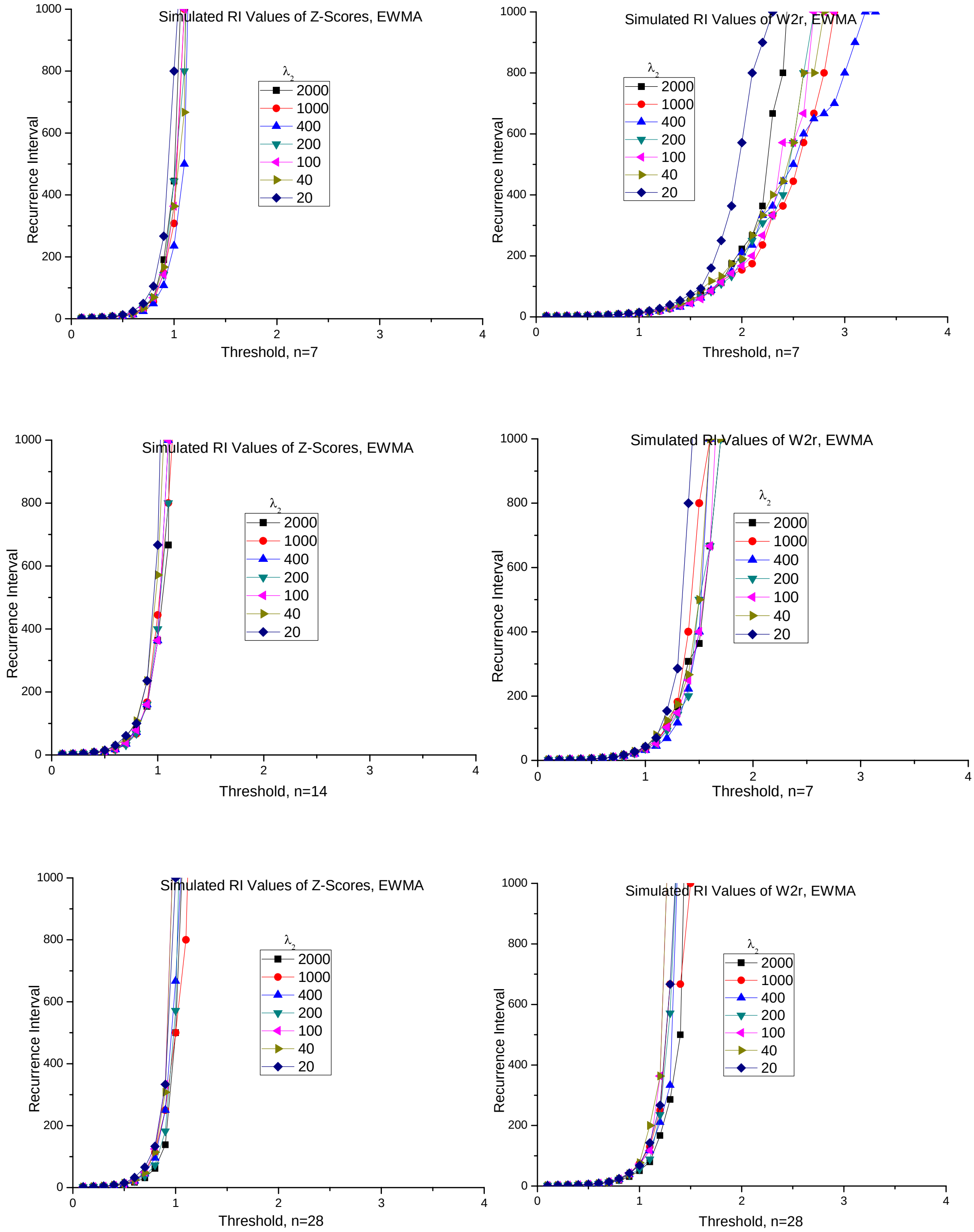

Figure 4-15: RI Thresholds for Adaptive Threshold Method Using MLE (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=200-EWMA

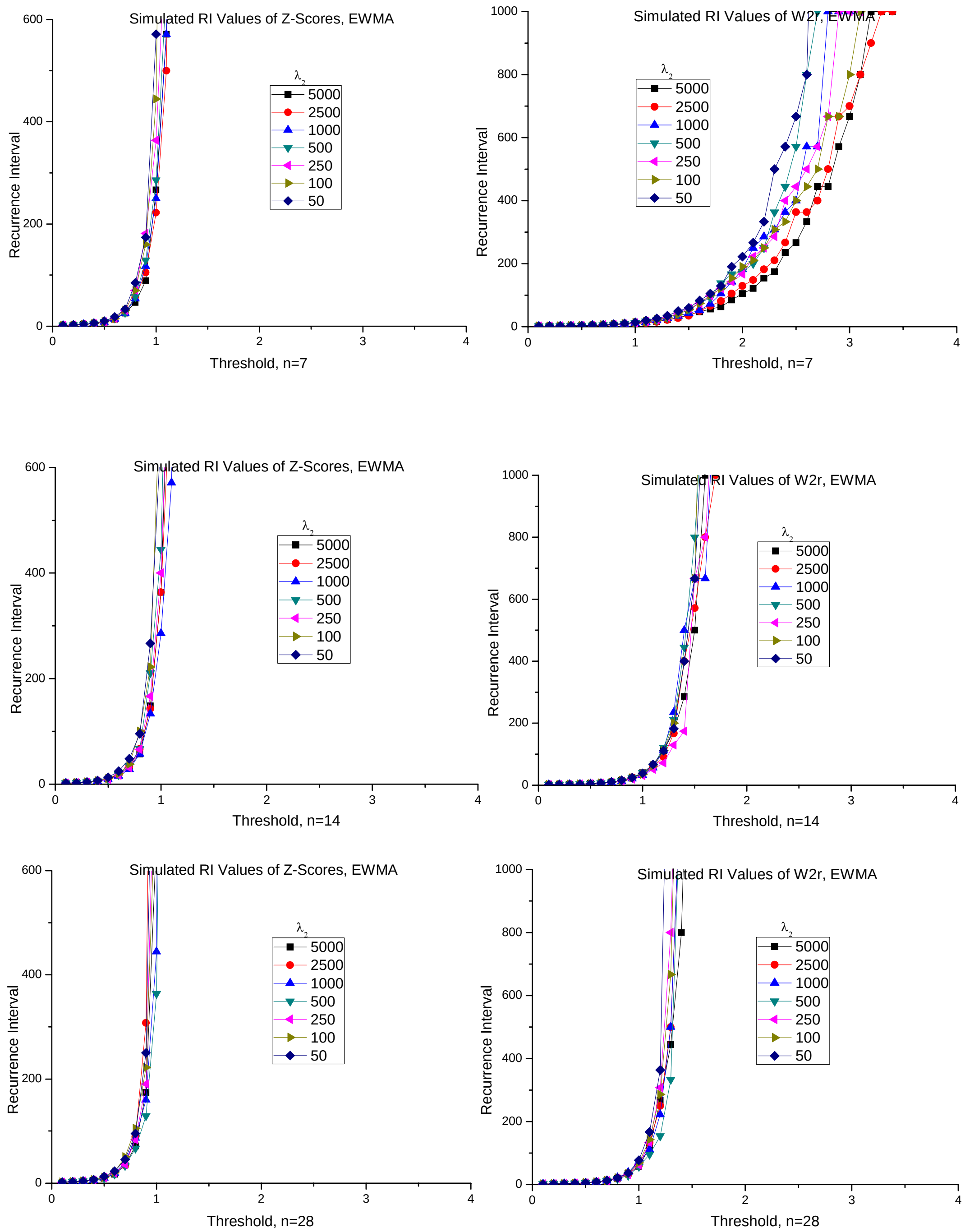

Figure 4-16: RI Thresholds for Adaptive Threshold Method Using MLE (left) and W2r (right) for Different Baselines-Conditional Binomial Counts, λ_1=500-EWMA

We further examined the performance of the adaptive threshold method assuming the Poisson parameters are known. This provides a limiting case as the value of n increases. Figures 4-17 to 4-21 show the RI threshold functions for this case across different baselines for the Shewhart (left) and EWMA (right) approaches given λ_1=10, 50, 100, 200, and 500, respectively. The RI threshold functions for the adaptive threshold method using MLE estimators, as shown in Figure 4-7 (left) to Figure 4-16 (left) for the Shewhart and EWMA approaches, were compared to those for the adaptive threshold method assuming known parameters. We observed that the Z-score threshold functions using MLE estimators are very close to the Z-scores threshold functions assuming known parameters. This result indicates that the MLE estimators for the conditional binomial distribution with Poisson inputs work very well for the adaptive threshold method. Generally the adaptive threshold method using known parameters or using MLE estimators works very well for large counts, but not so well for very small expected counts for both the Shewhart and EWMA approaches.

The RI threshold functions for the W2r method, as shown in Figure 4-7 (right) to Figure 4-16 (right) for the Shewhart and EWMA approaches, were also compared to those for the adaptive threshold method assuming known parameters. The Z-score threshold functions are considerably less variable than those of the W2r method as the parameters change.

Our RI analysis results show that, in either the EWMA or the Shewhart comparisons, the adaptive threshold method using known parameters or MLE estimators for the conditional binomial distribution outperforms the W2r method under the assumption of Poisson inputs.

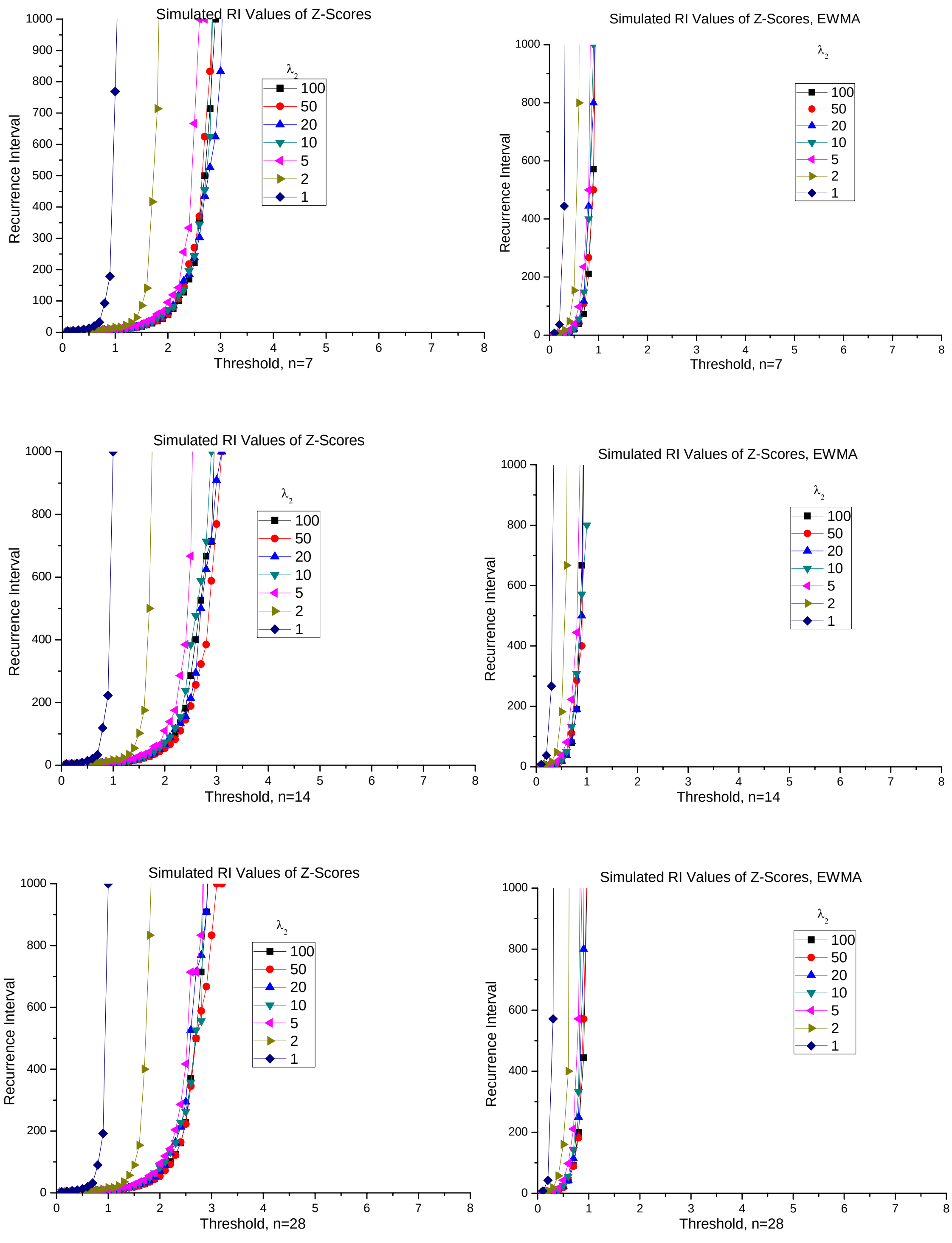

Figure 4-17: RI Thresholds for Adaptive Threshold Method Assuming Parameters Known for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=10

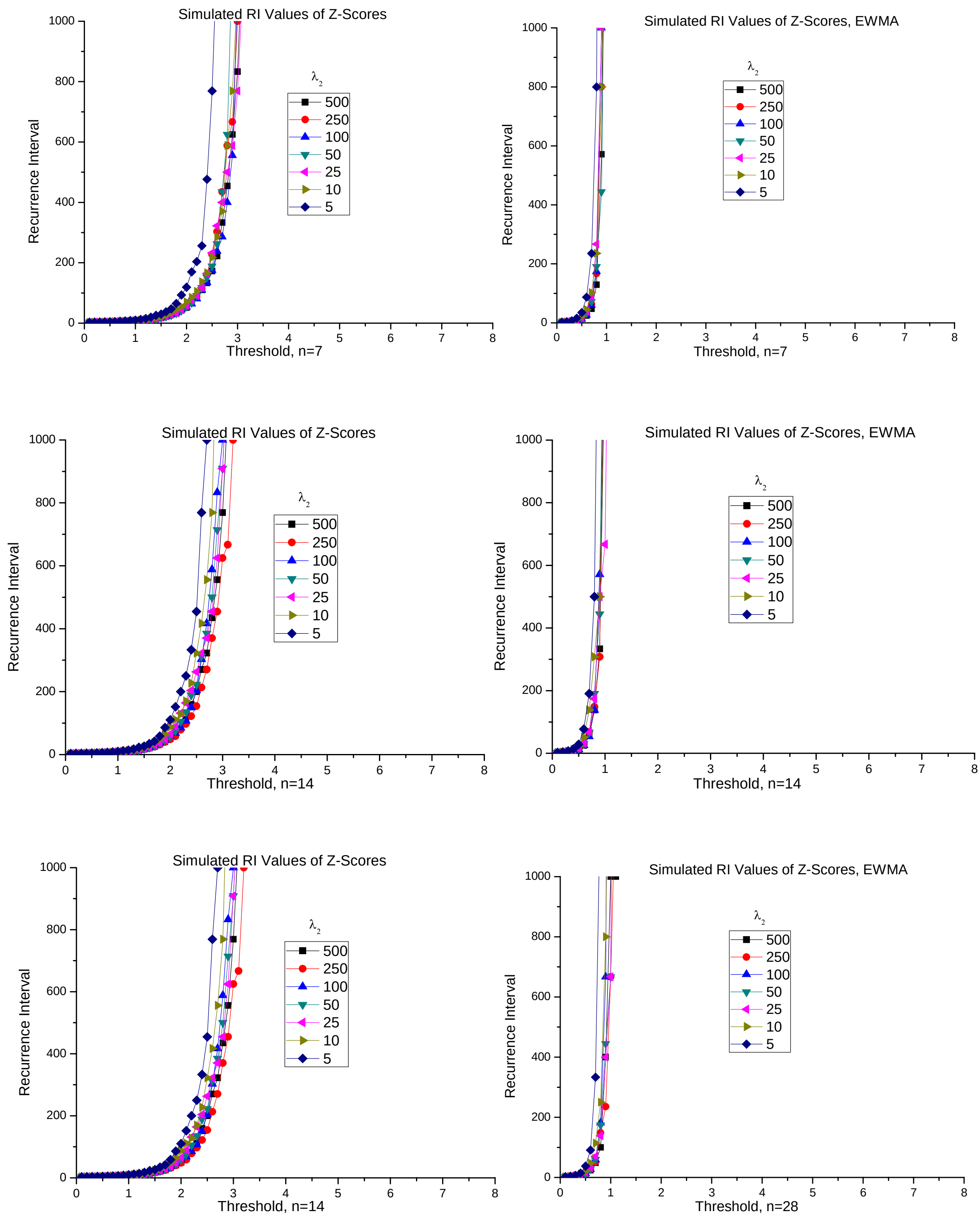

Figure 4-18: RI Thresholds for Adaptive Threshold Method Assuming Parameters Known for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=50

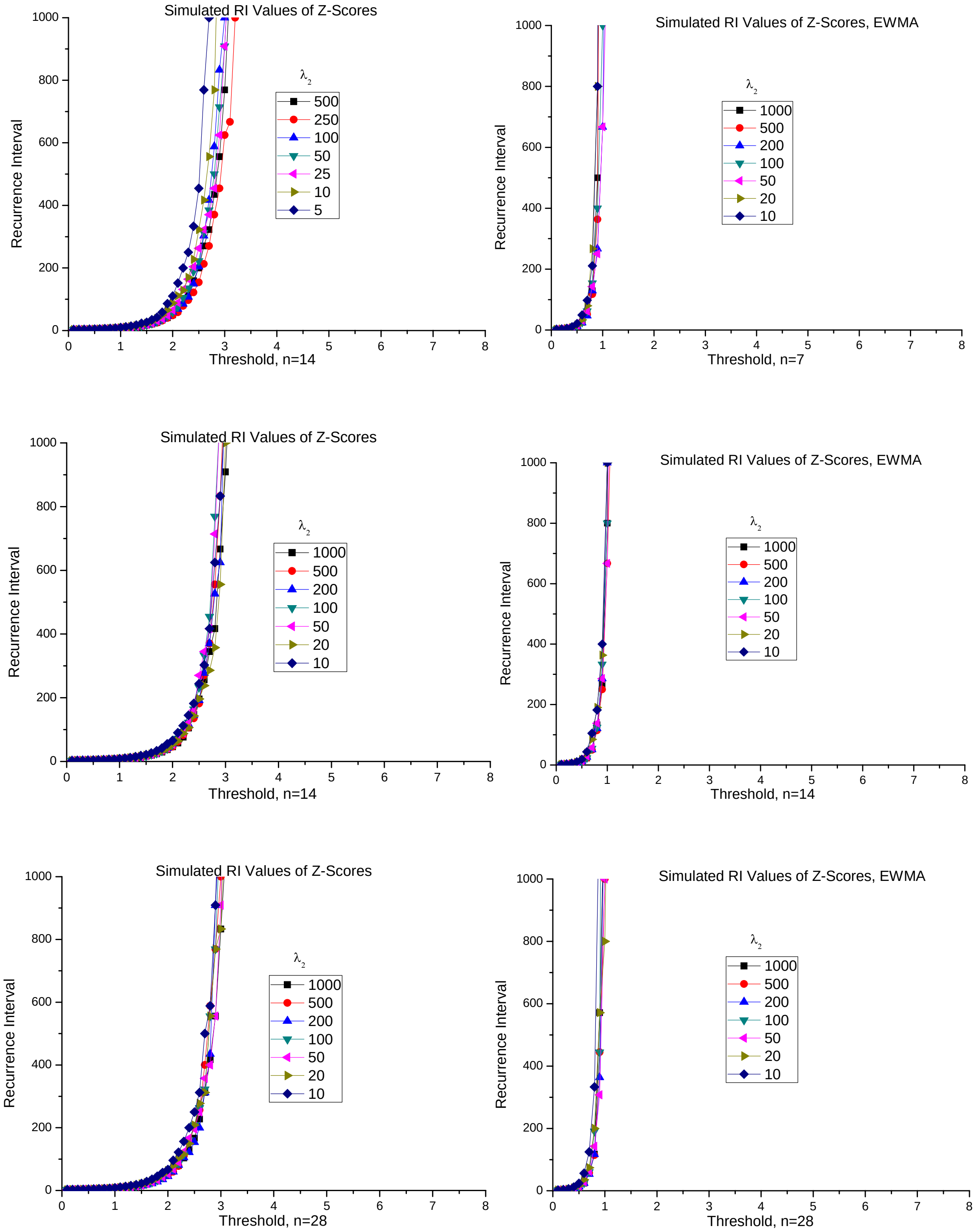

Figure 4-19: RI Thresholds for Adaptive Threshold Method Assuming Parameters Known for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=100

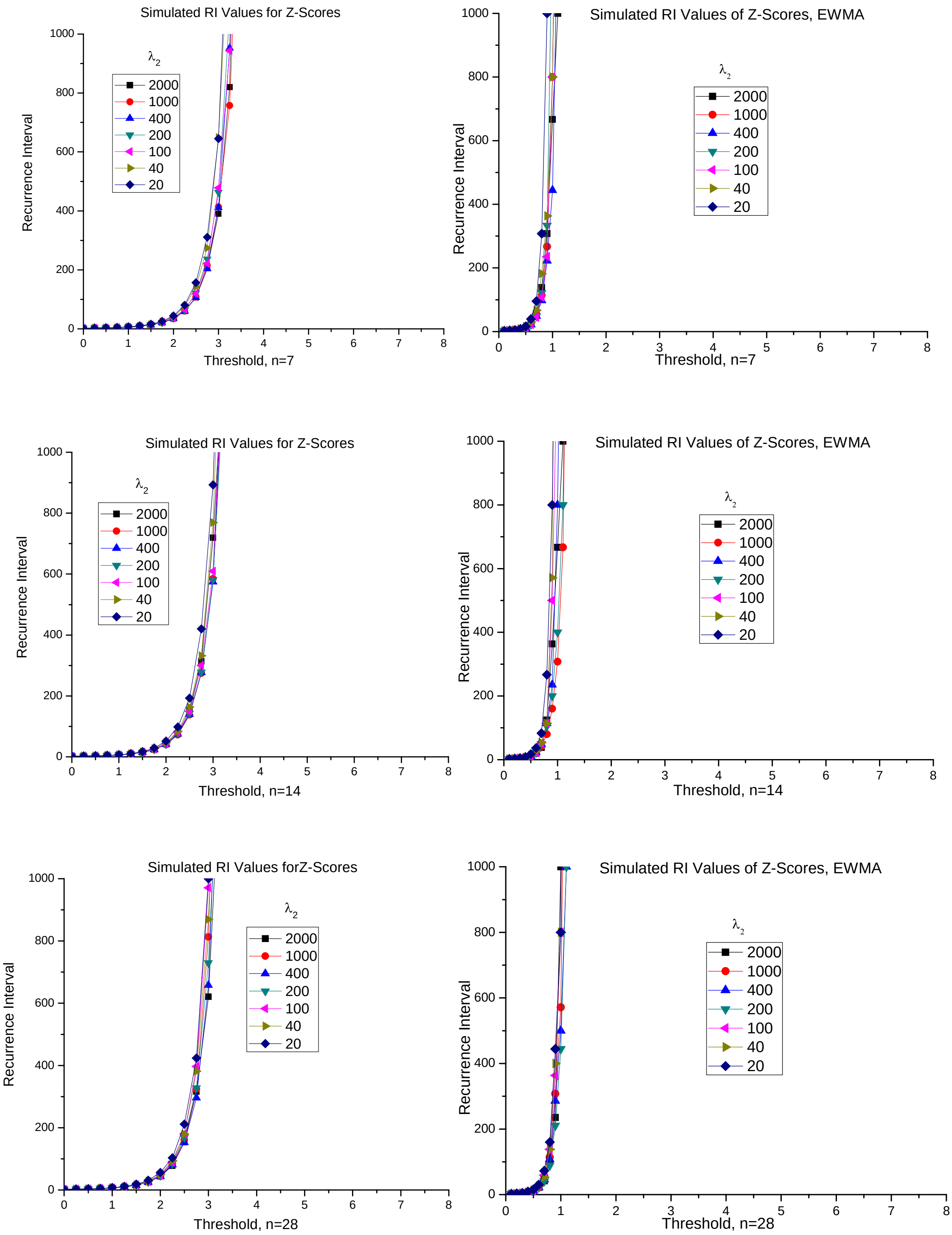

Figure 4-20: RI Thresholds for Adaptive Threshold Method Assuming Parameters Known for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=200

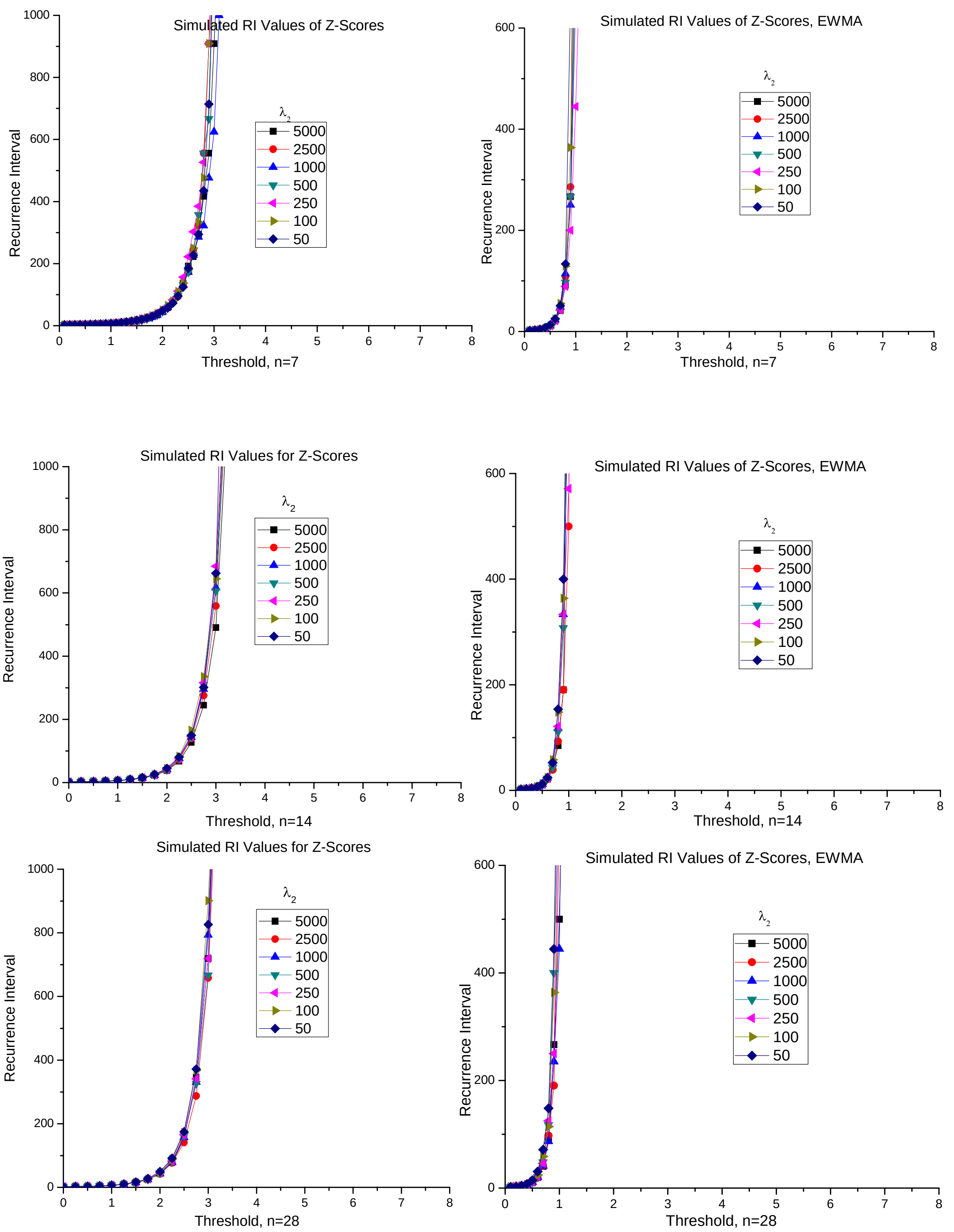

Figure 4-21: RI Thresholds for Adaptive Threshold Method Assuming Parameters Known for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=500

4.3.2 Comparison of W2r and Modified W2r Methods

A modified W2r method, i.e., W2r_1, was examined in terms of the RI threshold analysis for both Shewhart and EWMA approaches. Figures 4-22 to 4-26 show the RI threshold functions for this case across different baselines for the Shewhart (left) and EWMA (right) approaches, given λ_1=10, 50, 100, 200 and 500, respectively. There is some variability in the RI threshold curves for small counts with λ_1=10 and λ_1=50 for the W2r_1 method. Figures 4-25 and 4-26 show that the W2r_1 method works well for large counts with λ_1=200 and λ_1=500 as there is less variation in the corresponding RI threshold curves.

The RI threshold functions for the W2r method, as shown in Figure 4-7 (right) to Figure 4-16 (right) for the Shewhart and EWMA approaches, were compared to those for the W2r_1 method. Generally both methods work well for large expected counts while not very well for very small expected counts. We observed that the W2r_1 threshold functions become more robust to parameter changes when λ_1=10, compared to the W2r threshold functions. Overall the threshold functions of the W2r_1 method are somewhat less variable than those of the W2r method as the underlying parameters change.

Our RI analysis results show that, in either the EWMA or the Shewhart comparisons, the W2r_1 method outperforms the W2 rate method by giving more consistent values of the thresholds across the parameter space and baseline window lengths we used in our simulation study.

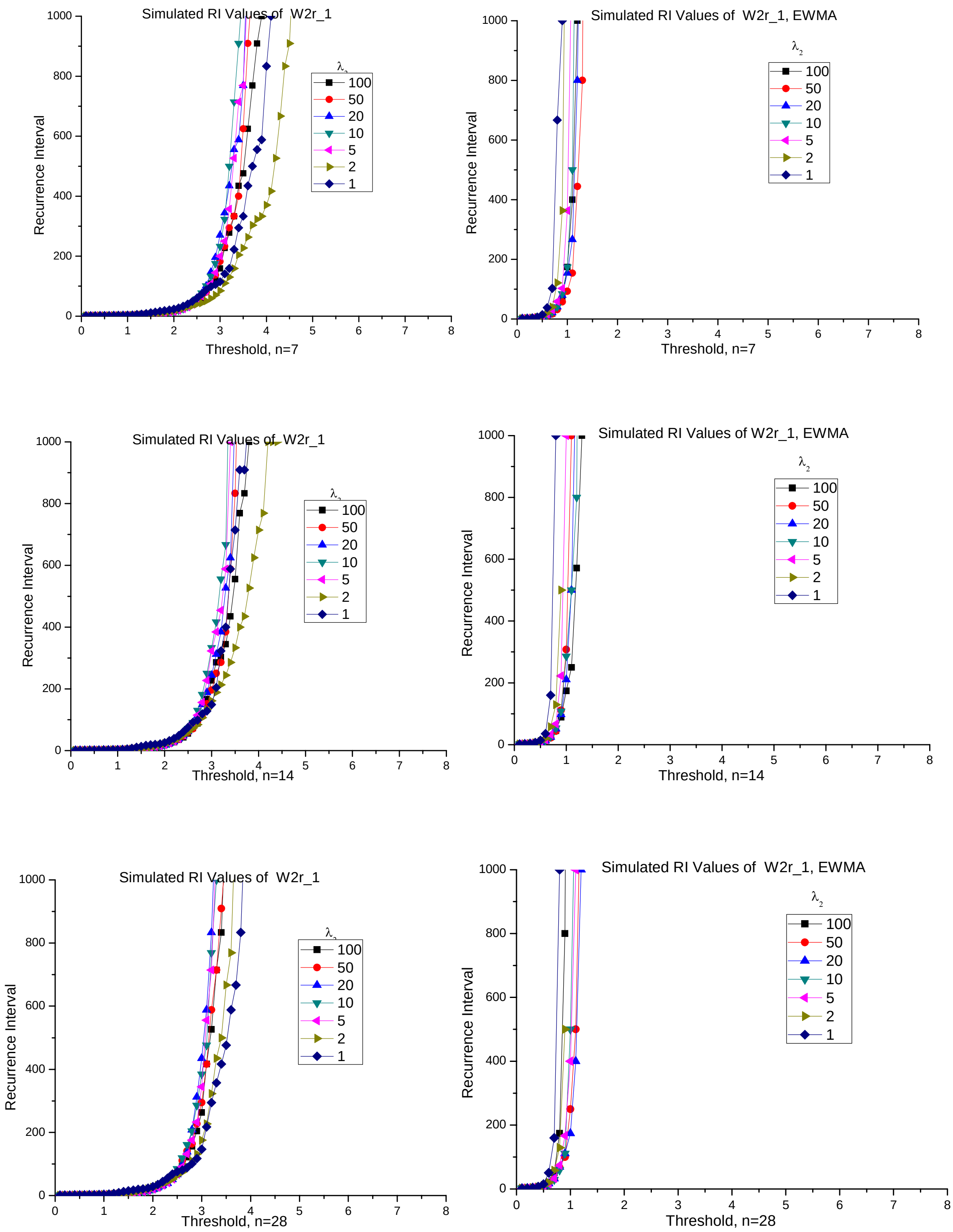

Figure 4-22: RI Thresholds for W2r_1 Method for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=10

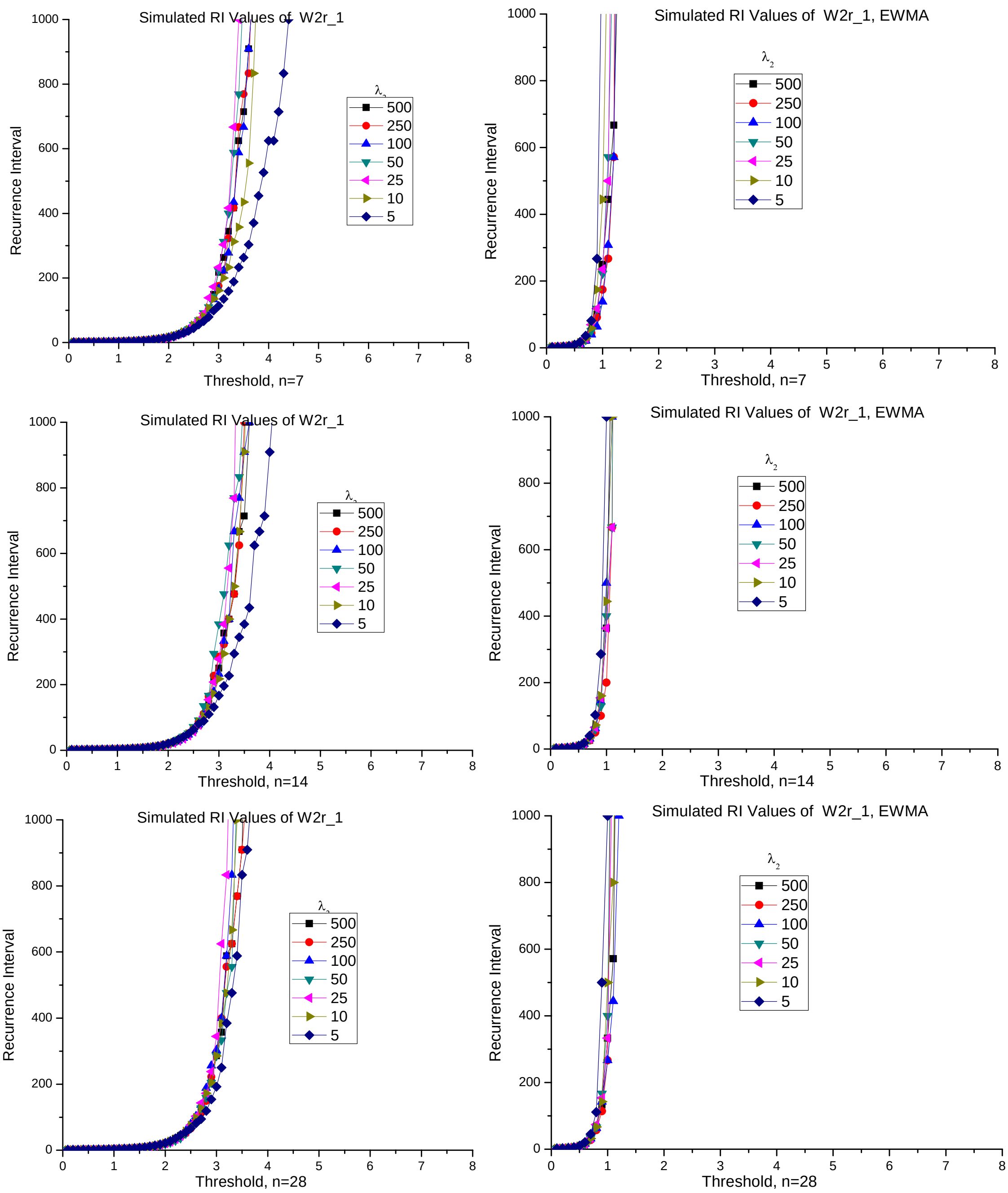

Figure 4-23: RI Thresholds for W2r_1 Method for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=50

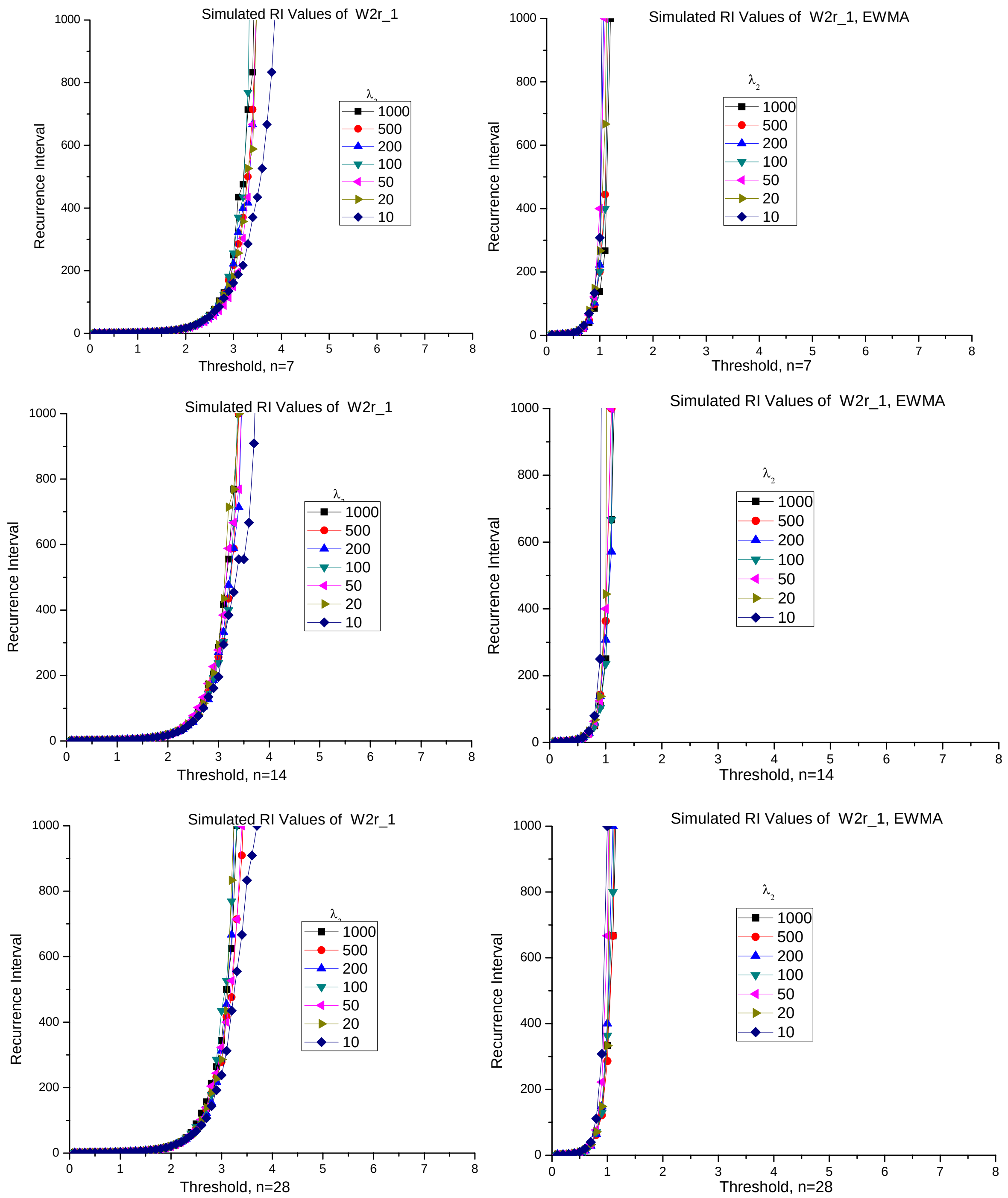

Figure 4-24: RI Thresholds for W2r_1 Method for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=100

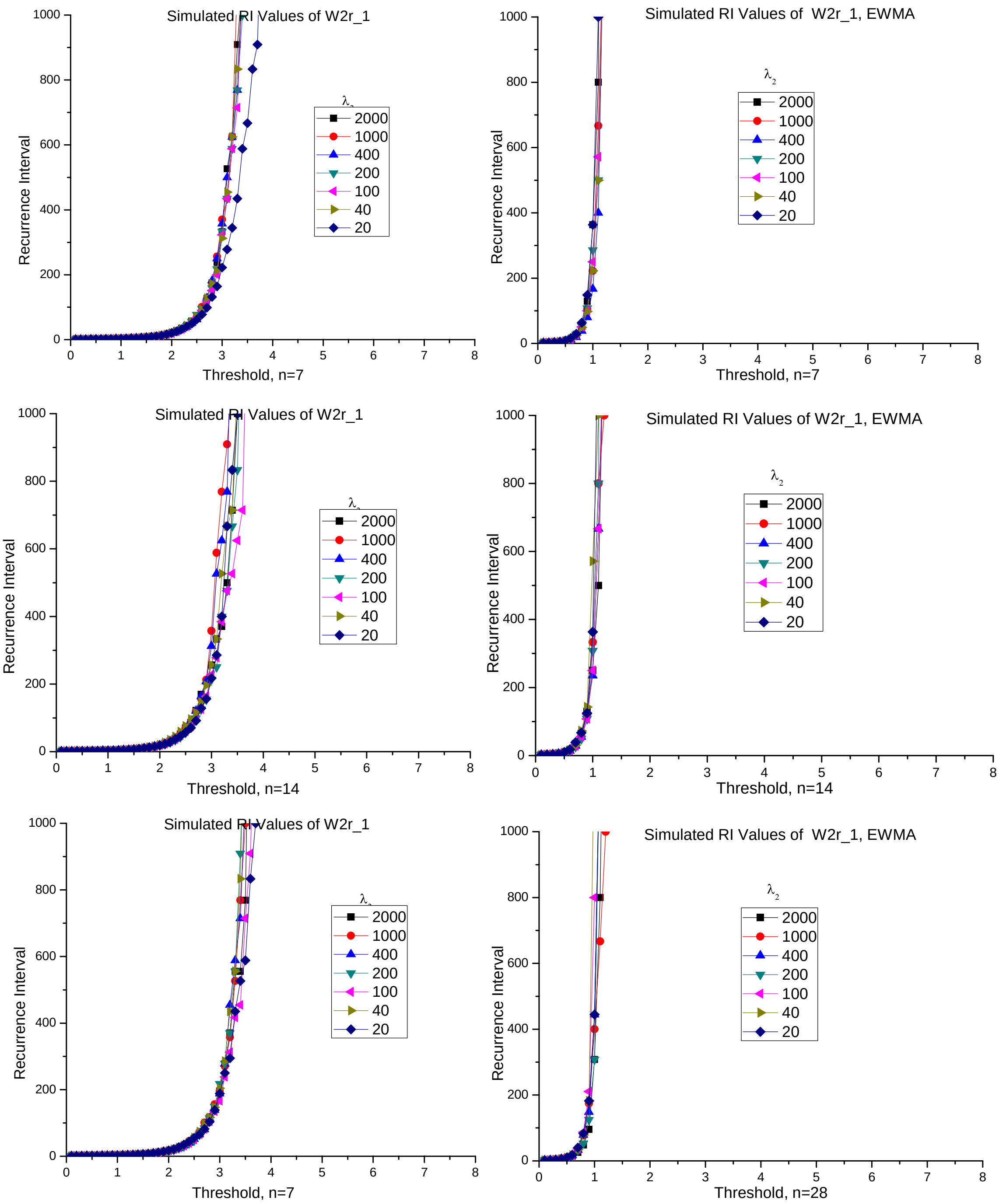

Figure 4-25: RI Thresholds for W2r_1 Method for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=200

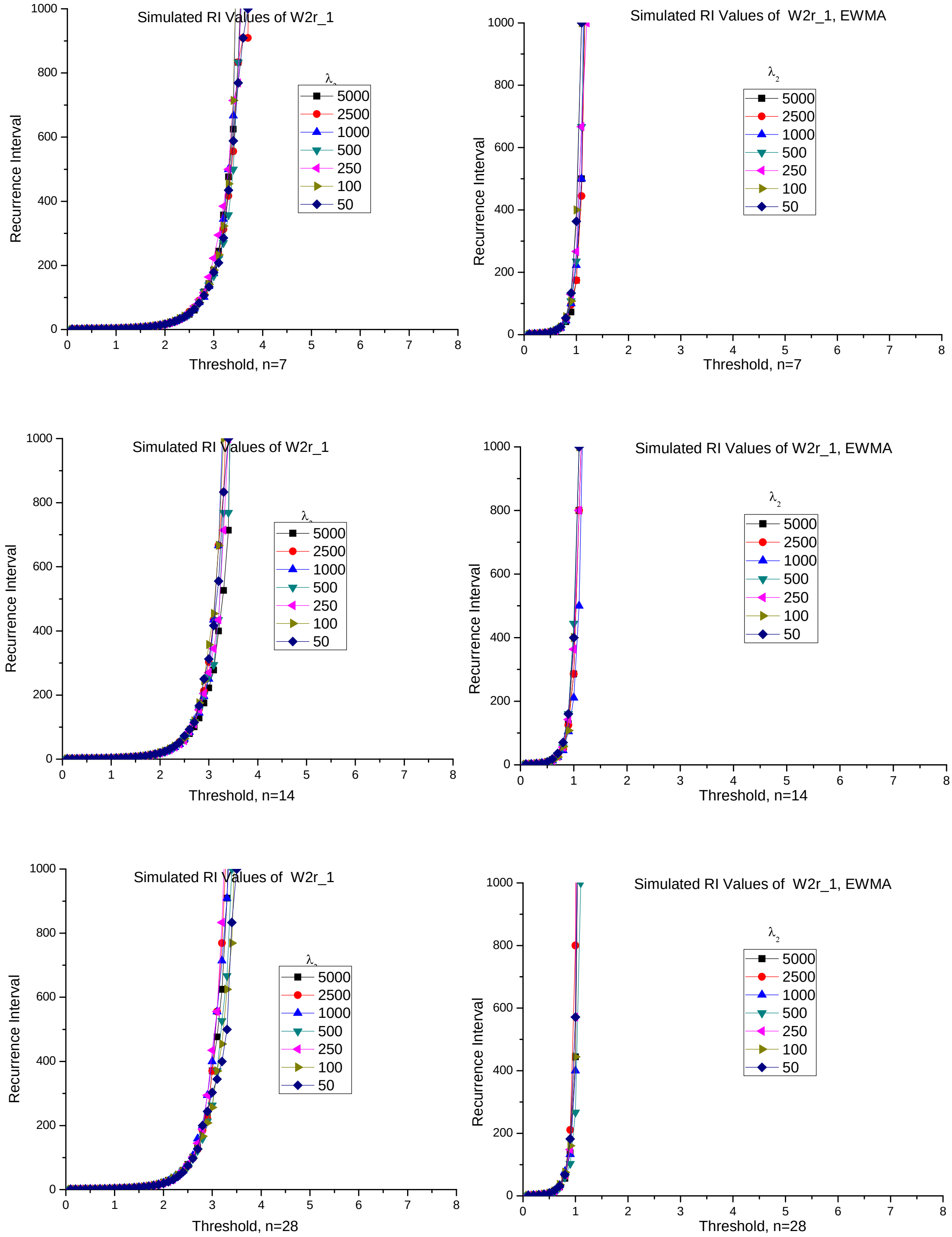

Figure 4-26: RI Thresholds for W2r_1 Method for Different Baselines-Conditional Binomial Counts-Shewhart (left) and EWMA (right), λ_1=500

4.4 Power Analysis

We examined the power to detect syndromic outbreaks for the adaptive threshold method, the W2r method, and the W2r_1 method. To make fair comparisons we have all the thresholds set to give a RI value of 500 days. Therefore, we expect that when no outbreak is present, we would expect one false alarm every 500 days. We also made comparisons with a RI value of 100 days.

The first methods compared were Shewhart-based methods, where if a daily statistic, i.e., W2r or Z-score, exceeds a particular threshold, a signal is given. The value of this threshold was derived from the results in Section 4.3, where for a given n and parameter set, thresholds were found so that the RI equals either 500 or 100.

The second methods we compared were the EWMA-based methods with a smoothing parameter of α=0.2. In all cases, we used the first ten weeks as the in-control data. The outbreak is then assumed to occur and be one week in length. We considered a percentage shift of size δ in the Poisson parameters λ_1 and λ_3 for this study which increases the average number of syndromic counts per weekday and weekend day over the course of the outbreak, respectively. The values used are δ= 0, .1, .2, .5, 1 and 2, where a zero shift indicates analysis when there is no outbreak. We considered $\lambda_1 = \lambda_2$ for the conditional binomial counts when there is no outbreak present. The increase in the Poisson parameter λ_1 or λ_3 is assumed to be (δ*100)% of its in-control parameter value.

4.4.1 Shewhart-based Methods

As discussed in Section 3.4, we take the Shewhart methods as special cases of the general EWMA methods in our study when α=1.0 is used. We used the *Power-Shewhart* algorithm described below to estimate the power values for the Shewhart-based methods in this section:

1. Simulate 1,000,000 days as the in-control data and calculate the statistic values of the Z-score, W2r, and modified W2r statistics using the estimated parameter values.
2. Compare the statistic values calculated in step 1 with a predetermined threshold (initially 0.2). Count the number of days signaling, i.e., any statistic value exceeding the threshold value.
3. Calculate the empirical RI as the reciprocal of the proportion of signals from the simulated 1,000,000 days.
4. Change the threshold from 0.2 to 9 with a step size of 0.2, and for each threshold value, repeat steps 2-3 to calculate its empirical RI.
5. Interpolate the RI-threshold curve to find the threshold values given the targeted RI values.
6. Simulate the in-control data of the first ten weeks for both syndrome counts and non-syndrome counts, and then inject outbreak data of one week in length. Calculate the statistic values of the one-sided EWMA of the W2r or Z-score values using the estimated parameter values. If the statistic value exceeds the threshold value from step 5, we get one signal.
7. Do the Monte Carlo simulation of step 6 100,000 times. Calculate the power as the proportion of the time that signals were obtained during the outbreak weeks.

Steps 1 to 5 of the *Power-Shewhart* algorithm are to set up the thresholds to determine whether a signal is given. Table 4-2 shows the threshold values we used in the power analysis. In Table 4-2, we denote by Z_MLE the adaptive threshold method using MLE estimators, and denote by Z_Par the adaptive threshold method assuming parameters are known. If the parameters are assumed to be known, then the length of the baseline, n, is no longer relevant.

Table 4-2: Threshold Values of Adaptive Threshold and W2r Methods-Conditional Binomial Case with Poisson Inputs-Shewhart

λ_1		RI = 100			RI = 500		
		n=7	14	28	n=7	14	28
10	W2r	4.16	3.75	3.50	5.38	4.76	4.35
	W2r_1	2.76	2.61	2.58	3.00	2.93	2.89
	Z_MLE	2.26	2.19	2.16	2.83	2.74	2.70
	Z_Par	2.14	2.14	2.14	2.67	2.67	2.67
50	W2r	4.27	3.70	3.49	5.79	4.67	4.30
	W2r_1	2.78	2.64	2.62	3.07	2.97	2.85
	Z_MLE	2.38	2.31	2.27	2.96	2.86	2.83
	Z_Par	2.20	2.20	2.20	2.77	2.77	2.77
100	W2r	4.28	3.72	3.47	5.83	4.76	4.29
	W2r_1	2.74	2.67	2.64	3.00	2.98	2.92
	Z_MLE	2.42	2.35	2.30	3.02	2.92	2.86
	Z_Par	2.26	2.26	2.26	2.83	2.83	2.83
200	W2r	4.28	3.66	3.44	5.80	4.71	4.24
	W2r_1	2.75	2.65	2.63	3.09	2.98	2.92
	Z_MLE	2.44	2.34	2.31	3.04	2.91	2.85
	Z_Par	2.27	2.27	2.27	2.82	2.82	2.82
500	W2r	4.24	3.70	3.45	5.85	4.70	4.23
	W2r_1	2.77	2.66	2.64	3.08	2.96	2.88
	Z_MLE	2.46	2.37	2.34	3.05	2.93	2.90
	Z_Par	2.29	2.29	2.29	2.84	2.84	2.84

We considered a percentage shift of size δ in the Poisson parameters λ_1 and λ_3 over the course of the outbreak. We first examined δ=0, where this zero shift indicates there is no outbreak. This no outbreak simulation was done to ensure that the proportion of signals should be approximately the same across parameter values and baseline lengths for the in-control case.

We expect the probabilities of an alarm when δ=0 and RI=500 to be around .014 for the Shewhart case as explained below:

1. For the Shewhart case, the probability that there is a signal in the in-control process is set to be the reciprocal of the given RI. So for RI=500 days, Prob(a signal per day | in-control) = 1/500 =.002.

2. We assume i.i.d. counts for the in-control data for the Shewhart case. For the outbreak of one week in length, the probability that there is at least one signal in the in-control process is calculated as Prob(at least one signal during the outbreak week| in-control) = 1-(1-Prob(a signal per day| in-control))^7 = 1- (1-.002)^7 =.014. Unless the parameters are assumed to be known, however, this value is only an approximation. The use of the moving window to obtain estimators introduces some dependence into the sequence of surveillance statistics.

The power values corresponding to δ=0 and RI=500 also provide an analytic tool for us to double check the thresholds we set up in Table 4-2 for the power analysis. Similarly we expect the probabilities of an alarm when δ=0 and RI=100 to be around 1- (1-.01)^7 =.068 for the Shewhart case.

4.4.1.1 Comparison of Adaptive Threshold and W2r Methods

The W2r method was compared to the adaptive threshold method using MLE estimators. Tables 4-3 and 4-4 show, for the Shewhart case, the power values for the adaptive threshold method and W2r methods when different baselines are used, i.e., 7-day, 14-day, and 28-day, given RI =500 and RI=100, respectively. We observed that the probabilities of an alarm when δ=0 are all around .014 in Table 4-3 and the probabilities of an alarm when δ=0 are all around .068 in Table 4-4. We also have the following conclusions regarding performance for both the W2r method and the adaptive threshold method:

(1) For a given δ, a given n and a given λ_1, the power values of the adaptive threshold method are at least as high as those of the W2r method.

(2) For a given shift δ and a given baseline, say n=7, it is clear that the higher the values of λ_1, the higher the power values are. This is because the absolute size of the shift in the parameter increases as λ_1 increases.

(3) For a given shift δ and a given λ_1, say λ_1=50, we observe that when n increases the power values increase.

(4) For the case of very small mean syndrome counts, say λ_1=10, given RI=500, the power values for both methods are very small when $\delta \leq$1.0. The power values are very small when $\delta \leq$0.5 given RI=100.

(5) For the case of a large shift, say $\delta \geq$2, except for the W2r method given RI=500, n=7 and λ_1=10, the power values for both methods are all very high. Given RI=100, for the case of a large shift, say $\delta \geq$2, the power values for both methods are all close to 1.0 across parameter values and baselines.

(6) For a given δ and n, say δ =0.5 and n=7, the power values of the adaptive threshold method are close to 1.0 when λ_1=100, while those of the W2r method are nearly 1.0 only when λ_1 is as high as 200.

Table 4-3: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart (RI=500)

δ	n	W2r					Adaptive Threshold				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0155	0.0139	0.0132	0.0151	0.0132	0.0162	0.0137	0.0139	0.0135	0.0130
	14	0.0143	0.0142	0.0151	0.0128	0.0125	0.0143	0.0148	0.0134	0.0139	0.0149
	28	0.0142	0.0154	0.0147	0.0152	0.0160	0.0143	0.0144	0.0132	0.0170	0.0143
0.1	7	0.0178	0.0212	0.0250	0.0361	0.0712	0.0258	0.0448	0.0619	0.1082	0.2846
	14	0.0179	0.0250	0.0358	0.0610	0.1340	0.0259	0.0519	0.0750	0.1495	0.3010
	28	0.0194	0.0303	0.0452	0.0881	0.2225	0.0263	0.0520	0.0852	0.1715	0.4161
0.2	7	0.0246	0.0407	0.0621	0.1212	0.3260	0.0413	0.1161	0.2114	0.4483	0.9110
	14	0.0291	0.0675	0.1020	0.2361	0.6039	0.0430	0.1400	0.2645	0.5732	0.9699
	28	0.0305	0.0910	0.1728	0.3870	0.8306	0.0451	0.1531	0.3176	0.6508	0.9854
0.5	7	0.0724	0.2192	0.4380	0.9420	0.9879	0.1324	0.6265	0.9590	1.0000	1.0000
	14	0.0954	0.4178	0.7199	0.9660	1.0000	0.1569	0.7416	0.9790	1.0000	1.0000
	28	0.1338	0.6001	0.9092	0.9990	1.0000	0.1709	0.8085	0.9938	1.0000	1.0000
1.0	7	0.2090	0.7743	0.9663	0.9998	1.0000	0.4379	0.9983	1.0000	1.0000	1.0000
	14	0.3624	0.9666	0.9999	1.0000	1.0000	0.5295	1.0000	1.0000	1.0000	1.0000
	28	0.5288	0.9981	1.0000	1.0000	1.0000	0.5948	1.0000	1.0000	1.0000	1.0000
2.0	7	0.7617	0.9999	1.0000	1.0000	1.0000	0.9300	1.0000	1.0000	1.0000	1.0000
	14	0.9086	1.0000	1.0000	1.0000	1.0000	0.9797	1.0000	1.0000	1.0000	1.0000
	28	0.9863	1.0000	1.0000	1.0000	1.0000	0.9936	1.0000	1.0000	1.0000	1.0000

Table 4-4: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart (RI=100)

δ	n	W2r					Adaptive Threshold				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0682	0.0631	0.0599	0.0663	0.0658	0.0681	0.0689	0.0649	0.0661	0.0656
	14	0.0651	0.0662	0.0636	0.0718	0.0633	0.0642	0.0695	0.0640	0.0757	0.0688
	28	0.0677	0.0645	0.0657	0.0681	0.0698	0.0641	0.0688	0.0674	0.0680	0.0678
0.1	7	0.0773	0.0879	0.1049	0.1381	0.2475	0.1045	0.1670	0.2247	0.3305	0.6105
	14	0.0796	0.1070	0.1358	0.2138	0.3746	0.1091	0.1814	0.2572	0.4074	0.7056
	28	0.0863	0.1163	0.1654	0.2611	0.5069	0.1130	0.1955	0.2762	0.4355	0.7593
0.2	7	0.0979	0.1501	0.2173	0.3498	0.6765	0.1508	0.3383	0.5102	0.7762	0.9899
	14	0.1089	0.2147	0.3244	0.5539	0.8809	0.1635	0.3827	0.5906	0.8655	0.9984
	28	0.1258	0.2680	0.4391	0.7066	0.9736	0.1688	0.4160	0.6483	0.9068	0.9999
0.5	7	0.2081	0.5080	0.7659	0.9544	1.0000	0.3522	0.8881	0.9948	1.0000	1.0000
	14	0.2709	0.7284	0.9367	0.9986	1.0000	0.4095	0.9449	0.9995	1.0000	1.0000
	28	0.3554	0.8724	0.9909	1.0000	1.0000	0.4413	0.9685	1.0000	1.0000	1.0000
1.0	7	0.5029	0.9547	0.9986	1.0000	1.0000	0.7403	1.0000	1.0000	1.0000	1.0000
	14	0.6601	0.9976	1.0000	1.0000	1.0000	0.8292	1.0000	1.0000	1.0000	1.0000
	28	0.8177	1.0000	1.0000	1.0000	1.0000	0.8775	1.0000	1.0000	1.0000	1.0000
2.0	7	0.9178	1.0000	1.0000	1.0000	1.0000	0.9915	1.0000	1.0000	1.0000	1.0000
	14	0.9844	1.0000	1.0000	1.0000	1.0000	0.9990	1.0000	1.0000	1.0000	1.0000
	28	0.9994	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figures 4-27 to 4-31 show the power values for n=7, 14, and 28 when λ_1 =10, 50, 100, 200, and 500, separately. Those figures support all the results we observed from Tables 4-3 and 4-4. Figure 4-27 shows that for the case λ_1 =10, the adaptive threshold method with n=7 works better than the W2r method with n=14. Figure 4-30 shows that for the case λ_1 =200, the adaptive threshold method with n=14 works better than the W2r method with n=28. We further observed that for the cases of λ_1 =50, 100, and 500, the adaptive threshold method with n=7 works even better than the W2r method with n=28, as shown in Figures 4-28, 4-29 and 4-31, respectively. It is very clear that there is a huge improvement in going from the W2r method to the adaptive threshold method.

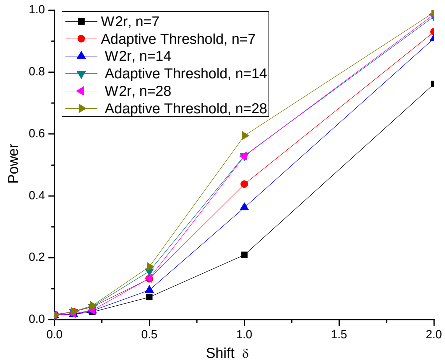

Figure 4-27: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs -Shewhart - λ_1=10, RI=500

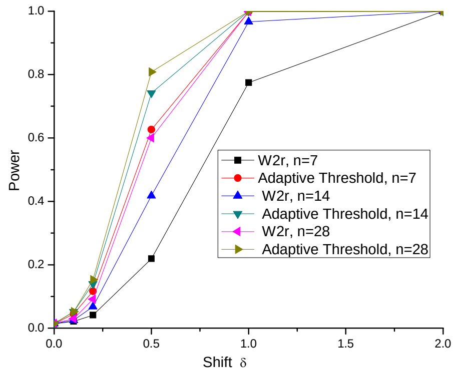

Figure 4-28: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=50, RI=500

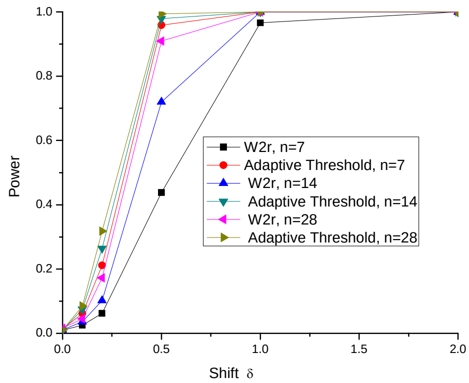

Figure 4-29: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=100, RI=500

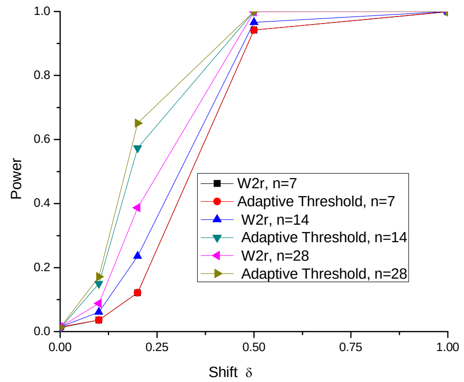

Figure 4-30: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=200, RI=500

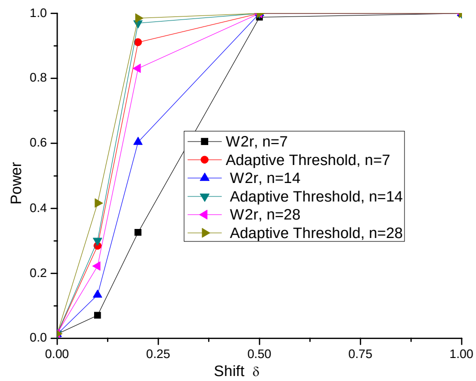

Figure 4-31: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=500, RI=500

We further examined the performance of the adaptive threshold method assuming the parameters are known and compared that to the MLE-based adaptive threshold method. Table 4-5 shows the power values for these adaptive threshold methods across different baselines given RI =500 for the Shewhart case. We observed that the probabilities of an alarm when δ=0 are all around .014. For a given shift δ>0, Table 4-5 leads to some conclusions regarding performance of the adaptive threshold methods as given below.

(1) For a given δ, λ_1 and baseline *n*, the power values of the adaptive threshold method using known parameters is at least as high as those of the MLE-based adaptive threshold method with

n=28.

(2) For a given δ and baseline n, it is clear that the higher the values of λ_1, the higher the power values are.

(3) For the case of very small mean syndrome counts, say λ_1=10, the power values for both methods are small when $\delta \leq$1.0. For the case of a large shift, say $\delta \geq$2, generally the power values for both methods are very high.

Our power analysis results show that for the Shewhart approach there is a huge improvement in going from the W2r method to the adaptive threshold method using MLE estimators or using known in-control parameters for the conditional binomial distribution.

Table 4-5: Power Analysis for Adaptive Threshold Method Using Known Parameters and Using MLE Estimators-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart (RI=500)

δ	n	Adaptive Threshold Using MLE Estimators					Adaptive Threshold Using Known Parameters				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0162	0.0137	0.0139	0.0135	0.0130	0.0126	0.0125	0.0154	0.0148	0.0136
	14	0.0143	0.0148	0.0134	0.0139	0.0149	0.0157	0.0144	0.0131	0.0132	0.0150
	28	0.0143	0.0144	0.0132	0.0170	0.0143	0.0132	0.0150	0.0138	0.0169	0.0149
0.1	7	0.0258	0.0448	0.0619	0.1082	0.2846	0.0262	0.0553	0.0902	0.1772	0.4815
	14	0.0259	0.0519	0.0750	0.1495	0.3010	0.0311	0.0604	0.0903	0.1787	0.4862
	28	0.0263	0.0520	0.0852	0.1715	0.4161	0.0326	0.0632	0.1088	0.2064	0.4901
0.2	7	0.0413	0.1161	0.2114	0.4483	0.9110	0.0474	0.1658	0.3484	0.7212	0.9962
	14	0.0430	0.1400	0.2645	0.5732	0.9699	0.0539	0.1791	0.3559	0.7246	0.9963
	28	0.0451	0.1531	0.3176	0.6508	0.9854	0.0574	0.1921	0.3597	0.7541	0.9968
0.5	7	0.1324	0.6265	0.9590	1.0000	1.0000	0.1903	0.8676	0.9980	1.0000	1.0000
	14	0.1569	0.7416	0.9790	1.0000	1.0000	0.1925	0.8862	0.9987	1.0000	1.0000
	28	0.1709	0.8085	0.9938	1.0000	1.0000	0.2074	0.8986	0.9984	1.0000	1.0000
1.0	7	0.4379	0.9983	1.0000	1.0000	1.0000	0.6802	1.0000	1.0000	1.0000	1.0000
	14	0.5295	1.0000	1.0000	1.0000	1.0000	0.7073	1.0000	1.0000	1.0000	1.0000
	28	0.5948	1.0000	1.0000	1.0000	1.0000	0.7081	1.0000	1.0000	1.0000	1.0000
2.0	7	0.9300	1.0000	1.0000	1.0000	1.0000	0.9993	1.0000	1.0000	1.0000	1.0000
	14	0.9797	1.0000	1.0000	1.0000	1.0000	0.9993	1.0000	1.0000	1.0000	1.0000
	28	0.9936	1.0000	1.0000	1.0000	1.0000	0.9996	1.0000	1.0000	1.0000	1.0000

4.4.1.2 Comparison of W2r and Modified W2r Methods

A modified W2r method, i.e., W2r_1, was compared to the W2r method in terms of the power analysis for the Shewhart approaches. Table 4-6 shows the power values for the W2r_1 and W2r methods when different baselines are used given RI =500 for the Shewhart case. We observed that the probabilities of an alarm when δ=0 are all around .014 in Table 4-6. For a given shift δ >0, Table 4-6 leads to the following conclusions regarding performance for both the W2r and W2r_1 methods:

(1) For a given δ, λ_1 and baseline *n*, the power values of the W2r_1 method are at least as high as those of the W2r method.

(2) For a given δ and baseline *n*, it is clear that the higher values of λ_1 lead to higher power values.

(3) For a given δ and λ_1, we observe that when *n* increases, the power values increase.

(4) For the case of very small mean syndrome counts, say λ_1=10, the power values for both methods are very small when the shift $\delta \leq$1.0. For the case of a large shift, say $\delta \geq$2, the power values for both methods are all very high.

Table 4-6: Power Analysis for W2r and W2r_1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart (RI=500)

δ	n	W2r					W2r_1				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0155	0.0139	0.0132	0.0151	0.0132	0.0105	0.0143	0.0143	0.0103	0.0133
	14	0.0143	0.0142	0.0151	0.0128	0.0125	0.0170	0.0188	0.0143	0.0107	0.0190
	28	0.0142	0.0154	0.0147	0.0152	0.0160	0.0182	0.0245	0.0198	0.0153	0.0202
0.1	7	0.0178	0.0212	0.0250	0.0361	0.0712	0.0288	0.0425	0.0707	0.0885	0.2773
	14	0.0179	0.0250	0.0358	0.0610	0.1340	0.0293	0.0537	0.0785	0.1287	0.4203
	28	0.0194	0.0303	0.0452	0.0881	0.2225	0.0337	0.0978	0.1070	0.1438	0.5170
0.2	7	0.0246	0.0407	0.0621	0.1212	0.3260	0.0407	0.1175	0.2372	0.3887	0.9053
	14	0.0291	0.0675	0.1020	0.2361	0.6039	0.0548	0.1668	0.2737	0.5013	0.9792
	28	0.0305	0.0910	0.1728	0.3870	0.8306	0.0552	0.2297	0.3673	0.6078	0.9942
0.5	7	0.0724	0.2192	0.4380	0.9420	0.9879	0.0963	0.6352	0.9455	0.9997	1.0000
	14	0.0954	0.4178	0.7199	0.9660	1.0000	0.1580	0.7612	0.9810	1.0000	1.0000
	28	0.1338	0.6001	0.9092	0.9990	1.0000	0.1902	0.8868	0.9962	1.0000	1.0000
1.0	7	0.2090	0.7743	0.9663	0.9998	1.0000	0.3658	0.9977	1.0000	1.0000	1.0000
	14	0.3624	0.9666	0.9999	1.0000	1.0000	0.5275	0.9998	1.0000	1.0000	1.0000
	28	0.5288	0.9981	1.0000	1.0000	1.0000	0.6030	0.9998	1.0000	1.0000	1.0000
2.0	7	0.7617	0.9999	1.0000	1.0000	1.0000	0.9067	1.0000	1.0000	1.0000	1.0000
	14	0.9086	1.0000	1.0000	1.0000	1.0000	0.9728	1.0000	1.0000	1.0000	1.0000
	28	0.9863	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figures 4-32 to 4-36 show the power values for n=7, 14, and 28 when λ_1 =10, 50, 100, 200, and 500, respectively. These figures support the conclusions we reached from Tables 4-5 and 4-6. Figure 4-33 shows that for the case λ_1 =50, the W2r_1 method with n=7 works better than the W2r method with n=14, and the W2r_1 method with n=14 works better than the W2r method with n=28. We observed that for the cases of λ_1 =100, 200 and 500, the W2r_1 method with n=7 works even better than the W2r method with n=28, as shown in Figures 4-34 to 4-36. It is very clear that there is a huge improvement in going from W2r method to the W2r_1 method for the Shewhart approach.

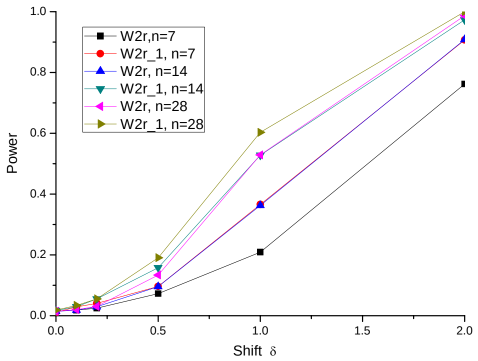

Figure 4-32: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=10, RI=500

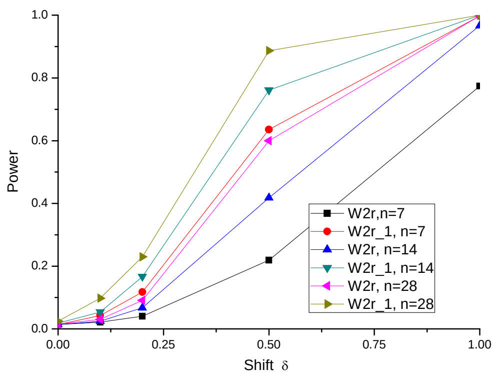

Figure 4-33: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=50, RI=500

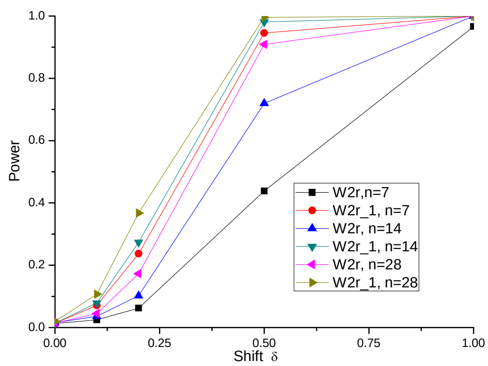

Figure 4-34: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=100, RI=500

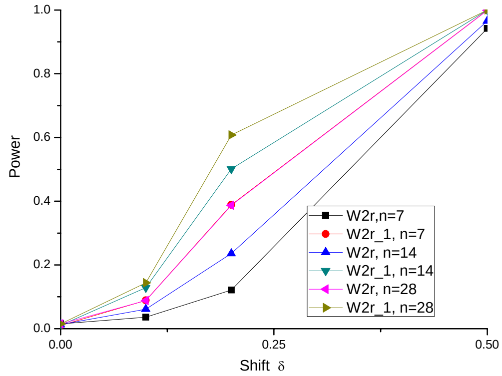

Figure 4-35: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=200, RI=500

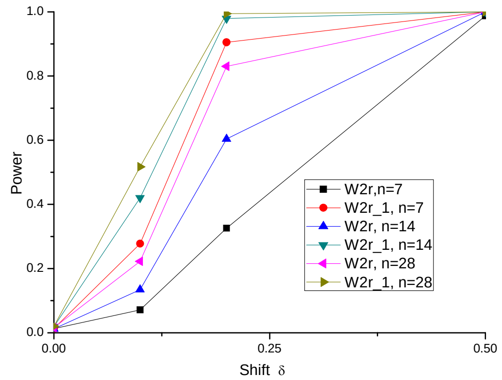

Figure 4-36: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart - λ_1=500, RI=500

4.4.2 One-Sided EWMA-based Methods

The algorithm we used to estimate the power values for the EWMA-based methods was similar to the *Power-Shewhart* algorithm used for the Shewhart-based methods. The only difference was that in step 1 we calculated the EWMA values for the W2r, W2r_1 or Z-score statistics with a smoothing constant of α=0.2. Table 4-7 shows the threshold values we used in the following power analysis. Note that we denote by Z_MLE the adaptive threshold method based on MLE estimators, and denote by Z_Par the adaptive threshold method assuming in-control parameters are known.

Table 4-7: Threshold Values of Adaptive Threshold and W2r Methods-Conditional Binomial Case with Poisson Inputs-EWMA

λ_1		RI = 100			RI = 500		
		n=7	14	28	n=7	14	28
10	W2r	1.97	1.86	1.78	2.26	2.12	2.01
	W2r_1	1.44	1.38	1.35	1.12	1.08	1.04
	Z_MLE	0.77	0.73	0.70	0.96	0.92	0.87
	Z_Par	0.65	0.65	0.65	0.82	0.82	0.82
50	W2r	2.09	1.86	1.77	2.52	2.17	2.00
	W2r_1	1.43	1.38	1.35	1.11	1.10	1.04
	Z_MLE	0.85	0.82	0.79	1.04	1.00	0.96
	Z_Par	0.74	0.74	0.74	0.90	0.90	0.90
100	W2r	2.09	1.86	1.77	2.49	2.18	2.07
	W2r_1	1.42	1.38	1.34	1.11	1.09	1.05
	Z_MLE	0.87	0.84	0.81	1.06	1.03	0.99
	Z_Par	0.76	0.76	0.76	0.93	0.93	0.93
200	W2r	2.11	1.86	1.77	2.64	2.13	2.01
	W2r_1	1.44	1.39	1.35	1.14	1.08	1.06
	Z_MLE	0.88	0.85	0.82	1.08	1.04	1.01
	Z_Par	0.78	0.78	0.78	0.95	0.95	0.95
500	W2r	2.12	1.85	1.77	2.62	2.14	2.06
	W2r_1	1.44	1.37	1.35	1.15	1.11	1.05
	Z_MLE	0.90	0.86	0.84	1.09	1.06	1.01
	Z_Par	0.80	0.80	0.80	0.98	0.98	0.98

4.4.2.1 Comparison of Adaptive Threshold and W2r Methods

Tables 4-8 and 4-9 give, for the one-sided EWMA case, the power values for the adaptive threshold and W2r methods when different baselines are used, i.e., n=7, 14, and 28, given RI =500 and RI=100, respectively. We first examined δ=0 to ensure that the proportion of signals should be approximately the same across parameter values and baseline lengths for the in-control case. We observed from Table 4-8 that when δ=0 and RI=500, the proportion of signals are approximately around 0.01 for the two underlying methods. Unlike the Shewhart case shown in Table 4-3, we do not use 1- $(1-.002)^7$=0.014 when δ=0 and RI=500 since the daily statistic values for the one-sided EWMA case are not independent. From Table 4-8 and Table 4-9, when δ=0 the proportion of signals is higher for the Shewhart methods. The signaling thresholds chosen for both the Shewhart and EWMA methods were such that if we are in-control, we expect one false alarm in 500 days. However, the EWMA statistics are autocorrelated, so we are more likely to have successive signals than with the Shewhart method. Therefore, the δ=0 case has the same number of days signal for the Shewhart and EWMA methods, but since the EWMA method likely signals an outbreak in clusters, it signals an outbreak in fewer weeks than the Shewhart method. This point was also discussed in Szarka, Gan, and Woodall [13].

Table 4-8 and Table 4-9 lead to the following conclusions regarding the performance for the W2r and adaptive threshold methods for the EWMA approach:

(1) For a given δ, λ_1 and baseline n, the power values of the adaptive threshold method are at least as high as those of the W2r method.

(2) For a given δ and baseline n, the higher values of λ_1 lead to higher power values.

(3) For a given δ and λ_1, when n increases the power values increase.

(4) For the case of very small mean syndrome counts, say λ_1=10, the power values for the W2r method are very small when the shift $\delta \leq 1.0$, whereas those for the adaptive threshold method are small when $\delta \leq 0.5$

(5) In the case of a large shift, say $\delta \geq 2$, the power values for both methods are all very high across parameter values and baseline lengths.

(6) For a given δ and n, say δ =0.5 and n=7, the power values of the adaptive threshold method are close to 1.0 when λ_1=100, while those of the W2r method are nearly 1.0 only if $\lambda_1 \geq 500$.

Table 4-8: Power Analysis for Adaptive Threshold and W2r Methods- Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA (RI=500)

δ	n	W2r					Adaptive Threshold				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0090	0.0091	0.0105	0.0074	0.0052	0.0115	0.0109	0.0111	0.0110	0.0114
	14	0.0115	0.0083	0.0076	0.0102	0.0091	0.0104	0.0129	0.0113	0.0118	0.0099
	28	0.0099	0.0108	0.0067	0.0092	0.0065	0.0110	0.0106	0.0105	0.0119	0.0128
0.1	7	0.0140	0.0135	0.0216	0.0343	0.0728	0.0255	0.0602	0.1128	0.2209	0.5828
	14	0.0144	0.0198	0.0292	0.0716	0.2151	0.0272	0.0730	0.1320	0.2928	0.7035
	28	0.0149	0.0280	0.0337	0.0987	0.3039	0.0306	0.0858	0.1596	0.3462	0.7973
0.2	7	0.0227	0.0363	0.0775	0.1316	0.4701	0.0523	0.2112	0.4518	0.7986	0.9980
	14	0.0264	0.0648	0.1464	0.4126	0.8805	0.0575	0.2751	0.5546	0.8947	1.0000
	28	0.0279	0.1134	0.2227	0.6168	0.9770	0.0681	0.3461	0.6533	0.9470	1.0000
0.5	7	0.0941	0.3178	0.6738	0.9093	1.0000	0.2332	0.9139	0.9991	1.0000	1.0000
	14	0.1211	0.6239	0.9306	0.9998	1.0000	0.3042	0.9684	1.0000	1.0000	1.0000
	28	0.1700	0.8648	0.9939	1.0000	1.0000	0.3808	0.9902	1.0000	1.0000	1.0000
1	7	0.3942	0.9335	0.9985	1.0000	1.0000	0.7258	1.0000	1.0000	1.0000	1.0000
	14	0.5542	0.9983	1.0000	1.0000	1.0000	0.8537	1.0000	1.0000	1.0000	1.0000
	28	0.7490	1.0000	1.0000	1.0000	1.0000	0.9291	1.0000	1.0000	1.0000	1.0000
2	7	0.9193	1.0000	1.0000	1.0000	1.0000	0.9963	1.0000	1.0000	1.0000	1.0000
	14	0.9874	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4-9: Power Analysis for Adaptive Threshold and W2r Methods- Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA (RI=100)

δ	n	W2r					Adaptive Threshold				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0425	0.0388	0.0374	0.0386	0.0364	0.0478	0.0496	0.0502	0.0488	0.0468
	14	0.0423	0.0460	0.0420	0.0430	0.0383	0.0488	0.0473	0.0468	0.0505	0.0448
	28	0.0391	0.0427	0.0440	0.0386	0.0456	0.0441	0.0492	0.0449	0.0505	0.0494
0.1	7	0.0516	0.0578	0.0770	0.1086	0.2493	0.0954	0.1784	0.2766	0.4477	0.7977
	14	0.0522	0.0805	0.1088	0.1900	0.4447	0.0989	0.2000	0.3187	0.5340	0.8799
	28	0.0556	0.0895	0.1370	0.2458	0.5905	0.1009	0.2228	0.3538	0.5937	0.9229
0.2	7	0.0753	0.1238	0.2146	0.3964	0.8055	0.1607	0.4327	0.6904	0.9339	1.0000
	14	0.0822	0.2040	0.3649	0.6627	0.9739	0.1749	0.5051	0.7797	0.9714	1.0000
	28	0.0879	0.2607	0.4977	0.8217	0.9979	0.1853	0.5738	0.8436	0.9876	1.0000
0.5	7	0.2097	0.6028	0.8852	0.9916	1.0000	0.4617	0.9786	1.0000	1.0000	1.0000
	14	0.2674	0.8593	0.9893	1.0000	1.0000	0.5468	0.9937	1.0000	1.0000	1.0000
	28	0.3466	0.9565	0.9999	1.0000	1.0000	0.6052	0.9986	1.0000	1.0000	1.0000
1	7	0.6030	0.9901	1.0000	1.0000	1.0000	0.8919	1.0000	1.0000	1.0000	1.0000
	14	0.7621	1.0000	1.0000	1.0000	1.0000	0.9525	1.0000	1.0000	1.0000	1.0000
	28	0.8887	1.0000	1.0000	1.0000	1.0000	0.9788	1.0000	1.0000	1.0000	1.0000
2	7	0.9734	1.0000	1.0000	1.0000	1.0000	0.9997	1.0000	1.0000	1.0000	1.0000
	14	0.9979	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figures 4-37 to 4-41 show the power values for the EWMA case for n=7, 14, and 28 when λ_1 =10, 50, 100, 200, and 500, respectively, and RI=500. Those figures support the conclusions we reached from Table 4-8 and Table 4-9. Figure 4-37 shows that when δ =1.0 and 2.0, the adaptive threshold method with n=7 works better than the W2r method with n=14, whereas when δ =0.5, the adaptive threshold method with n=7 works even better than the W2r method with n=28. As shown in Figures 4-38 to 4-41, the adaptive threshold method with n=7 works at least as well as the W2r method with n=28 for any positive values of δ in our study.

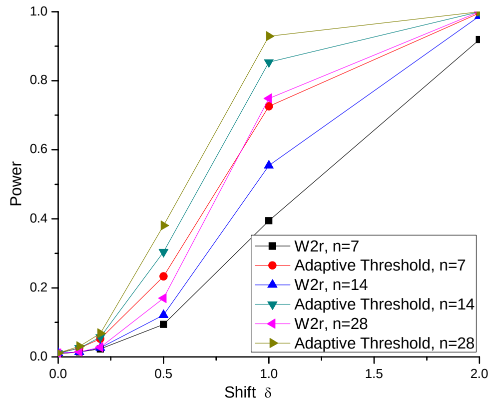

Figure 4-37: Power Analysis for Adaptive Threshold and W2r Methods -Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA- λ_1=10, RI=500

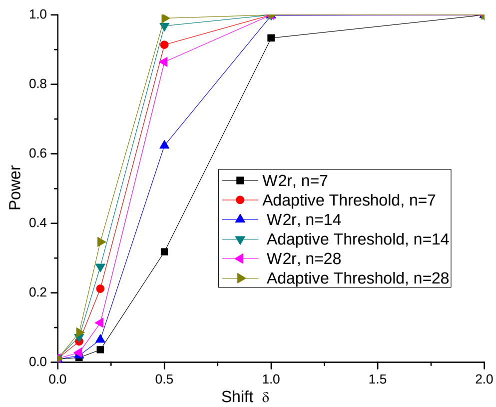

Figure 4-38: Power Analysis for Adaptive Threshold and W2r Methods -Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA- λ_1=50, RI=500

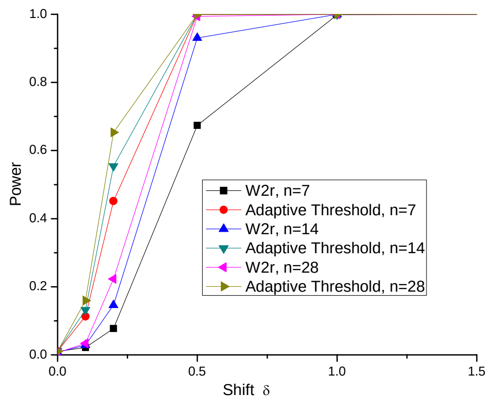

Figure 4-39: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA- λ_1=100, RI=500

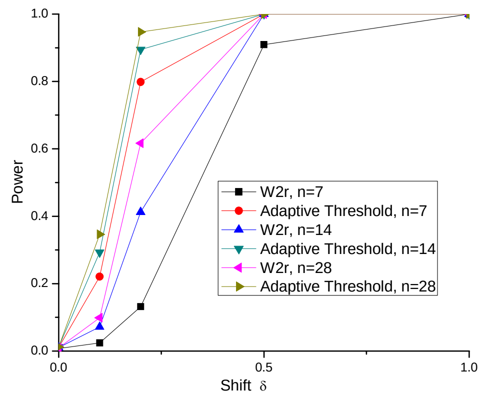

Figure 4-40: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA- λ_1=200, RI=500

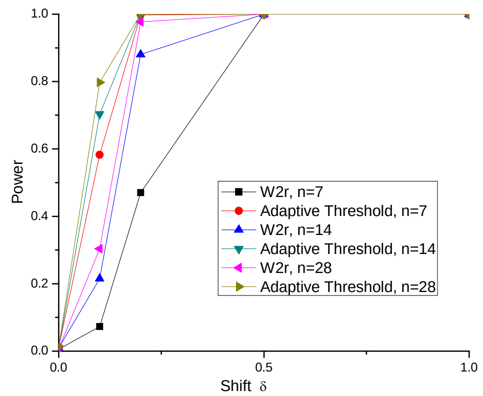

Figure 4-41: Power Analysis for Adaptive Threshold and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA- λ_1=500, RI=500

We further examined the performance of the adaptive threshold method under the assumption that the in-control Poisson parameters are known compared to that of the MLE-based adaptive threshold method. Table 4-10 shows, for the EWMA case, the power values for the adaptive threshold methods across different baselines given RI =500. Table 4-10 leads to the following conclusions:

(1) For a given δ, λ_1 and baseline n, the power values of the adaptive threshold method using known parameters is at least as high as those of the MLE-based adaptive threshold method with

n=28., but not considerably higher.

(2) For a given δ and baseline n, it is clear that the higher the values of λ_1, the higher the power values are.

(3) For the case of very small mean syndrome counts, say λ_1=10, the power values for both methods are small when $\delta \leq 1.0$. For the case of a large shift, say $\delta \geq 2$, generally the power values for both methods are very high.

Our power analysis results show that, in either the EWMA or the Shewhart comparisons, the performance of the adaptive threshold method using known parameters or MLE estimates for the conditional binomial distribution is much better than that of the W2r method under the assumption of Poisson inputs. As a limiting case when the value of n increases, the adaptive threshold method using known in-control parameters works better than the MLE-based adaptive threshold method with n=28, as expected. By considering the n=28 case compared to the known parameters case, we can see how much improvement could be obtained by increasing the size of the baseline ever further.

Table 4-10: Power Analysis for Adaptive Threshold Method Using Known Parameters and Using MLE Estimators-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA (RI=500)

δ	n	Adaptive Threshold Using MLE Estimators					Adaptive Threshold Using Known Parameters				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0115	0.0109	0.0111	0.0110	0.0114	0.0109	0.0105	0.0131	0.0109	0.0099
	14	0.0104	0.0129	0.0113	0.0118	0.0099	0.0122	0.0126	0.0111	0.0142	0.0099
	28	0.0110	0.0106	0.0105	0.0119	0.0128	0.0111	0.0124	0.0117	0.0110	0.0116
0.1	7	0.0255	0.0602	0.1128	0.2209	0.5828	0.0320	0.1012	0.2035	0.4338	0.8821
	14	0.0272	0.0730	0.1320	0.2928	0.7035	0.0325	0.1130	0.2069	0.4401	0.8822
	28	0.0306	0.0858	0.1596	0.3462	0.7973	0.0329	0.1189	0.2128	0.4404	0.8934
0.2	7	0.0523	0.2112	0.4518	0.7986	0.9980	0.0807	0.4218	0.7743	0.9826	1.0000
	14	0.0575	0.2751	0.5546	0.8947	1.0000	0.0902	0.4337	0.7777	0.9857	1.0000
	28	0.0681	0.3461	0.6533	0.9470	1.0000	0.0950	0.4406	0.7812	0.9858	1.0000
0.5	7	0.2332	0.9139	0.9991	1.0000	1.0000	0.4863	0.9983	1.0000	1.0000	1.0000
	14	0.3042	0.9684	1.0000	1.0000	1.0000	0.4933	0.9987	1.0000	1.0000	1.0000
	28	0.3808	0.9902	1.0000	1.0000	1.0000	0.4986	0.9987	1.0000	1.0000	1.0000
1.0	7	0.7258	1.0000	1.0000	1.0000	1.0000	0.9805	1.0000	1.0000	1.0000	1.0000
	14	0.8537	1.0000	1.0000	1.0000	1.0000	0.9822	1.0000	1.0000	1.0000	1.0000
	28	0.9291	1.0000	1.0000	1.0000	1.0000	0.9825	1.0000	1.0000	1.0000	1.0000
2.0	7	0.9963	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

4.4.2.2 Comparison of W2r and Modified W2r Methods

The W2r_1 method was compared to the W2r method in terms of the power analysis for the EWMA approach. Table 4-11 shows the power values for this case when different baselines are used, given RI =500. For a given shift δ >0, Table 4.11 shows some main conclusions regarding performance for both the W2r and W2r_1 methods.

(1) For a given shift δ, a given n and a given λ_1, the power values of the W2r_1 method are at least as high as the corresponding values of the W2r method.

(2) For a given shift δ and a given baseline, it is clear that the higher the parameter values of λ_1, the higher the power values are.

(3) For a given δ and λ_1, we observe that when n increases, the power values increase.

(4) For the case of very small mean syndrome counts, say λ_1=10, the power values for both methods are very small when the shift $\delta \leq 1.0$. For the case of a large shift, say $\delta \geq 2$, the power values for both methods are all very high.

Table 4-11: Power Analysis for Modified W2r and W2r Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA (RI=500)

δ	n	W2r					W2r_1				
		In-control $\lambda_1 = \lambda_2$					In-control $\lambda_1 = \lambda_2$				
		10	50	100	200	500	10	50	100	200	500
0	7	0.0090	0.0091	0.0105	0.0074	0.0052	0.0043	0.0055	0.0155	0.0103	0.0053
	14	0.0115	0.0083	0.0076	0.0102	0.0091	0.0077	0.0103	0.0078	0.0037	0.0055
	28	0.0099	0.0108	0.0067	0.0092	0.0065	0.0112	0.0115	0.0105	0.0130	0.0078
0.1	7	0.0140	0.0135	0.0216	0.0343	0.0728	0.0195	0.0430	0.1452	0.2197	0.4873
	14	0.0144	0.0198	0.0292	0.0716	0.2151	0.0238	0.0590	0.1493	0.3231	0.6002
	28	0.0149	0.0280	0.0337	0.0987	0.3039	0.0278	0.1020	0.1755	0.3498	0.7667
0.2	7	0.0227	0.0363	0.0775	0.1316	0.4701	0.0435	0.1747	0.5065	0.7958	0.9958
	14	0.0264	0.0648	0.1464	0.4126	0.8805	0.0553	0.2377	0.5413	0.8395	0.9993
	28	0.0279	0.1134	0.2227	0.6168	0.9770	0.0678	0.3557	0.6658	0.9482	0.9998
0.5	7	0.0941	0.3178	0.6738	0.9093	1.0000	0.1408	0.8787	0.9992	1.0000	1.0000
	14	0.1211	0.6239	0.9306	0.9998	1.0000	0.2798	0.9620	0.9997	1.0000	1.0000
	28	0.1700	0.8648	0.9939	1.0000	1.0000	0.3673	0.9907	1.0000	1.0000	1.0000
1.0	7	0.3942	0.9335	0.9985	1.0000	1.0000	0.5827	1.0000	1.0000	1.0000	1.0000
	14	0.5542	0.9983	1.0000	1.0000	1.0000	0.8240	1.0000	1.0000	1.0000	1.0000
	28	0.7490	1.0000	1.0000	1.0000	1.0000	0.9175	1.0000	1.0000	1.0000	1.0000
2.0	7	0.9193	1.0000	1.0000	1.0000	1.0000	0.9897	1.0000	1.0000	1.0000	1.0000
	14	0.9874	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figures 4-42 to 4-46 show the power values for n=7, 14, and 28 when λ_1 =10, 50, 100, 200, and 500, respectively. Figure 4-42 shows that for the case λ_1 =10, the W2r_1 method with n=14 works better than the W2r method with n=28. We observed that for the cases of λ_1=50, 100, 200 and 500, the W2r_1 method with n=7 works even better than the W2r method with n=28, as shown in Figures 4-43 to 4-46. It is very clear that there is a large improvement in going from the W2r method to the W2r_1 method for the EWMA control chart methods.

Our power analysis results show that, in either the EWMA or the Shewhart comparisons, the performance of the W2r_1 method is much better than that of the W2r method under the assumption of Poisson inputs.

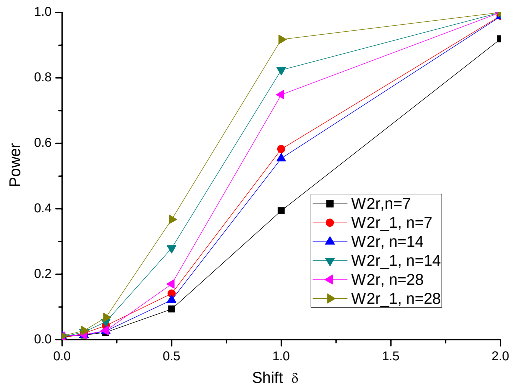

Figure 4-42: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA - λ_1=10, RI=500

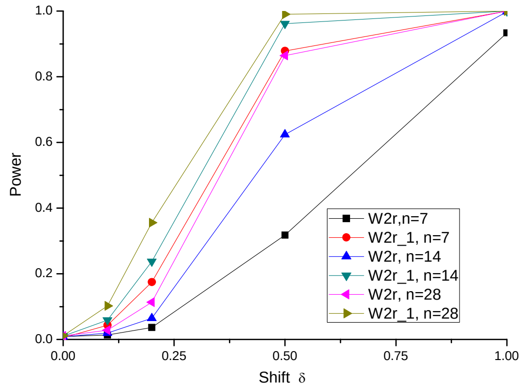

Figure 4-43: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA - λ_1=50, RI=500

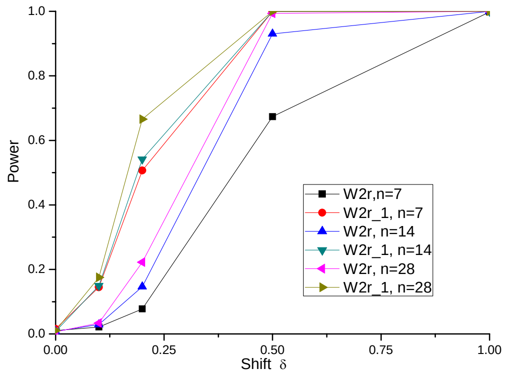

Figure 4-44: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA - λ_1=100, RI=500

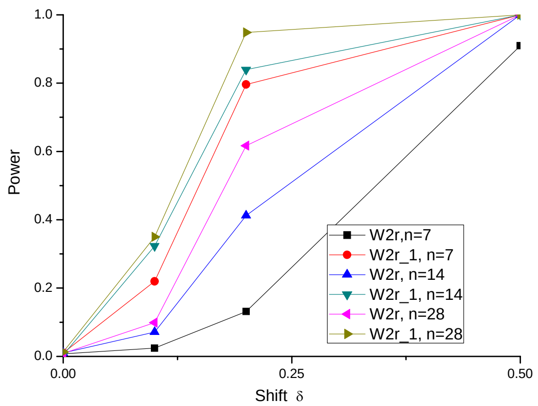

Figure 4-45: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA - λ_1=200, RI=500

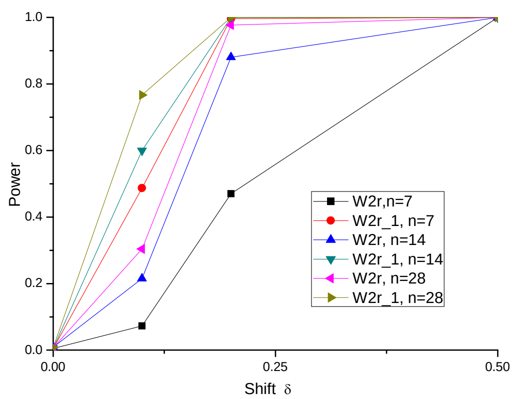

Figure 4-46: Power Analysis for W2r and W2r _1 Methods-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA - λ_1=500, RI=500

4.4.3 Comparison of Shewhart and EWMA Approaches

We compared the Shewhart and EWMA approaches for the adaptive threshold method with Poisson inputs. We observed that for the adaptive threshold method, the EWMA approach with n=7 works at least as well as the Shewhart approach with n=28 for all values of δ>0 and λ_1 in our study. Figure 4-47 shows an example of the power values of the adaptive threshold method for both control chart methods for n=7, 14, and 28 when λ_1 =10 and RI=500. It is very clear that in terms of the power analysis, the EWMA approach is superior to the Shewhart approach in all cases for the adaptive threshold method.

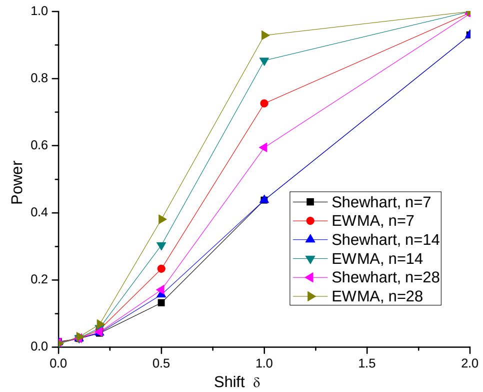

Figure 4-47: Power Analysis for Adaptive Threshold Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart - λ_1=10, RI=500

We also compared the Shewhart and EWMA approaches for the W2r method with Poisson inputs. We observed that for the W2r method, the EWMA approach is generally better than the Shewhart method with a few exceptions in terms of the power analysis. Figure 4-48 shows an example that when λ_1 =10 the EWMA approach with n=7 works better than the Shewhart approach with n=14 for all values of δ in our study. Figure 4-49 shows that, however, when λ_1=200 and n=7, the power value of the Shewhart approach is 0.0361 with δ =0.1 or 0.9420 with δ =0.5, whereas the power value of the EWMA approach is 0.0343 with δ =0.1 or 0.9093 with δ

=0.5. These power values are also shown in Table 4-3 and Table 4-8.

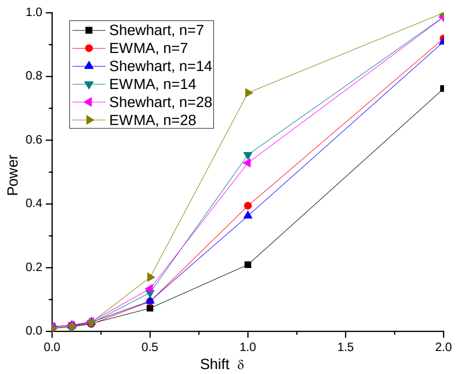

Figure 4-48: Power Analysis for W2r Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=10, RI=500

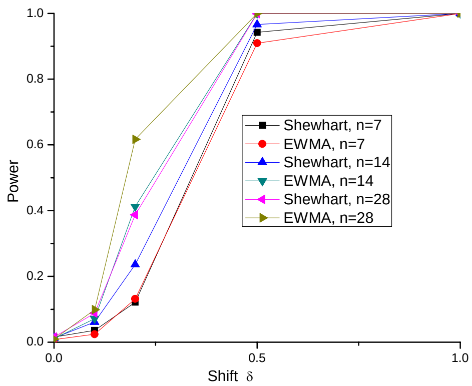

Figure 4-49: Power Analysis for W2r Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=200, RI=500

Figures 4-50 to 4-54 show the power values of the W2r_1 method for both the Shewhart and EWMA approaches for *n*=7, 14, and 28 when λ_1 =10, 50, 100, 200, and 500, given RI=500. We observed that the EWMA approach is generally better than the Shewhart approach in the power analysis for the W2r_1 method. Figures 4-52 and 4-53 show that the EWMA approach with *n*=7 works better than the Shewhart approach with *n*=14 or *n*=28 for all values of δ when λ_1=100 and 200. There are some exceptions, however, when λ_1=10. Figure 4-50 shows that when λ_1=10 and

δ =0.1 the power value of the Shewhart approach is 0.0288 with n=7, or 0.0293 with n=14, or 0.0337 with n=28, whereas the power value of the EWMA approach is 0.0195 with n=7, or 0.0238 with n=14, or 0.0278 with n=28. These power values are also shown in Table 4-6 and Table 4-11.

Our power analysis results show that for Poisson inputs, the EWMA approach was superior to the Shewhart approach in all the cases for the adaptive threshold method, whereas the performance of the EWMA method was better than that of the Shewhart approach in most cases for the W2r and modified W2r methods.

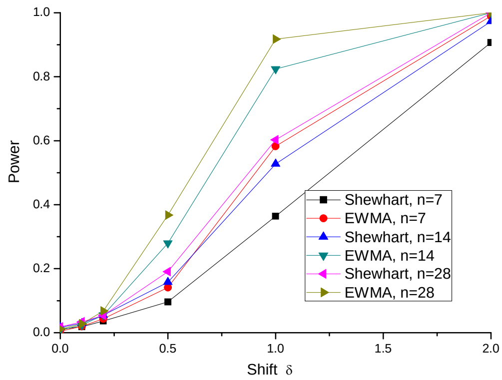

Figure 4-50: Power Analysis for W2r_1-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=10, RI=500

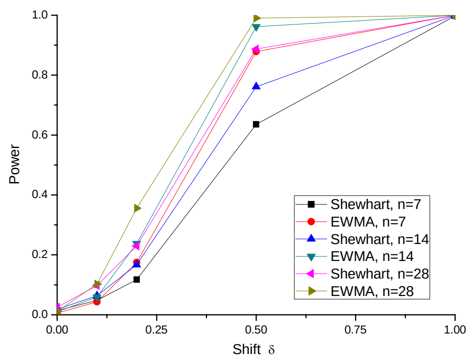

Figure 4-51: Power Analysis for W2r_1 Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=50, RI=500

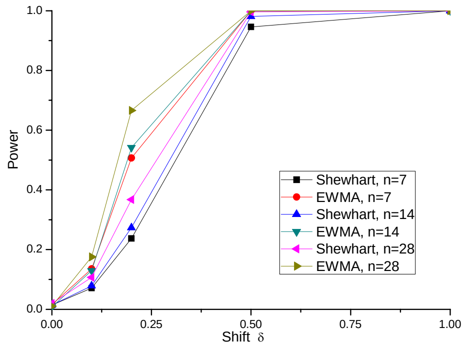

Figure 4-52: Power Analysis for W2r_1 Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=100, RI=500

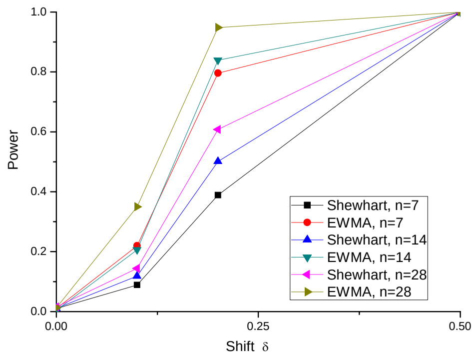

Figure 4-53: Power Analysis for W2r_1 Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=200, RI=500

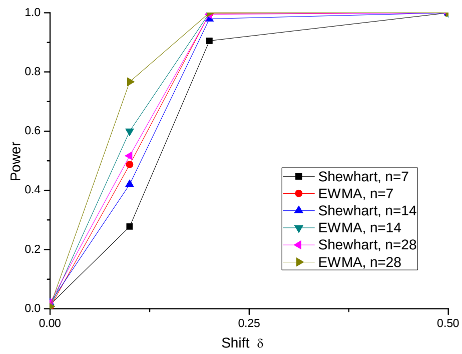

Figure 4-54: Power Analysis for W2r_1 Method-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA vs. Shewhart- λ_1=500, RI=500

4.5 Weekend Effects

Health care facilities typically have a lower number of visits on weekends and this feature can be reflected in the simulations. It is of interest to determine how the RI threshold functions are affected when these data are pooled together to form an empirical distribution of monitoring statistics. We examined how the W2r method and the adaptive threshold method based on the conditional binomial distribution perform where the average Poisson count is significantly lower on weekends with the results reported in this section.

Tokars et al. [6] used the Department of Defense (DoD) data with a strong day-of-week effect; 16%–21% of total weekly visits occurred per day on weekdays, and only 3%–4% of visits occurred per day on weekend days and holidays. Their hospital emergency department data had a minimal day-of-week effect: 14%–16% of visits occurred per day on weekdays, and 14%–15% of visits occurred per day on weekend days. For simplicity, we assumed the syndrome counts per day on weekdays are 5 times of those per day on weekends, i.e., we let $\lambda_1 = 5\lambda_3$ and $\lambda_2 = 5\lambda_4$. We used the parameters listed in Table 4-1. The ratio of λ_1 and λ_2 was varied from 0.1 to 10.

4.5.1 RI Threshold Function Analysis

Figures 4-55 and 4-56 show the RI threshold functions using the Shewhart approach for the W2r and adaptive threshold methods when different baselines are used, given λ_1=10 and λ_1=100, respectively. Figure 4-55 shows there is considerable variability in the RI threshold curves for small counts with λ_1=10 for both methods. Figure 4-56 shows that generally both methods work better for counts with λ_1=100 as there is less variation in the corresponding RI threshold curves. It is clearly seen that the Z-score threshold functions are more robust to parameter changes than the W2r threshold functions for all of the different baseline periods, as shown in Figure 4-56.

Figures 4-57 and 4-58 show the RI threshold functions using the EWMA approach for the W2r and adaptive threshold methods when different baselines are used, given λ_1=10 and λ_1=100, respectively. The adaptive threshold RI threshold functions are more robust to parameter changes than the W2r threshold functions for all baseline periods.

It appears that for both the Shewhart and EWMA approaches the adaptive threshold method outperforms the W2 rate method by having less variation in the required threshold values across the parameter space. These RI threshold function analysis results for the W2 method and adaptive threshold method are similar to what we found in Section 4.1 when the average counts of the Poisson inputs are assumed to be the same on weekdays and weekend days.

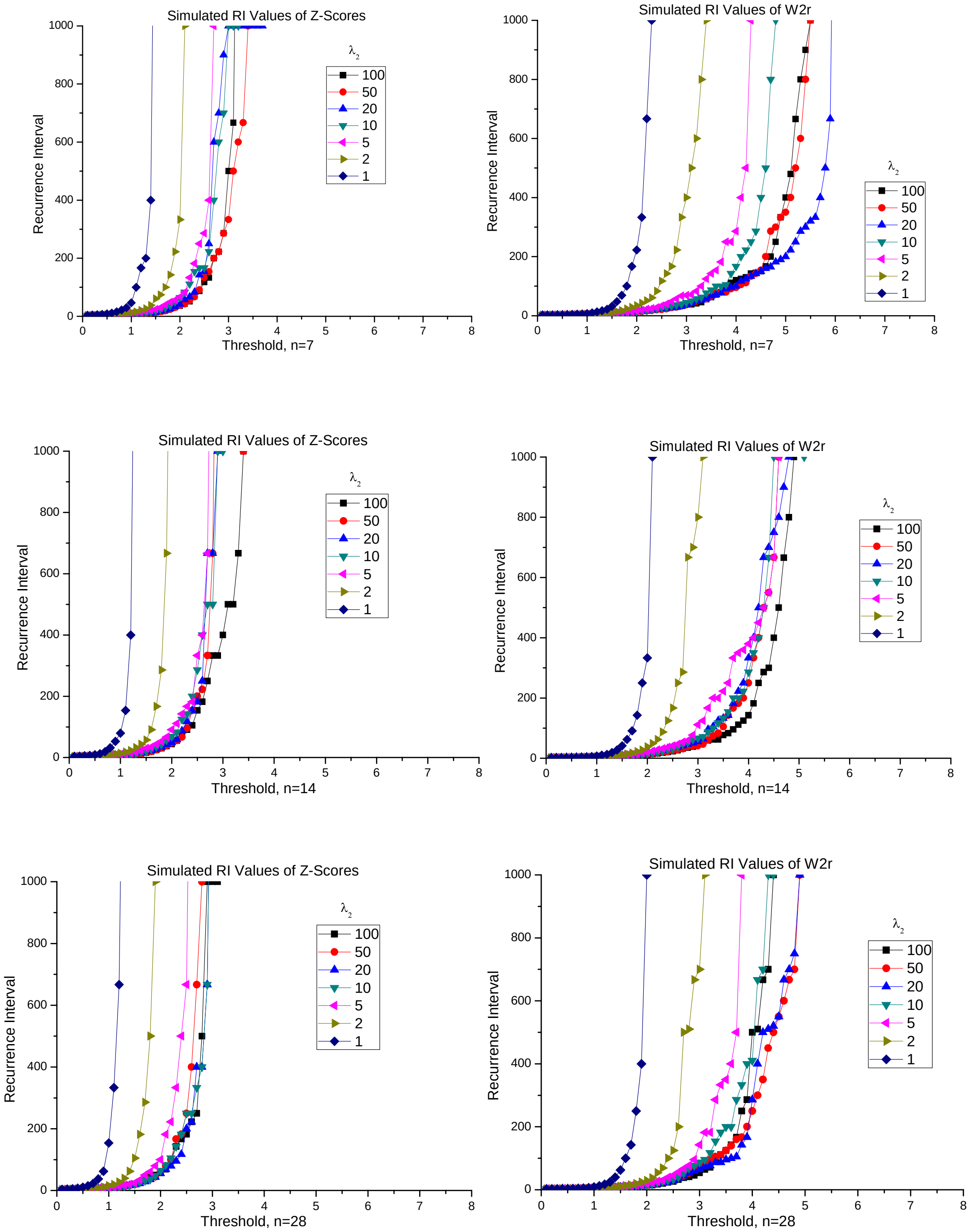

Figure 4-55: RI Threshold Functions Reflecting Weekend Effects-Conditional Binomial Counts, λ_1=10 -Shewhart

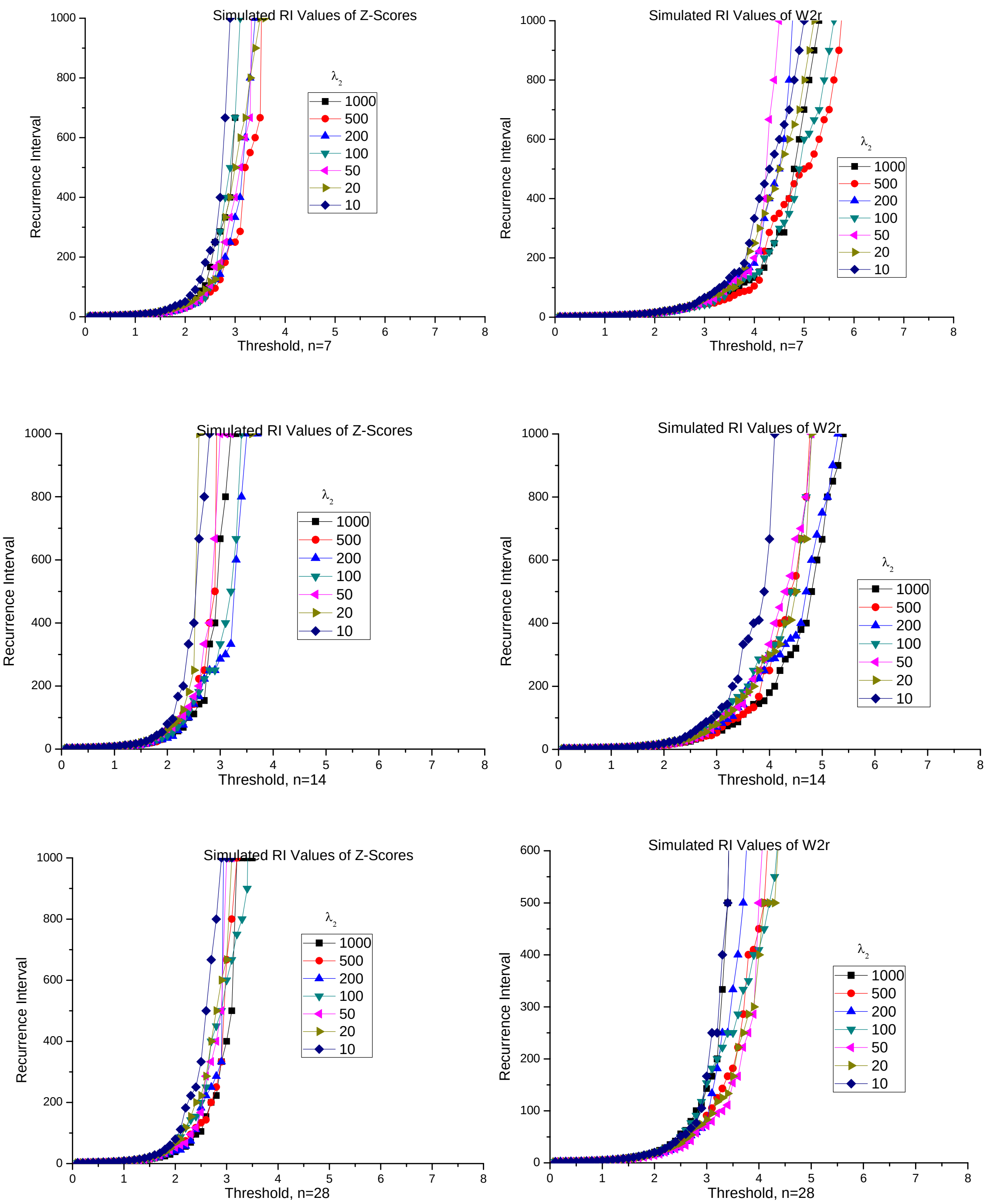

Figure 4-56: RI Threshold Functions Reflecting Weekend Effects-Conditional Binomial Counts, λ_1=100-Shewhart

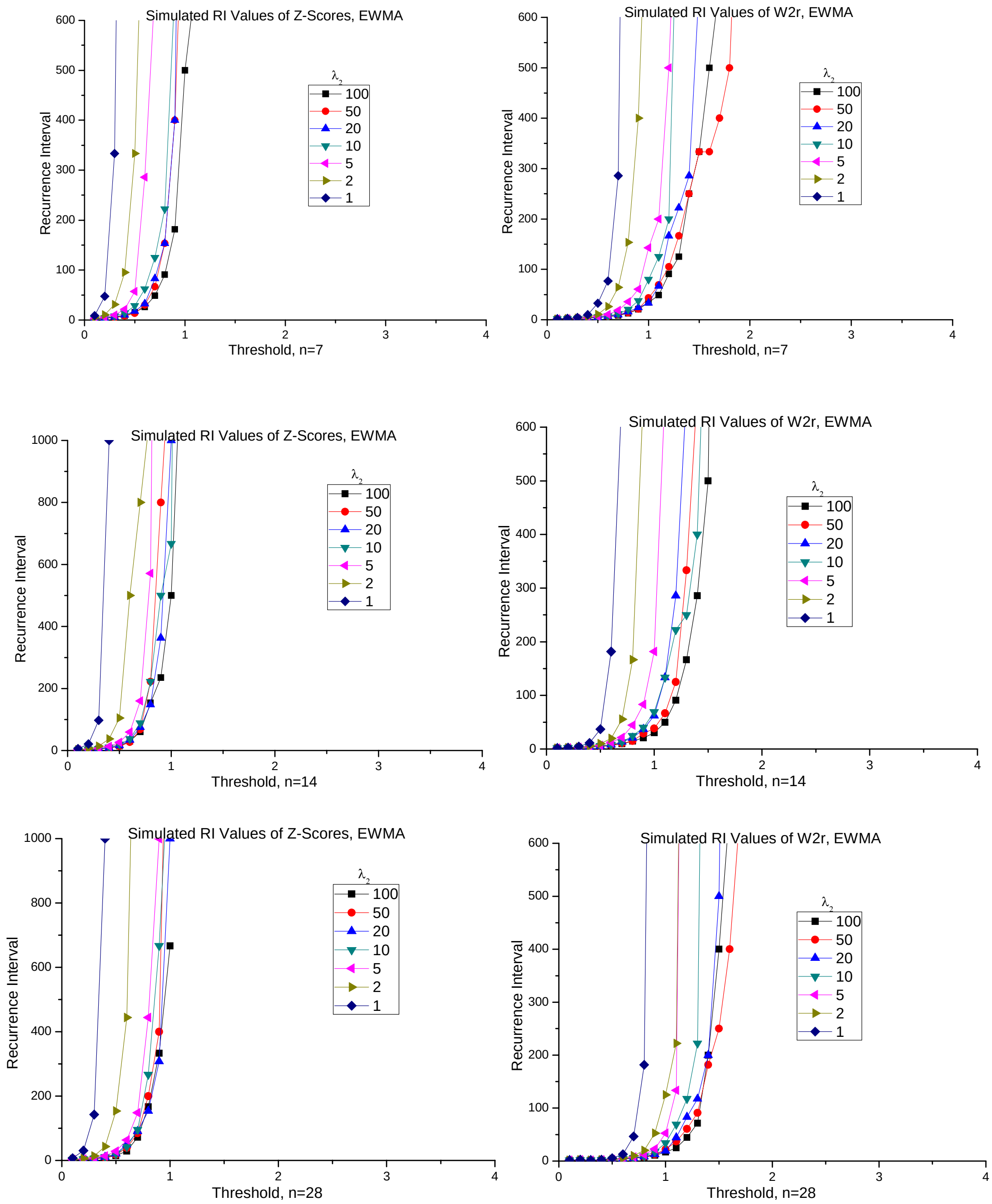

Figure 4-57: RI Threshold Functions Reflecting for Weekend Effects-Conditional Binomial Counts, λ_1=10- EWMA

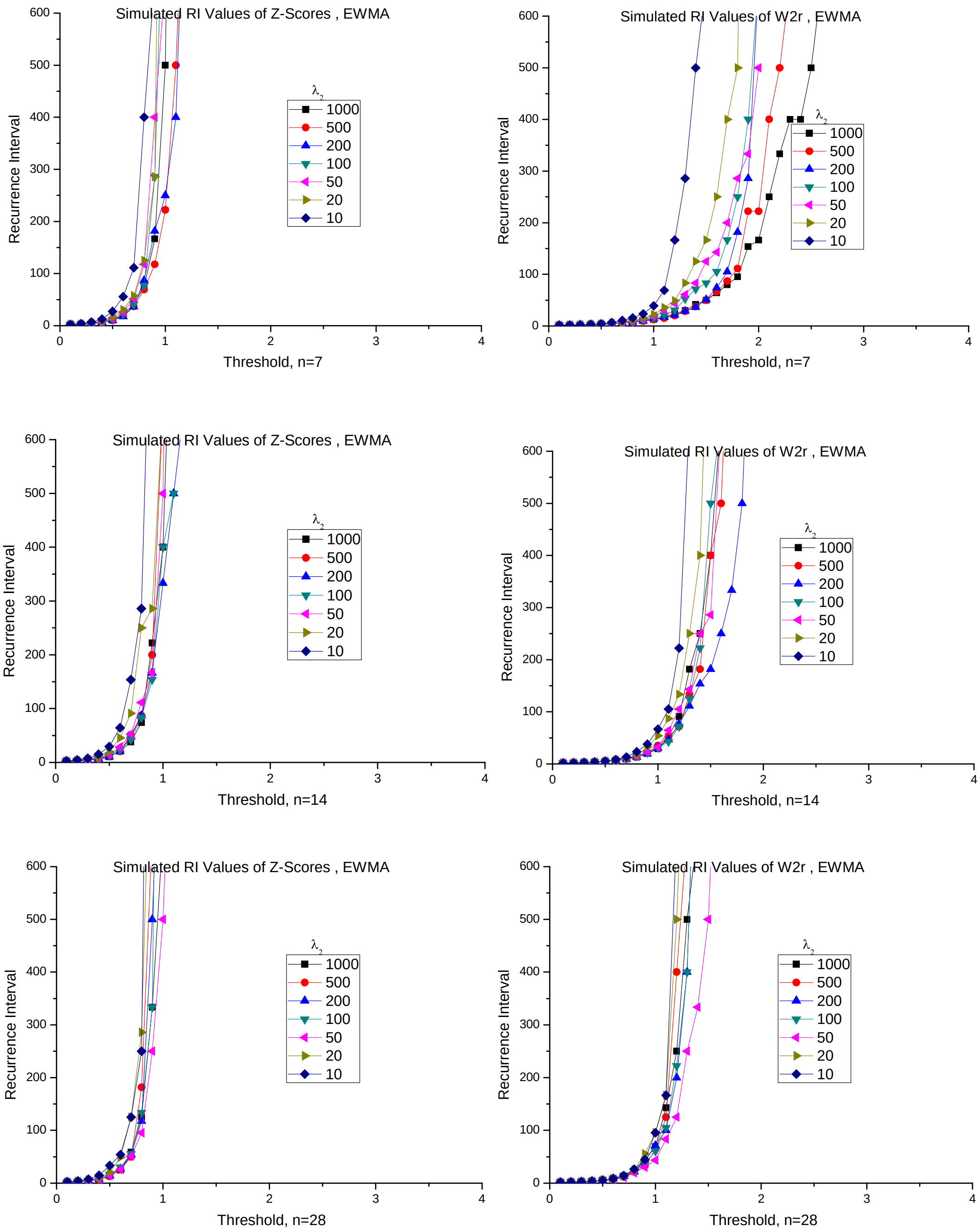

Figure 4-58: RI Threshold Functions Reflecting for Weekend Effects-Conditional Binomial Counts, λ_1=100-EWMA

4.5.2 Power Analysis

To study the power of the approaches when there were weekend effects, we used the first ten weeks as the in-control data. The outbreak is then assumed to occur and be one week in length. We considered a percentage shift of size δ in the Poisson parameters λ_1 and λ_3 which increases the average number of syndromic counts per weekday and weekend day over the course of the outbreak, respectively. The values used are δ= 0, .1, .2, .5, 1, and 2, where a zero shift indicates analysis when there is no outbreak. We considered $\lambda_1 = \lambda_2 = 5\lambda_3 = 5\lambda_4$ for the conditional binomial counts when there is no outbreak present. The increase in the Poisson parameter λ_1 is assumed to be $(\delta * 100)\%$ of the in-control parameter value. Correspondingly the increase in λ_3 is also assumed to be $(\delta * 100)\%$ of the in-control parameter value. Table 4-12 shows the threshold values we used for the surveillance methods in the power analysis.

Table 4-12: Threshold Values of Adaptive Threshold and W2r Methods-Conditional Binomial Case with Poisson Inputs-Weekend Effects (RI=500)

λ_1		Shewhart			EWMA		
		n=7	14	28	*n*=7	14	28
10	W2r	5.53	5.10	4.63	1.42	1.47	1.31
	Z-Score	2.96	2.76	2.73	0.83	0.80	0.78
50	W2r	6.16	5.03	4.64	1.51	1.37	1.23
	Z-Score	2.91	2.88	2.85	0.93	0.94	0.93
100	W2r	6.18	5.27	4.61	1.58	1.46	1.29
	Z-Score	3.11	2.99	2.79	1.00	0.95	0.96
200	W2r	6.74	5.03	4.66	1.66	1.51	1.28
	Z-Score	3.10	2.92	2.85	1.03	0.97	0.97
500	W2r	6.18	5.25	4.90	1.58	1.35	1.43
	Z-Score	3.16	2.98	3.01	1.04	1.02	0.99

Table 4-13 shows the power values for the adaptive threshold and W2r methods with the

Shewhart approach across different baselines given RI =500. We can observe that the probabilities of a signal during a week with no outbreak, i.e., δ=0, are all around .014 as shown in Table 4-13. Figure 4-59 shows the power values given λ_1=100 (left) and 200 (right). Table 4-13 and Figure 4-59 lead to the following conclusions regarding performance for both methods based on the Shewhart approach:

(1) For a given shift δ, a given n and a given λ_1, the power values of the adaptive threshold method are at least as high as those of the W2r method.

(2) For a given shift δ and a given baseline n, it is clear that higher values of λ_1 lead to higher power values.

(3) For a given shift δ and a given λ_1, we observe that when n increases the power values increase.

(4) It is also noted that δ=2 is a huge shift when λ_1=100 or λ_1=200.

The power values for the adaptive threshold and W2r methods with weekend effects as shown in Table 4-13 were also compared to those in Table 4-4. We observed that the power values for these surveillance methods are a little lower when there is a weekend effect for the Shewhart approach.

Table 4-13: Power Analysis for Weekend Effect-Transient Shift in Conditional Binomial Case with Poisson Inputs-Shewhart (RI=500)

δ	n	W2r		Adaptive Threshold	
		In-control $\lambda_1 = \lambda_2$		In-control $\lambda_1 = \lambda_2$	
		100	200	100	200
0	7	0.0097	0.0091	0.0097	0.0095
	14	0.0110	0.0124	0.0106	0.0148
	28	0.0116	0.0101	0.0138	0.0154
0.1	7	0.0209	0.0207	0.0496	0.0982
	14	0.0242	0.0598	0.0537	0.1243
	28	0.0321	0.0855	0.0840	0.1417
0.2	7	0.0539	0.1206	0.1717	0.3854
	14	0.0966	0.2076	0.2091	0.5105
	28	0.1450	0.3492	0.2944	0.5667
0.5	7	0.4315	0.7421	0.8767	0.9981
	14	0.7013	0.9612	0.9473	1.0000
	28	0.9079	0.9993	0.9845	1.0000
1	7	0.9639	0.9980	1.0000	1.0000
	14	0.9994	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000

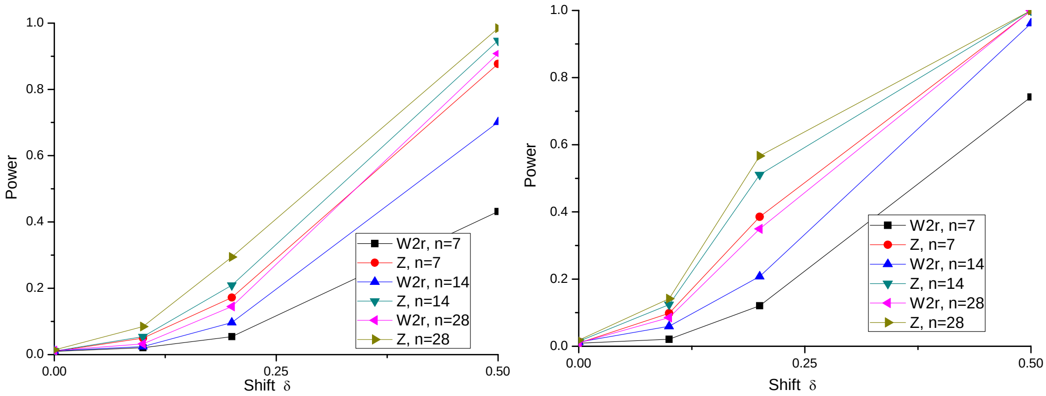

Figure 4-59: Power Analysis for Weekend Effects-Shewhart- λ_1 =100(left) and λ_1 =200 (right), RI=500

Table 4-14 shows the power values for the adaptive threshold and W2r method based on the EWMA approach across different baselines given RI =500. Figure 4-60 shows the power values given λ_1 =100 (left) and 200 (right). Table 4-14 and Figure 4-60 lead to the following conclusions regarding performance for these methods based on the EWMA approach:

(1) For a given δ, n and λ_1, the power values of the adaptive threshold method are at least as high as those of the W2r method.

(2) For a given δ and baseline n, it is clear that higher values of λ_1 lead to higher power values.

(3) For a given δ and λ_1, we observe that when n increases the power values increase.

(4) It is also noted that δ=2 is a huge shift when λ_1=100 or λ_1=200.

The power values for the adaptive threshold and W2r methods with weekend effects as shown in Table 4-14 were also compared to those in Table 4-8. We observed that the power values for these surveillance methods are a little lower when there is a weekend effect for the EWMA approach.

Table 4-14: Power Analysis for Weekend Effect-Transient Shift in Conditional Binomial Case with Poisson Inputs-EWMA (RI=500)

δ	n	W2r		Adaptive Threshold	
		In-control $\lambda_1 = \lambda_2$		In-control $\lambda_1 = \lambda_2$	
		100	200	100	200
0	7	0.0136	0.0168	0.0175	0.0144
	14	0.0175	0.0148	0.0154	0.0157
	28	0.0154	0.0178	0.0124	0.0137
0.1	7	0.1276	0.1816	0.1297	0.2286
	14	0.1411	0.2427	0.1469	0.2874
	28	0.1464	0.4128	0.1488	0.5007
0.2	7	0.3725	0.5668	0.4401	0.7400
	14	0.4987	0.7604	0.5332	0.8544
	28	0.6607	0.8833	0.6676	0.8874
0.5	7	0.9616	0.9983	0.9957	1.0000
	14	0.9964	1.0000	0.9995	1.0000
	28	1.0000	1.0000	1.0000	1.0000
1	7	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000

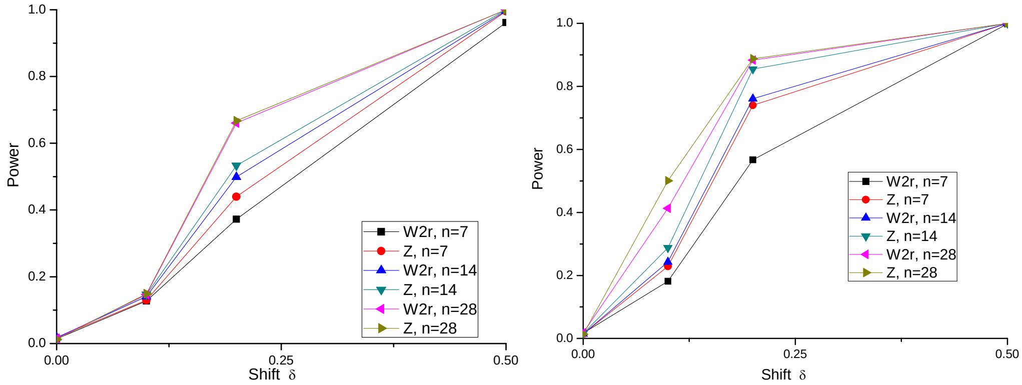

Figure 4-60: Power Analysis for Weekend Effects-EWMA- λ_1=100(left) and λ_1=200 (right), RI=500

Figures 4-61 and 4-62 show the power values of the adaptive threshold and W2r methods respectively for both the Shewhart and EWMA approaches. We observed that for the adaptive threshold and W2r methods, the EWMA approach with n=7 works at least as well as the Shewhart approach with n=28 for any positive values of δ in our study.

Our power analysis results show that for the cases where the average Poisson count is significantly lower on weekends, the EWMA approach is superior to the Shewhart approach in all cases for the adaptive threshold method. For the W2r method, the EWMA approach works better than the Shewhart approach in all cases. These results for the W2r and adaptive threshold methods are similar to what we reported in Section 4.4.3 when we assumed the average Poisson counts are the same as on the weekdays as on the weekend days. It is also noted that the power values for these surveillance methods are a little lower when there is a weekend effect.

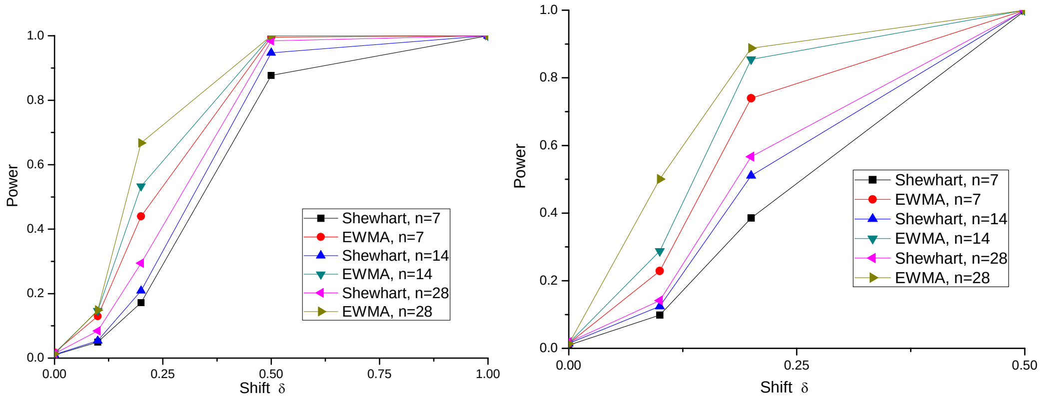

Figure 4-61: Power Analysis of Adaptive Threshold Method with Weekend Effects-EWMA vs. Shewhart- λ_1=100(left) and λ_1=200 (right), RI=500

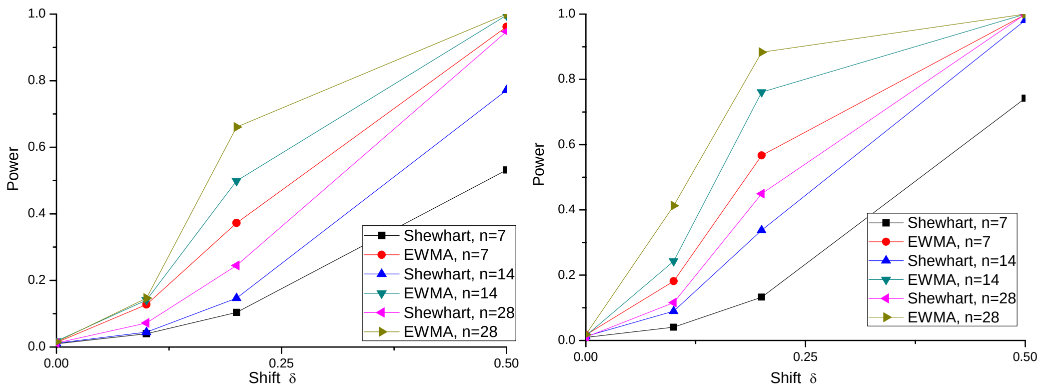

Figure 4-62: Power Analysis of W2r Method with Weekend Effects-EWMA vs. Shewhart - λ_1=100(left) and λ_1=200 (right), RI=500

Chapter 5 Performance Evaluation of Adaptive Threshold and W2r Methods with Negative Binomial Inputs

In this chapter, we report the results of a simulation study for several surveillance methods when the input data streams are assumed to be not Poisson distributed, as in Chapter 4, but negative binomial distributed. The negative binomial distribution is often more realistic than the Poisson distribution for public health data. Public count data are often overdispersed compared to the Poisson distribution. We consider the adaptive threshold methods underlying conditional binomial and conditional negative binomial distributions, the W2r method, and a modified W2r method. We examine the performance of these methods in terms of the RI threshold function analysis and a power analysis for both Shewhart and one-sided EWMA chart approaches.

5.1 Simulation Plan

5.1.1 In-control Data

We assumed weekday and weekend counts each follow independent negative binomial distributions. More precisely, we assumed the syndrome counts in weekdays follows a negative binomial distribution with the parameters (r_1, p_1), the non-syndrome counts in weekdays follows a negative binomial distribution with the parameters (r_2, p_2), the syndrome counts in weekends follows a negative binomial distribution with the parameters (r_3, p_3), and the non-syndrome counts in weekends follows a negative binomial distribution with the parameters (r_4, p_4). For simplicity, we assume $r_1 = r_3$, $r_2 = r_4$, $p_1 = p_3$, and $p_2 = p_4$ in this section. We used the parameter combinations for eight cases as listed in Table 5-1 for the negative binomial inputs. These include moderate to extreme overdispersion compared to the Poisson distribution. The variance-to-the-mean ratio ranges from 2 to10 for these eight cases. We used the simulated in-control data to check the uniform distribution model for the in-control p-values for the adaptive threshold methods with the results reported in Section 5.2.1. We also did a RI threshold function analysis which is described in Section 5.3.

Table 5-1: Negative Binomial Parameters Used in Chapter 5

Case	(r_1, p_1),	(μ_1, σ_1^2)	(r_2, p_2),	(μ_2, σ_2^2)
1	(100,.2)	(400,2000)	(50,.1)	(450, 4500)
2	(150,.3)	(350,1167)	(50, .1)	(450, 4500)
3	(150,.3)	(350,1167)	(100, .2)	(400,2000)
4	(150, .3)	(350,1167)	(150,.3)	(350,1167)
5	(200, .2)	(800,4000)	(200, .2)	(800,4000)
6	(150, .5)	(150,300)	(50, .1)	(450,4500)
7	(150, .5)	(150,300)	(100,.2)	(400,2000)
8	(150, .5)	(150,300)	(150,.5)	(150,300)

5.1.2 Outbreak Data

In Section 5.1.1 we briefly discussed the simulation of in-control data over time, where the distribution parameters stay constant. For outbreak data generation baseline data were first simulated from the in-control distributions for ten weeks, and then an outbreak lasting seven days was injected. This process was repeated 100,000 times for each parameter combination considered. For each of these transient shifts, we determined the proportion of times the various methods signaled during the outbreak. We used the simulated outbreak data to do the power analysis with the results given in Section 5.4.

5.2 Methods

5.2.1 Adaptive Threshold Methods

We first examine the performance of the adaptive threshold method based on the conditional negative binomial distribution. We then consider use of the conditional binomial distribution for the negative binomial inputs because use of this approach would be simpler to implement.

5.2.1.1 Conditional Negative Binomial Distribution

We first considered the conditional negative binomial distribution assuming the in-control values of the parameters are known. Figure 5-1 shows how the in-control p-values are distributed for the eight cases listed in Table 5-1. It can be seen that these in-control p-values are approximately uniformly distributed over (0,1).

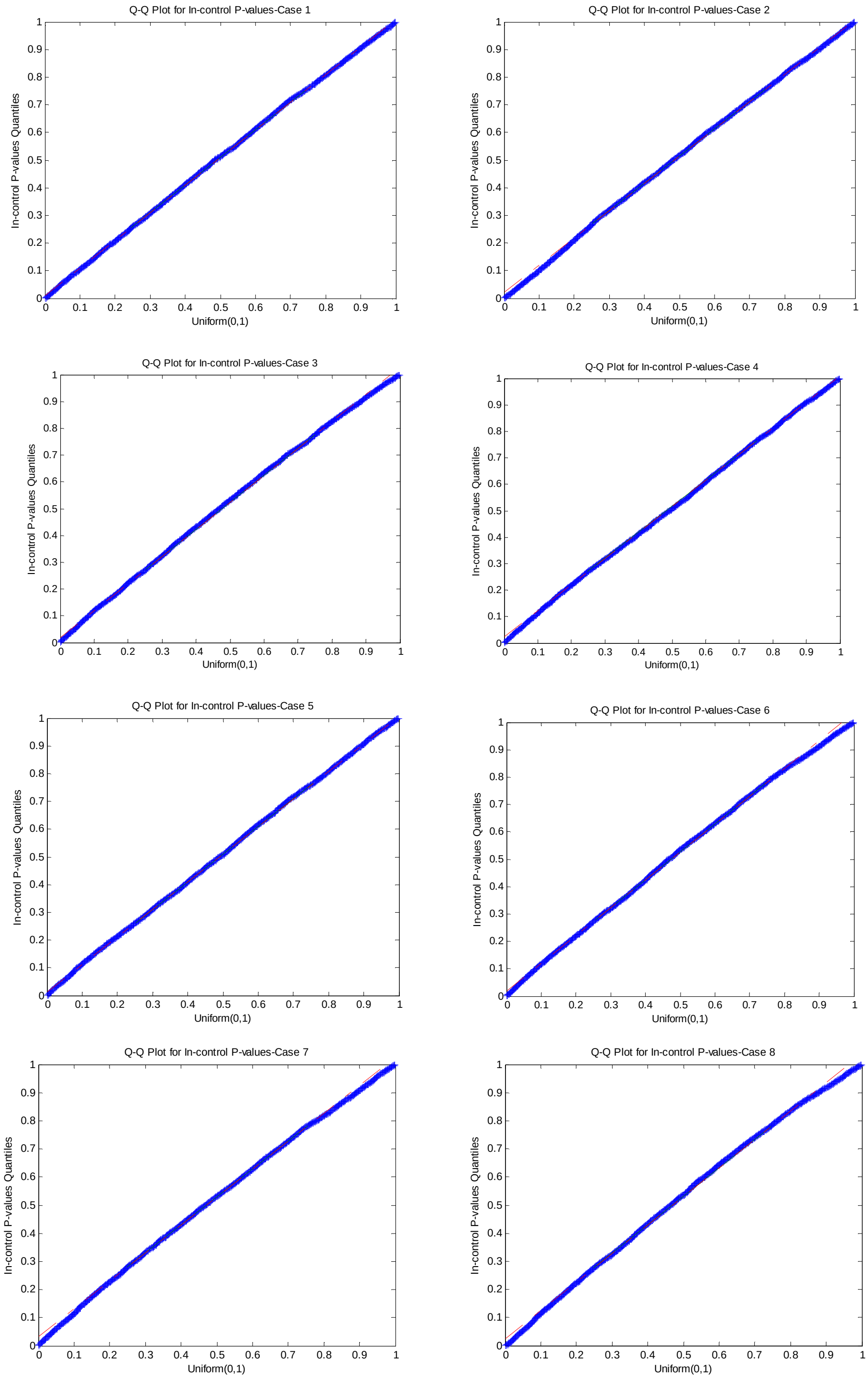

Figure 5-1: Q-Q Plots for In-control P-values for Adaptive Threshold Method with Known Parameters -Conditional Negative Binomial Distribution, *n*=7

We also examined the conditional negative binomial distribution using the method of moments (MOM) estimators. We showed in Section 3.2 that the MOM estimators of the negative binomial parameters for syndrome counts and nonsyndrome counts are

$$\hat{p}_{it} = \frac{\bar{x}_{it}}{\hat{\sigma}_{it}^2} \text{ and } \hat{r}_{it} = \frac{\bar{x}_{it}^2}{\hat{\sigma}_{it}^2 - \bar{x}_{it}} \; ; \; i = 1,2, \qquad (3.6)$$

where $\bar{x}_{it} = \frac{1}{n}\sum_{j=1}^{n} y_{ijt}$, $\hat{\sigma}_{it}^2 = \frac{1}{n}\sum_{j=1}^{n}(y_{ijt} - \bar{x}_{it})^2$, and y_{ijt}, i=1,2, j=1,2,…,n, correspond to the eligible baseline syndrome data (i=1) and nonsyndrome data (i=2) for day t. Clearly the domain of these parameters of the fitted negative binomial distributions is violated when $\bar{x}_{it} > \hat{\sigma}_{it}^2, i = 1,2.$ We developed an algorithm to compute the Z-scores using MOM estimators for the conditional negative binomial distribution, as discussed later in this section.

The values of z-scores for the conditional negative binomial distribution are determined by the p-values, whereas the p-values depend on six components, i.e., input data x_{1t}, x_{2t}, and MOM estimators for r_1, p_1, r_2, and p_2, which are denoted by $\hat{r}_1$, $\hat{p}_1$, $\hat{r}_2$, and $\hat{p}_2$. In order to improve the performance of these MOM estimators for our model, we used the *Z_Negative Binomial* algorithm to compute the z-score values as described in the following:

1. If a large syndromic count x_{1t} or a non-syndromic count x_{2t} that is beyond the .9999 quantile of the fitted negative binomial distribution, then it is replaced with a random count beyond the 99th percentile of the fitted distributions.
2. If $\bar{x}_{1t} \geq \hat{\sigma}_{1t}^2$, we set $\hat{\sigma}_{1t}^2 = 1.05\,\bar{x}_{1t}$; if $\bar{x}_{2t} \geq \hat{\sigma}_{2t}^2$, set $\hat{\sigma}_{2t}^2 = 1.05\,\bar{x}_{2t}$.
3. Compute the pmf given by Theorem 3.1 for the syndromic count x_{1t} given d_t by using the MOM estimators.
4. Calculate the p-values. If the p-value is less than 10^{-6}, then it is replaced with a random value between 10^{-6} and 10^{-5}. Compute a Z-score by taking the inverse standard normal CDF of one minus the p-value.

The first step in the above algorithm is to mitigate the problem of estimation for negative binomial inputs. Lambert and Liu [4] used a technique of outlier removal by replacing an extremely large outlier that is beyond .9999 quantile of the estimated distribution with a randomly generated count beyond the .99 quantile of the fitted distribution. The second step uses a technique implemented by Watkins et al. [26]. This is an ad-hoc method to give valid estimators. See also Szarka, Gan, and Woodall [13]. If the conditions from step 2 are used, then $\hat{p}_1 = \hat{p}_2 = 1/1.05 = .9524$, $\hat{r}_1 = 20\bar{x}_{1t}$, and $\hat{r}_2 = 20\bar{x}_{2t}$. Table 5-2 shows the proportion of the time $\bar{x}_{1t} \geq \hat{\sigma}^2_{1t}$ and $\bar{x}_{2t} \geq \hat{\sigma}^2_{2t}$, respectively. It is seen that the proportion is larger if n is small and p_1 (or p_2) is large.

Table 5-2: Proportion of the Time $\bar{x}_{1t} \geq \hat{\sigma}^2_{1t}$ and $\bar{x}_{2t} \geq \hat{\sigma}^2_{2t}$ for Negative Binomial Inputs

Case	(r_1, p_1)	7-days	14-days	28-days	(r_2, p_2)	7-days	14-days	28-days
1	(100,.2)	0.0159	0.0103	0.0097	(50,.1)	0.0062	0.0056	0.0052
2	(150,.3)	0.0307	0.0198	0.0164	(50, .1)	0.0064	0.0055	0.0054
3	(150,.3)	0.0314	0.0201	0.0167	(100, .2)	0.0160	0.0105	0.0102
4	(150, .3)	0.0313	0.0204	0.0166	(150,.3)	0.0305	0.0207	0.0169
5	(200, .2)	0.0158	0.0108	0.0107	(200, .2)	0.0158	0.0103	0.0099
6	(150, .5)	0.0718	0.0698	0.0513	(50, .1)	0.0064	0.0054	0.0054
7	(150, .5)	0.0715	0.0703	0.0516	(100,.2)	0.0159	0.0105	0.0100
8	(150, .5)	0.0710	0.0687	0.0522	(150,.5)	0.0722	0.0709	0.0518

For step 4 in the *Z_Negative Binomial* algorithm we used a technique of outlier removal for the extremely small p-values. We used the following example to explain the rationale behind step 4 of the *Z_Negative Binomial* algorithm. We simulated 21 days of in-control data assuming the parameter values are (r_1, p_1)= (100, 0.2), and (r_2, p_2)= (50, 0.1). Table 5-3 shows the MOM estimators and the Z-score values if we do not use the techniques of outlier removal listed in step 4. We observed the Z-score value calculated without step 4 takes the value of infinity at day 20. Figure 5-2 shows why we obtained such a value on the 20-th day. On that day, $d_{20}= x_{1,20}+x_{2,20}$

=578+458=1036. The p-value on day 20 is therefore $3.79*10^{-18}$. Since this p-value is extremely small, the corresponding Z-score value is shown as infinity.

Table 5-3: Example of Z-Score Values Using *Z_Negative Binomial* Algorithm without Step 4

Day	x_{1t}	x_{2t}	$\hat{r}_1$	$\hat{p}_1$	$\hat{r}_2$	$\hat{p}_2$	z	z_EWMA
1	444	450	0.000	0.000	0.000	0.000	0.000	0.000
2	367	527	8880.000	0.952	9000.000	0.952	-5.140	0.000
3	362	409	152.710	0.274	240.133	0.330	0.661	0.132
4	430	498	150.227	0.278	110.555	0.193	0.314	0.169
5	437	379	170.681	0.299	141.616	0.231	2.103	0.556
6	444	515	190.114	0.318	80.915	0.152	0.185	0.481
7	400	467	205.101	0.331	83.586	0.153	-0.443	0.296
8	465	439	248.475	0.376	100.724	0.178	1.489	0.535
9	426	358	185.959	0.309	97.349	0.174	1.513	0.731
10	388	445	335.240	0.442	74.442	0.145	-1.167	0.351
11	422	502	1158.044	0.731	80.690	0.154	-0.699	0.141
12	377	337	1134.225	0.727	78.816	0.151	-0.973	0.000
13	374	465	414.068	0.498	54.811	0.111	-1.669	0.000
14	413	317	332.403	0.449	68.608	0.137	1.240	0.248
15	383	406	341.426	0.455	42.788	0.095	-0.881	0.022
16	388	360	7752.842	0.951	43.415	0.097	-0.331	0.000
17	346	426	7842.857	0.952	43.781	0.098	-2.429	0.000
18	381	525	899.890	0.700	45.286	0.101	-0.847	0.000
19	380	500	7605.714	0.952	38.301	0.086	-0.440	0.000
20	578	458	7614.286	0.952	42.272	0.090	**Inf**	**Inf**
21	421	524	36.277	0.081	42.726	0.091	-0.831	**Inf**

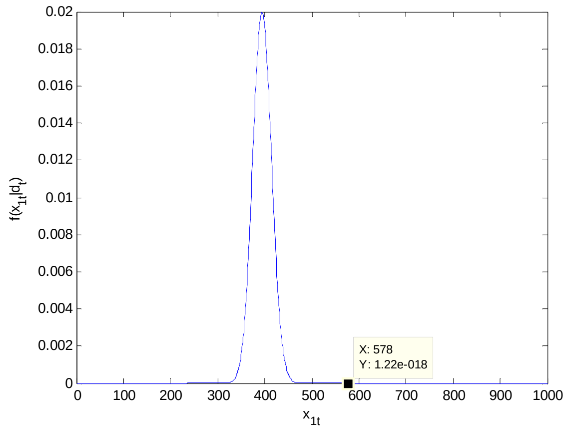

Figure 5-2: Example of the Probability Mass Function of $X_{1,20}$ Given d_{20} Using *Z_ Negative Binomial* Algorithm without Step 4

Figure 5-3 shows how the in-control adaptive threshold p-values based on the conditional negative binomial distribution are distributed given n=7. We observed there are heavier tails for these eight cases, compared to the Q-Q plots for the in-control p-values for the conditional negative binomial distribution assuming known parameters. It can be seen that, however, these in-control p-values are still approximately uniformly distributed over (0,1). The in-control p-values based on n=14 and 28 days are all more closely approximately uniformly distributed over (0,1). These results are not shown here.

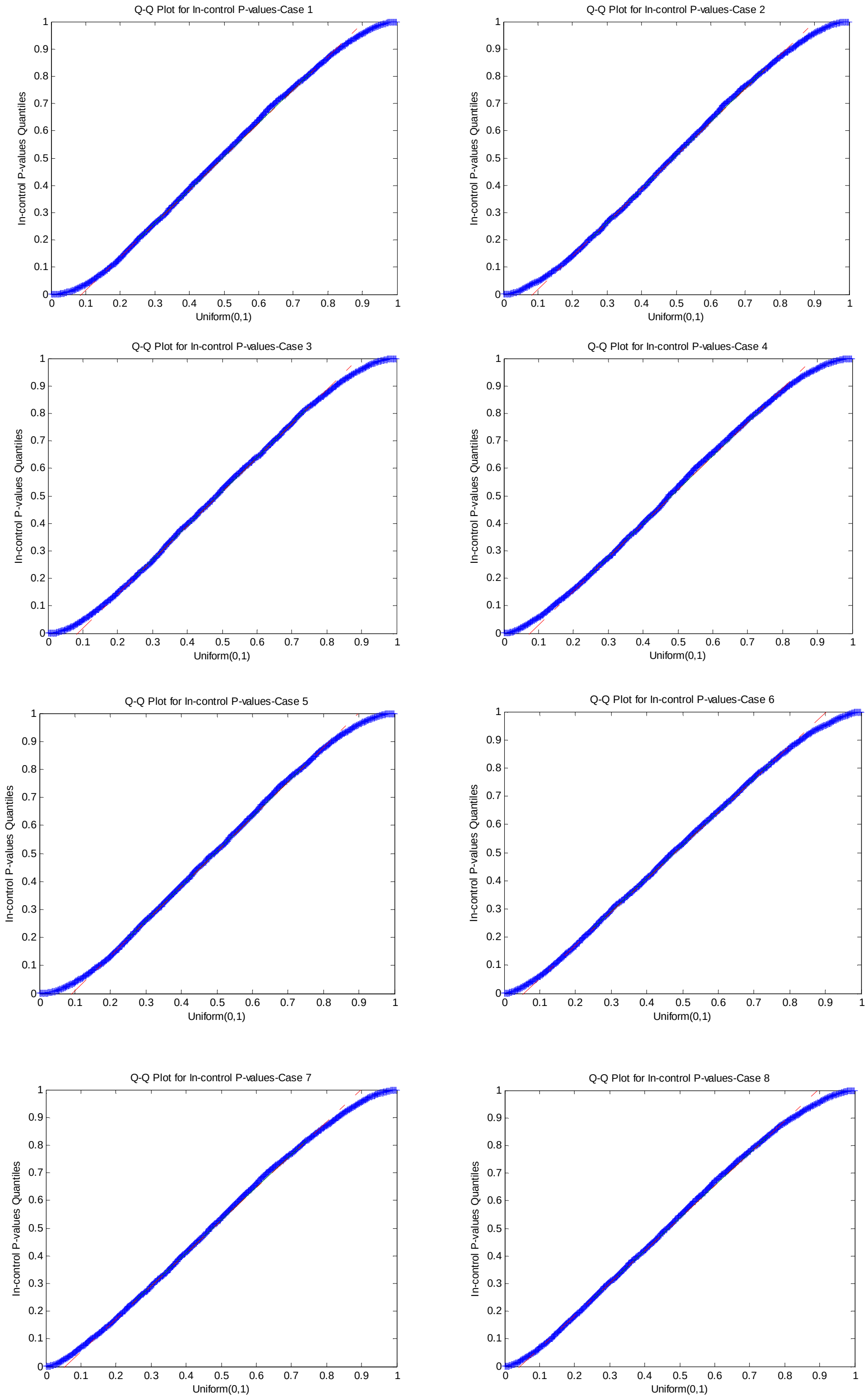

Figure 5-3: Q-Q Plots for In-control P-values for Adaptive Threshold Method Using MOM Estimators-Conditional Negative Binomial Distribution, n=7

5.2.1.2 Conditional Binomial Distribution

As shown in Chapter 4, the conditional binomial distribution works very well with the adaptive threshold method for the Poisson inputs. We considered whether the conditional binomial distribution works well for the negative binomial inputs. We first checked out whether the in-control p-values of the adaptive threshold methods are approximately distributed uniformly over [0,1] in this section. We used the *Z_Negative Binomial* algorithm to compute the z-score values as described in Section 5.2.1.1. The only differences are that we did not need step 2, and in step 3 we computed the pmf for the syndromic count by using the MLE estimators for the conditional binomial distribution as discussed in Section 4.2.1.

Figure 5-4 shows how the in-control p-values for the conditional binomial distribution with negative binomial inputs are distributed if n=7 for the eight cases. We observed that these in-control p-values seem to be far from the uniform (0,1) with heavier tails on the Q-Q plots, compared to the Q-Q plots for the p-values based on the conditional negative binomial distribution.

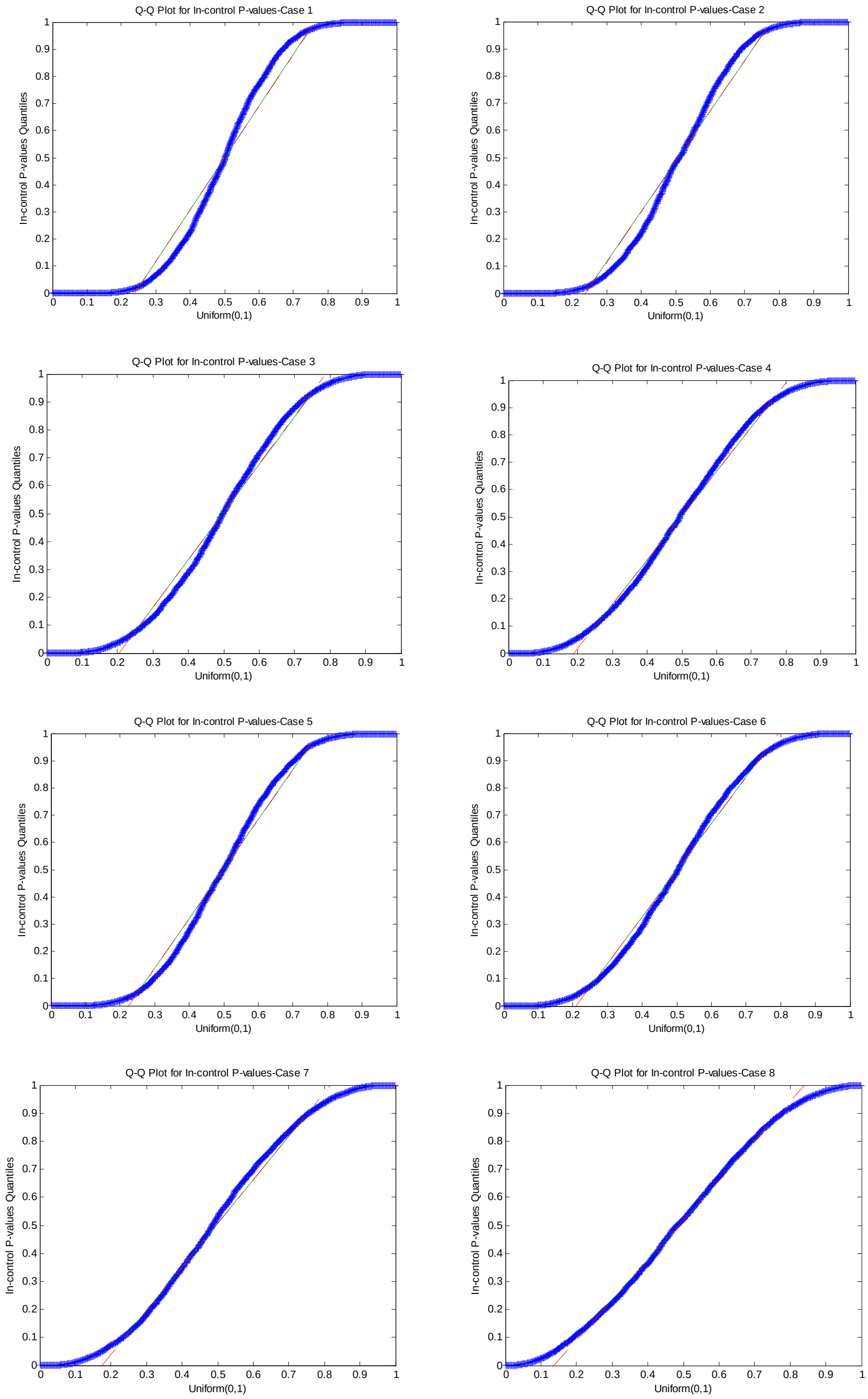

Figure 5-4: Q-Q Plots for In-control P-values for Adaptive Threshold Method-Conditional Binomial Distribution with Negative Binomial Inputs, *n*=7

5.2.2 W2r and Modified W2r Methods

We propose a modified W2r method in this section, W2r_2. The surveillance statistics of the W2r _2 method are defined as

$$W2r\_2(t) = \frac{x_{1t}-\hat{\mu}_t}{max\,(1,\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{it}-\hat{\mu}_{it})^2})}, \qquad (5.1)$$

where y_{it}, i=1,2,.., n, correspond to the eligible baseline data counts for day t, and n is the baseline window size, i.e., n=7, 14, or 28 days in our study. The estimator $\hat{\mu}_t$ is the average count from the baseline period, and $\hat{\mu}_{it}$ refers to the estimated mean count for day i in the baseline period, i=1,2,.., n. Similar to the mean absolute residual (MAR_t) for the W2r statistic in Eq. (2.4), the value of the denominator in Eq. (5.1) can never be less than one.

As we stated in Section 2.2, the definition of MAR_t does not reflect the total number of counts or visits at time t. The square root of the mean square deviation in the denominator defined in Eq. (5.1) doesn't do so either. The motivation behind the use of this estimator is an effort to obtain a better estimator of the process variance. Note that for the W2r and W2r_2 methods, we always use the estimator $\hat{\mu}_t = d_t \frac{\mathrm{BLS_t}}{\mathrm{BLV_t}}$.

5.3 RI Threshold Function Analysis

An empiric recurrence interval (RI), which is defined as the fixed number of time periods for which the expected number of false alarms is one, is used as one of the performance measures. In this section the RI threshold functions are compared for the adaptive threshold, the W2r, and the modified W2r methods. Again, if a single threshold is to be used for a given method, it is important that the non-outbreak performance not vary too much depending on the characteristics of the input data streams.

The first methods compared were Shewhart-based methods, and the second methods compared were the EWMA-based methods with a smoothing constant of α=0.2. The *RI-Shewhart* algorithm described in Section 4.3 was used to estimate the RI values for the Shewhart case. Again the algorithm we used to estimate the RI values for EWMA case was similar to the *RI-Shewhart* algorithm. The only difference is that in step 1, we calculated the EWMA of the statistic values for W2r, W2r_2, or Z-score with α=0.2.

5.3.1 Comparison of W2r Method and Adaptive Threshold Method with Conditional Binomial Distribution

For comparisons to the BioSense W2r RI functions, we used negative binomial inputs for the Shewhart-based methods. Figure 5-5 shows the RI threshold functions for the W2r method, compared to the BioSense values. The threshold functions for the W2r under our model and the BioSense function are quite close to one another, as shown in Figure 5-5.

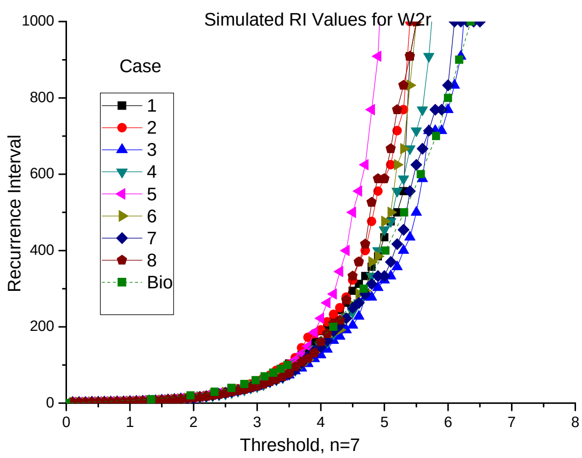

Figure 5-5: Thresholds Curves Based on RIs: Counts based on Negative Binomial Inputs

Figure 5-6 shows the RI threshold functions for the adaptive threshold Shewhart-based method based on conditional binomial distribution and the W2r method for negative binomial inputs when different baselines are used. We observed that the RI curves for the adaptive threshold method are less variable than those for the W2r method when *n*=7. As *n* increases, the Z-score threshold functions show almost the same robustness to the parameter changes, but the W2r threshold functions become more robust to the parameter changes.

Figure 5-7 shows the RI threshold functions for the adaptive threshold EWMA-based method and the W2r EWMA-based method for negative binomial inputs. We observed that RI threshold curves for the adaptive threshold method are more variable than those for the W2r method when *n*=7. As *n* increases, the Z-score threshold values show almost the same lack of robustness to the parameter changes, while the W2r threshold functions become more robust to the parameter changes. This illustrates, unfortunately, that the use of the conditional binomial distribution with the adaptive threshold method does not lead to good performance.

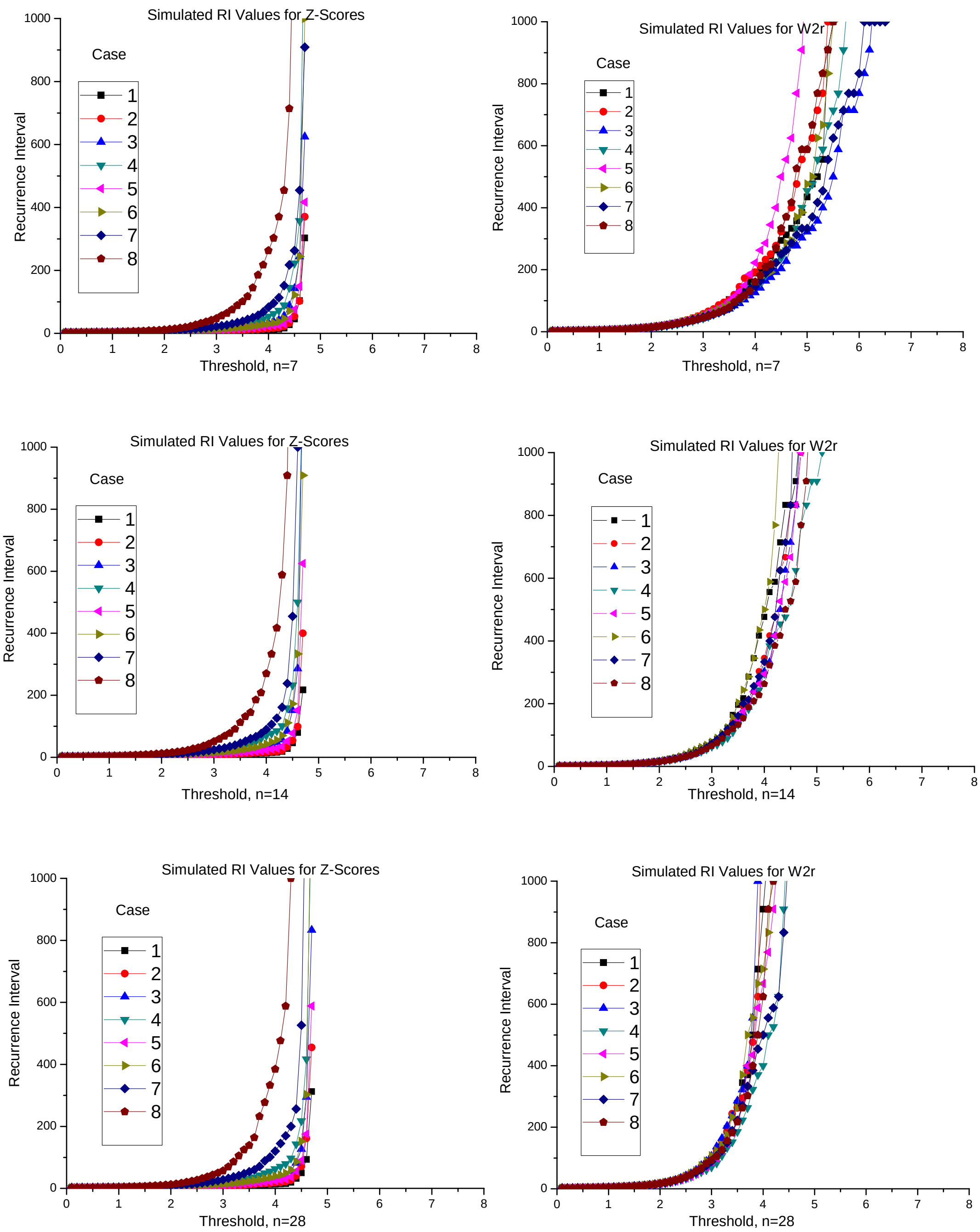

Figure 5-6: **RI Thresholds for Adaptive Threshold and W2r Methods for Different Baselines-Conditional Binomial Assumption with Negative Binomial Inputs-Shewhart**

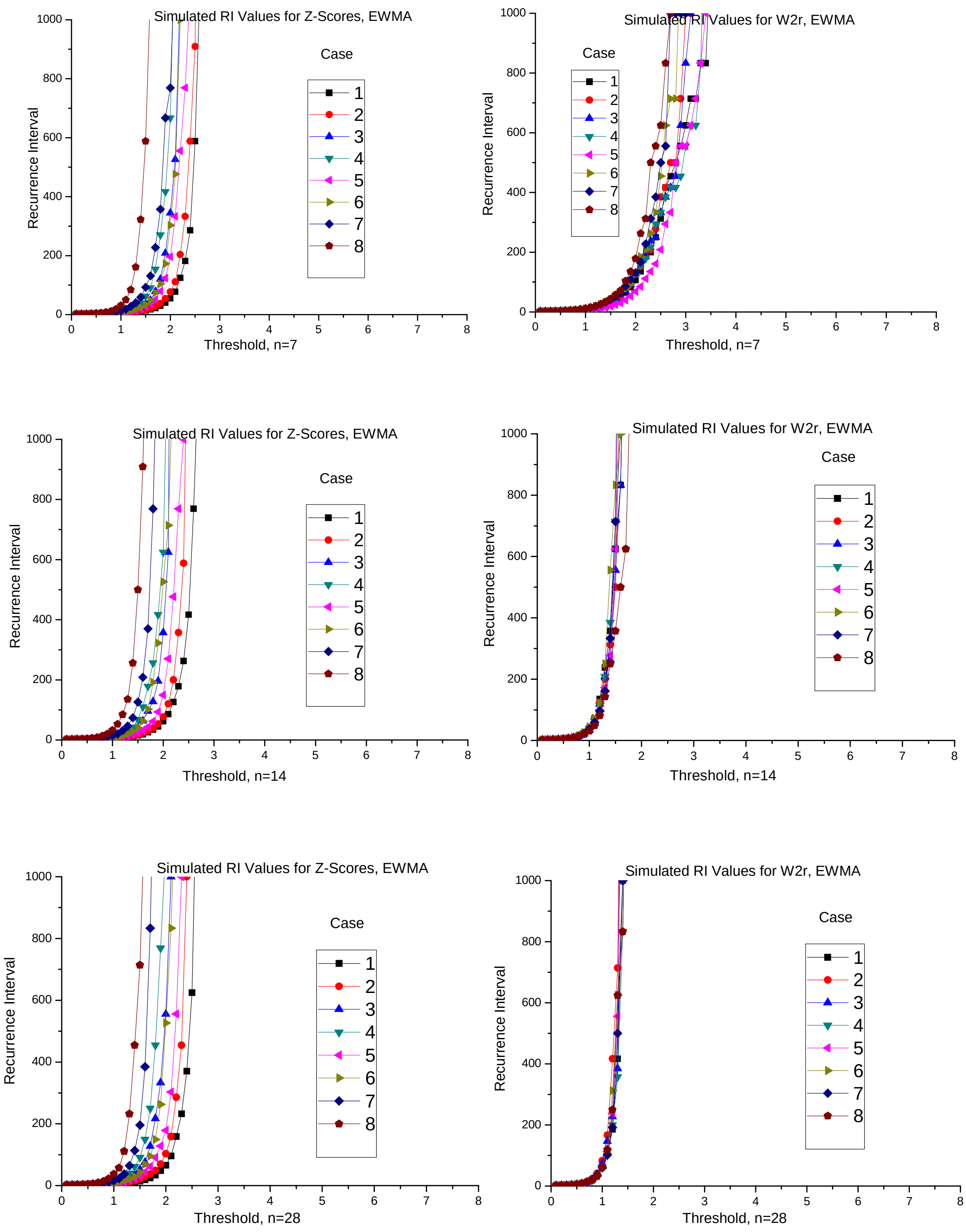

Figure 5-7: **RI Thresholds for Adaptive Threshold and W2r Methods for Different Baselines-Conditional Binomial Assumption with Negative Binomial Inputs-EWMA**

5.3.2 Comparison of W2r Method and Adaptive Threshold Method based on Conditional Negative Binomial Distribution

We first examined the adaptive threshold method assuming known parameters for the conditional negative binomial distribution. The RI threshold functions for the adaptive threshold method were compared to those for the W2r method, as shown in Figures 5-8 and 5-9 for the Shewhart and EWMA-based methods, respectively. Considering the set of parameters and a given window of historical data, the Z-score threshold functions are considerably less variable than those of the W2r method as the parameters change. It appears that the adaptive threshold method outperforms the W2r method by giving more consistent results across the parameter space and baseline window lengths for the conditional negative binomial distribution, as long as we assume that the negative binomial parameters are known.

We further examined the adaptive threshold method using MOM estimators for the conditional negative binomial distribution. Figure 5-10 (left) shows the RI threshold functions for the adaptive threshold method for the Shewhart case, and Figure 5-10 (right) shows the RI threshold functions for the EWMA case. We observed that, the adaptive threshold method threshold functions using the MOM estimators are close to the Z-score threshold functions assuming known parameters shown in Figure 5-8 (left) and Figure 5-9 (left) for both control chart approaches. This result indicates that the MOM estimators for the conditional negative binomial distribution work well for the adaptive threshold method. It is therefore clearly seen that the adaptive threshold method using MOM estimators for the conditional negative binomial distribution also tends to outperform the W2r method by giving more consistent threshold results across the parameter space for each baseline window length.

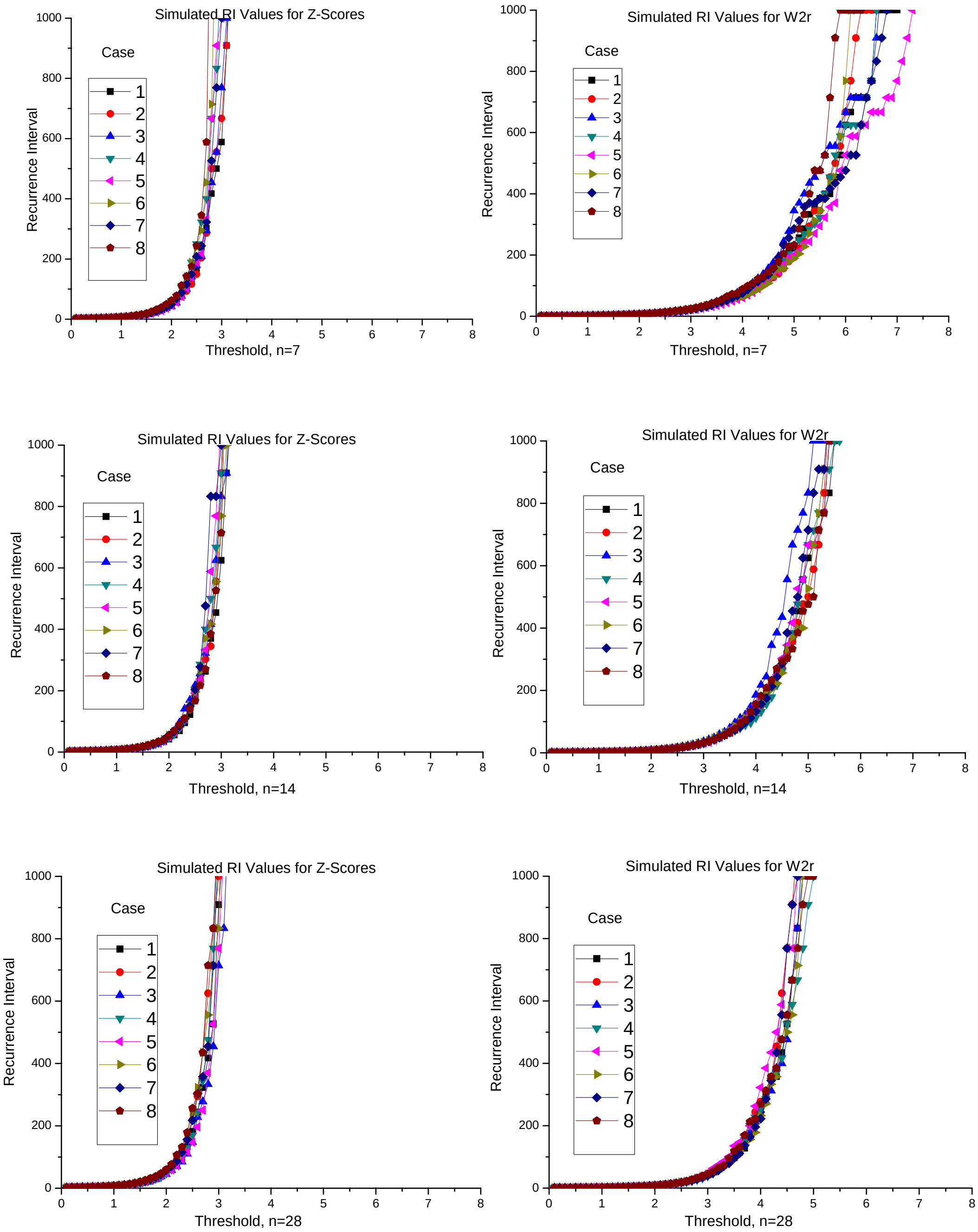

Figure 5-8: RI Thresholds for Adaptive Threshold Method Using Known Parameters and W2r Method for Different Baselines-Conditional Negative Binomial Distribution-Shewhart

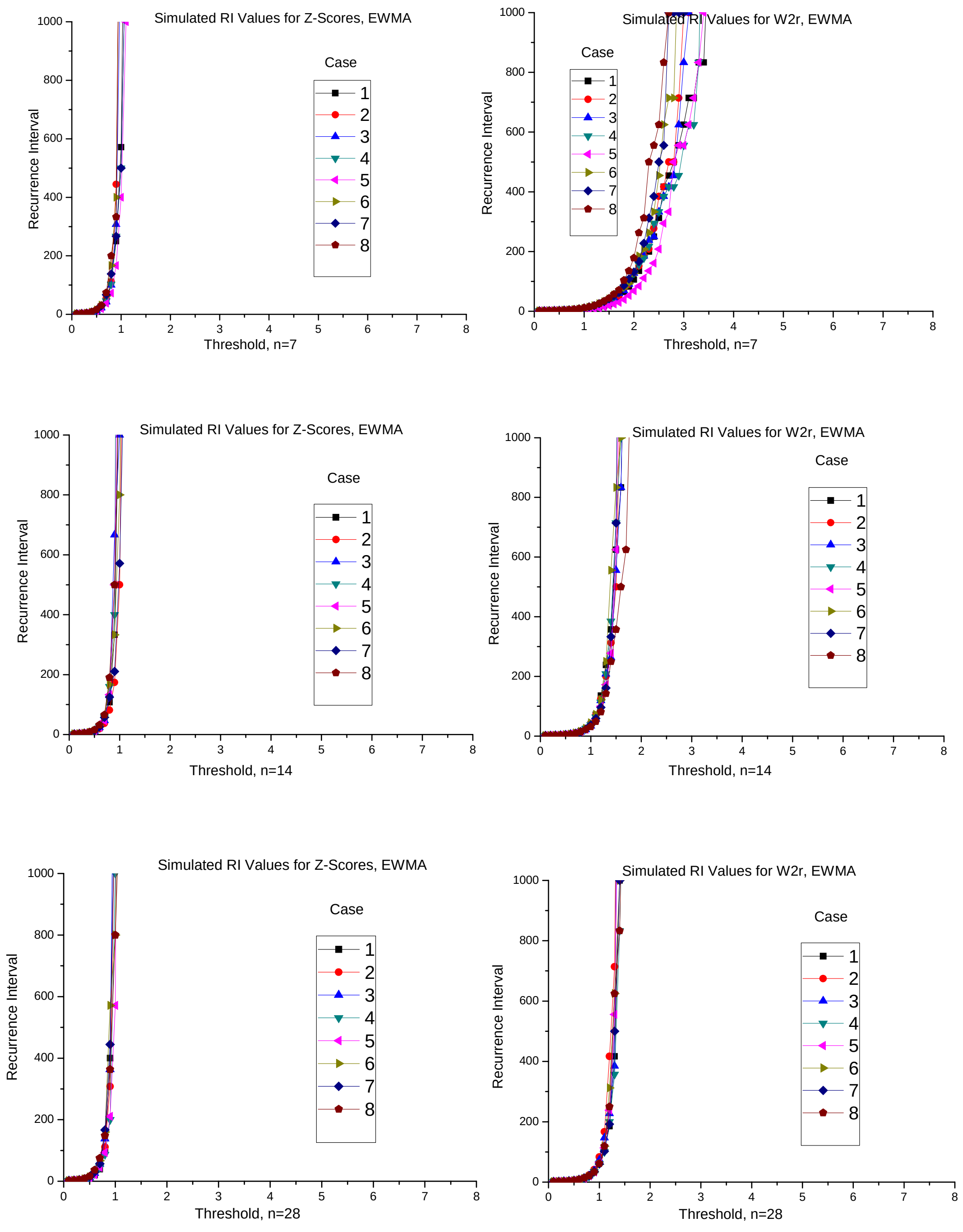

Figure 5-9: RI Thresholds for Adaptive Threshold Method Using Known Parameters and W2r Method for Different Baselines-Conditional Negative Binomial Distribution-EWMA

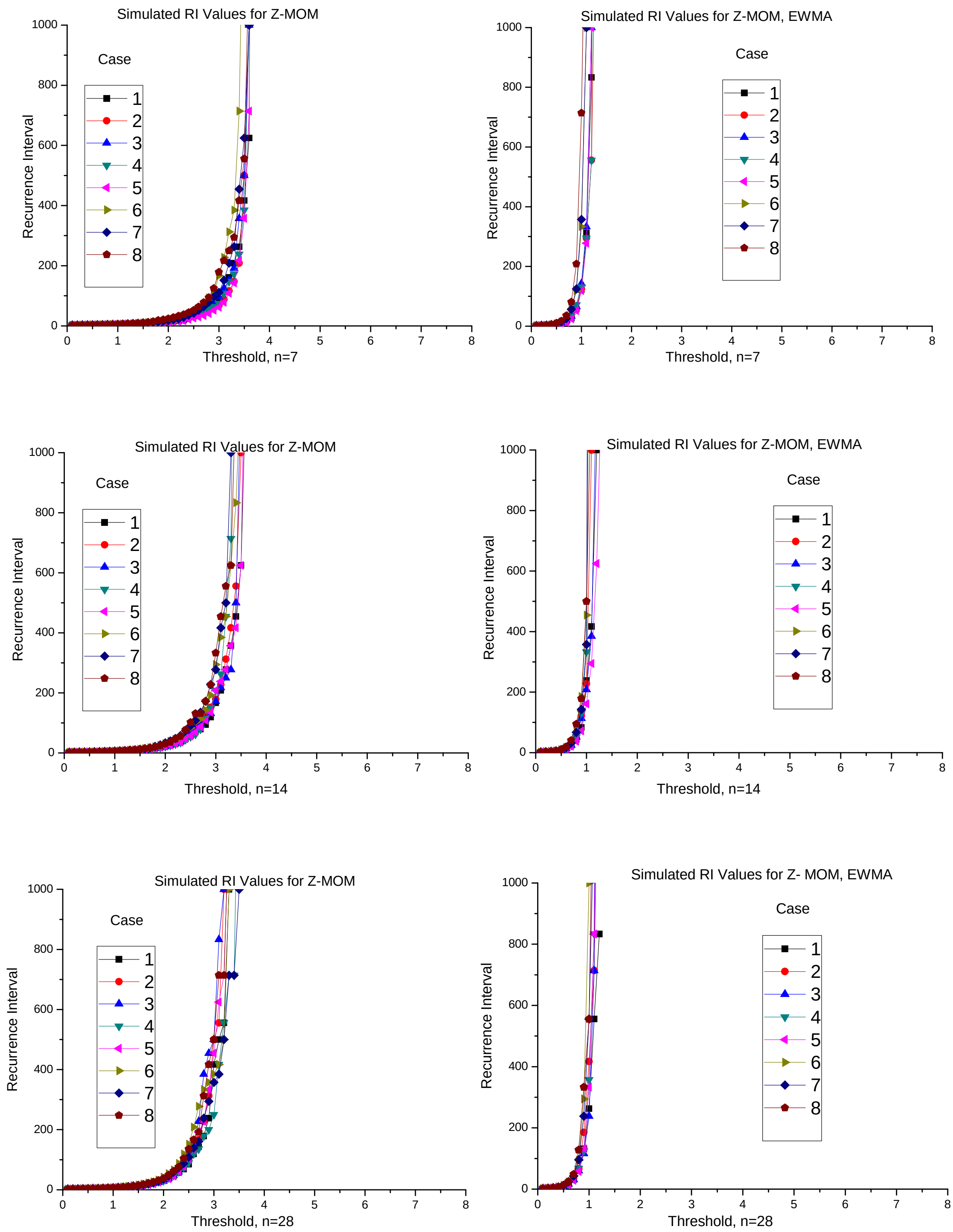

Figure 5-10: RI Thresholds for Adaptive Threshold Method Using MOM for Different Baselines - Conditional Negative Binomial Distribution-Shewhart (left) and EWMA (right) Methods

5.3.3 Comparison of W2r and Modified W2r Methods

We examined the performance of the W2r_2 method for the negative binomial inputs. Figure 5-11 (left) shows the RI threshold functions for this method for the Shewhart case and Figure 5-11 (right) shows the RI threshold functions for the EWMA case across different baseline assumptions. The W2r_2 method was then compared to the W2r method as shown in Figure 5-8 (right) and Figure 5-9 (right) for the Shewhart and EWMA approaches, respectively. We observed that, considering a given window of historical data, the threshold functions of the W2r_2 method are less variable than those of the W2r method as the parameters change in either the EWMA or the Shewhart comparisons. It appears that the W2r_2 method outperforms the W2 rate method by giving more consistent threshold results across the parameter space for each of the baseline window lengths we used in the simulation study.

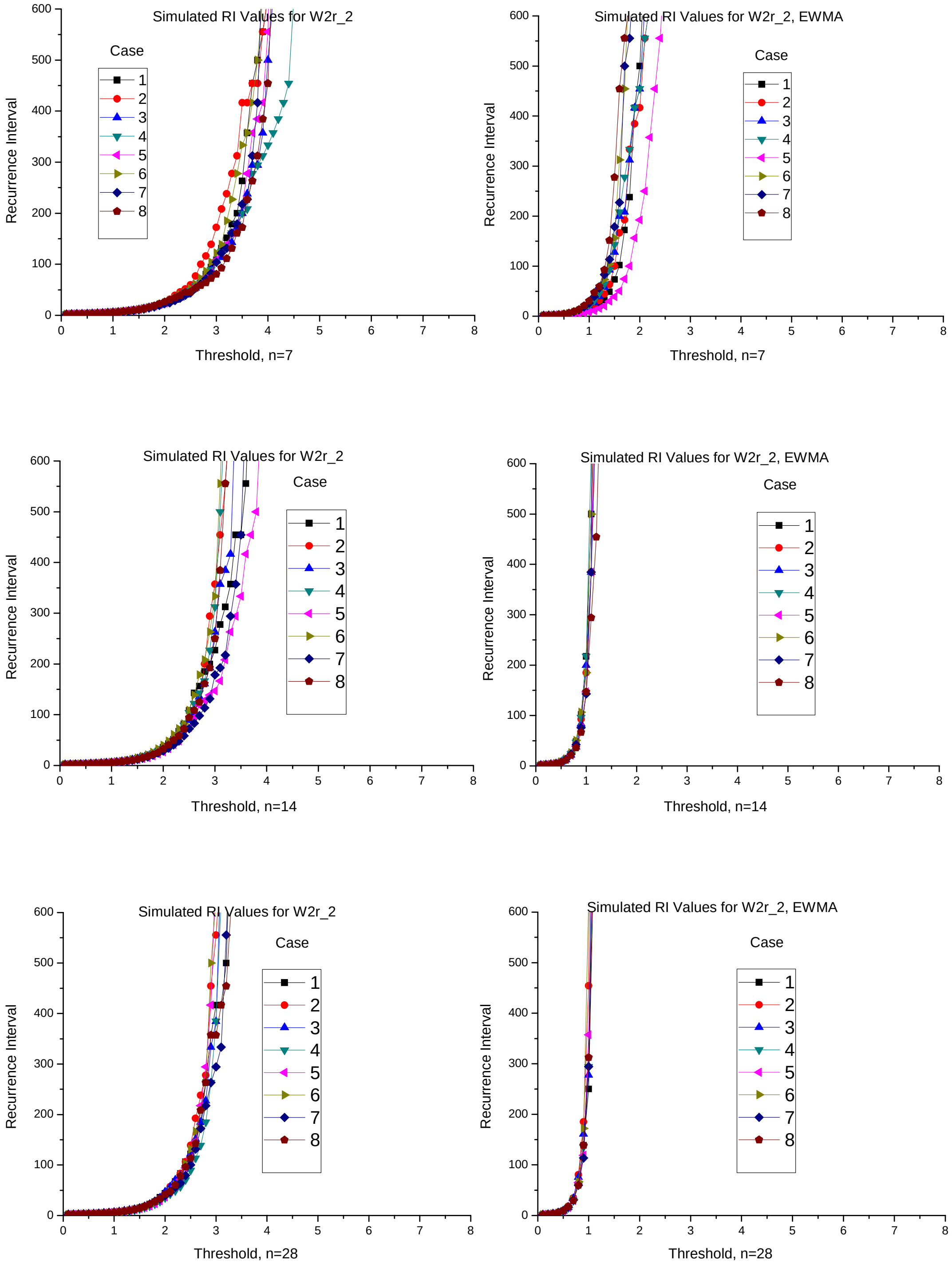

Figure 5-11: RI Thresholds for W2r_2 Method for Different Baselines-Conditional Negative Binomial Distribution-Shewhart (left) and EWMA (right) Methods

5.4 Power Analysis

We examined the power to detect syndromic outbreaks for the adaptive threshold and W2r methods. Similar to the power analysis in Section 4.4, we have all the methods set to have the same in-control false alarm rate in order to make fair comparisons. We used a RI of 500 days in order to expect one false alarm every 500 days when no outbreak is present.

The first methods compared were the Shewhart-based methods, and the second methods compared were the EWMA-based methods with α=0.2. For both approaches we again used the first ten weeks with the in-control data. The outbreak is then assumed to occur and be one week in length. We consider a percentage shift of size δ in the negative binomial parameters r_1 and r_3 for this study which will increase the average amount of syndromic counts per weekday and weekend day over the course of the outbreak, respectively. The values used are δ= 0, .1, .2, .5, 1 and 2, where a zero shift indicates analysis when there is no outbreak. The increase in the negative binomial parameter r_1 or r_3 is assumed to be ($\delta * 100$)% of its in-control value.

5.4.1 Shewhart-based Methods

We used the *Power-Shewhart* algorithm discussed in Section 4.4.1 to compute the power values in this section. Table 5-4 shows the threshold values we used in the power analysis. In Table 5-4, we denote by Z_CB the adaptive threshold method for the conditional binomial distribution, denote by Z_NB the adaptive threshold method assuming parameters are known for the conditional negative binomial distribution, and denote by Z_MOM the adaptive threshold method using MOM estimators for the conditional negative binomial distribution.

Table 5-4: Threshold Values of Adaptive Threshold and W2r Methods with Negative Binomial Input-Shewhart

	n	Case							
		1	2	3	4	5	6	7	8
W2r	7	5.04	4.82	5.02	4.92	5.26	5.01	5.19	5.22
	14	4.12	4.03	4.31	4.49	4.21	4.26	4.27	4.38
	28	3.75	3.72	3.84	3.93	3.98	3.81	3.82	3.80
W2r_2	7	4.07	3.90	4.07	4.03	4.32	4.09	4.22	4.19
	14	3.28	3.22	3.39	3.56	3.33	3.37	3.42	3.51
	28	2.98	2.96	3.05	3.17	3.17	3.04	3.06	3.02
Z_CB	7	7.40	7.20	6.10	5.60	6.74	6.10	5.20	4.10
	14	7.40	7.10	6.00	5.62	6.30	6.40	4.80	4.11
	28	7.30	7.00	5.82	5.22	6.71	5.92	4.90	3.96
Z_NB	7	2.80	2.83	2.81	2.84	2.84	2.79	2.76	2.75
	14	2.80	2.83	2.81	2.84	2.84	2.79	2.76	2.75
	28	2.80	2.83	2.81	2.84	2.84	2.79	2.76	2.75
Z_MOM	7	3.21	3.23	3.29	3.13	3.27	3.23	3.11	2.98
	14	3.15	3.03	3.10	3.10	3.10	3.04	3.07	2.91
	28	3.12	2.98	2.90	2.96	2.99	2.83	2.84	2.63

5.4.1.1 Comparison of W2r Method and Adaptive Threshold Method based on the Conditional Binomial Distribution

The W2r method was compared to the adaptive threshold method based on the conditional binomial distribution. Tables 5-5 and 5-6 show the power values for the two methods for the Shewhart approach. We first examined δ=0, where a zero shift indicates there is no outbreak, to ensure that the proportion of signals should be approximately the same across underlying parameter values and baseline lengths for the in-control case. We expect the probability of a signal for the Shewhart case with δ=0 and RI=500 to be around .014 as explained in Section 4.4.1. Tables 5-5 and 5-6 show that for a given shift δ in a given negative binomial case, the power values increase when n increases.

Table 5-5: Power Analysis for Adaptive Threshold Method-Transient Shift in Conditional Binomial Distribution with Negative Binomial Input-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0223	0.0163	0.0133	0.0100	0.0127	0.0173	0.0127	0.0173
	14	0.0127	0.0153	0.0147	0.0123	0.0183	0.0060	0.0207	0.0123
	28	0.0167	0.0157	0.0147	0.0173	0.0113	0.0120	0.0127	0.0180
.1	7	0.0632	0.0710	0.0464	0.0652	0.0619	0.0522	0.0574	0.0621
	14	0.0950	0.0894	0.0606	0.0912	0.0751	0.0696	0.0750	0.0686
	28	0.0790	0.0919	0.0913	0.0913	0.0854	0.0703	0.0857	0.0930
.2	7	0.1749	0.1803	0.1773	0.1670	0.3729	0.1621	0.1789	0.1585
	14	0.1850	0.1887	0.2088	0.2201	0.4046	0.1894	0.2192	0.1760
	28	0.2481	0.2454	0.2561	0.2559	0.5336	0.2494	0.2469	0.2619
.5	7	0.6613	0.7597	0.9207	0.9563	0.9987	0.7337	0.8693	0.8817
	14	0.7340	0.8843	0.9730	0.9750	0.9997	0.7373	0.9600	0.9180
	28	0.8170	0.9390	0.9890	0.9967	1.0000	0.8680	0.9687	0.9700
1	7	0.9660	0.9943	1.0000	1.0000	1.0000	0.9993	1.0000	1.0000
	14	0.9873	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	0.9963	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	0.9887	0.9980	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	0.9923	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	0.9960	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-6: Power Analysis for W2r with Negative Binomial Inputs-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0165	0.0133	0.0113	0.0125	0.0125	0.0109	0.0173	0.0167
	14	0.0160	0.0195	0.0135	0.0128	0.0160	0.0183	0.0103	0.0165
	28	0.0155	0.0188	0.0185	0.0115	0.0130	0.0165	0.0193	0.0138
.1	7	0.0406	0.0467	0.0348	0.0405	0.0327	0.0333	0.0400	0.0492
	14	0.0446	0.0526	0.0484	0.0578	0.0424	0.0502	0.0539	0.0577
	28	0.0724	0.0607	0.0781	0.0541	0.0515	0.0788	0.0551	0.0660
.2	7	0.0883	0.0980	0.0831	0.0774	0.0772	0.0944	0.0884	0.0882
	14	0.1799	0.1641	0.1628	0.1822	0.1487	0.1698	0.1717	0.1566
	28	0.2441	0.2178	0.2182	0.2420	0.2537	0.2891	0.2273	0.2436
.5	7	0.4187	0.4280	0.4660	0.5490	0.6880	0.3820	0.5335	0.3003
	14	0.7047	0.7795	0.8075	0.8413	0.8905	0.6905	0.8040	0.5642
	28	0.7255	0.9273	0.9692	0.9477	0.9138	0.8473	0.9372	0.8925
1	7	0.9353	0.9517	0.9715	0.9858	0.9575	0.9275	0.9785	0.8960
	14	0.9983	0.9998	0.9998	1.0000	1.0000	0.9970	1.0000	0.9968
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The W2r method was compared in another way to the adaptive threshold methods based on the conditional binomial distribution for the Shewhart approach. Table 5-7 shows the percentage increases of the power values of the adaptive threshold method compared to the W2r method. The values in Table 5-7 were computed by using the formula for a percentage increase, i.e., (Power_adaptive threshold – Power_W2r)/ Power_W2r)*100%. Note that these power values are given in Tables 5-5 and 5-6. We observed that for a given shift δ and a given n the adaptive threshold method works better than the W2r method with a few exceptions. Figure 5-12 shows an example of the power analysis for the two methods for Case 2. This figure supports the results we reached from Table 5-7.

Table 5-7: Percentage Increase of the Power Values for Adaptive Threshold Method Compared to W2r Method -Transient Shift in Conditional Binomial Distribution with Negative Binomial Inputs-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	35.62	52.03	33.33	60.99	89.30	56.76	43.50	26.22
	14	113.00	69.96	25.21	57.79	77.12	38.65	39.15	18.89
	28	9.12	51.40	16.90	68.76	65.83	-10.79	55.54	40.91
.2	7	98.07	83.98	113.36	115.76	383.03	71.72	102.38	79.71
	14	2.83	14.99	28.26	20.80	172.09	11.54	27.66	12.39
	28	1.64	12.67	17.37	5.74	110.33	13.73	8.62	7.51
.5	7	57.94	77.50	97.58	74.19	45.16	92.07	62.94	193.61
	14	4.16	13.44	20.50	15.89	12.26	6.78	19.40	62.71
	28	12.61	1.26	2.04	5.17	9.43	2.44	3.36	8.68
1	7	3.28	4.48	2.93	1.44	4.44	7.74	2.20	11.61
	14	-1.10	0.02	0.02	0.00	0.00	0.30	0.00	0.32
	28	-0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	-1.13	-0.20	0.00	0.00	0.00	0.00	0.00	0.00
	14	-0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	-0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00

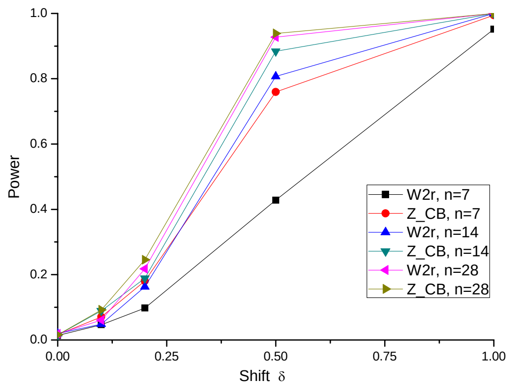

Figure 5-12: Power Analysis for Adaptive Threshold Method with Conditional Binomial Distribution and W2r Methods for Different Baselines-Transient Shift in Counts with Negative Binomial Inputs-Shewhart, RI=500 -Case 2

5.4.1.2 Comparison of W2r and Adaptive Threshold Methods with Conditional Negative Binomial Distribution

We examined the performance of the adaptive threshold methods based on the conditional negative binomial distribution with the Shewhart approach. Tables 5-8 and 5-9 show the power values for the adaptive threshold method assuming the negative binomial parameters are known and using MOM estimators of the parameters for the Shewhart case, respectively. We first examined whether the probabilities of signals for the Shewhart case with δ=0 were all around .014 in the two tables. Tables 5-8 and 5-9 show that for a given shift δ in a given case, when n increases the power values increase. Note that the power values of the adaptive threshold method assuming parameters are known are only of interest in seeing how well the baseline with n=28 works. We observe that the power values of the adaptive threshold method assuming known parameters with n=14 is at least as high as those of the MOM-based adaptive threshold method with n=28 for the eight cases.

Table 5-8: Power Analysis for Adaptive Threshold Method Assuming Parameters Known - Transient Shift in Conditional Negative Binomial Distribution-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0123	0.0120	0.0120	0.0105	0.0145	0.0210	0.0140	0.0103
	14	0.0145	0.0130	0.0123	0.0103	0.0118	0.0142	0.0183	0.0180
	28	0.0118	0.0140	0.0155	0.0185	0.0188	0.0182	0.0122	0.0165
.1	7	0.0895	0.0784	0.0874	0.0792	0.0950	0.0847	0.0906	0.0867
	14	0.1221	0.1158	0.1234	0.1252	0.1334	0.1286	0.1150	0.1141
	28	0.1708	0.1514	0.1677	0.1610	0.1715	0.1443	0.1584	0.1503
.2	7	0.3162	0.2974	0.3024	0.2805	0.2858	0.2875	0.2999	0.2875
	14	0.3698	0.3313	0.3550	0.3485	0.3379	0.3671	0.3639	0.3605
	28	0.5147	0.5614	0.5467	0.5343	0.5404	0.5350	0.5067	0.5346
.5	7	0.9983	1.0000	1.0000	0.9945	1.0000	1.0000	1.0000	0.9377
	14	0.9995	1.0000	1.0000	0.9920	1.0000	1.0000	1.0000	0.9535
	28	0.9998	1.0000	1.0000	0.9968	1.0000	1.0000	1.0000	0.9757
1	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-9: Power Analysis for Adaptive Threshold Method Using MOM Estimators-Transient Shift in Conditional Negative Binomial Distribution-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0132	0.0132	0.0136	0.0189	0.0175	0.0193	0.0171	0.0232
	14	0.0179	0.0121	0.0125	0.0107	0.0132	0.0143	0.0257	0.0364
	28	0.0189	0.0125	0.0189	0.0207	0.0161	0.0100	0.0175	0.0129
.1	7	0.0550	0.0541	0.0531	0.0568	0.0480	0.0658	0.0631	0.0599
	14	0.0798	0.0744	0.0871	0.0791	0.0750	0.0895	0.0784	0.0759
	28	0.1176	0.1132	0.1110	0.1202	0.1332	0.1225	0.1099	0.1131
.2	7	0.1984	0.2071	0.1880	0.1790	0.1870	0.2030	0.2035	0.2041
	14	0.2582	0.2626	0.2533	0.2721	0.2409	0.2664	0.2481	0.2553
	28	0.3305	0.3307	0.2988	0.3440	0.3325	0.3187	0.3163	0.3159
.5	7	0.6450	0.7400	0.7250	0.6850	0.7800	0.7650	0.7501	0.5250
	14	0.7674	0.8450	0.8150	0.8450	0.8950	0.7900	0.8523	0.6100
	28	0.7800	0.9350	0.9710	0.9490	0.9200	0.8600	0.9456	0.8950
1	7	0.9600	0.9850	0.9750	0.9550	0.9509	0.9950	0.9900	0.9220
	14	0.9700	1.0000	0.9800	0.9700	0.9700	1.0000	0.9950	0.9450
	28	0.9900	1.0000	1.0000	0.9850	1.0000	1.0000	1.0000	0.9850
2	7	1.0000	1.0000	1.0000	0.9600	1.0000	1.0000	1.0000	0.9965
	14	1.0000	1.0000	1.0000	0.9600	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The W2r method was compared to the adaptive threshold methods using MOM estimators based on the conditional negative binomial distribution for the Shewhart approach. Table 5-10 shows the percentage increases of the power values of the adaptive threshold method compared to the W2r method. The values of the percentage increases in Table 5-10 are computed by using the formula, i.e., (Power_adaptive threshold – Power_W2r)/ Power_W2r)*100%. Note that these power values are given in Tables 5-6 and 5-9. We observed that for a given δ and baseline n the adaptive threshold method works better than the W2r method with only a few exceptions. When δ = 1.0 in Case 1 with n=14 and n=28, Case 3 with n=14, Cases 4-5, and Case 7 with n=14, and Case 8 with n=14 and n=28 only, the power values of the W2r method are slightly higher than those of the adaptive threshold method. When δ = 2.0 with a given n, the power values of the W2r method are at least as high as those of the adaptive threshold method. Figures 5-22 to 5-25 show the power values of these methods for the eight cases. These figures support the conclusions we reached from Table 5-10.

Table 5-10: Percentage Increase of the Power Values for Adaptive Threshold Method Using MOM Estimators Compared to W2r Method -Transient Shift in Conditional Negative Binomial Distribution-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	0.84	0.74	1.83	1.63	1.53	3.25	2.31	1.07
	14	3.52	2.18	3.87	2.13	3.26	3.93	2.45	1.82
	28	4.52	5.25	3.29	6.61	8.17	4.37	5.48	4.71
.2	7	11.01	10.91	10.49	10.16	10.98	10.86	11.51	11.59
	14	7.83	9.85	9.05	8.99	9.22	9.66	7.64	9.87
	28	8.64	11.29	8.06	10.20	7.88	2.96	8.90	7.23
.5	7	22.63	41.20	25.90	13.60	9.20	38.30	21.66	22.47
	14	6.27	6.55	0.75	0.37	0.45	9.95	4.83	4.58
	28	5.45	0.77	0.18	0.13	0.62	1.27	0.84	0.25
1	7	2.47	3.33	0.35	-3.08	-0.66	6.75	1.15	2.60
	14	-2.83	0.02	-1.98	-3.00	-3.00	0.30	-0.50	-5.18
	28	-1.00	0.00	0.00	-1.50	0.00	0.00	0.00	-1.50
2	7	0.00	0.00	0.00	-4.00	0.00	0.00	0.00	-0.35
	14	0.00	0.00	0.00	-4.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

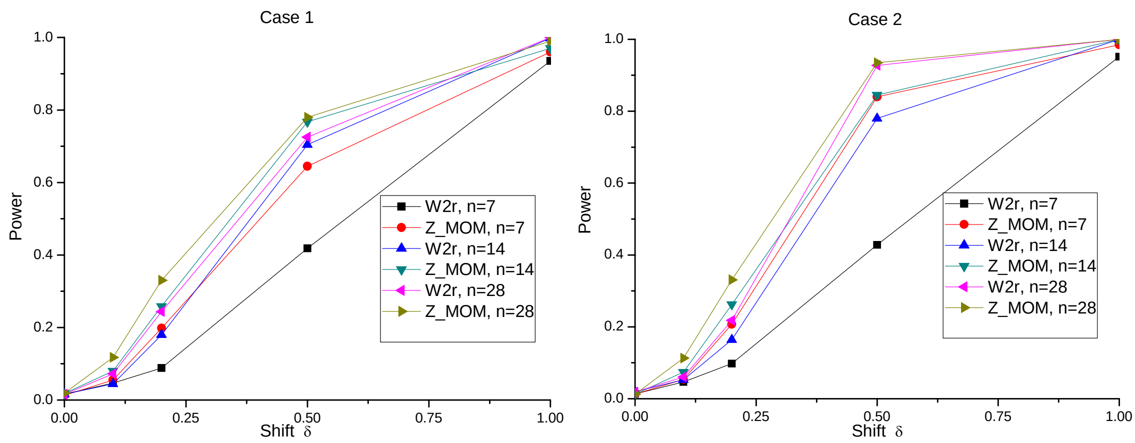

Figure 5-13: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-Shewhart, Case 1 and Case 2

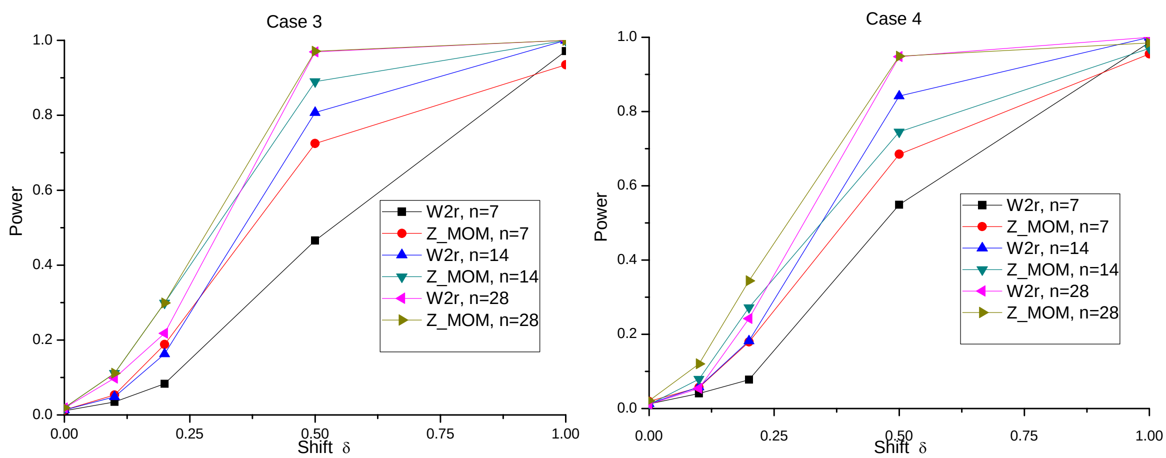

Figure 5-14: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-Shewhart, Case 3 and Case 4

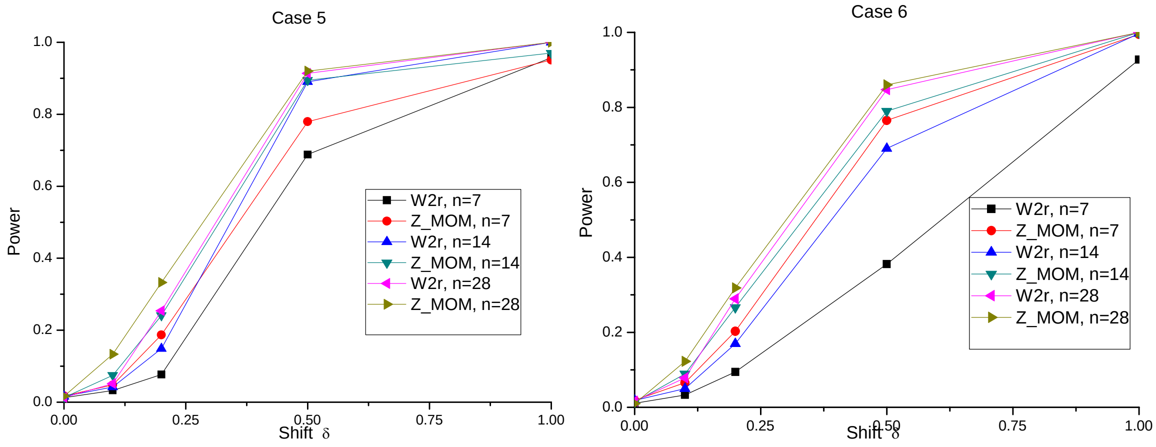

Figure 5-15: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-Shewhart, Case 5 and Case 6

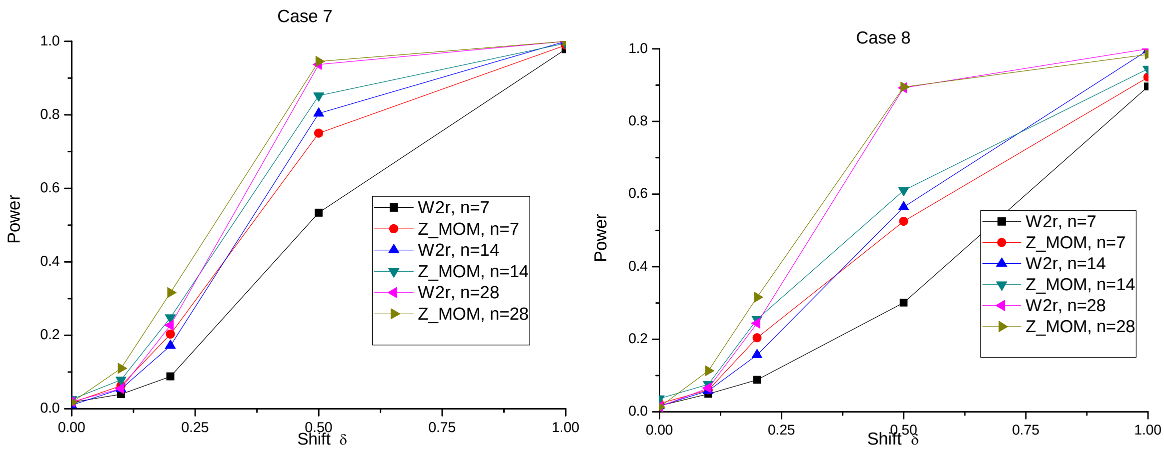

Figure 5-16: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-Shewhart, Case 7 and Case 8

5.4.1.3 Comparison of W2r and Modified W2r Methods

Table 5-11 gives the power values for the W2r_2 methods for the Shewhart case given RI =500. We observed that the probabilities of signals when δ=0 are all around .014. We also observed that for a given shift δ in a given negative binomial case, the power values increase as n increases.

Table 5-11: Power Analysis for W2r_2 Method with Negative Binomial Input-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0188	0.0173	0.0187	0.0160	0.0118	0.0125	0.0158	0.0128
	14	0.0115	0.0118	0.0107	0.0095	0.0170	0.0118	0.0115	0.0113
	28	0.0158	0.0105	0.0130	0.0110	0.0120	0.0105	0.0142	0.0185
.1	7	0.0485	0.0543	0.0690	0.0700	0.0680	0.0427	0.0508	0.0537
	14	0.0570	0.0565	0.0697	0.0780	0.1165	0.0522	0.0580	0.0635
	28	0.0743	0.0665	0.0912	0.0870	0.1185	0.0878	0.0857	0.0968
.2	7	0.1003	0.1422	0.1643	0.1987	0.2258	0.1003	0.1318	0.1248
	14	0.1893	0.1705	0.2248	0.2185	0.4078	0.1380	0.1880	0.1548
	28	0.2500	0.2265	0.3200	0.3277	0.4715	0.2925	0.2732	0.2903
.5	7	0.5633	0.6518	0.7628	0.8425	0.9150	0.5360	0.6485	0.6538
	14	0.7920	0.8518	0.9300	0.9307	0.9980	0.7495	0.8790	0.8510
	28	0.9350	0.9580	0.9852	0.9940	1.0000	0.9100	0.9748	0.9710
1	7	0.9872	0.9972	0.9990	0.9998	1.0000	0.9875	0.9963	0.9952
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-12 shows the percentage increases of the power values of the W2r_2 method compared to the W2r method. The values of the percentage increases in Table 5-12 are computed by using the formula, i.e., (Power_W2r_2 – Power_W2r)/ Power_W2r)*100%. Note that these power values are given in Tables 5-6 and 5-11. Table 5-12 shows that, for a given shift δ and baseline n in a given negative binomial case, the power values of the W2r_2 method are always as high as those of the W2r method with only a few exceptions. We observed that the power value of the W2r method is 3.18% higher than that of the W2r_2 method given δ =0.2 and n=14 days for Case 6. We also observed that the power value of the W2r method is 0.18% higher than that of the W2r_2 method given δ =0.2 and n= 14 days for Case 8. Figures 5-17 to 5-20 show how the W2r method is compared to the W2r_2 method for power analysis for the eight cases. These figures support the conclusions we obtained from Table 5-12. Clearly there is an improvement in going from the W2r method to the W2r_2 method for the Shewhart approach.

Table 5-12: Percentage Increase of the Power Values for W2r_2 Method Compared to W2r Method with Negative Binomial Input-Shewhart (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	0.19	0.76	3.42	2.95	3.53	0.94	1.08	0.45
	14	1.24	0.39	2.13	2.02	7.41	0.20	0.41	0.58
	28	2.62	0.58	1.31	3.29	6.70	0.90	3.06	3.08
.2	7	1.20	4.42	8.12	12.13	14.86	0.59	4.34	3.66
	14	0.94	0.64	6.20	3.63	25.91	-3.18	1.63	-0.18
	28	0.59	0.87	10.18	8.57	21.78	0.34	4.59	4.67
.5	7	14.46	22.38	29.68	29.35	22.70	15.40	11.50	35.35
	14	8.73	7.23	12.25	8.94	10.75	5.90	7.50	28.68
	28	20.95	3.07	1.60	4.63	8.62	6.27	3.76	7.85
1	7	5.19	4.55	2.75	1.40	4.25	6.00	1.78	9.92
	14	0.17	0.02	0.02	0.00	0.00	0.30	0.00	0.32
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

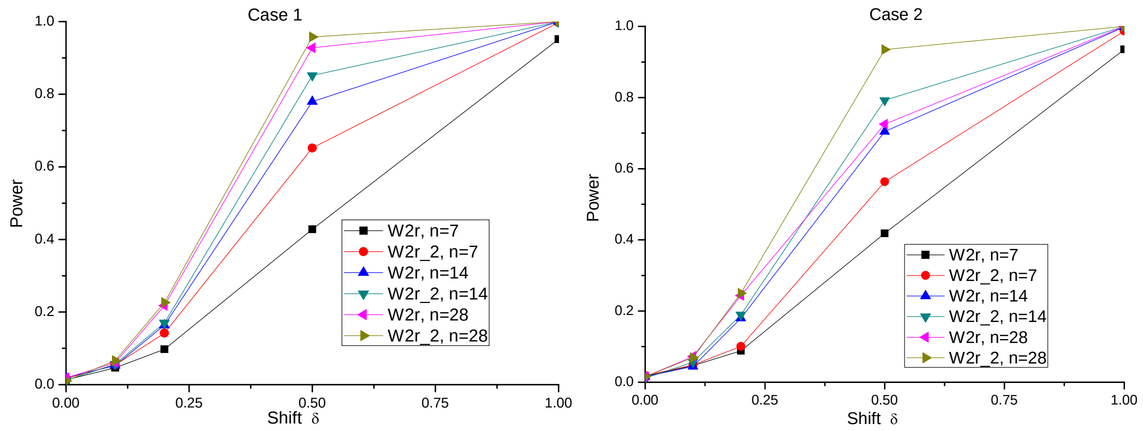

Figure 5-17: Power Analysis for W2r and W2r_2 Methods for Different Baselines- Transient Shift in Counts with Negative Binomial Inputs -Shewhart, Case 1 and Case 2

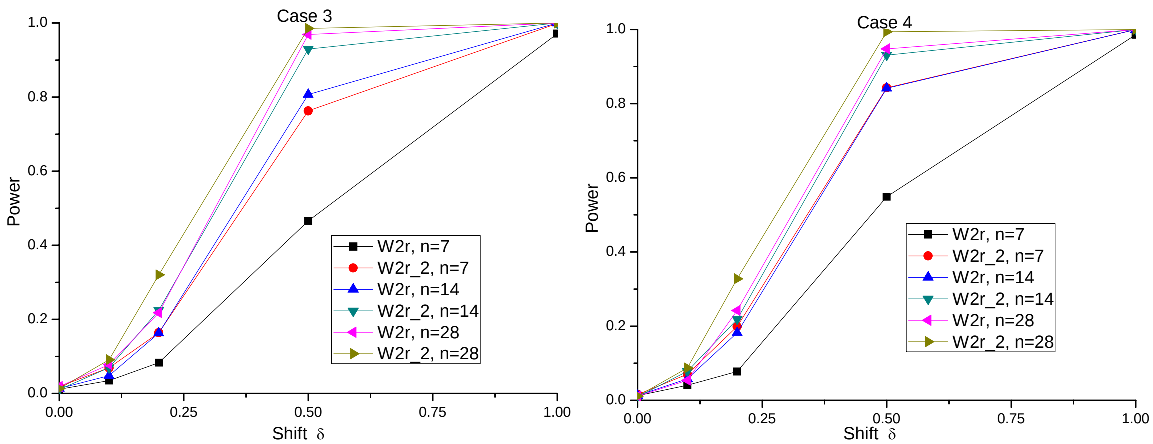

Figure 5-18: Power Analysis for W2r and W2r_2 Methods for Different Baselines- Transient Shift in Counts with Negative Binomial Inputs -Shewhart, Case 3 and Case 4

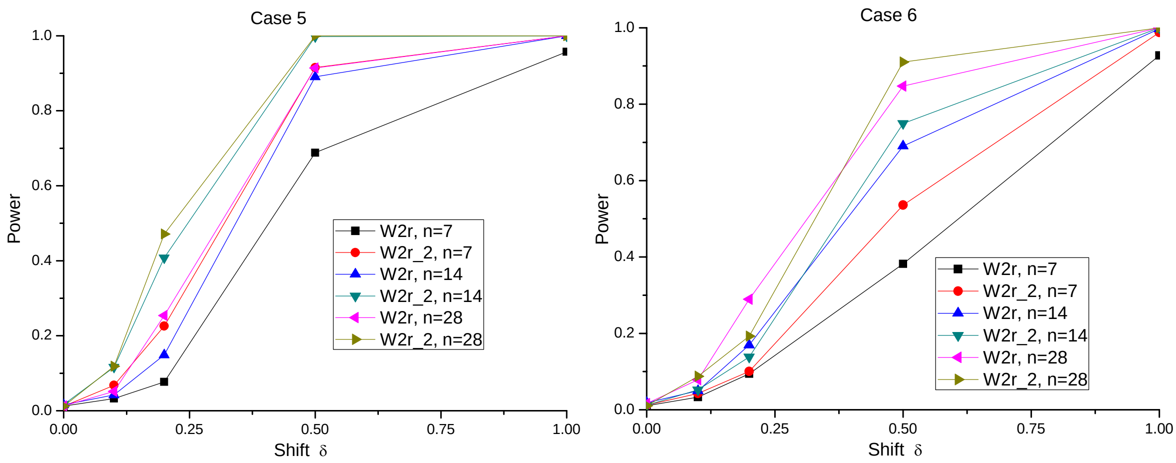

Figure 5-19: **Power Analysis for W2r and W2r_2 Methods for Different Baselines- Transient Shift in Counts with Negative Binomial Inputs -Shewhart, Case 5 and Case 6**

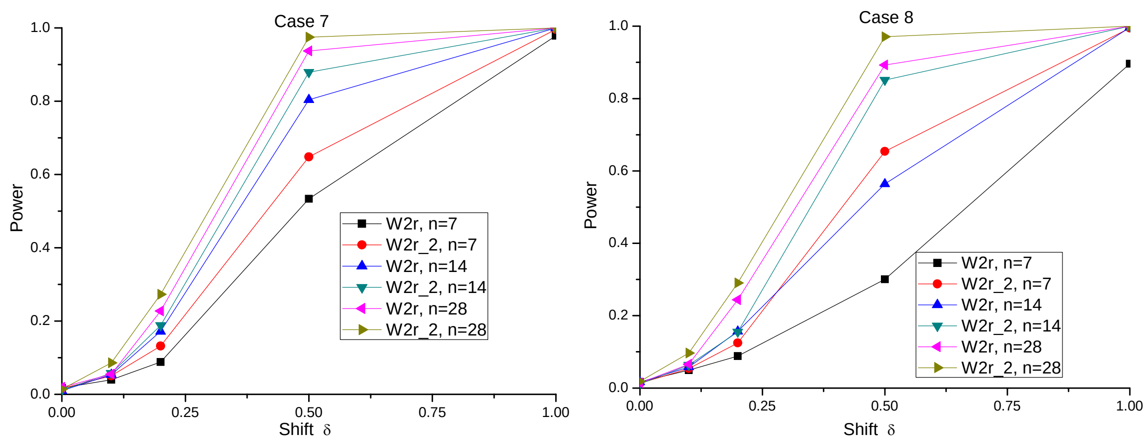

Figure 5-20: Power Analysis for W2r and W2r_2 Methods for Different Baselines- Transient Shift in Counts with Negative Binomial Inputs -Shewhart, Case 7 and Case 8

5.4.2 One-Sided EWMA-based Methods

The *Power-Shewhart* algorithm discussed in Section 4.4.1 was used to estimate the power values in this section. The only difference is that we calculated EWMA statistics based on the Z-scores, the W2r statistics, and the W2r_2 statistics with α=0.2 in step 1. Table 5-13 shows the threshold values we used in the power analysis.

Table 5-13: Threshold Values of Adaptive Threshold and W2r Methods with Negative Binomial Input-EWMA

	n	Case							
		1	2	3	4	5	6	7	8
W2r	7	2.80	2.90	2.94	2.64	3.54	2.60	2.60	2.51
	14	1.46	1.40	1.36	1.47	1.42	1.46	1.40	1.47
	28	1.22	1.27	1.30	1.30	1.42	1.24	1.34	1.30
W2r_ 2	7	2.08	1.92	1.95	1.97	2.22	1.88	1.84	1.80
	14	1.11	1.10	1.15	1.14	1.17	1.15	1.14	1.15
	28	1.03	1.00	1.05	1.04	1.04	1.05	1.05	1.04
Z_CB	7	3.60	3.30	2.80	2.20	3.10	2.70	1.91	1.62
	14	3.30	3.20	3.00	2.60	2.70	2.51	1.87	1.48
	28	3.50	2.80	2.70	2.50	2.90	2.60	2.10	1.40
Z_NB	7	0.95	0.95	0.95	0.94	0.96	0.93	0.93	0.90
	14	0.95	0.95	0.95	0.94	0.96	0.93	0.93	0.90
	28	0.95	0.95	0.95	0.94	0.96	0.93	0.93	0.90
Z_MOM	7	0.99	1.00	0.96	1.04	1.00	0.96	0.97	0.94
	14	1.01	0.99	0.94	0.98	1.00	0.93	0.92	0.89
	28	0.99	1.00	0.96	1.04	1.00	0.96	0.97	0.94

5.4.2.1 Comparison of W2r Method and Adaptive Threshold Method based on the Conditional Binomial Distribution

The W2r method was compared to the adaptive threshold method based on the conditional binomial distribution under the EWMA approach. Tables 5-14 and 5-15 give the power values for the two methods with the EWMA approach given RI=500. For a given δ in a given negative binomial case, the power values increase when n increases, as shown in Tables 5-14 and 5-15.

Table 5-14: Power Analysis for Adaptive Threshold Method-Transient Shift in Conditional Binomial Distribution with Negative Binomial Input-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0187	0.0137	0.0110	0.0090	0.0260	0.0090	0.0093	0.0097
	14	0.0128	0.0173	0.0093	0.0097	0.0260	0.0080	0.0093	0.0133
	28	0.0195	0.0283	0.0163	0.0093	0.0217	0.0100	0.0130	0.0137
.1	7	0.1410	0.0873	0.0603	0.0650	0.0692	0.0628	0.0587	0.0580
	14	0.1870	0.0937	0.0966	0.0921	0.0909	0.0936	0.0918	0.1077
	28	0.2097	0.1680	0.1357	0.1535	0.1330	0.1361	0.1343	0.1373
.2	7	0.2280	0.1877	0.2138	0.1978	0.3904	0.1821	0.2111	0.2620
	14	0.3460	0.2463	0.2878	0.2339	0.5863	0.1962	0.2955	0.4273
	28	0.3707	0.4337	0.4104	0.2939	0.6202	0.2624	0.4113	0.5470
.5	7	0.8613	0.9033	0.9713	0.9970	1.0000	0.8297	0.9903	0.9797
	14	0.9670	0.9663	0.9843	0.9973	1.0000	0.9570	0.9983	0.9977
	28	0.9760	0.9963	0.9990	0.9997	1.0000	0.9673	0.9973	1.0000
1	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-15: Power Analysis for W2r Method with Negative Binomial Input-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0107	0.0095	0.0155	0.0083	0.0092	0.0093	0.0085	0.0090
	14	0.0115	0.0090	0.0067	0.0118	0.0050	0.0130	0.0143	0.0095
	28	0.0095	0.0147	0.0163	0.0133	0.0083	0.0080	0.0095	0.0092
.1	7	0.0513	0.0485	0.0715	0.0578	0.0630	0.0435	0.0495	0.0377
	14	0.0745	0.0953	0.0850	0.1195	0.1313	0.0755	0.0874	0.0505
	28	0.0765	0.1260	0.1513	0.1655	0.2122	0.0848	0.1153	0.0978
.2	7	0.1467	0.1390	0.2418	0.2017	0.2712	0.1352	0.1729	0.1135
	14	0.2442	0.3320	0.3343	0.4615	0.5405	0.2678	0.3600	0.2225
	28	0.2460	0.4660	0.5965	0.6342	0.5784	0.3060	0.3869	0.4160
.5	7	0.7035	0.7525	0.9063	0.9063	0.9648	0.7090	0.7945	0.6885
	14	0.9535	0.9703	0.9880	0.9958	0.9705	0.9515	0.9875	0.9512
	28	0.9788	0.9963	0.9920	0.9998	0.9980	0.9865	0.9980	0.9980
1	7	0.9970	0.9985	0.9998	1.0000	1.0000	0.9980	1.0000	0.9970
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The W2r method was compared to the adaptive threshold methods based on the conditional binomial distribution for the EWMA approach. Table 5-16 shows the percentage increases of the power values of the adaptive threshold method compared to the W2r method. Note that these power values are given in Tables 5-14 and 5-15. We observed that for a given shift δ and baseline n the adaptive threshold method works better than the W2r method with only a few exceptions. Figure 5-21 shows an example of the power analysis for the two methods for Case 1.

We observed that the adaptive threshold method works better than the W2r method for Case 1. There is one exception, however, when δ =0.5 and n=28 the power value of the W2r method is 0.9788, whereas the power value of the adaptive threshold method is 0.9670 for Case 1.

Our power analysis results show that, under either the EWMA or the Shewhart approach, the adaptive threshold method based on the conditional binomial distribution outperformed the W2r method in most cases under the assumption of negative binomial inputs. Still, we would prefer to use the adaptive threshold method based on the conditional negative binomial approach since the RI threshold function performance is so much better.

Table 5-16: **Percentage Increase of the Power Values for Adaptive Threshold Method Compared to W2r Method -Transient Shift in Conditional Binomial Distribution with Negative Binomial Inputs-EWMA (RI=500)**

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	174.85	80.00	-15.66	12.46	9.84	44.37	18.59	53.85
	14	151.01	-1.68	13.65	-22.93	-30.77	23.97	5.03	113.27
	28	174.12	33.33	-10.31	-7.25	-37.32	60.50	16.48	40.39
.2	7	55.42	35.04	-11.58	-1.93	43.95	34.69	22.09	130.84
	14	41.69	-25.81	-13.91	-49.32	8.47	-26.74	-17.92	92.04
	28	50.69	-6.93	-31.20	-53.66	7.23	-14.25	6.31	31.49
.5	7	22.43	20.04	7.17	10.01	3.65	17.02	24.64	42.29
	14	1.42	-0.41	-0.37	0.15	3.04	0.58	1.09	4.89
	28	-0.29	0.00	0.71	-0.01	0.20	-1.95	-0.07	0.20
1	7	0.30	0.15	0.02	0.00	0.00	0.20	0.00	0.30
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

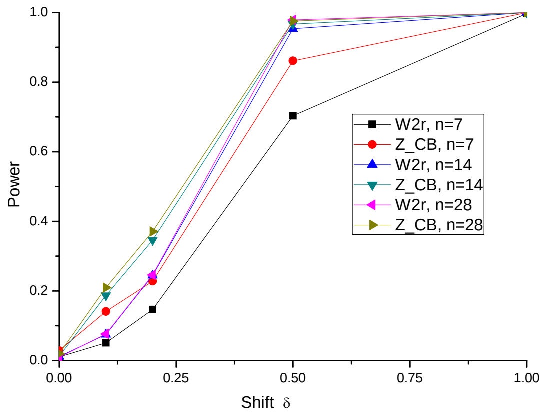

Figure 5-21: Power Analysis for Adaptive Threshold Method based on Conditional Binomial Distribution and W2r Method for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA, RI=500-Case 1

5.4.2.2 Comparison of W2r and Adaptive Threshold Methods based on the Conditional Negative Binomial Distribution

Tables 5-17 and 5-18 show the power values for the adaptive threshold method using known parameters and MOM estimators for the EWMA approach, respectively. We observed that for a given shift δ in a given negative binomial case, when n increases the power values increase. We also observe that the power values of the adaptive threshold method using known parameters with n=14 is at least as high as those of the MOM-based adaptive threshold method with n=28 for the eight cases.

Table 5-17: Power Analysis for Adaptive Threshold Method Assuming Parameters Known-Transient Shift in Conditional Negative Binomial Distribution-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0120	0.0097	0.0138	0.0142	0.0093	0.0100	0.0093	0.0193
	14	0.0097	0.0158	0.0052	0.0178	0.0160	0.0118	0.0217	0.0103
	28	0.0135	0.0155	0.0102	0.0085	0.0070	0.0178	0.0155	0.0147
.1	7	0.2215	0.2580	0.2068	0.2095	0.3053	0.3023	0.2000	0.2020
	14	0.2272	0.4010	0.2467	0.2362	0.3055	0.3045	0.3229	0.2548
	28	0.2275	0.4028	0.2488	0.2553	0.3192	0.3375	0.3672	0.2638
.2	7	0.8070	0.9327	0.8710	0.7848	0.8915	0.9098	0.8128	0.7027
	14	0.8677	0.9615	0.8745	0.7888	0.9085	0.9100	0.9177	0.7205
	28	0.8685	0.9678	0.8810	0.7955	0.9185	0.9233	0.9290	0.7347
.5	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-18: Power Analysis for Adaptive Threshold Method Using MOM Estimators - Transient Shift in Conditional Negative Binomial Distribution-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0504	0.0186	0.0304	0.0243	0.0329	0.0107	0.0193	0.0282
	14	0.0361	0.0246	0.0354	0.0186	0.0311	0.0164	0.0218	0.0214
	28	0.0157	0.0150	0.0236	0.0182	0.0189	0.0207	0.0236	0.0239
.1	7	0.1607	0.1336	0.1507	0.1300	0.1714	0.1046	0.1171	0.1150
	14	0.1671	0.2436	0.2386	0.1400	0.2400	0.2004	0.1839	0.1211
	28	0.1939	0.2593	0.2454	0.1682	0.2496	0.2511	0.2464	0.1554
.2	7	0.3575	0.3093	0.3150	0.2600	0.3261	0.3179	0.3229	0.2554
	14	0.4593	0.6107	0.5368	0.5121	0.5682	0.6114	0.5400	0.4854
	28	0.5179	0.7364	0.6414	0.6668	0.5804	0.7804	0.7200	0.5211
.5	7	0.7504	0.7626	0.9086	0.9077	0.9659	0.8236	0.8614	0.7929
	14	0.9575	0.9782	0.9893	0.9980	0.9980	0.9871	0.9879	0.9539
	28	0.9823	0.9969	0.9992	0.9998	0.9998	1.0000	0.9982	0.9986
1	7	0.9875	0.9970	0.9995	0.9998	0.9999	1.0000	0.9990	0.9989
	14	0.9910	1.0000	0.9998	1.0000	1.0000	1.0000	0.9998	0.9995
	28	0.9935	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The W2r method was compared to the adaptive threshold methods using MOM estimators based on the conditional negative binomial distribution for the EWMA approach. Table 5-19 shows the percentage increases of the power values of the adaptive threshold method compared to the W2r method. These power values are given in Tables 5-15 and 5-18. We observed that for a given shift δ and baseline n, the adaptive threshold method works better than the W2r method with only a few exceptions. For a given baseline and a given δ = 1.0 or 2.0 in Cases 1 to 3, Case 5, Case 7, and Case 8 (only with n=7), the power values of the W2r method are at least as high as those of the adaptive threshold method. Figures 5-22 to 5-25 show the power values of the surveillance methods for the eight cases. These figures support the conclusions we obtained from Table 5-19.

Table 5-19: Percentage Increase of the Power Values for Adaptive Threshold Method Using MOM Estimators Compared to W2r Method -Transient Shift in Conditional Negative Binomial Distribution-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	10.94	8.51	7.92	7.22	10.84	6.11	6.76	7.73
	14	9.26	14.83	15.36	2.05	10.87	12.49	9.65	7.06
	28	11.74	13.33	9.41	0.27	3.74	16.63	13.11	5.76
.2	7	21.08	17.03	7.32	5.83	5.49	18.27	15.00	14.19
	14	21.51	27.87	20.25	5.06	2.77	34.36	18.00	26.29
	28	27.19	27.04	4.49	3.26	0.20	47.44	33.31	10.51
.5	7	4.69	1.01	0.23	0.14	0.11	11.46	6.69	10.44
	14	0.40	0.79	0.13	0.22	2.75	3.56	0.04	0.27
	28	0.35	0.06	0.73	0.00	0.18	1.35	0.02	0.06
1	7	-0.95	-0.15	-0.03	0.02	-0.01	0.20	-0.10	0.19
	14	-0.90	-0.20	-0.02	0.00	0.00	0.00	-0.02	-0.05
	28	-0.65	-0.02	0.00	0.00	0.00	0.00	0.00	-0.02
2	7	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

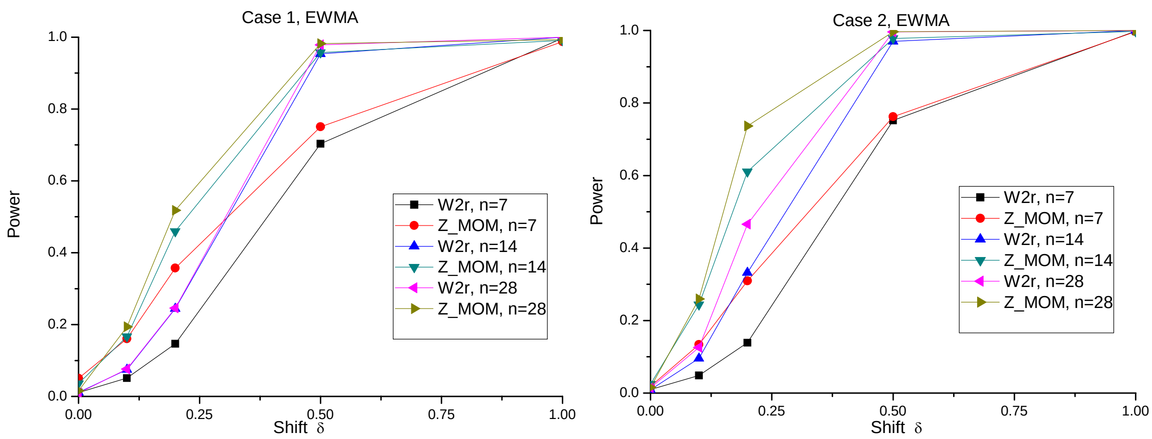

Figure 5-22: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA, Case 1 and Case 2

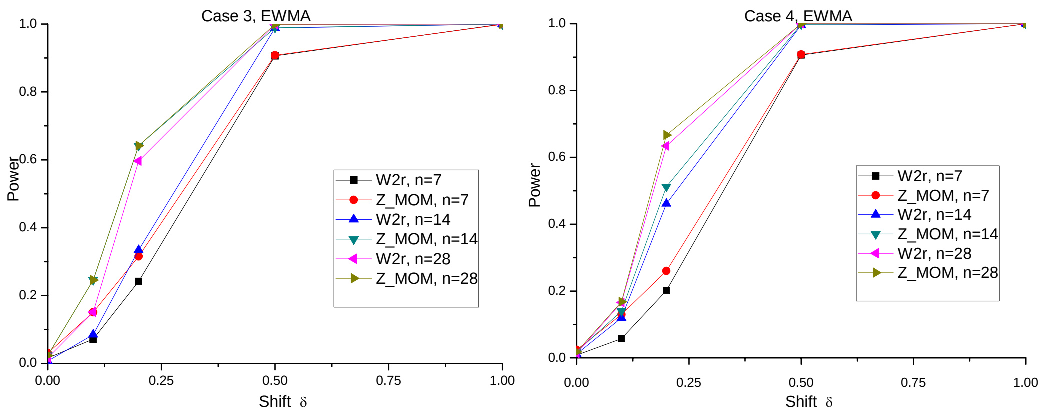

Figure 5-23: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA, Case 3 and Case 4

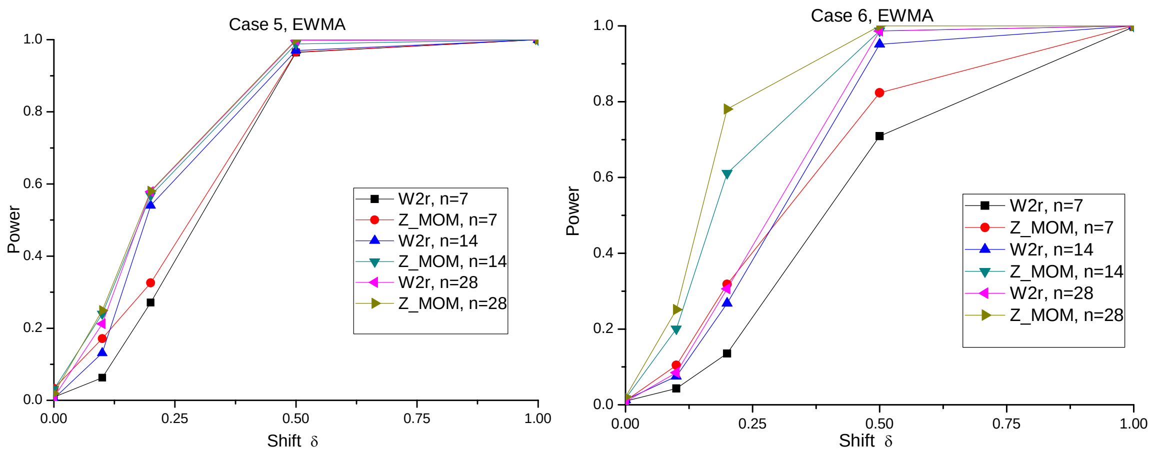

Figure 5-24: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA, Case 5 and Case 6

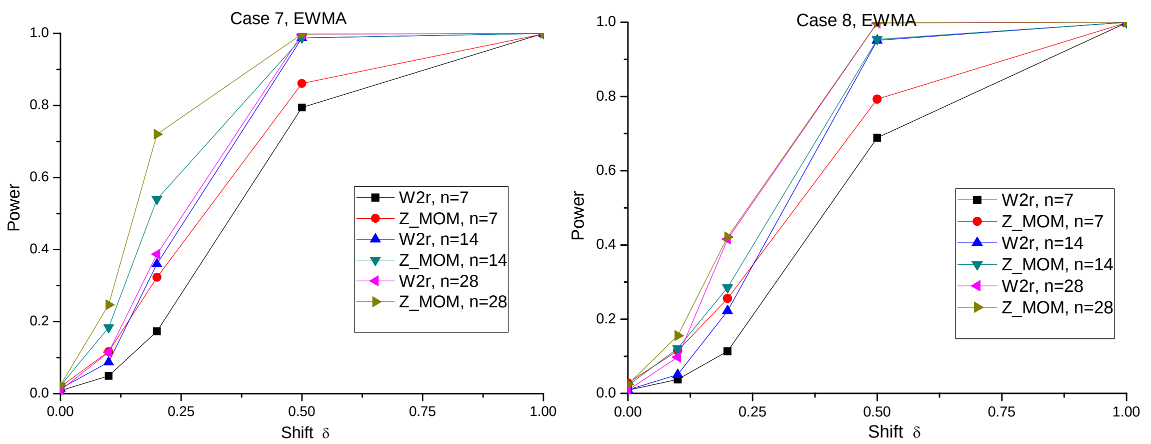

Figure 5-25: Power Analysis for W2r and Adaptive Threshold Methods for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA, Case 7 and Case 8

5.4.2.3 Comparison of W2r and Modified W2r Methods

Table 5-20 gives the power values for the W2r_2 method for the EWMA approach when different baselines are used, given RI =500. Table 5-20 shows that for a given δ in a given negative binomial case, when n increases the power values increase.

Table 5-20: Power Analysis for W2r_2 Method with Negative Binomial Input-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
0	7	0.0093	0.0120	0.0143	0.0093	0.0125	0.0095	0.0093	0.0143
	14	0.0140	0.0118	0.0105	0.0098	0.0155	0.0083	0.0133	0.0088
	28	0.0107	0.0118	0.0115	0.0120	0.0115	0.0085	0.0120	0.0170
.1	7	0.0880	0.0925	0.1085	0.1430	0.1580	0.0730	0.1183	0.1022
	14	0.0965	0.1013	0.1278	0.1538	0.2042	0.0785	0.1213	0.1075
	28	0.1208	0.1417	0.1698	0.2180	0.3117	0.1025	0.1598	0.1603
.2	7	0.3348	0.3570	0.4852	0.5063	0.5175	0.3500	0.3980	0.4013
	14	0.3678	0.4193	0.5288	0.6265	0.7675	0.4247	0.4678	0.4393
	28	0.4725	0.5570	0.6620	0.7542	0.9143	0.7497	0.7908	0.6013
.5	7	0.7930	0.7587	0.9242	0.9298	0.9720	0.7990	0.8025	0.8137
	14	0.9923	0.9965	0.9995	1.0000	1.0000	0.9905	0.9995	0.9985
	28	0.9992	0.9995	1.0000	1.0000	1.0000	0.9988	1.0000	1.0000
1	7	0.9998	0.9992	1.0000	1.0000	1.0000	0.9988	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	28	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 5-21 shows the percentage increases of the power values of the W2r_2 method compared to the W2r method. Note that these power values are given in Tables 5-15 and 5-20.

Table 5-21 shows that, for a given δ and baseline n in a given negative binomial case, the power values of the W2r_2 method are always as high as those of the W2r method with a few exceptions. We observed that the power value of the W2r method is 0.02% higher than that of the W2r_2 method given δ =1.0 and n= 7 days for Case 1. Figure 5-26 shows an example of the power analysis for the two methods for Case 1. This figure supports the conclusions we reached from Table 5-21.

Our power analysis results show that, in either the EWMA or the Shewhart comparisons, the W2r_2 method outperformed the W2r method in nearly all cases for negative binomial inputs.

Table 5-21: Percentage Increase of the Power Values for W2r_2 Method Compared to W2r Method with Negative Binomial Input-EWMA (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	3.67	4.40	3.70	8.52	9.50	2.95	6.88	6.45
	14	2.20	0.60	4.28	3.43	7.29	0.30	3.39	5.70
	28	4.43	1.57	1.85	5.25	9.95	1.77	4.45	6.25
.2	7	18.81	21.80	24.34	30.46	24.63	21.48	22.51	28.78
	14	12.36	8.73	19.45	16.50	22.70	15.69	10.78	21.68
	28	22.65	9.10	6.55	12.00	33.59	14.37	20.39	18.53
.5	7	8.95	0.62	1.79	2.35	0.72	9.00	0.80	12.52
	14	3.88	2.62	1.15	0.42	2.95	3.90	1.20	4.73
	28	2.04	0.32	0.00	0.02	0.20	1.23	0.20	0.20
1	7	-0.02	0.07	0.02	0.00	0.00	0.08	0.00	0.30
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

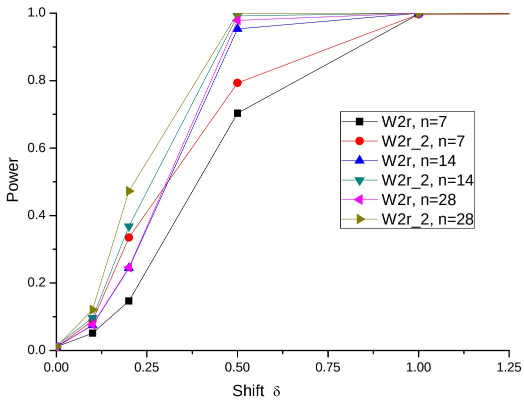

Figure 5-26: Power Analysis for W2r and W2r_2 Methods for Different Baselines- Transient Shift in Counts with Negative Binomial Inputs-EWMA, RI=500-Case 1

5.4.3 Comparison of Shewhart and EWMA Approaches

The EWMA approach was compared to the Shewhart approach for the adaptive threshold methods using MOM estimators based on the conditional negative binomial distribution. Table 5-22 shows the percentage increases of the power values of the adaptive threshold method for the EWMA approach compared to those for the Shewhart approach. The values of the percentage increase reported in Table 5-22 were computed by using the formula, i.e., (Power_EWMA – Power_Shewhart)/ Power_ Shewhart) *100%. These power values are given in Tables 5-9 and 5-18. We observed that for a given δ and baseline n, the EWMA is clearly superior to the Shewhart approach for the adaptive threshold method based on MOM estimators. Figure 5-27 shows an example of the power values of the adaptive threshold method for both control chart approaches for n=7, 14, and 28 days for Case 1, given RI=500. It is very clear that in terms of the power analysis, the EWMA is superior to the Shewhart approach for the adaptive threshold method using MOM estimators for the conditional negative binomial distribution.

Table 5-22: Percentage Increase of the Power Values for EWMA-based Adaptive Threshold Method Compared to Shewhart-based Adaptive Threshold Method-Transient Shift in Conditional Negative Binomial Distribution (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	192.18	146.95	183.80	128.87	257.08	58.97	85.58	91.99
	14	109.40	227.42	173.94	76.99	220.00	123.91	134.57	59.55
	28	64.88	129.06	121.08	39.93	87.39	104.98	124.20	37.40
.2	7	80.19	49.35	67.55	45.25	74.39	56.60	58.67	25.13
	14	77.89	132.56	111.92	88.20	135.87	129.50	117.65	90.13
	28	56.70	122.68	114.66	93.84	74.56	144.87	127.63	64.96
.5	7	16.34	3.05	25.32	32.51	23.83	7.66	14.84	51.03
	14	24.77	15.76	21.39	18.11	11.51	24.95	15.91	56.38
	28	25.94	6.62	2.90	5.35	8.67	16.28	5.56	11.58
1	7	2.86	1.22	2.51	4.69	5.15	0.50	0.91	8.34
	14	2.16	0.00	2.02	3.09	3.09	0.00	0.48	5.77
	28	0.35	0.00	0.00	1.52	0.00	0.00	0.00	1.50
2	7	0.00	0.00	0.00	4.17	0.00	0.00	0.00	0.35
	14	0.00	0.00	0.00	4.17	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

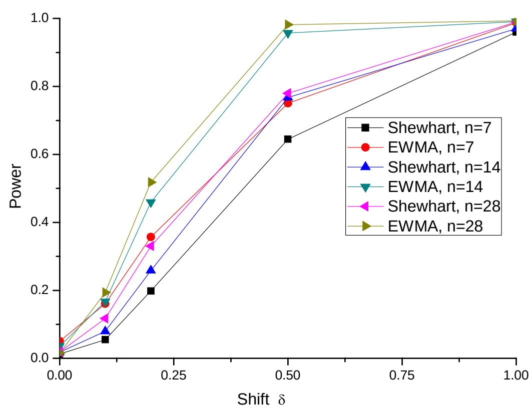

Figure 5-27: Power Analysis for Adaptive Threshold Method Using MOM Estimators for Different Baselines-Transient Shift in Conditional Negative Binomial Distribution-EWMA vs. Shewhart, RI=500-Case 1

The EWMA approach was compared to the Shewhart approach for the W2r method. Table 5-23 shows the percentage increases of the power values of the W2r method for the EWMA approach compared to those for the Shewhart approach. These power values are given in Tables 5-6 and 5-15. We observed that for a given δ and baseline n, the EWMA is clearly superior to the Shewhart approach for the W2r method with a few exceptions. Figure 5-28 shows an example of the power values of the W2r method for both control chart approaches for n=7, 14, and 28 days for Case 1, given RI=500. It is very clear that in terms of the power analysis, the EWMA is superior to the Shewhart approach for the W2r method for negative binomial inputs.

Table 5-23: Percentage Increase of the Power Values for EWMA-based W2r Method Compared to Shewhart-based W2r Method -Transient Shift in Negative Binomial Inputs (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	26.35	3.85	105.46	42.72	92.66	30.63	23.75	-23.37
	14	67.04	81.18	75.62	106.75	209.67	50.40	62.15	-12.48
	28	5.66	107.58	93.73	205.91	312.04	7.61	109.26	48.18
.2	7	66.14	41.84	190.97	160.59	251.30	43.22	95.59	28.68
	14	35.74	102.32	105.34	153.29	263.48	57.71	109.67	42.08
	28	0.78	113.96	173.37	162.07	127.99	5.85	70.22	70.77
.5	7	68.02	75.82	94.48	65.08	40.23	85.60	48.92	129.27
	14	35.31	24.48	22.35	18.36	8.98	37.80	22.82	68.59
	28	34.91	7.44	2.35	5.50	9.21	16.43	6.49	11.82
1	7	6.60	4.92	2.91	1.44	4.44	7.60	2.20	11.27
	14	0.17	0.02	0.02	0.00	0.00	0.30	0.00	0.32
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

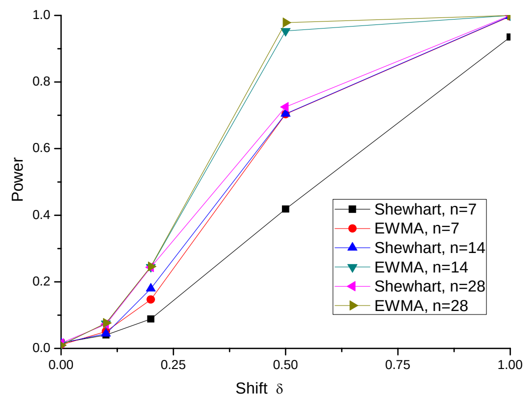

Figure 5-28: Power Analysis for W2r Method for Different Baselines -Transient Shift in Counts with Negative Binomial Inputs-EWMA vs. Shewhart, RI=500-Case 1

The EWMA approach was compared to the Shewhart approach for the W2r_2 method. Table 5-24 shows the percentage increases of the power values of the W2r_2 method for the EWMA approach compared to those for the Shewhart approach. These power values are given in Tables 5-11 and 5-20. We observed that for a given δ and baseline n, the EWMA is clearly superior to the Shewhart approach for the W2r_2 method. Figure 5-29 shows an example of the power values of the W2r method for both control chart approaches for n=7, 14, and 28 days for Case 1, given RI=500. It is very clear that in terms of the power analysis, the EWMA is superior to the Shewhart approach for the W2r_2 method for the negative binomial inputs.

Our power analysis results show that for negative binomial inputs, the EWMA was clearly superior to the Shewhart approach in all the cases for the adaptive threshold and W2r_2 methods, whereas the performance of the EWMA method was far better than that of the Shewhart approach in most cases for the W2r method.

Table 5-24: Percentage Increase of the Power Values for EWMA-based W2r_2 Method Compared to Shewhart-based W2r_2 Method-Transient Shift in Negative Binomial Inputs (RI=500)

δ	n	Case							
		1	2	3	4	5	6	7	8
.1	7	81.44	70.35	57.25	104.29	132.35	70.96	132.87	90.32
	14	69.30	79.29	83.36	97.18	75.28	50.38	109.14	69.29
	28	62.58	113.08	86.18	150.57	163.04	16.74	86.46	65.60
.2	7	233.80	151.05	195.31	154.81	129.19	248.95	201.97	221.55
	14	94.29	145.92	135.23	186.73	88.21	207.75	148.83	183.79
	28	89.00	145.92	106.88	130.15	93.91	156.31	189.46	107.13
.5	7	40.78	16.40	21.16	10.36	6.23	49.07	23.75	24.46
	14	25.29	16.99	7.47	7.45	0.20	32.15	13.71	17.33
	28	6.87	4.33	1.50	0.60	0.00	9.76	2.59	2.99
1	7	1.28	0.20	0.10	0.02	0.00	1.14	0.37	0.48
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

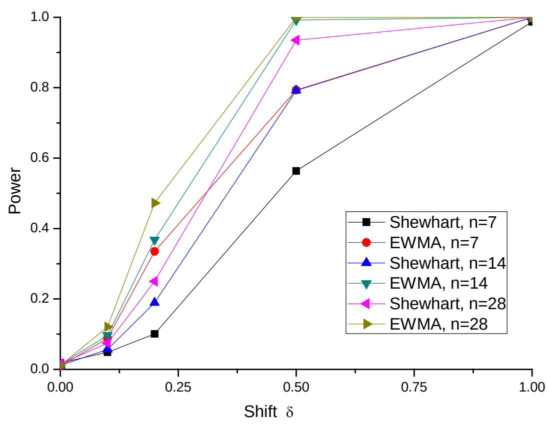

Figure 5-29: Power Analysis for W2r_2 Method for Different Baselines-Transient Shift in Counts with Negative Binomial Inputs-EWMA vs. Shewhart, RI=500-Case 1

Chapter 6 Conclusions

The main aim of my research was to evaluate the performance of the W2 rate method and the adaptive threshold methods, using both the Shewhart and EWMA approaches, for applications to public health surveillance. We developed new methods and revised existing methods for improved performance.

In this dissertation we reviewed the EARS W2 methods the CDC uses in its BioSense program. These methods provide simple tools for biosurveillance based on a moving window approach. The W2 method adjusts for total daily visits with the rate algorithm W2r. The W2 rate method was not based on any parametric model, but evaluated using real data from the Early Aberration Reporting System (EARS). We considered an alternative surveillance method which uses underlying parametric models to fit the data. We showed that this approach can lead to better performance. An adaptive threshold method was introduced using Z-scores calculated from the conditional probability distributions under the assumption of Poisson or negative binomial input data streams. The Z-score threshold functions were generally more tightly grouped than the W2r threshold functions over different sets of parameters and baseline lengths. The adaptive threshold methods also had higher power in detecting outbreaks. The improved performance of the adaptive threshold methods was likely due to implicitly or explicitly estimating the variance of the syndromic count by conditioning on the total number of visits each day. The use of an EWMA feature with either the W3r or the adaptive threshold method led to significantly improved performance.

The adaptive threshold method was first implemented by using a conditional binomial distribution based on independent Poisson inputs, and then implemented by using a conditional negative binomial distribution based on independent negative binomial inputs. The MLE estimators were used for the conditional binomial distribution based on Poisson inputs. The threshold functions for the adaptive threshold method obtained using MLE estimators match quite closely the functions assuming parameters are known for Poisson inputs. It is clearly seen that the MLE estimators for the conditional binomial distribution with Poisson inputs work very well for the adaptive threshold method.

The MOM estimators were given for the conditional negative binomial distribution and we discussed the problems that may arise in the estimation. We showed that the adaptive threshold method is more difficult to implement in the negative binomial case since we have more restrictions on the parameter space compared to the Poisson case. When using a short baseline to construct the estimators, it is more likely to lead to problems. In general, we recommend a longer baseline than the current use of a baseline length of n = 7. It is clearly seen, however, that the MOM estimators for the conditional negative binomial distribution worked well for the adaptive threshold method in our simulation study. Also, the negative binomial-based method works well with Poisson input data streams (simulation results not shown in this dissertation).

We reported that the definition of the mean absolute residual (MAR_t) for the W2r statistic does not reflect the total number of counts or visits at time t. A modified W2r method, called the W2r_1 method, was proposed based on Poisson inputs. The W2r_1 method replaces the MAR_t with the standard deviation based on the conditional binomial distribution, and therefore reflects the total number of visits at time t. Another modified W2r method, called the W2r_2 method, was proposed for use with negative binomial inputs. Note that the denominator of the W2r_2 method still does not reflect the total number of counts or visits at time t, but does lead to improved performance. The RI threshold functions of the modified W2r methods were compared to those of the W2r method over different sets of parameters and baseline lengths for both the Poisson and the negative binomial inputs.

The threshold functions for the W2r method obtained using our simulations matched quite closely the BioSense functions for the Poisson and negative binomial inputs. For Poisson inputs the threshold functions for the adaptive threshold method were compared to the W2r threshold functions over different sets of parameters and baseline lengths for both the Shewhart and

EWMA approaches. The RI function analysis showed the adaptive threshold method was more robust to different underlying parameters than the W2r threshold functions when the non-outbreak average counts of the Poisson inputs are assumed to be the same as on the weekends. The adaptive threshold method outperformed the W2r method by giving more consistent results across the parameter space and baseline window lengths for both the Shewhart and EWMA approaches in the Poisson case. Robustness to changes in parameter values, however, is much more important than robustness to different baseline lengths since different tables of thresholds can be given for different values of the baseline window lengths. The RI function analysis for Poisson inputs also showed the W2r_1 method outperformed the W2 rate method by giving more consistent results across the parameter space and baseline window lengths.

For negative binomial inputs the adaptive threshold method was first implemented by using a conditional binomial distribution. The RI function analysis for this case showed the conditional binomial distribution for the adaptive threshold method does not work nearly as well with the negative binomial inputs. It works much better for the Poisson inputs. The adaptive threshold method was then implemented by using the conditional negative binomial distribution for negative binomial inputs. The RI threshold function analysis for this case showed the adaptive threshold method was more robust to different parameters than the W2r RI threshold functions across the parameter space and baseline window lengths for both the Shewhart and EWMA approaches. The RI threshold function analysis for the negative binomial inputs also showed the W2r_2 method outperformed the W2 rate method by giving more consistent results across the parameter space and baseline window lengths.

Fraker et al. [27] noted that we should be wary of the use of the RI metric. If a process over time signals multiple times in a row, in practice this would typically be considered to be one alarm. However, in this BioSense application, all monitoring statistic values are used without regard to order, and the empiric RI is calculated based strictly on the percentage of days with the surveillance statistic values beyond a given threshold. As an example, suppose we have 100 observations over time. If we signal at observation 27, 45, 79, 89, and 96, the five signals would give us a RI of 20. Suppose we consider the same scenario, but observations 93-97 signal. Under the traditional RI approach, again we have a RI of 20, but there is only one signal event. It may be that large W2 values are found in clusters in the actual BioSense data, which would make the RI metric less meaningful. Another metric to use in place of the RI may be the average time between signals, also used by Fraker et al. [27]. This issue is also discussed in Szarka, Gan, and Woodall [13].

A separate simulation study was performed to test how well the W2r and adaptive threshold methods signal an outbreak that lasts seven days for different magnitudes of outbreaks for both the Poisson inputs and the negative binomial inputs. Shewhart and EWMA approaches were used for the power analyses. Our power analyses show that the EWMA approach with a reflecting lower boundary may be very useful for detecting syndromic outbreaks. The EWMA approach was clearly superior to the Shewhart approach in all the cases for the adaptive threshold method, whereas the performance of the EWMA method was far better than that of the Shewhart approach in most cases for the W2r and the modified W2r methods in the power analyses. In either the EWMA or the Shewhart comparisons, the adaptive threshold method for the conditional binomial distribution uniformly outperformed the W2r method for Poisson inputs. The adaptive threshold method for the conditional negative binomial distribution generally worked better than the W2r method for negative binomial inputs. The power analysis also showed that in either the EWMA or the Shewhart comparisons, the W2r_1 method uniformly outperformed the W2 rate method for Poisson inputs, and the W2r_2 method outperformed the W2 rate method for negative binomial inputs in nearly all cases.

Health care facilities typically have a lower number of visits on weekends. We therefore also examined how the W2r and the adaptive threshold methods based on the conditional binomial distribution perform with Poisson inputs where the average count is significantly lower on weekends. The RI threshold function analysis showed that for both the Shewhart and EWMA approaches the adaptive threshold method outperformed the W2 rate method by giving more consistent results across the parameter space and baseline window lengths. The power analysis results showed that the EWMA approach was superior to the Shewhart approach in all cases for the adaptive threshold method. For the W2r method, the EWMA approach worked better than the Shewhart approach in nearly all cases. These results for the W2r and adaptive threshold methods are similar to what we reported in Sections 4.3 and 4.4 when we assumed the average Poisson counts are the same as on the weekends.

In my future research the following topics will be explored:

(1) In a simulation study that includes outbreak data, it would be important not only to know how often the surveillance method signals, but if it does so quickly. Besides the power analysis, we could explore the timeliness analysis for both the Shewhart and EWMA approaches. We can consider the average time to signal given that a signal occurs. In both the power and timeliness analyses, we could also use outbreak data other than one week in length in the simulation study.

(2) We could further examine the estimation problems for the negative binomial inputs. Other than MOM estimators, we iteratively calculated the MLE estimators by combining the *Newton-Raphson* algorithm and the *Z-Negative Binomial* algorithm proposed in this dissertation. Our preliminary RI threshold function study of the adaptive threshold method shows that the MLE estimators work for the conditional negative binomial distribution (these results are not shown in this dissertation). Further work is needed on algorithms to better improve parameter estimation for the conditional negative binomial distribution.

(3) While the recurrence interval is an essential component of the current EARS method, it is appropriate to consider metrics such as the average time between signals events from Fraker et al. [27].

(4) We have asked the authors of Tokars et al. [6] to test the adaptive threshold methods based on the conditional negative binomial distribution on BioSense data. Our results show that

using n=14 or n=28 instead of n=7 results in much more consistent non-outbreak performance across various distributions of data streams. This supports in another way the recommendation of Tokars et al. [6] to use a longer baseline. The EWMA approach with the adaptive threshold method is the most promising. Also, it is important to see how the negative binomial-based method works with Poisson data. It should work well, but we need to check this.

References

1. Real-Time Hospital Data User Guide: Application Version 2.11. http://www. cdc.gov/BioSense/application.htm. [November 2006]. [not accessed online].

2. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response early aberration reporting system (EARS). *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 2003; **80**(2)(Suppl. 1)i89–i96.

3. BioSense Technical Overview of Data Collection, Analysis, and Reporting. http://www.cdc.gov/BioSense/files/BioSense Techn Overview 102908 webpage.pdf. [June 2008].[accessed online 26 May 2009].

4. Lambert D, Liu C. Adaptive thresholds: monitoring streams of network counts. *Journal of the American Statistical Association* 2006; **101**(473):78–88.

5. Kleinman KP, Abrams A, Mandl K, Platt R. Simulation for assessing statistical methods of biologic terrorism surveillance. *Morbidity and Mortality Weekly Report* 2005; **54**(Suppl.)101–108.

6. Tokars JI, Burkom H, Xing J, English R, Bloom S, Cox K, Pavlin JA. Enhancing time-series detection algorithms for automated biosurveillance. *Emerging Infectious Diseases* 2009; **15**(4):533–539.

7. BioSense User Guide. http://www.cdc.gov/BioSense/files/CDC BioSense User Guide VA DoD LabCorp v2.05.pdf. [August 2006]. [accessed online 26 May 2009].

8. Fricker Jr RD, Hegler BL, Dunfee DA. Comparing syndromic surveillance detection methods: EARS versus a CUSUM-based methodology. *Statistics in Medicine* 2008; **27**:3407–3429.

9. Hutwagner LC, Thompson WW, Seeman GM, Treadwell T. A simulation model for assessing aberration detection methods used in public health surveillance for systems with limited baselines. *Statistics in Medicine* 2005; **24**:543–550.

10. Hutwagner L, Browne T, Seeman GM, Fleischauer AT. Comparing aberration detection methods with simulated data. *Emerging Infectious Diseases* 2005; **11**(2):314–316.

11. Zhu Y, Wang W, Atrubin D, Wu Y. Initial evaluation of the early aberration reporting system - Florida. *Morbidity and Mortality Weekly Report* 2005; **54**(Suppl.):123–130.

12. Watkins RE, Eagleson S, Veenedaal B, Wright G, Plant AJ. Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. *BMC Medical Informatics and Decision Making* 2008; **8**:37.

13. Szarka JL, Gan L, Woodall WH. Comparison of the Early Aberration Reporting System (EARS) W2 method to an adaptive threshold method. *Statistics in Medicine*, 2010. (to appear)

14. Przyborowski J, Wilenski H. Homogeneity of results in testing samples from Poisson series. *Biometrika* 1940; **31**(3-4): 313-323.

15. Jain GC, Consul PC. A Generalized Negative Binomial Distribution. *Journal on Applied Mathematics*, 1971; **21**(4): 501-513.

16. Roberts SW. Control chart tests based on geometric moving averages. *Technometrics* 1959; **1**(3):239–250.

17. Crowder SV. A simple method for studying run-length distribution of exponentially weighted moving average charts. *Technometrics* 1987; **29**(4): 401-407.

18. Crowder SV. Average run lengths of exponentially weighted moving average control charts. *Journal of Quality Technology* 1987; **19**(3): 161-164.

19. Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics* 1990; **32**(1): 1-12.

20. Montgomery DC. *Introduction to Statistical Quality Control* (4th edn). Wiley, 2001.

21. Woodall WH, Mahmoud MA. The inertial properties of quality control charts. *Technometrics* 2005; **47**(4): 425-436.

22. Crowder SV, Hamilton MD. An EWMA for monitoring a process standard deviation. *Journal of Quality Technology* 1992; **24**(1):12–21.

23. Champ CW, Woodall WH, Mohsen HA. A generalized quality control procedure. *Statistics and Probability Letters.* 1991; **11**: 211-218.

24. Quesenberry CP. SPC Q-Charts for Start-Up Processes and Short or Long Runs. *Journal of*

Quality Technology, 1991; **23**: 213–224.

25. Simulation Data Sets for Comparison of Aberration Detection Methods. http://emergency.cdc.gov/surveillance/ears/datasets.asp. [16 April 2004]. [accessed online 26 May 2009].

26. Watkins RE, Eagleson S, Veenedaal B, Wright G, Plant AJ. R code for early aberration reporting system (EARS) and negative binomial cusum (NBC) algorithms. http://www.biomedcentral.com/content/supplementary/1472-6947-8-37-S1.pdf. [accessed online 26 May 2009].

27. Fraker SE, Woodall WH, Mousavi S. Performance metrics for surveillance schemes. *Quality Engineering* 2008; **20**:451–464.