Research Article

Statistics in Medicine

Received 14 July 2013, Accepted 11 March 2014 Published online 1 April 2014 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6161

Continuous sequential boundaries for vaccine safety surveillance

Rongxia Li,*† Brock Stewart, Eric Weintraub and Michael M. McNeil

Various recently developed sequential methods have been used to detect signals for post-marketing surveillance in drug and vaccine safety. Among these, the maximized sequential probability ratio test (MaxSPRT) has been used to detect elevated risks of adverse events following vaccination using large healthcare databases. However, a limitation of MaxSPRT is that it only provides a time-invariant flat boundary. In this study, we propose the use of time-varying boundaries for controlling how type I error is distributed throughout the surveillance period. This is especially useful in two scenarios: (i) when we desire generally larger sample sizes before a signal is generated, for example, when early adopters are not representative of the larger population; and (ii) when it is desired for a signal to be generated as early as possible, for example, when the adverse event is considered rare but serious. We consider four specific time-varying boundaries (which we call critical value functions), and we study their statistical power and average time to signal detection. The methodology we present here can be viewed as a generalization or flexible extension of MaxSPRT. Published 2014. This article is a U.S. Government work and is in the public domain in the USA.

Keywords: critical value; adverse events; drug and vaccine surveillance; sequential methods; maximized sequential probability ratio test (MaxSPRT)

1. Introduction

Before a new vaccine is licensed and released on the market for public use, pre-licensure clinical trials are conducted to ensure its efficacy and safety. However, it is often difficult or impossible to detect the full safety spectrum of a vaccine during clinical trials. This is especially true for rare events—such as Guillain–Barré syndrome (GBS), narcolepsy, or intussusception—because of limited sample size and the selected healthy population typically enrolled in these trials [1, 2]. In addition, the public's tolerance to adverse events following vaccination is generally low because vaccines, unlike therapeutic products, are routinely administered to healthy people, mostly infants and young children. Therefore, it is very important to conduct post-marketing vaccine safety surveillance to continuously monitor any potential adverse events following vaccination to ensure proper actions are taken in a timely manner.

Large electronic databases combining patient outcome information and vaccination records can be employed to establish a vaccine adverse event surveillance system. For example, the Vaccine Safety Datalink (VSD) is a post-marketing vaccine safety surveillance project sponsored by the Centers for Disease Control and Prevention (CDC) that is built upon electronic databases from nine integrated health care organizations. Also, the Food and Drug Administration initiated the Sentinel System to track the safety of vaccines, drugs, and medical devices in the post-marketing setting. In contrast to clinical trials, these surveillance systems utilize observational data on millions of individuals and collect detailed information about drug/vaccine exposures and adverse event occurrences at a relatively low cost.

Immunization Safety Office, Division of Healthcare Quality Promotion, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA, U.S.A.

**Correspondence to: Rongxia Li, MS D-26, Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA 30333, U.S.A.*

†E-mail: vwo3@cdc.gov

Sequential methods have increasingly been used to conduct near real-time safety monitoring of vaccines in the post-marketing setting. Wald [3, 4] proposed a sequential probability ratio test (SPRT) in the 1940s wherein a likelihood ratio test statistic is calculated and compared with pre-specified upper and lower limits to determine whether the surveillance should stop with a conclusion of rejecting or accepting the null hypothesis. This method has been applied to VSD vaccine safety surveillance studies in recent years [5]. However, the SPRT method can only be applied to a simple alternative hypothesis, for example, $H_a : \rho = \rho_0$, which limits its usage in drug and vaccine safety surveillance. To overcome this limitation, Kulldorff *et al.* [2] developed a maximized SPRT (MaxSPRT), wherein a composite alternative hypothesis, for example, $H_a : \rho > \rho_0$, is allowed in forming the likelihood ratio test statistic. Kulldorff presented two designs for MaxSPRT: a version comparing current rates to a historical comparison using a Poisson-based likelihood and a version comparing exposed and unexposed time periods using a binomial-based likelihood. A conditional MaxSPRT (CMaxSPRT) has also been developed to address variability in the historical data used for the Poisson-based MaxSPRT [1].

Since its inception, the MaxSPRT approach has been applied in several vaccine safety projects. For example, MaxSPRT was used to identify an elevated risk of febrile seizure in infants following the measles, mumps, rubella, and varicella (MMRV) vaccine compared with those who received the measles, mumps, rubella (MMR) vaccine and the varicella vaccine separately [6]. This finding led directly to changes in the national vaccine policy recommendations made by the CDC's Advisory Committee on Immunization Practices. The MaxSPRT method was also used to detect signals of GBS [7] and other neurologic, allergic, and cardiac adverse events [8] following the 2009–2010 H1N1 and seasonal influenza vaccine. Other examples of using MaxSPRT in vaccine safety surveillance can be found in [9, 10].

In contrast to group sequential methods, where an analysis is performed at discrete time intervals after a certain amount of data has accumulated, the continuous sequential (e.g., MaxSPRT), sometimes referred to as fully sequential approach, has benefits of continuous monitoring of adverse events and conducting analyses as often as possible, thus allowing a signal to be generated in a timely manner. Group sequential methods are more suitable for small sample size clinical trial studies with aims of assessing product efficacy rather than safety because of its delayed generation of early signals. One advantage of group sequential methods is their flexible stopping boundaries, such as Pocock's [11] constant nominal significance level, O'Brien and Fleming's [12] increasing nominal significance level, and others [13–17]. DeMets and Lan [18, 19] generalized and extended the group sequential boundary approach to a more flexible context and referred to it as the 'alpha spending approach'.

A flat boundary is used in the MaxSPRT method, meaning the critical value used to reject the null is constant throughout the surveillance. In some situations, subject-matter considerations might dictate that a flat boundary is less desirable than a time-varying boundary, for example, when the population varies over the surveillance period. However, to our knowledge, there have been no time-varying boundaries proposed for continuous sequential tests. In this paper, we propose an approach similar to the group sequential alpha-spending function approach, but tailored to continuous sequential testing in the context of vaccine safety. In particular, different boundary shapes are created that represent more flexible type I error spending strategies based on MaxSPRT. Although the Poisson MaxSPRT will be used throughout the paper, the same concept can also be applied to binomial MaxSPRT.

This paper is structured as follows: in Section 2, we briefly describe the concept and formulation of MaxSPRT; in Section 3, we introduce alpha-spending functions and sequential boundaries in a group sequential context; in Section 4, we propose critical value functions for Poisson MaxSPRT; in Section 5, we present the generated critical values, power, and earliest time to detect signals for four of our proposed critical value functions; we then provide two illustrative examples in Section 6 using the critical value function approach and finally present a discussion in Section 7.

2. Maximized sequential probability ratio test

A MaxSPRT is based on a maximum log-likelihood test statistic with a composite alternative hypothesis. Assuming the number of adverse events at time t that have occurred within a pre-defined time-window following vaccination, N_t, follows a Poisson process with known mean μ_t under a null hypothesis $H_0 : \rho = 1$, the MaxSPRT log-likelihood ratio test statistic, LLR_t, with an alternative hypothesis $H_A : \rho > 1$, is defined as follows:

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

$$
\begin{aligned}
LLR_t = \log(LR_t) &= \log\left(\max_{H_A} \frac{P(N_t = n_t|H_A)}{P(N_t = n_t|H_0)}\right) \\
&= \log\left(\max_{\rho>1} \frac{e^{-\rho\mu_t}(\rho\mu_t)^{n_t}/n_t!}{e^{-\mu_t}(\mu_t)^{n_t}/n_t!}\right) \\
&= \log\left(\max_{\rho>1} e^{-(\rho-1)\mu_t}\theta^{n_t}\right) \\
&= I(n_t \geqslant \mu_t)\left[\mu_t - n_t + n_t \log\frac{n_t}{\mu_t}\right],
\end{aligned}
\tag{1}
$$

where LR_t is likelihood ratio at time t; ρ is relative risk, the parameter of interest, which has the restriction of being $\geqslant 1$; n_t is a realization of the variable N_t; and the maximum likelihood estimate of ρ is $\frac{n_t}{\mu_t}$. Kulldorf *et al.* [2] proposed the use of the following stopping rule:

$$
T = \min\left(\hat{T}, t^*\right), \tag{2}
$$

where t^* is a pre-defined surveillance period length, which can be expressed as expected or observed number of events, or as calendar time, while $\hat{T}$ is defined as

$$
\hat{T} = \inf(t > 0 : LLR_t \geqslant V), \tag{3}
$$

where V is a time-invariant constant, also called the critical value. At each time point t, the log-likelihood test statistic LLR_t is calculated and compared with a pre-defined critical value V. If LLR_t exceeds the critical value, then we reject the null hypothesis; otherwise, the surveillance continues until it reaches the pre-specified upper limit t^*. Note that MaxSPRT can be viewed as a special case of sequential generalized likelihood ratio (GLR) tests defined specifically on Poisson or binomial probability distributions without a lower boundary [1, 20]. Lack of a lower boundary means H_0 cannot be accepted before the surveillance reaches the upper limit t^*. Compared with the standard GLR tests, MaxSPRT exhibited a larger power but also required a larger sample size regardless of the true parameter value [20]. However, the large sample size requirement is generally not a major concern for the observational databases because the cost of accruing additional data is minimal.

For a given type I error probability α, the critical value in the stopping rule (2) can be obtained by solving for the largest positive V that satisfy $pr_{H_0}(LLR_T \geqslant V) = \alpha$, i.e., $V = \sup(v > 0 : f(v, \alpha) = 0)$, where $f(v, \alpha) = pr_{H_0}(LLR \geqslant v) - \alpha$. These error probabilities can be computed by recursive numerical integration using an iterative Markov chain approach [2] or Monte Carlo simulations [1]. Then a linear or nonlinear interpolation approach or any optimization root-finding routine can be used to identify a critical value given an alpha value.

3. Group sequential boundaries and alpha-spending functions

A group sequential boundary is formed by a series of critical values pre-defined at each interim analysis to preserve the overall type I error rate maintained at a specific level. Pocock [11] proposed a group sequential method that defined a constant critical value at each equally spaced interim analysis. O'Brien and Fleming [12] developed another group sequential procedure in which the critical values decrease with an increase in the number of interim analyses. O'Brien and Fleming's sequential design effectively increases the threshold for rejecting the null hypothesis in the early stages of studies when sample sizes are small and allows for easier rejection during the later stages when more samples are accumulated. In general, these group sequential boundaries require the number of interim analyses to be fixed in advance and at equally spaced intervals between each interim analysis [16]. The alpha-spending function approach was first described by DeMets and Lan [17, 18] and was aimed at generalizing various group sequential boundaries to create more flexibility in designing and controlling the overall type I error rate in group sequential clinical trials. Sequential boundaries can be produced from a wide array of possible alpha-spending functions similar to earlier group sequential boundary designs but with fewer restrictions [21]. Several common alpha-spending functions are listed below, including the widely used Pocock type

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* **2014,** 33 3387–3397

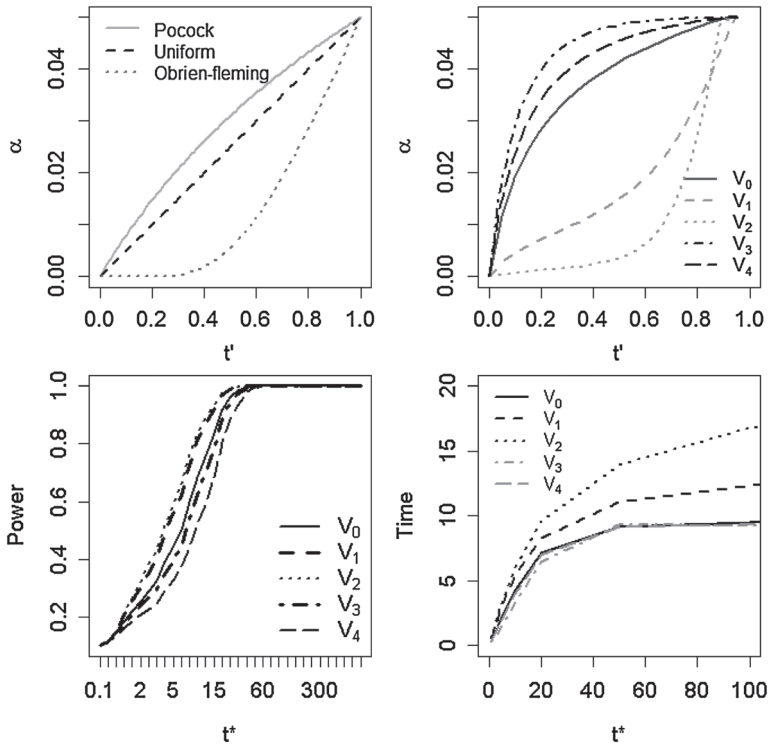

Figure 1. Typical alpha-spending function curves (top left) used in group sequential methods, alpha-spending curves of V_0–V_4 (top right), statistical power using V_0–V_4 when relative risk = 2 (bottom left), and the average time of V_0–V_4 to reject H_0 when relative risk = 2 (bottom right).

and O'Brien-Fleming type.

$$\alpha_1(t') = 2 - 2\Phi\left(Z_{\alpha/2}/\sqrt{t'}\right) \qquad \text{O'Brien-Fleming type} \qquad (4)$$

$$\alpha_2(t') = \alpha \ln(1 + (e-1)t') \qquad \text{Pocock type} \qquad (5)$$

$$\alpha_3(t') = at'^{\theta}, \text{ for } \theta > 0 \qquad \text{uniform.} \qquad (6)$$

In Equation (4), Φ denotes the standard normal cumulative distribution function. In all three equations, t' represents a relative time that denotes a fraction of total information available at the k^{th} interim analysis at time $t = t_k$. All three equations meet conditions such that $\alpha(0) = 0$ and $\alpha(1) = \alpha$. Equation (6) is uniform across 0 to t' when $\theta = 1$. These three alpha-spending functions are illustrated in Figure 1 (top left).

4. Critical value functions for Poisson MaxSPRT

In this study, we propose critical value functions for the Poisson MaxSPRT continuous sequential method. To have a time-varying boundary, the stopping rule (3) is modified as

$$\hat{T} = \inf(t > 0 : LLR_t \geq V(t')), \qquad (7)$$

where t' is the relative time that can be defined in various ways so long as it represents a fraction of the total surveillance period. In this study, t' is defined as a ratio between μ_t— the expected number of adverse events at time point t and t^*— the upper limit of the expected number of adverse events when the surveillance ends. t^* is generally set as the expected maximum number of adverse events that would occur during the entire surveillance period if no elevated risk is observed. The value of t^* can be derived from the historical data. Similar to alpha-spending functions, critical value functions are defined to allocate the type I error throughout the surveillance, all the while maintaining an overall type I error rate at

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

a pre-specified level. In contrast to alpha-spending functions where an alpha value is explicitly defined through a function $\alpha = f(t)$ and stopping boundaries are obtained by solving this function, critical value functions $v = V(t)$ provide a direct definition of boundary shape over time t. A critical value function $V(t', c)$ is composed of a constant component c and a time-varying component t'. We propose four critical value functions here to illustrate the range of possible choices that could be developed:

$$V_1(t') = c_1\ e^{-t'} \tag{8}$$

$$V_2(t') = c_2\theta^{t'},\ \theta = 0.2 \tag{9}$$

$$V_3(t') = c_3\theta^{t'},\ \theta = 2 \tag{10}$$

$$V_4(t') = c_4 \log(e + t'), \tag{11}$$

where $t' = \mu_t/t^*$, and c_1 to c_4 are constant parameters. All four functions meet the condition of $V(t' = 0) = c$. In general, there are no strict rules in terms of constructing critical value functions; in theory, any monotone function—even a step function—can be used as a critical value boundary. However, some realistic considerations need to be taken into account when we design a critical value function for real applications. To construct the aforementioned four critical value functions, the first step was to decide whether an increasing or decreasing function was desired. For example, V_1 and V_2 are decreasing functions, and V_3 and V_4 are increasing functions. Compared with the original flat boundary V_0, a decreasing function (V_1, V_2) means it is less likely to reject the null hypothesis at the beginning because of larger critical values, in other words, it is more conservative; conversely, an increasing function (V_3, V_4) means it is more likely to reject the null hypothesis at the beginning because of smaller critical values. Another realistic consideration is that the critical value at time $t' = 0$ should not be zero, because we do not want to trigger a signal for any $LLR > 0$ close to the beginning of surveillance. This consideration excludes all the linear, quadratic, and cubic functions commonly used for alpha-spending functions. In addition, we used the flat boundary V_0 as a reference point so that the starting critical values ($t' = 0$) were not too far off from V_0. We have done this so as to avoid the non-convergence problem. Note that V_2 and V_3 are essentially the same type of function with different parameter values θ. The value of θ can be set as any value greater than zero but less than 1 in the function of V_2 to ensure the decreasing trend; a smaller θ in V_2 shows a rapid decreasing speed in terms of critical values, which means more type I error is spent towards the end of the surveillance when critical values are low; similarly, the value of θ can be set as any value greater than 1 in the function of V_3 to ensure the increasing trend; however, when θ is too large, for example, >10, the convergence is difficult to attain. Therefore, we recommend $1 < \theta \leqslant 10V_3$.

5. Critical values, power, and timeliness

We used the similar numerical iterative approach as in [2] that is based on exact methods to calculate critical values. We first developed an algorithm to calculate type I error probability α, for a given critical value and an upper limit t^*. The algorithm is based on Poisson log-likelihood function described earlier in Equation (1). The specific algorithm is also presented in the Appendix algorithm 1. We then used a numerical Brent [22] root-finding optimization routine to find the values for the constant parameters in each critical value function for a pre-specified α, for example, 0.05. Table I shows calculated values of the constant component c in each of four critical value functions based on 0.05 type I error rate. Readers can easily compute the true critical values given μ_t and t^*. For example, if the expected Poisson mean at time t is $\mu_t = 0.2$ and the upper limit t^* is 5, we can obtain a critical value $V_1 = c_1 e^{-\mu_t/t^*} = 4.390092e^{-0.2/5} = 4.217954$. Similar tables for $\alpha = 0.01$ and $\alpha = 0.001$ are available upon request.

Figure 1 (top right) shows how α is distributed from the beginning to the end of the surveillance period for each of four critical value functions and the original Poisson MaxSPRT (expressed as V_0). Functions of V_1 and V_2 have a slow increase initially but rapidly increase in the later stages. Conversely, V_3 and V_4 increase rapidly during the early stage of surveillance, but there is a slowing down as more samples are collected.

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

Table I. Values for the constant component c in different critical value functions based on Poisson MaxSPRT ($\alpha = 0.05$).

t^*	c_1	c_2	c_3	c_4
0.1	2.044069	2.044069	2.044069	2.044069
0.2	2.266893	2.266893	2.266893	2.266893
0.5	2.694399	2.728036	2.490008	2.588212
1	3.122208	3.227530	2.635550	2.737780
1.5	3.380957	3.614013	2.720890	2.840241
2	3.587209	4.012429	2.783489	2.912427
2.5	3.787892	4.464598	2.831571	2.965968
3	3.997402	4.561893	2.870966	3.008000
4	4.133295	5.046222	2.933658	3.072644
5	4.390092	5.324411	2.982585	3.123311
6	4.442170	5.911193	3.022698	3.166197
8	4.668864	6.174868	3.085735	3.230935
10	4.831027	6.606456	3.134832	3.281175
12	4.889554	6.907241	3.174834	3.321844
15	5.012898	7.148300	3.223563	3.371567
20	5.142281	7.513498	3.286179	3.433908
25	5.236568	7.737337	3.334355	3.481446
30	5.303198	7.885104	3.373467	3.520325
40	5.395645	8.119302	3.434563	3.579965
50	5.476074	8.238967	3.481407	3.625379
60	5.508858	8.373239	3.519281	3.661972
80	5.584140	8.476278	3.578223	3.718709
100	5.628304	8.558943	3.623198	3.761919
120	5.662947	8.630289	3.659448	3.796438
150	5.702387	8.678375	3.703165	3.837993
200	5.748196	8.757515	3.758445	3.890431
250	5.779197	8.799796	3.800454	3.930108
300	5.804038	8.836083	3.834201	3.961939
400	5.841079	8.881620	3.886387	4.011055
500	5.867026	8.916706	3.925959	4.048226
600	5.886838	8.940856	3.957703	4.078014
800	5.916339	8.969530	4.006715	4.123930
1000	5.937798	8.990306	4.043831	4.158662

Power was calculated in a similar fashion to the type I error calculation. Rather than finding all probabilities to reject null hypothesis under H_0, we summed up all probabilities to reject null hypothesis under a specific alternative hypothesis H_1, for example, $\rho = 2$. The specific algorithm is presented in the Appendix algorithm 2. eTables 1–3 (Supporting information) present statistical power for each of the four critical value functions at a significance level of 0.05 for relative risk 1.2, 2, and 5, respectively. Figure 1 (bottom left) shows a comparison of statistical power when the relative risk is 2. V_1 and V_2 exhibit greater power than the original V_0, whereas V_3 and V_4 show less power than V_0. Similar trends were detected when relative risks are 1.2 and 5, respectively.

Timeliness in terms of signal detection is another important feature in surveillance studies. We present in eTables 4–6 (Supporting information) the expected time to reject the null hypothesis given the null will be rejected and the expected time to end the surveillance regardless of whether the null hypothesis is rejected or not for V_0–V_4 under different relative risks. Figure 1 (bottom right) shows that V_3 and V_4 use less time compared with V_0 to generate a signal for most of the upper limits of the number of expected adverse events when the relative risk equals to 2.

6. Examples

6.1. Febrile seizure in the 2010–2011 influenza vaccine safety surveillance

The trivalent inactivated influenza vaccine (TIV) is recommended for routine administration to children older than 6 months. However, the increased risk of febrile seizure is a major concern for children

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

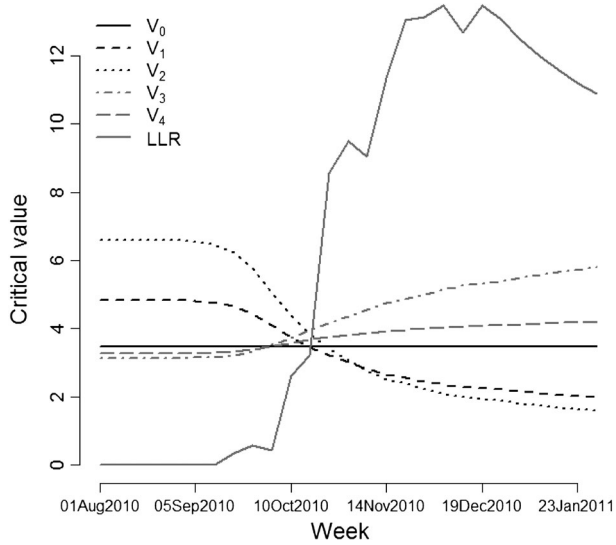

Figure 2. Poisson maximized sequential probability ratio test results for febrile seizures following inactivated influenza vaccine during the 2010–2011 influenza season using the original MaxSPRT flat critical values (V_0) and four critical value functions.

younger than 5 years, 0–1 day after TIV vaccination. The CDC-sponsored VSD established surveillance to monitor whether TIV is positively associated with an elevated risk of febrile seizure in this young age group. Both binomial MaxSPRT and CMaxSPRT were used in the original analysis [23]. The sequential monitoring was conducted from 1 August 2010 to 5 February 2011 for children ages 6–59 months following the administration of their first dose TIV. A statistical signal for seizures was identified during the week of 14 November 2010 using the CMaxSPRT approach. In this section, we present results of our re-analysis of the same data using the MaxSPRT method and our four proposed critical value functions.

Figure 2 shows the sequential log-likelihood ratio curve and critical value boundaries, including the original flat MaxSPRT boundary and the boundaries from our four proposed critical value functions ($V_1 - V_4$). V_1 and V_2 allow a larger value of the log-likelihood ratio test statistic before it signals in the beginning stage of surveillance when sample size is small, while V_3 and V_4 have values slightly lower than the flat MaxSPRT boundary at the beginning and then gradually increase the threshold values as surveillance progresses. Interestingly, for all five critical value boundaries, a statistical signal was detected during the same week of 24 October 2010. This may be explained by the nature of the uptake of the influenza vaccine. The majority of vaccine doses were administered by the end of October or the beginning of November, which corresponds to the time when most adverse events occur.

6.2. Febrile seizure in MMRV vaccine safety surveillance

The MMRV vaccine was licensed for use in 2005 for children aged 1 to 12 years. This combination vaccine was intended to replace two separate shots—the MMR vaccine and the varicella vaccine. Beginning in February 2006, safety surveillance was established in VSD to monitor the potential for an increased risk of adverse events following MMRV compared with MMR. In the week of 22 October 2006, a signal was detected using the Poisson MaxSPRT and found that MMRV was associated with an approximately twofold increased risk of febrile seizure in children 12 to 23 months compared with children who were administered MMR [6]. In this study, we re-analyzed these data using both the original MaxSPRT method and our proposed critical value function method. We present results in this section. Note that because our data are complete without any data delay issues, the results are slightly different from the original analysis even when we use the same MaxSPRT method.

We were able to detect increased risk signals using both the original MaxSPRT critical value and each of our proposed critical value functions, with an upper limit of 250, the same value used in the original analysis. The expected number of adverse events was calculated based on the historical 2000–2005 MMR data. The difference lies in the time when the signal was first detected. In our re-analysis, V_3 and V_4 detected a signal in the same week as that using the original flat critical value 4.112234, whereas V_1 detected a signal 2 weeks later and V_2 did not detect a signal until 38 weeks later (Figure 3).

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

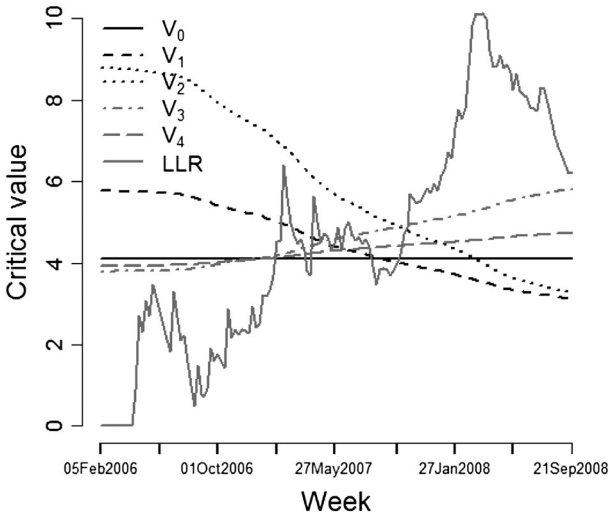

Figure 3. Poisson maximized sequential probability ratio test results for febrile seizures following measles, mumps, rubella, and varicella vaccine from February 2006 to September 2008 using the original MaxSPRT flat critical values (V_0) and four critical value functions.

7. Discussion

The MaxSPRT sequential method has been a useful tool for vaccine safety surveillance. Its success has been demonstrated in several VSD studies [7–10, 23–25]. This study serves as an extension or modification of MaxSPRT, which gives researchers greater flexibility to control how they spend type I error according to the features of their specific application.

For example, if an adverse event is extremely rare but may cause severe harmful consequences, researchers may allow an early signal to be more probable. This means more type I error is spent at the beginning of surveillance even if it increases the chance of a false-positive signal compared with alternative strategies. Conversely, if an adverse event is common but is considered not serious, e.g., mild local injection site reactions (redness and non-severe pain), we might want to be more conservative and wait for sample size to grow larger before we make a conclusion to reject the null hypothesis. This is because early adopters of a new drug or vaccine generally differ from later adopters (i.e., do not represent well the entire population); thus, a minimum sample size has to be attained before a signal is generated to reduce false alarms. Vaccines, unlike other pharmaceutical products, are routinely administered to a large population to prevent catastrophic or epidemic diseases, and their significant benefits most often outweigh less serious adverse events. Thus, minimizing false alarms is especially meaningful in vaccine safety surveillance, because too many false signals impose a negative impact on the public's confidence in vaccines and hinder fully utilization of benefits of vaccination.

In this study, we proposed four critical value functions. V_1 and V_2 serve well in the first situation, and V_3 and V_4 can be used in the second situation. It is occasionally beneficial to define a step function to embed different function forms in different stages of the surveillance. In particular, a spending function could be defined that has a step to disallow signaling prior to achieving a certain sample size, which is especially useful in situations where a delayed signal is desired. A delayed start boundary commonly used in many vaccine safety surveillance projects is essentially a special case of a step function, in which the first step function is set as positive infinity, and the second step function is set as a constant V_0. We also need to be aware that the earliest time of signal detection depends not only on the shape of critical value boundaries but also on the log-likelihood ratio curve. It is likely that a conservative boundary (larger critical values in the beginning) detects a signal earlier than a less conservative boundary (smaller critical values in the beginning), because the time of signaling is when the *LLR* curve intersects the critical value boundary curve. Note there is a trade-off between early time to signal and less power. In situations where neither the earliest time to signal nor the maximum power is pursued, the original flat boundary V_0 can be used as a middle ground between signal time and power. In addition, the flat boundary V_0 is easier to implement compared with time-varying boundaries.

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397

From a statistical point of view, an appropriately selected time-varying stopping boundary may exhibit desirable features, such as greater power and reduced sample size. For example, in the aforementioned febrile seizure case, although all the five functions including the constant V_0 generated a signal at the same time, V_1 and V_2 showed slightly greater power (Supporting information eTables 1–3) than the flat boundary. Although the group sequential method can provide a variety of stopping boundaries in clinical trial settings, only recently have there been some initial efforts made to adapting the group sequential method to drug or vaccine safety surveillance settings [26–28]. Further discussion and research work need to be performed in this area, and comparisons between continuous sequential and group sequential methods needs to be carried out. In this study, we borrow the concept of alpha-spending functions from group sequential methods and used them in a continuous setting to demonstrate that different time-varying boundaries can also be created when using a continuous sequential MaxSPRT method.

One limitation of this study is that because we do not explicitly define alpha-spending functions, it is not obvious how to obtain cumulative type I error rates at each stage of the surveillance. In general, there is no closed analytic form to derive alpha value or alpha functions from critical value functions. If researchers would like to know a specific type I error rate at one time point, they would have to run iterative computer programs to obtain it. However, we can provide such a computer program upon request. In addition, efficient algorithms to generate time-varying critical values should be developed to reduce the computation cost.

In conclusion, this study provides public health researchers and scientists greater flexibility to control the design and performance of their safety surveillance studies. An unlimited number of critical value functions can be implemented to form various sequential boundaries (not limited to Pocock or O'Brien and Flemming) in the surveillance process. In a similar fashion to the alpha-spending functions allowing for expansion and generalization of group sequential boundaries, critical value functions may generalize continuous MaxSPRT sequential boundaries and provide a greater number of options to distribute the type I error rate over the course of surveillance. Although this study focuses on the Poisson-based MaxSPRT in a historical control design, we are currently working on similar methods that can be applied to other types of distributions such as binomial MaxSPRT. Further studies should be conducted with close attention paid to in which situation one specific (type of) critical value function(s) may be more appropriate than another.

Appendix

Algorithm 1: An algorithm to compute Type I error.

```
Function TypeIError(t*, c, f)
    n ← 1
    p ← 0
    t ← 1
    μ_t ← 0
    while μ_t < t* do
        V = f(c, μ_t / t*)
        μ_t = −n * lambertW(−e^{−1−V/n})
        If (μ_t ≥ t*) μ_t = t*
        If t=1 Then
            π = Prob_Poisson (X_t=1 | μ_t)
            p= π
        Else
            For (i in 0:n-2)
                π = Prob_Poisson (X_t=n | μ_t , X_{t-1}=i)
                p=p+ π
            End for
        End if
        n=n+1
        t= t+1
    End while
    return p
End function
```


Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* **2014,** 33 3387–3397

Algorithm 2: An algorithm to compute power.

```
Function Power(t*, c, ρ, f)
    n ← 1
    p ← 0
    t ← 1
    μ_t ← 0
    while μ_t < t* do
        V = f(c, μ_t / t*)
        μ_t = −n * lambertW(−e^{−1−V/n})
        If (μ_t ≥ t*) μ_t = t*
        If t=1 Then
            π = Prob_Poisson(X_t=1 | ρ * μ_t)
            p= π
        Else
            For (i in 0:n-2)
                π = Prob_Poisson(X_t=n | ρ * μ_t, X_{t-1}=i)
                p=p+ π
            End for
        End if
        n=n+1
        t= t+1
    End while
    return p
End function
```

Acknowledgements

We thank the Vaccine Safety Datalink Project for providing us with data sets of febrile seizures following influenza and MMR vaccines. We also thank Frank DeStefano and David Kleinbaum for their valuable comments.

References

1. Li L, Kulldorff M. A conditional maximized sequential probability ratio test for pharmacovigilance. *Statistics in Medicine* 2009; **29**:284–295.
2. Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. *Sequential Analysis* 2011; **30**:58–78.
3. Wald A. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics* 1945; **16**:117–186.
4. Wald A. *Sequential Analysis*. Wiley: New York, 1947.
5. Davis RL, Kolczak M, Lewis E, Nordin J, Goodman M, Shay DK, Platt R, Black S, Shinefield H, Chen RT. Active surveillance of vaccine safety: a system to detect early signs of adverse events. *Epidemiology* 2005; **16**:336–341.
6. Klein NP, Fireman B, Yih WK, Lewis E, Kulldorff M, Baxter P, Hambidge S, Nordin J, Naleway A, Belongia EA, Lieu T, Baggs J, Weintraub E. Measles–mumps–rubella–varicella combination vaccine and the risk of febrile seizures. *Pediatrics* 2010; **126**(1):1–8.
7. Greene SK, Kulldorff M, Lewis EM, Li R, Yin R, Weintraub ES, Fireman BH, Lieu TA, Nordin JD, Glanz JM, Baxter R, Jacobsen SJ, Broder KR, Lee GM. Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink project. *American Journal of Epidemiology* 2009; **171**:177–188.
8. Lee GM, Greene SK, Weintraub ES, Baggs J, Kulldorff M, Fireman BH, Baxter R, Jacobsen SJ, Irving S, Daley MF, Yin R, Naleway A, Nordin J, Li L, McCarthy N, Vellozzi C, DeStefano F, Lieu TA. H1N1 and seasonal influenza vaccine safety in the Vaccine Safety Datalink project. *American Journal of Preventative Medicine* 2011; **41**:121–128.
9. Belongia EA, Irving SA, Shui IM, Kulldorff M, Lewis E, Yin R, Lieu TA, Weintraub E, Yih WK, Li R, Baggs J. Real-time surveillance to assess risk of intussusception and other adverse events after pentavalent, bovine-derived rotavirus vaccine. *The Pediatric Infectious Disease Journal* 2010; **29**:1–5.
10. Greene SK, Kulldorff M, Yin R, Yih WK, Lieu TA, Weintraub ES, Lee GM. Near real-time vaccine safety surveillance with partially accrued data. *Pharmacoepidemiology and Drug Safety* 2011; **20**:583–590.
11. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 1977; **64**:191–199.
12. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979; **35**:549–556.
13. Fleming TR, Harrington DP, O'Brien PC. Designs for group sequential tests. *Controlled Clinical Trials* 1984; **5**:348–362.
14. Wang SK, Tsiatis AA. Approximately optimal one-parameter boundaries for group sequential trials. *Biometrics* 1987; **43**:193–200.

15. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observations of each patient. I. Introduction and design. *British Journal of Cancer* 1976; **34**:585–612.
16. Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika* 1983; **70**:659–663.
17. Kittelson JM. A unifying family of group sequential test designs. *Biometrics* 1999; **55**:874–882.
18. DeMets DL, Lan KKG. Interim analysis: the alpha spending function approach. *Statistics in Medicine* 1994; **13**:1341–1352.
19. DeMets DL, Lan KKG. The alpha spending function approach to interim data analyses. In *Recent Advances in Clinical Trial Design and Analysis*, Thall PF (ed.). Kluwer Academic Publishers: Boston, MA, 1995.
20. Shih MC, Lai TL, Heyse JF, Chen J. Sequential generalized likelihood ratio tests for vaccine safety evaluation. *Statistics in Medicine* 2010; **29**:2698–2708.
21. DeMets DL. Sequential designs in clinical trials. *Cardiac Electrophysiology Review* 1998; **2**:57–60.
22. Brent R. *Algorithms for Minimization Without Derivatives*. Prentice-Hall: Englewood Cliffs, NJ, 1973.
23. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink project, 2010-2011. *Vaccine* 2012; **30**:2024–2031.
24. Lieu TA, Kulldorff M, Davis RL, Lewis EM, Weintraub E, Yih K, Yin R, Brown JS, Platt R. Real-time vaccine safety surveillance for the early detection of adverse events. *Medical Care* 2007; **45**:S89–S95.
25. Gee J, Naleway A, Shui I, Baggs J, Yin R, Li R, Kulldorff M, Lewis E, Fireman B, Daley MF, Klein NP, Weintraub ES. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the Vaccine Safety Datalink. *Vaccine* 2011; **29**:8279–8284.
26. Cook A, Tiwari RC, Wellman RD, Heckbert SR, Li L, Heagerty P, Marsh T, Nelson JC. Statistical approaches to group sequential monitoring of postmarket safety surveillance data: current state of the art for use in the Mini-Sentinel pilot. *Pharmacoepidemiology and Drug Safety* 2012; **21**(S1):72–81.
27. Zhao S, Cook A, Jackson L, Nelson J. Statistical performance of group sequential methods for observational post-licensure medical product safety surveillance: a simulation study. *Statistics and its Interface* 2008; **0**:1–11.
28. Nelson JC, Yu O, Dominguez-Islas CP, Cook AJ, Peterson D, Greene SK, Yih WK, Daley MF, Jacobsen SJ, Klein NP, Weintraub ES, Broder KR, Jackson LA. Adapting group sequential methods to observational postlicensure vaccine safety surveillance: results of a pentavalent combination DTaP–IPV–Hib vaccine safety study. *American Journal of Epidemiology* 2013; **177**(2):131–141.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

Published 2014. This article is a U.S. Government work and is in the public domain in the USA. *Statist. Med.* ***2014,*** 33 3387–3397