Journal of Clinical Epidemiology 113 (2019) 114−122

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Applying sequential surveillance methods that use regression adjustment or weighting to control confounding in a multisite, rare-event, distributed setting: Part 2 in-depth example of a reanalysis of the measles-mumps-rubella-varicella combination vaccine and seizure risk

Andrea J. Cook[a,b,*], Robert D. Wellman[a], Tracey Marsh[a,b], Azadeh Shoaibi[c], Ram Tiwari[d], Michael Nguyen[e], Denise Boudreau[a], Eric S. Weintraub[f], Lisa Jackson[a], Jennifer C. Nelson[a,b]

[a]Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA
[b]Department of Biostatistics, University of Washington, Seattle, WA, USA
[c]Center for Biologics Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA
[d]Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA
[e]Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA
[f]Division of Health Care Quality Promotion, Immunization Safety Office, Centers for Disease Control and Prevention, Atlanta, GA, USA

Accepted 5 April 2019; Published online 2 May 2019

Abstract

Objective: In-depth example of two new group sequential methods for postmarket safety monitoring of new medical products.

Study Design and Setting: Existing trial-based group sequential approaches have been extended to adjust for confounders, accommodate rare events, and address privacy-related constraints on data sharing. Most adaptations have involved design-based confounder strategies, for example, self-controlled or exposure matching, while analysis-based approaches like regression and weighting have received less attention. We describe the methodology of two new group sequential approaches that use analysis-based confounder adjustment (GS GEE) and weighting (GS IPTW). Using data from the Food and Drug Administration's Sentinel network, we apply both methods in the context of a known positive association: the measles-mumps-rubella-varicella vaccine and seizure risk in infants.

Results: Estimates from both new approaches were similar and comparable to prior studies using design-based methods to address confounding. The time to detection of a safety signal was considerably shorter for GS IPTW, which estimates a risk difference, compared to GS GEE, which provides relative estimates of excess risk.

Conclusion: Future group sequential safety surveillance efforts should consider analysis-based confounder adjustment techniques that evaluate safety signals on the risk difference scale to achieve greater statistical power and more timely results.

Keywords: Active surveillance; Distributed databases; Electronic health record (EHR); Group sequential analysis; Rare events; Vaccine safety

1. Introduction

Proactive safety surveillance of newly marketed drugs and vaccines is a national public health priority [1,2]. To address this need, "big data" networks have linked vast amounts of electronic health record data across multiple health care organizations and insurers [3−6]. This allows group sequential monitoring of large cohorts of health care enrollees to assess suspected safety concerns, often in real time as new drug or vaccine uptake occurs [7−9]. Group sequential methods involve routine estimation and testing of drug-outcome or vaccine-outcome associations over time, which can lead to earlier identification of excess risk compared with a traditional one-time analysis. Developing and applying group sequential methodologies for observational safety studies using electronic health record data is an active and relatively new area of research. Details of the history of group sequential methods developed in the context of postmarket safety surveillance using electronic health record data and their performance via simulation evaluation can be found in Part 1 by Nelson et al. of a related article in this issue (Insert hyperlink).

* Corresponding author. Biostatistics Unit, Kaiser Permanente Washington Health Research Institute, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, USA. Tel.: +1-206-287-4257; fax: +1 206-287-2871.

E-mail address: Andrea.J.Cook@kp.org (A.J. Cook).

https://doi.org/10.1016/j.jclinepi.2019.04.019
0895-4356/

What is new?

Key Findings

- Analysis-based confounder adjustment methods, such as regression adjustment and weighting, are a viable alternative to design-based approaches for sequential safety monitoring.
- Estimating and conducting inference with a risk difference measure in the rare event setting may be more stable than estimating a relative risk.

What this adds to what was known?

- Risk difference measures in the rare event setting with flexible confounding control are available and can be implemented in a distributed data setting.

What is the implication and what should change now?

- Applying methods with different confounder adjustment strategies and different targets for inference in a real-world setting showcases that regression adjustment based risk difference methods may be preferable to traditional designed-based relative risk estimates.

In this article, we present an in-depth example safety evaluation that compares and illustrates the real-world implementation in practice of two new analysis-based sequential methods developed for observational electronic health record and claims data studies. The first is a group sequential approach that utilizes generalized estimating equations (GS GEE) to directly adjust for confounders [10]. The second utilizes inverse probability of treatment weighted regression (GS IPTW) with propensity score−based weights [11]. We demonstrate the use of both methods using data on the combination measles-mumps-rubella-varicella vaccine and seizure risk from the FDA's Sentinel network, a vaccine-outcome pair previously observed to have a positive association [12,13]. Effect estimates and the time to detection of this documented safety signal are compared, and advantages of analysis-based methods are discussed. The ultimate goal of this work is to increase awareness of sequential approaches with analysis-based confounder adjustment and demonstrate how they can be effectively used in practice to monitor vaccine safety using observational electronic health record data.

2. Materials and methods

2.1. Surveillance design, population, and data

Data from the FDA Mini-Sentinel pilot project were obtained to conduct mock safety surveillance for the measles-mumps-rubella-varicella combination vaccine; originally licensed in 2005. An observational cohort was assembled consisting of children aged 11 years through 23 months who received either the combination measles-mumps-rubella-varicella vaccine or separate injections of the measles-mumps-rubella and varicella vaccines. The primary scientific question was whether the measles-mumps-rubella-varicella vaccine was associated with elevated risk of febrile seizure compared with same-day receipt of the measles-mumps-rubella and varicella vaccines. The surveillance period covered 3.5 years, from February 2006 to July 2008. The eligible population was drawn from subjects who were enrolled at one of four Mini-Sentinel Data Partners that maintain databases with demographics, immunizations, and International Classification of Diseases 9 Clinically modified diagnoses assigned to inpatient, emergency department, and outpatient visits. Febrile seizure was defined as in previous studies [12] by the first occurrence of International Classification of Diseases 9 Clinically modified codes 345 or 780.3 in the 7−10 days after vaccination. Febrile seizure events occurring in either an inpatient or emergency department setting were included. To avoid inadvertently including follow-up visits for febrile seizures that occurred before vaccination, children with a history of seizure in the past 180 days in any care setting were excluded. Institutional review board approval was not required.

2.2. Sequential design

As new doses of measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines were administered within the surveillance cohort, routine sequential hypothesis tests that evaluated the association between vaccine receipt and risk of febrile seizure were conducted separately with the GS GEE and GS ITPW methods. Both methods specified a one-sided hypothesis for elevated febrile seizure risk among measles-mumps-rubella-varicella vaccinees compared with recipients of measles-mumps-rubella plus varicella vaccines on the same day. Further details on the specific forms of these hypotheses and their corresponding test statistics are provided in the following sections.

The first sequential analysis for each method occurred 12 months after the beginning of surveillance (February 2006). Thereafter, 10 quarterly analyses were planned (i.e., at 15, 18, 21, 24, 27, 30, 33, 36, 39, and 42 months). The threshold for a potential safety signal was specified in advance and was the same at each sequential analysis. This threshold, also referred to as a Pocock-style stopping boundary [14], was selected to maintain the overall false-positive rate (type 1 error) at 0.05 across all planned analyses. For both the GS GEE and GS IPTW methods, each consecutive analysis was performed until either the value of the test statistic exceeded the boundary value or the end of surveillance was reached. If at any of the planned analyses the test statistic exceeded the signaling threshold, the null hypothesis of equivalent febrile seizure risk was rejected and surveillance was discontinued. If the final planned analysis was reached and the test statistic did not exceed the signaling threshold, then the surveillance ends and we failed to reject the null hypothesis of equivalent febrile seizure risk.

Because both GS GEE and GS IPTW were designed for use with rare adverse event outcomes, both use a nonparametric permutation approach to calculate signaling thresholds. This is in contrast to many standard sequential methods from clinical trials that rely on large sample assumptions [15]. Specifically, data are simulated under the respective null hypotheses and the unifying boundary approach [16] is used to compute the threshold. This methodology requires that investigators specify the following sequential design features in advance: (1) the desired number of analyses, (2) the timing of analyses (based on observed, or expected sample size, at each analysis time), (3) a total expected sample size to be accumulated by the end of surveillance (typically selected to achieve the desired power for the study), and (4) the shape of the signaling threshold over time, such as flat [14] or nonlinear decreasing [17].

The specification details of the sequential design yield different sequential signaling thresholds, and these differences reflect the trade-off between power and the average amount of calendar time it takes to detect a signal if one is present [18]. For instance, a flat boundary (which is lower at earlier time points) may signal earlier and require a relatively lower elevated risk to do so. But it will have less power at later analyses compared with a threshold that is conservative early on (i.e., has a higher boundary at earlier time points) and decreases over time. For the seizure and measles-mumps-rubella-varicella vaccine example, a flat boundary was used in combination with a delay (of 12 months) before the first analysis. This strategy has been used in previous studies [19] to avoid early false-positive signaling based on small amounts of data and to conserve statistical power for later tests.

2.3. Sequential analyses

Two group sequential methods, GS GEE and GS IPTW, were implemented using the sequential testing schedule described in the previous section. GS GEE uses regression adjustment for confounders. GS IPTW uses inverse probability of treatment weighting to address confounding, with weights based on propensity scores. An overview of both these methods is provided in the following two sections. Further technical detail have been published elsewhere [10,11].

2.3.1. Group sequential GEE (with regression adjustment for confounding)

GS GEE is a flexible, robust approach that uses a GEE framework to conduct multivariable regression analyses in a group sequential fashion. Because febrile seizure is an acute outcome occurring in a short window after vaccination, and the measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines are one-time exposures, both were treated as binary variables (as opposed to time-to-event, for instance). Logistic regression with GEE was used to estimate an adjusted odds ratio (OR), which, because febrile seizure is a rare outcome, approximates the relative risk (RR). Note that to classify febrile seizures as binary and ensure equal follow-up time between individuals, vaccinees were not included in the analysis until the 7- to 10-day-risk window was fully observed. Categorical confounders included in each regression model were age group (11−12 months, 13−14 months, 15−16 months, 17−19 months, and 20−23 months), sex, and data partner (4 sites).

More generally, and similar to a one-time analysis using GEE, GS GEE with multiple analysis times can accommodate more complicated data scenarios, and is not limited to binary outcomes and exposures. For example, with longer-term exposures (e.g., a drug used chronically for many months or years) and outcomes that may require a longer follow-up period (e.g., myocardial infarction), Poisson regression can be used in the GS GEE framework to estimate an adjusted rate ratio that accounts for variable duration of exposure (i.e., person-time contributions) across subjects. In other words, the GEE framework used by GS GEE facilitates the use of any type of generalized linear model (GLM) and corresponding link function.

Regardless of the outcome type, GS GEE calculates a score test statistic and compares this standardized test statistic to a preset signaling threshold. A score test statistic was used for computational efficiency when determining sequential monitoring boundaries. For the seizure and measles-mumps-rubella-varicella vaccine example, the score statistic was used to conduct group sequential tests of the hypothesis that febrile seizure risk among measles-mumps-rubella-varicella vaccine and measles-mumps-rubella plus varicella vaccine recipients is equivalent (i.e., H_0: OR = 1 vs. H_A: OR > 1).

2.3.2. Group sequential IPTW (with propensity score weights to adjust for confounding)

Like GS GEE, GS IPTW was developed for use with cohort designs comparing two independent exposure groups. Unlike GS GEE, the current GS IPTW method is only appropriate for binary exposures and outcomes, for example, a single point-in-time or short-term exposure and an outcome occurring acutely within a short, prespecified risk window. Future methods work could extend GS IPTW to more complex data settings.

GS IPTW uses a group sequential framework and propensity score−based weights to conduct sequential, site-stratified, weighted regression analyses of an adverse event outcome. Specifically, for the seizure and measles-mumps-rubella-varicella example, we first fit a propensity score model (logistic regression) to estimate the probability of measles-mumps-rubella-varicella receipt conditional on confounders (age group and sex) at each contributing data partner site. Using the estimated propensity scores, an inverse probability of treatment weight was calculated for each individual and trimmed at the 0.025 and 0.975 quantiles for weight stability. A weighted linear regression model of seizure risk and measles-mumps-rubella-varicella vaccine receipt was then fit to estimate a site-specific, adjusted risk difference (RD) and estimated variance that incorporates the variability of the propensity score weights as detailed by Lunceford et al. [20]. The site-specific RDs and variances were then combined via meta-analytic methods to calculate an overall stratified RD, variance, and corresponding standardized Wald test statistic (RD divided by standard error of the RD). GS IPTW then uses the Wald test statistic and the predetermined signaling thresholds to perform group sequential tests of the hypothesis that febrile seizure risk among measles-mumps-rubella-varicella vaccine recipients and same-day recipients of measles-mumps-rubella plus varicella is equivalent on the RD scale (i.e., H_0: RD = 0 vs. H_A: RD > 0). Similar to the GS GEE analyses, individuals were not included in the GS IPTW analyses until their entire outcome risk window (10 days) was fully observed.

3. Results

3.1. Surveillance population and vaccine uptake

Fig. 1 shows the uptake patterns for the measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines by week of surveillance. Uptake of the measles-mumps-rubella-varicella vaccine was slow initially, with very few observed doses administered during the first year of surveillance. However, by week 90, the number of measles-mumps-rubella-varicella doses exceeded the number of doses of measles-mumps-rubella plus varicella. Uptake of both vaccines increased steadily thereafter. Fig. 1 also shows the timing of the group sequential analyses by study week. Based on observed uptake, the first analysis at 12 months included 15,558 measles-mumps-rubella plus varicella vaccinees and 2,796 measles-mumps-rubella-varicella vaccinees. Subsequent quarterly looks added between 6,000 and 12,000 newly vaccinated children, about 4,000 to 8,000 of whom were vaccinated with measles-mumps-rubella-varicella.

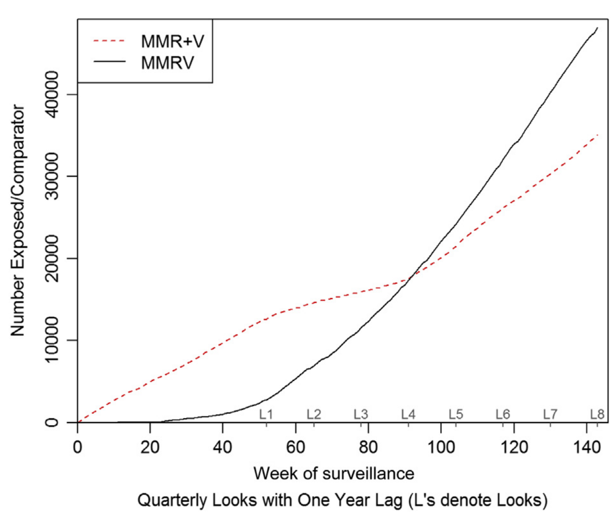

Fig. 1. Number of measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines received by week of sequential surveillance February 2006–December 2008. Abbreviations: GS GEE, group sequential estimating equations; GS IPTW, group sequential inverse probability treatment weighting; MMRV, measles-mumps-rubella-varicella; MMR+V, same day separate injections of measles-mumps-rubella and varicella vaccines.

The top half of Table 1 shows the baseline demographic characteristics among vaccinated children accrued into the surveillance cohort as of December 2008 when a safety signal was detected with GS GEE (i.e., among 48,233 measles-mumps-rubella-varicella vaccinees and 35,137 measles-mumps-rubella plus varicella vaccinees). The bottom half of Table 1 shows the demographics of the surveillance cohort in May of 2007 when the GS IPTW method detected the safety signal (i.e., among 14,633 measles-mumps-rubella-varicella vaccinees and 6,970 measles-mumps-rubella plus varicella vaccinees). The distributions of age, sex, and site were relatively similar at both time points. In particular, uptake was comparable between recipients of measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines over time by sex, but it differed by age group and data site. Consistent with age-based vaccination recommendations, nearly 90% of vaccinations were given to the three youngest age groups (11–16 months of age). Across contributing data sites, over 81% of measles-mumps-rubella-varicella doses were administered at a single site (site 15), whereas the distribution of measles-mumps-rubella plus varicella vaccinations was fairly even across sites.

3.2. GS GEE surveillance results

The top half of Table 2 shows the main group sequential surveillance results for the GS GEE analysis. One row of data is presented for each sequential analysis that was conducted. In total, eight analyses were performed before a safety signal was identified. Specifically, at analysis 8 in December 2008 (day 1,001 or 2.75 years after surveillance start), the observed score test statistic exceeded the preset threshold, indicating an increased risk of febrile seizure among measles-mumps-rubella-varicella vaccinees.

At the time of signal detection, 45 of the 48,233 measles-mumps-rubella-varicella recipients had experienced a febrile seizure compared with 13 of the 35,137 recipients of measles-mumps-rubella plus varicella vaccines. After adjusting for age, sex, and site, the adjusted seizure rates at look 8 were 3.8 and 9.1 seizures per 10,000 doses among measles-mumps-rubella plus varicella and measles-mumps-rubella-varicella vaccinees, respectively, yielding an adjusted odds

Table 1. Number (%) of doses by baseline characteristics, overall and among measles-mumps-rubella plus varicella and measles-mumps-rubella-varicella vaccine recipients from surveillance (February 2006 until a safety signal was detected [in December 2008 for GS GEE and in May 2007 for GS IPTW])

Baseline characteristics	Overall	MMR+V	MMRV
GS GEE surveillance cohort, signal detected December 2008			
Overall	83,370	35,137	48,233
Age			
11–12 mo	45,105 (54)	21,207 (60)	23,898 (50)
13–14 mo	16,752 (20)	6,101 (17)	10,651 (22)
15–16 mo	13,024 (16)	4,798 (14)	8,226 (17)
17–19 mo	5,717 (7)	2,030 (6)	3,687 (8)
20–23 mo	2,772 (3)	1,001 (3)	1,771 (4)
Sex			
Male	42,600 (51)	18,052 (51)	24,548 (51)
Female	40,770 (49)	17,085 (49)	23,685 (49)
Site			
2	10,992 (13)	8,730 (25)	2,262 (5)
4	8,088 (10)	7,686 (22)	402 (1)
15	46,655 (56)	7,321 (21)	39,334 (82)
16	17,635 (21)	11,400 (32)	6,235 (13)
GS IPTW surveillance cohort, signal detected may 2007			
Overall	21,603	14,633	6,970
Age			
11–12 mo	11,330 (52)	8,363 (57)	2,967 (43)
13–14 mo	4,277 (20)	2,730 (19)	1,547 (22)
15–16 mo	3,705 (17)	2,096 (14)	1,609 (23)
17–19 mo	1,565 (7)	974 (7)	591 (9)
20–23 mo	726 (3)	470 (3)	256 (4)
Sex			
Male	11,069 (51)	7,523 (51)	3,546 (51)
Female	10,534 (49)	7,110 (49)	3,424 (49)
Site			
4	3,600 (17)	3,574 (24)	21 (<1)
15	9,967 (46)	3,857 (26)	6,110 (88)
16	8,036 (37)	7,202 (49)	834 (12)

Abbreviations: GS GEE, group sequential estimating equations; GS IPTW, group sequential inverse probability treatment weighting; MMRV, measles-mumps-rubella- varicella; MMR+V, same day separate injections of measles-mumps-rubella and varicella vaccine.

ratio of 2.37 (sequentially adjusted P-value = 0.03). Although the signal was not statistically significant until look 8, the results in Table 2 show that, beginning at analysis 1, the adjusted odds ratio showed consistent elevated risk.

The top half of Table 3 provides data site—specific results. Odds ratios ranged from 1.88 to 3.08 across sites, where there was enough information to feasibly compute an odds ratio. Although the magnitude of these effects varied, the results demonstrate that febrile seizure risk was consistently elevated across contributing data sites. The uptake of measles-mumps-rubella-varicella vaccine at site 4 was relatively low (N = 402), and no febrile seizure events were observed; thus, a site-specific adjusted odds ratio could not be computed.

3.3. GS IPTW surveillance results

The bottom half of Table 2 presents the main group sequential surveillance results for the GS IPTW analyses. Comparing with GS GEE, the GS IPTW method detected a safety signal earlier, with the standardized test statistic crossing the signaling threshold at the second look on day 455, or 1.25 years after the beginning of surveillance (sequential adjusted P-value = 0.029).

At the time the signal was detected, 10 of the 6,970 measles-mumps-rubella-varicella vaccinees had experienced a seizure compared with 7 of the 14,633 measles-mumps-rubella plus varicella vaccinees. After adjusting at each site for age and sex using IPTW, the overall adjusted

Table 2. Results of GS GEE (top half) and GS IPTW (bottom half) sequential analyses comparing seizure risk among measles-mumps-rubella-varicella and measles-mumps-rubella plus varicella vaccines recipients

GS GEE look:	Days	MMR+V *N*	MMR+V outcome (%)	MMRV *N*	MMRV outcome (%)	MMR+VAdj[a] %Out	MMRV Adj[a] %Out	Adj[a] OR	Score test	Signaling threshold	Error spent	Signal
1:	364	12,652	5 (0.040)	2,796	5 (0.179)	0.043	0.131	3.05	1.84	3.41	0.000	No
2:	455	14,633	7 (0.048)	6,970	10 (0.143)	0.047	0.150	3.20	2.79	4.33	0.012	No
3:	546	15,968	7 (0.044)	11,577	14 (0.121)	0.044	0.119	2.69	2.79	4.49	0.012	No
4:	637	17,502	7 (0.040)	17,321	18 (0.104)	0.045	0.093	2.09	2.12	3.68	0.017	No
5:	728	21,491	8 (0.037)	24,195	21 (0.087)	0.042	0.079	1.86	2.04	3.87	0.019	No
6:	819	26,169	10 (0.038)	32,123	29 (0.090)	0.041	0.085	2.07	3.68	3.80	0.026	No
7:	910	30,385	12 (0.039)	40,326	36 (0.089)	0.045	0.081	1.79	2.51	3.60	0.029	No
8:	1,001	35,137	13 (0.037)	48,233	45 (0.093)	0.038	0.091	2.37	5.78	3.83	0.038	Yes
GS IPTW look:	Days	MMR+V *N*	MMR+V outcome (%)	MMRV *N*	MMRV outcome (%)	MMR+V Adj[b] %Out	MMRV Adj[b] %Out	Adj[b] %RD	Wald Test[c]	Signaling threshold	Error spent	Signal
1:	364	12,652	5 (0.040)	2,796	5 (0.179)	0.043	0.074	0.031	0.794	1.230	0.024	No
2:	455	14,633	7 (0.048)	6,970	10 (0.143)	0.053	0.157	0.104	3.053	2.153	0.024	Yes
3:	546	15,968	7 (0.044)	11,577	14 (0.121)	0.048	0.125	0.077	2.734			
4:	637	17,502	7 (0.040)	17,321	18 (0.104)	0.031	0.100	0.069	2.983			
5:	728	21,491	8 (0.037)	24,195	21 (0.087)	0.031	0.083	0.052	2.713			
6:	819	26,169	10 (0.038)	32,123	29 (0.090)	0.028	0.088	0.060	3.374			
7:	910	30,385	12 (0.039)	40,326	36 (0.089)	0.043	0.082	0.039	2.070			
8:	1,001	35,137	13 (0.037)	48,233	45 (0.093)	0.038	0.091	0.053	3.009			

Abbreviations: GS GEE, group sequential estimating equations; GS IPTW, group sequential inverse probability treatment weighting; MMRV, measles-mumps-rubella-varicella; MMR+V, same day separate injections of measles-mumps-rubella and varicella vaccines; Adj % Out, adjusted percent outcome; OR, odds ratio; RD, risk difference.

Surveillance was conducted in February 2006 until a safety signal was detected (which was in December 2008 for GS GEE and May 2007 for GS IPTW).

[a] Adjusted for sex, age group, site, and indicator for analysis period (GS GEE); sequential *P*-value = 0.03.

[b] Adjusted for sex, age group, and indicator for analysis period (GS ITPW); sequential *P*-value = 0.03.

[c] Wald test is the adjusted RD divided by the standard error of the adjusted RD.

seizure rates were 5.3 and 15.7 seizures per 10,000 doses for measles-mumps-rubella plus varicella and measles-mumps-rubella-varicella vaccinees, respectively, yielding an adjusted RD of 10.2 per 10,000 doses. This equates to one extra seizure per 980 children exposed to the measles-mumps-rubella-varicella vaccine instead of the measles-mumps-rubella plus varicella vaccines. The trend of elevated risk was observed across all analyses even after signaling relatively early compared with GS GEE at look 8.

Data site–specific results are shown in the bottom half of Table 3. The RDs ranged from −2.8 to 17.7 per 10,000 doses across data partner sites, and the RD was only elevated above at two sites (sites 15 and 16). Site 4 had very little measles-mumps-rubella-varicella uptake and only one observed seizure outcome by analysis 4, yielding little influence on the overall results.

4. Conclusions

We compared the ability and timeliness of two new sequential methods using analysis-based confounder adjustment techniques (GS GEE and GS IPTW) to detect a previously documented safety signal between receipt of the measles-mumps-rubella-varicella vaccine and the risk of seizure among infants 11 to 23 months of age. Both GS IPTW and GS GEE successfully found evidence of elevated risk of seizure among recipients of the combination measles-mumps-rubella-varicella vaccine compared with recipients of two separate injections of measles-mumps-rubella plus varicella vaccines.

Using the same sequential testing plan, the GS IPTW method signaled at the second planned analysis after 1.25 years and 6,970 doses of measles-mumps-rubella-varicella vaccine. The GS GEE method did not signal until the eighth analysis after 2.75 years and 48,233 doses of measles-mumps-rubella-varicella vaccine.

The magnitudes of the effect estimates obtained from both the GS IPTW and GS GEE analyses were similar. The RD estimated at the analysis at which GS IPTW signaled was 10.2 seizures per 10,000 doses, or an approximate adjusted RR of 2.96 (15.7 per 10,000 doses/5.3 per 10,000 doses). The adjusted odds ratio estimated from the

Table 3. Site-specific results of GS GEE and GS IPTW sequential analyses comparing the rates of seizure among measles-mumps-rubella-varicella vaccine and measles-mumps-rubella plus varicella vaccines recipients from the mini-sentinel surveillance population

Site	MMR+V *N*	MMR+V Outcome (%)	MMRV *N*	MMRV Outcome (%)	Adj %Out	Adj %Out	Adj OR (SE)[a]
2	8,730	4 (0.046)	2,262	2 (0.088)	0.063	0.118	1.88 (2.38)
4	7,686	1 (0.013)	402	0 (0.000)	0.013	0.000	0.00 (Inf)
15	7,321	3 (0.041)	39,334	35 (0.089)	0.041	0.088	2.15 (1.83)
16	11,400	5 (0.044)	6,235	8 (0.128)	0.051	0.156	3.08 (1.77)
Site	*N*	Outcome (%)	*N*	Outcome (%)	%Out	%Out	%RD (%SE)[a]
4	3,574	1 (0.028)	26	0 (0.000)	0.028	0.000	−0.028 (0.028)
15	3,857	2 (0.052)	6,110	8 (0.131)	0.053	0.146	0.093 (0.060)
16	7,202	4 (0.056)	834	2 (0.240)	0.064	0.241	0.177 (0.052)

Abbreviations: Adj %Out, adjusted percent outcome; GS GEE, group sequential estimating equations; GS IPTW, group sequential inverse probability treatment weighting; MMRV, measles-mumps-rubella-varicella; MMR+V, same day separate injections of measles-mumps-rubella and varicella.

Sequential Surveillance was conducted in February 2006 until a safety signal was detected (which was in December 2008 for GS GEE and November 2007 for GS IPTW).

[a] Adjusted for sex and age group.

GS GEE analysis at the time of signal was 2.37, or an approximate adjusted RD of 5.3 per 10,000 doses (9.1 minus 3.8 per 10,000 doses). Despite the comparability of the estimates, GS IPTW detected the signal with approximately 85% fewer doses of measles-mumps-rubella-varicella (6,970 vs. 48,233 doses) and one and half years earlier (1.25 vs. 2.75 years) than GS GEE.

Results of the current analyses were also comparable with findings from two previously published studies [12,13]. A VSD study, conducted within multiple managed care cohorts, sequentially monitored the safety of the measles-mumps-rubella-varicella vaccine when uptake began in February 2006 using a continuous sequential monitoring approach called the Poisson maximized likelihood ratio test, or MaxSPRT [21]. Observed seizure risk was compared to expected estimates of age-, sex- and site-adjusted risk from historical controls that received measles-mumps-rubella vaccine with or without receipt of varicella vaccine between 2000 and 2006. Using a definition of febrile seizure based on the same diagnostic codes as our study, preliminary results suggested an elevated seizure risk in a 0- to 42-day postvaccination window after 43,353 administered doses of measles-mumps-rubella-varicella vaccine (adjusted RR~2; exact RR not reported) (Center for Disease Control and Prevention-Morbidity and Mortality Weekly Report) [22]. Follow-up analyses compared risk in a 7- to 10-day postvaccination window among measles-mumps-rubella-varicella vaccine recipients to 314,599 recipients of separate measles-mumps-rubella plus varicella vaccines in the concurrent time period. After validation of diagnostic coded febrile seizures using medical charts the estimated adjusted odds ratio was 2.3 and the estimated RD was 5 seizures per 10,000 vaccinated (Center for Disease Control and Prevention-Morbidity and Mortality Weekly Report) [22]. Final multivariable regression analyses, conducted after a larger cohort of 83,107 measles-mumps-rubella-varicella vaccinees and 376,354 measles-mumps-rubella plus varicella vaccinees was available, and, using a 7- to 10-day-risk window, yielded an RR of 2.0 and RD of 4.3 per 10,000 doses [12].

Jacobsen et al. [13] conducted a measles-mumps-rubella-varicella safety study among 12—23 month-olds from a single managed care organization. Investigators matched 31,298 children who received the measles-mumps-rubella-varicella vaccine from February 2006 through June 2007 to 31,298 historical recipients of separate measles-mumps-rubella plus varicella vaccines from November 2003 through January 2006 based on age, sex, and calendar date of vaccination. Diagnostic coded febrile seizures occurring 5—12 days after vaccination were adjudicated by a committee using the Brighton criteria combined with chart-documented fever [23]. This study found an RR of 2.2 comparing measles-mumps-rubella-varicella vaccinees with historical measles-mumps-rubella plus varicella vaccinees and an RD of 3.8 per 10,000 vaccinations.

Despite differences in the design and analytic methods, our current findings using FDA Sentinel data agree with these prior results. All evaluations suggested a similar increased risk of seizure among measles-mumps-rubella-varicella recipients, with relative measures of odds or risk (odds ratios or RRs) 2—3 and RD measures of about 4—10 per 10,000 doses. The effect estimates for GS GEE and GS IPTW were slightly higher than the others (RR: 2.37 and 2.96; RD: 5.3 and 10.4 per 10,000 doses, respectively). This is not surprising because these estimates are computed at the time when a signal occurs rather than at the end of a prespecified study period. To compare both methods at look 8, the GS GEE and GS IPTW methods have comparable estimates (RR: 2.37 and 2.39; RD: 5.3 and 5.3 per 10,000 doses, respectively). However, sequential estimates that are conditional on an observed signal are known to be biased upward and correction methods are typically applied [24]. Furthermore, the previous studies used adjudicated outcomes yielding lower febrile seizure prevalence rates relative to our nonadjudicated outcomes which could modify estimated results. Among the three sequential evaluations, the GS IPTW method, which bases its signaling criteria on the risk difference, found the elevated risk most rapidly (after 6,970 doses). Both the VSD Poisson MaxSPRT and the GS GEE method, which base their signaling criteria on a relative measure of risk (i.e., either an OR or RR), required over four times as much information (about 45,000 measles-mumps-rubella-varicella doses) to signal. This may be due to the very rare event setting in which a relative measure can only be calculated when both intervention arms have an event which, when using an exact test such as permutation which all methods apply, can be relatively less stable than an RD.

In addition to potentially improving the power and timeliness with which a signal can be detected, GS IPTW has several other advantages. Confounding by site is accounted for via stratification which effectively accommodates interactions between data sites and confounders. GS IPTW has also been shown to be as efficient as a nonstratified estimate when no site interaction with other baseline confounders exists [11]. Use of the propensity score enables adjustment for a relatively large number of confounders at each data site. Moreover, differences in the variability of the estimated propensity score across sites are properly accounted for. Finally, RDs are often conceptually appealing to policy-makers because they can be readily translated into the number of excess events because of the new vaccine or drug.

As with any propensity score approach, GS IPTW requires an adequate number of exposed and unexposed patients to allow estimation. It may also be necessary to trim or restrict the cohort to avoid including patients with extreme propensity scores which result in large weights that may unduly inflate the variance. Because GS IPTW can detect signals with less data than other methods, it may be necessary to include more sites or wait longer to accrue more patients at each site to ensure stable site-specific estimates of the RD. In the measles-mumps-rubella-varicella analysis, for example, substantial elevation in risk was only observed at one site, with RD estimates at the other two sites based on fewer than 10 events. GS IPTW is currently designed for one-time or short-term exposures with acutely occurring adverse events. With extensions, it could be developed to be applicable for chronically used drugs and outcomes requiring long-term follow-up.

The GS GEE method also offers a flexible application option. With its foundation in GLMs, GS GEE can be used for both one-time exposures with acute-onset events (logistic regression), and chronically used drugs with longer-term follow-up (Poisson regression). The regression framework also allows for flexible adjustment for individual confounders or balancing scores, for example, categorical or nonlinear functions of the propensity score. Like GS IPTW, GS GEE does not restrict to a subset of the available data the way a matched analysis might, thus making use of all the information available in the data. Cohort restriction can reduce power, which is especially important when the outcome of interest is rare. A limitation of GS GEE in the rare event setting is that it requires an adequate number of patients with and without an adverse event to fit the outcome model. Furthermore, owing to issues of model fit with rare events, the ability to include large numbers of confounders may be limited. This problem can be largely avoided, however, through the use of propensity scores or other balancing scores to reduce the dimensionality of the confounders.

The objective of this research was to increase awareness and demonstrate the feasibility of two new group sequential approaches to safety surveillance. Both methods presented use analysis-based confounder adjustment strategies and are custom-designed to accommodate rare events in a distributed data setting. Specifically, the use of GS GEE and GS IPTW methods was illustrated using an example vaccine-event pair with a known association and data from the FDA's Sentinel data network. This proof-of-concept application demonstrated the practicality of implementing these methods in claims data settings and their relative advantages compared with existing approaches. In particular, this work suggests that future sequential observational studies of drug or vaccine safety should consider methods that use analysis-based confounder adjustment techniques and, in particular, methods such as GS IPTW that quantify risk on the difference scale, as such approaches have the potential to offer greater statistical power and faster evaluation of pressing safety questions. Furthermore, simulation studies comparing approaches across a wide array of settings would be valuable to assess when difference estimates should be applied.

Acknowledgments

This work was funded by a subcontract with America's Health Insurance Plans (AHIP) under contract 200-2002-00732 from the Centers for Disease Control and Prevention and by the Sentinel System, a project sponsored by the U.S. Food and Drug Administration (FDA) to support monitoring the safety of FDA-regulated medical products. The Sentinel System is one piece of the Sentinel Initiative, a multifaceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services Contract number HHSF223201400030I.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the US Food and Drug Administration.

References

[1] System, C.o.t.A.o.t.U.D.S.. In: Baciu A, Stratton K, Burke SP, editors. The future of drug safety: promoting and protecting the health of the public. Washington, DC: The National Academies Press; 2007.

[2] Food and Drug Administration Amendments Act of 2007. In: U.S.P.L., 110th Cong. 2008. Available at https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/food-and-drug-administration-amendments-act-fdaaa-2007. Accessed June 13, 2019.

[3] Chen RT, DeStefano F, Davis RL, Jackson LA, Thompson RS, Mullooly JP, et al. The Vaccine Safety Datalink: immunization research in health maintenance organizations in the USA. Bull World Health Organ 2000;78(2):186−94.

[4] Baggs J, Gee J, Lewis E, Fowler G, Benson P, Lieu T, et al. The Vaccine Safety Datalink: a model for monitoring immunization safety. Pediatrics 2011;127(Suppl 1):S45−63.

[5] Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the sentinel system–a national resource for evidence development. N Engl J Med 2011;364:498−9.

[6] Nguyen M, Ball R, Midthun K, Lieu TA. The Food and drug administration's post-licensure rapid immunization safety monitoring program: strengthening the federal vaccine safety enterprise. Pharmacoepidemiol Drug Saf 2012;21(Suppl1):291−7.

[7] Fireman B, Toh S, Butler MG, Go AS, Joffe HV, Graham DJ, et al. A protocol for active surveillance of acute myocardial infarction in association with the use of a new antidiabetic pharmaceutical agent. Pharmacoepidemiol Drug Saf 2012;21(Suppl 1):282−90.

[8] Yih WK, Kulldorff M, Fireman BH, Shui IM, Lewis EM, Klein NP, et al. Active surveillance for adverse events: the experience of the Vaccine Safety Datalink project. Pediatrics 2011;127(Suppl 1): S54−64.

[9] Carnahan RC, Gagne JJ, Nelson JC, Fireman B, Wang S, Shoaibi A, et al. Mini-sentinel prospective routing observational monitoring program tools (PROMPT): Rivaroxaban surveillance 2015. Available at: https://www.sentinelinitiative.org/sites/default/files/Drugs/Assessments/Mini-Sentinel_PROMPT_Rivaroxaban-Surveillance-Plan.pdf. Accessed June 13, 2019.

[10] Cook AJ, Wellman RD, Marsh T, Tiwari RC. Group sequential method for observational data by using generalized estimating equations: application to Vaccine Safety Datalink. J R Stat Soc Ser C Appl Stat 2014;64(2):319−38.

[11] Cook AJ, Wellman RD, Marsh T, Tiwari R, Nguyen M, Russek-Cohen E, et al. Statistical methods for estimating causal risk differences in the distributed data setting for postmarket safety outcomes 2012. Available at: https://www.sentinelinitiative.org/sites/default/files/Methods/Mini-Sentinel_PRISM_Statistical-Methods-for-Estimating-Causal-Risk-Differences.pdf. Accessed June 13, 2019.

[12] Klein NP, Fireman B, Yih WK, Lewis E, Kulldorff M, Ray P, et al. Measles-mumps-rubella-varicella combination vaccine and the risk of febrile seizures. Pediatrics 2010;126(1):e1−8.

[13] Jacobsen SJ, Ackerson BK, Sy LS, Tran TN, Jones TL, Yao JF, et al. Observational safety study of febrile convulsion following first dose MMRV vaccination in a managed care setting. Vaccine 2009; 27(34):4656−61.

[14] Pocock SJ. Group sequential methods in the design and analysis of clincial trials. Biometrika 1977;64(2):191−9.

[15] Lan K, Demets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70(3):659−63.

[16] Kittelson JM, Emerson SS. A unifying family of group sequential test designs. Biometrics 1999;55:874−82.

[17] O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35:549−56.

[18] Emerson SS, Kittelson JM, Gillen DL. Frequentist evaluation of group sequential clinical trial designs. Stat Med 2007;26:5047−80.

[19] Nelson JC, Yu O, Dominguez-Islas CP, Cook AJ, Peterson D, Greene SK, et al. Adapting group sequential methods to observational postlicensure vaccine safety surveillance: results of a pentavalent combination DTaP-IPV-Hib vaccine safety study. Am J Epidemiol 2013;177:131−41.

[20] Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat Med 2004;23:2937−60.

[21] Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. Seq Anal 2011;30(1):58−78.

[22] Center for Disease Control and Prevention-Morbidity and Mortality Weekly Report. Available at: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5903a1.htm.

[23] Bonhoeffer J, Menkes J, Gold MS, de Souza-Brito G, Fisher MC, Halsey N, et al. Generalized convulsive seizure as an adverse event following immunization: case definition and guidelines for data collection, analysis, and presentation. Vaccine 2004;22(5−6):557−62.

[24] Fan XF, DeMets DL, Lan KK. Conditional bias of point estimates following a group sequential test. J Biopharm Stat 2004;14(2): 505−30.