MINI-SENTINEL METHODS

QUANTIFYING THE IMPACT OF TIME-VARYING BASELINE RISK ADJUSTMENT IN THE SELF-CONTROLLED RISK INTERVAL DESIGN

Prepared by: Lingling Li, PhD,[1] Martin Kulldorff, PhD,[2] Estelle Russek-Cohen, PhD,[3] Alison Tse Kawai, ScD,[1] Wei Hua, MD, PhD[3]

Author Affiliations: 1. Department of Population Medicine, Harvard Pilgrim Health Care Institute and Harvard Medical School, Boston, MA; 2. Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; 3. Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research, FDA, Rockville, MD

October 26, 2015

Mini-Sentinel is a pilot project sponsored by the U.S. Food and Drug Administration (FDA) to inform and facilitate development of a fully operational active surveillance system, the Sentinel System, for monitoring the safety of FDA-regulated medical products. Mini-Sentinel is one piece of the Sentinel Initiative, a multi-faceted effort by the FDA to develop a national electronic system that will complement existing methods of safety surveillance. Mini-Sentinel Collaborators include Data and Academic Partners that provide access to health care data and ongoing scientific, technical, methodological, and organizational expertise. The Mini-Sentinel Coordinating Center is funded by the FDA through the Department of Health and Human Services (HHS) Contract number HHSF223200910006I.

Mini-Sentinel Methods

Quantifying the Impact of Time-Varying Baseline Risk Adjustment in the Self-Controlled Risk Interval Design

Table of Contents

I. INTRODUCTION

The self-controlled risk interval (SCRI) design,[1-3] a variant of the self-controlled case series (SCCS) design,[4,5] is being increasingly used in vaccine and drug safety studies to examine the association between an acute exposure (e.g., vaccination, antibiotics use) and a selected adverse event (AE) of interest that occurs shortly after exposure.[2,6-9] In SCRI, each individual serves as his/her own control and contributes a risk interval following the exposure and a control interval that can be either before the exposure or after the end of the risk interval. Both SCCS and SCRI implicitly adjust for fixed covariates such as sex, race, genetic factors, chronic health conditions, and geographic region, and thus are preferred over cohort or case-control designs when fixed-covariate confounding is a major concern. Please refer to the paper by Baker et al.[3] for a comprehensive discussion on SCRI, SCCS, and other related study designs.

Explicit adjustment needs to be made for SCRI if there are time-varying confounders such as age and/or seasonality. Age is an important time-varying confounder in vaccine safety studies among young children, who receive vaccines at particular ages according to the immunization schedule.[10] The risks of certain AEs (e.g., febrile seizures, intussusceptions) vary substantially during early years of life.[6,11] In particular, among infants, the risks can vary substantially within a typical follow-up period of several weeks. Without proper adjustment, we may miss a true vaccine-AE association or detect a false positive, both of which may result in unwanted consequences for public health. Age becomes much less important in studies among adults as we do not expect AE risks to vary substantially during a period of several weeks or months. Then seasonality and evolving health status (among elders or subjects with poor health) may be more relevant. Note that seasonality matters when both exposure and AE incidence are seasonal. Throughout this paper, we use age as an example. We expect the results to be applicable to other time-varying confounders as well.

There are two approaches to adjust for time-varying confounders in SCRI: a random adjustment approach (SCRI-r) and a fixed adjustment approach (SCRI-f). The SCRI-f approach has been used in the FDA sponsored Mini-Sentinel Post-Licensure Rapid Immunization Safety Monitoring (MS-PRISM) Rotavirus Vaccines and Intussusception study.[6] The SCRI-r approach has been used in the MS-PRISM Influenza and Febrile Seizures study.[12] SCRI-r adjusts for time-varying baseline risks and estimates the exposure-associated relative risk (RR) simultaneously, taking into consideration the variation in both the study sample and the baseline risk estimates; however, it requires access to individual-level data from an unexposed baseline sample. This may not be feasible in multi-site studies due to data confidentiality and privacy concerns or unavailability of raw data used in published risk estimates. The alternative approach, SCRI-f, only requires an estimated baseline risk function; however, it does not account for the uncertainty in the baseline risk estimates and thus may lead to an underestimated variance estimate for the RR and consequently to a confidence interval (CI) that is too narrow.

Despite its theoretical advantages, the use of SCRI-r is limited practically due to the unavailability of individual-level data for the baseline sample in many study settings. Thus it is important to evaluate the performance of SCRI-f and identify scenarios in which it can be used as an alternative. We conducted a

comprehensive simulation study to assess and compare the performance of four SCRI analyses: the unadjusted SCRI analysis that ignores time-varying confounding (SCRI-u), SCRI-f, SCRI-r, and SCRI-g, which refers to the gold-standard SCRI analysis adjusting for time-varying confounding using the true baseline risks. In real-life applications, the true baseline risk function is always unknown. We implement SCRI-g simply for the purpose of calibrating the performance of the other SCRI analyses as it represents the optimal but unattainable gold-standard. We designed the simulation studies using empirical data from the MS-PRISM Rotavirus Vaccines and Intussusception study in children 5-36.9 weeks of age[6] and the MS-PRISM Influenza Vaccines and Febrile Seizures study in children 6-59 months of age.[12] The overall goal of the simulation study is to develop practical guidelines on the selection of appropriate SCRI analyses in the presence of time-varying confounding and to demonstrate and discuss the potential caveats in implementation of these analyses.

II. METHODS

Next, we introduce the four SCRI analyses in a general vaccine safety study setting with age being the time-varying confounder and then introduce the simulation study design and parameter specifications.

Vaccinated individuals with at least one AE occurring in either the risk or control interval are informative in the analyses to estimate the RR. We refer to those individuals as the SCRI sample. The baseline sample may come from a different study population that did not receive this specific vaccine of interest, or come from the same study population as the SCRI sample but with the person-time in the risk interval for vaccinated individuals excluded. The design typically precludes more than one AE per person by only counting incident events that are the first diagnosis in X days, where X is usually larger than the post-exposure risk and control intervals combined; however, the design does allow the analysis of recurrent AEs. For simplicity, we consider first-over AE in our simulation study.

The *unadjusted SCRI analysis* (SCRI-u) fits a logistic regression model to the SCRI sample to estimate the RR. The dependent variable is whether or not the AE occurred in the risk interval. The regression model has no independent variable, so it is the estimate of the intercept that is of interest. A pre-specified offset term is used that is defined as the log of the ratio between the risk and control interval lengths.[6] In the absence of time-varying confounding and vaccine effect (RR=1), it reflects the log odds of having the AE occurring in the risk interval. This may be the analysis of choice when the impact of age on AE risk is negligible during the observation period. When baseline risks vary greatly between the risk and control intervals the RR estimate is likely to be biased. The same RR and variance estimates can be obtained by fitting a conditional Poisson regression model to the SCRI sample.[2,3]

Mathematically, suppose there are n_s subjects in the SCRI sample and for each subject i, $i = 1, \ldots, n_s$. Let L_i and C_i denote the lengths of the risk and control intervals respectively. For each day s inside either the risk or control interval, let $X_{s,i}$ denote individual i's age, $E_{s,i}$ denote the exposure status ($E_{s,i} = 1$ for days inside the risk interval and $E_{s,i} = 0$ for days inside the control interval), and $Y_{s,i}$ denote whether an AE occurred on day s. Since we assume one individual contributes at most one AE,

let $Z_i \equiv \sum_{s=1}^{L_i} Y_{s,i}$ denotes whether the AE occurred in the risk interval. Let $r_0(x)$ denote the true baseline risk function as a function of age $X = x$. Let RR_0 denote the true exposure-associated RR.

In SCRI-u, we fit the following intercept-only logistic regression model to the SCRI sample,

$logit(\Pr(Z_i = 1)) = \alpha$ with an offset term of $\log\left(\frac{L_i}{C_i}\right)$.

The exponent of the estimated intercept, $\exp(\hat{\alpha})$ is the RR estimate by SCRI-u.

The *gold-standard SCRI analysis* (SCRI-g) is very similar to SCRI-u with the exception that the offset term is defined as the log of the ratio between the true cumulative risk of having the AE occurring in the risk interval and the true cumulative risk of having the AE occurring in the control interval, i.e., $log\left(\frac{\sum_{s=1}^{L_i} r_0(X_{s,i})}{\sum_{s=L_i+1}^{L_i+C_i} r_0(X_{s,i})}\right)$.

The *SCRI analysis with fixed adjustment* (SCRI-f) is very similar to SCRI-g with the exception that the true cumulative risks for the risk and control intervals are replaced by the estimated cumulative risks for the risk and controls intervals respectively, i.e., the offset term is $log\left(\frac{\sum_{s=1}^{L_i} \hat{r}_0(X_{s,i})}{\sum_{s=L_i+1}^{L_i+C_i} \hat{r}_0(X_{s,i})}\right)$. Here $\hat{r}_0(x)$ denotes an estimated baseline risk function. The baseline risk estimates can be obtained from a published study or from fitting a Poisson regression model to an unexposed baseline sample. The potential limitation of SCRI-f is its failure to incorporate the uncertainty in baseline risk estimates that may lead to an underestimated variance estimate and a CI that is too narrow. For some analyses, the MS-PRISM Rotavirus Vaccines and Intussusception study[6] used the SCRI-f approach to adjust for age, using an external baseline risk estimate curve obtained from the literature,[13] but the validity of that approach has not been evaluated to date.

The *SCRI analysis with random adjustment* (SCRI-r) is technically very different from the other three SCRI analyses introduced above. It fits a joint conditional Poisson regression model to the pooled data from the SCRI and baseline samples to adjust for time-varying AE risks and estimate the exposure-associated RR simultaneously. Specifically, the unit of analysis is person-day. The dependent variable is whether or not an AE occurred on that day. The independent variables include the exposure status on that day (inside vs. outside the risk interval), explanatory variables of age (e.g., age, age^2), and the categorical subject ID for the individuals in the SCRI sample to denote their baseline incidence AE rates. This approach was developed after the MS-PRISM Rotavirus Vaccines and Intussusception study[6] was completed but has been implemented in the MS-PRISM Influenza Vaccines and Febrile Seizures study.[12]

Intuitively, the (unconditional) Poisson regression applied to the (unexposed) baseline sample estimates the baseline risk curve; the conditional Poisson regression applied to the SCRI sample (by including in the regression model the dummy variables for the n_s individuals in the SCRI sample) estimates the exposure-associated RR while adjusting for age-induced time-varying risk. By conducting one joint regression to the pooled data from the baseline and SCRI samples, we estimate baseline risk curve and

RR simultaneously. Thus, we account for variations from both samples such that the standard error (s.e.) for the RR estimate is correctly estimated.

Similar to SCRI-f, SCRI-r requires that (i) the baseline sample is comparable to the SCRI sample in terms of the baseline AE risks and (ii) the imposed parametric model for baseline risks estimation reflects the true relationship between the AE risk and age. In contrast to SCRI-f, SCRI-r appropriately accounts for the uncertainty in the estimated baseline risks, and yields valid point and interval estimates for RR when conditions (i) and (ii) mentioned above are satisfied.

III. SIMULATION STUDIES

We conducted two simulation studies, one mimicking the MS-PRISM Rotavirus Vaccines and Intussusception study[6] and the other mimicking the MS-PRISM Influenza Vaccines and Febrile Seizures study.[12] We designate the former as the primary simulation study because it presented a more interesting scenario in that the magnitude of the time-varying baseline risk was greater, and thus it allowed us to better assess the performance of the four SCRI analyses.

A. THE PRIMARY SIMULATION STUDY

1. Study Design

We designed our primary simulation study using empirical data from the MS-PRISM Rotavirus Vaccines and Intussusception study in children 5-36.9 weeks of age.[6]

For simplicity, we assumed in the simulation study all children entered the cohort at 7 weeks of age and had continuous enrollment throughout the first year of life to guarantee complete and accurate data ascertainment in exposure and AE status and dates. We assumed each child had an 80% probability of receiving the vaccine. The vaccination rate turned out to be irrelevant because we fixed the number of AEs in both the SCRI and baseline samples. While RotaTeq vaccine is administered in a series of 3 doses, in this simulation study, we assumed each child would receive at most one vaccine dose, which we do not anticipate would lead to loss of generalizability to other settings. For vaccinated children, the distribution of vaccination time was specified based on the empirical vaccine dose distributions from the RotaTeq all-dose analysis. In the US, the recommended ages for RotaTeq vaccination are 2, 4, and 6 months. The empirical distribution of RotaTeq doses observed in the MS-PRISM Rotavirus Vaccines and Intussusception study[6] is highly consistent with the guideline. Figure 1 presents the distribution of vaccination time among vaccinated children used in the simulation study.

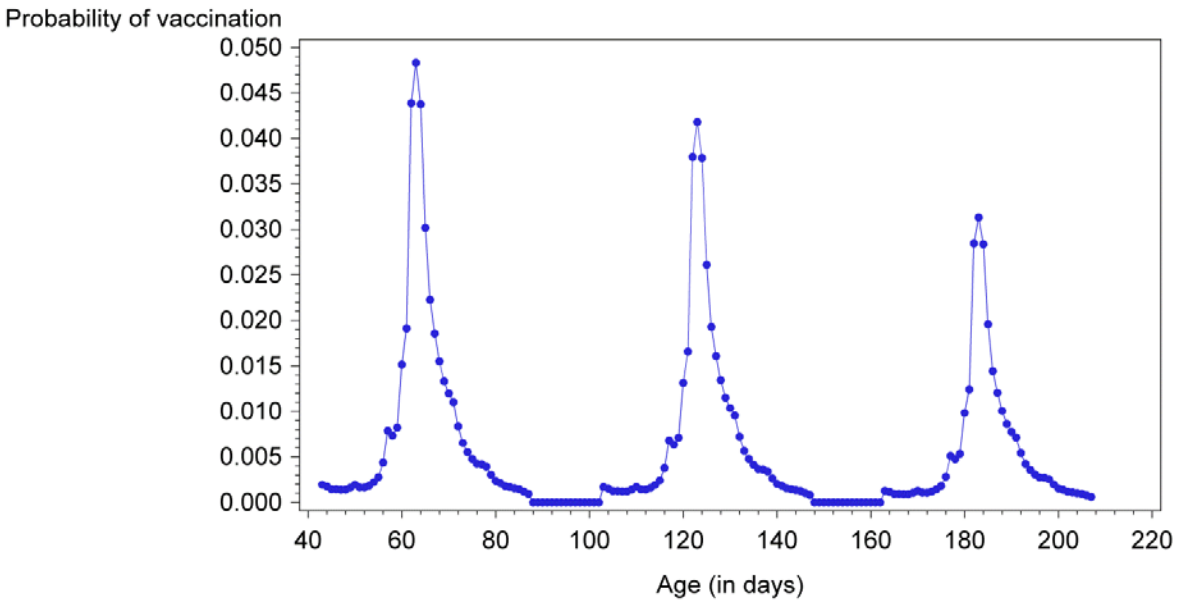

Figure 1: Distribution of vaccination by age

We considered three scenarios with different characteristics regarding the "true" baseline risks. In scenario 1, the baseline risk did not vary with age, and thus SCRI-g and SCRI-u are equivalent. The intention for this scenario was to assess the potential efficiency loss in SCRI-f and SCRI-r when adjustment for time-varying confounding is unnecessary. In scenario 2, the baseline AE incidence rates (per person-day) were extrapolated from the published age-specific (weekly) background rates extracted by Tate et al.[13] from the Healthcare Cost and Utilization Project's (HCUP) U.S. hospital-discharge data for 11 years during which no rotavirus vaccine was used. The scale of the incidence rates in the simulation study differed from that in the Tate study, but the shape of the curve was largely preserved. Note that for the adjustment of time-varying risks due to age, only the relative incidence rates (the shape of the curve) and not the absolute incidence rates matter. For each vaccinated child, the AE risks for the risk interval were elevated from the baseline risks by a range of RR's (1.0, 3.0, and 5.0). We only counted the first AE for each child. In scenario 3, we assessed the impact of site heterogeneity on the performance of the SCRI analyses. Specifically, we assumed each child had an equal chance of being from one of two sites. The baseline incidence rates for individuals from site 1 were the same as those in scenario 2, while the baseline incidence rates for individuals from site 2 were elevated by a constant of 30/100,000 per person-day. Site heterogeneity is common in multi-site studies due to differences in population characteristics, healthcare utilization patterns, or coding practice. While the specified between-site difference is larger than what one may expect to encounter in real-life applications, the motivation for this scenario was to assess the robustness of the methods under extreme scenarios. Figure 2 denotes the true baseline risks (solid line) and the estimated baseline risks (dotted curve) obtained by fitting a quadratic model to a baseline sample of 10,000 events.

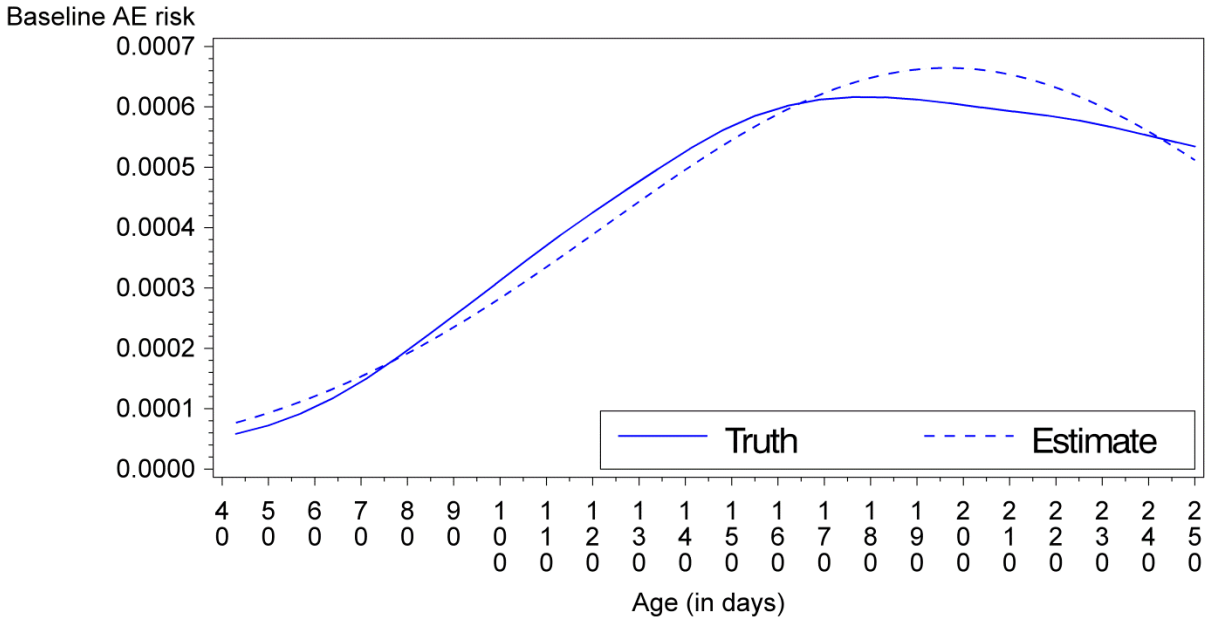

Figure 2: True and estimated baseline risks in scenario 2

In Figure 3, similarly as in Figure 2, the solid lines denote the true site-specific baseline risk curves and the dotted lines denote the estimated baseline risk curves obtained by fitting a quadratic model to a baseline sample of 10,000 events. The red curves are for site 1 and the blue curves are for site 2. For each vaccinated child, the AE risks were elevated from the baseline risks for the days that were inside the specified risk interval following vaccination by the specified RR. We only count the first AE for each child.

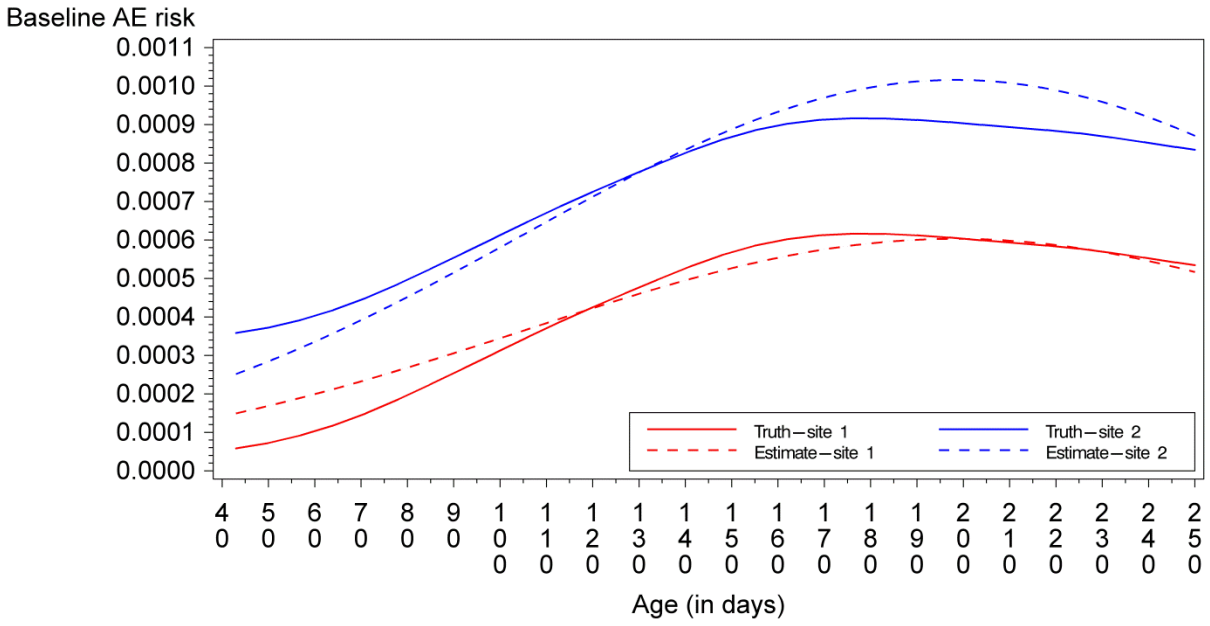

Figure 3: True and estimated baseline risks in scenario 3 with between-site heterogeneity

For all scenarios, the SCRI sample consists of children who were vaccinated and had an AE in either the risk or control interval. The baseline sample consists of all person-days for the unvaccinated children as well as the person-days outside the risk interval for the vaccinated children.

We varied the number of individuals in the SCRI sample, n_s, from 20 to 200; for each given n_s, we varied the number of AE in the baseline sample, n_b, from as small as $0.2n_s$ to $2n_s$; for each given (n_s, n_b), we varied the RR to be 1.0, 3.0, or 5.0. While the baseline sample is typically larger than the SCRI sample, or it would not likely be used, for completeness it is important to evaluate both smaller and larger values. We considered two risk intervals: 1-7 days and 1-21 days post vaccination and a control interval of 22-42 days post vaccination. For each data setting, we conducted 5,000 simulation replications to assess the performance on bias, variance, and CI coverage.

For SCRI-f and SCRI-r, we used the same explanatory variables of age to estimate baseline risks, i.e., (age, age^2) in scenarios 1 and 2 and (age, age^2, *I*(site 2)) in scenario 3. The imposed working models are expected to deviate from the true baseline risk functions in all three scenarios, particularly in scenario 3. The imposed working model assumes that the *log* incidence rates differ between the two sites by a constant and that the incidence rates themselves differ by a constant on the original scale. However, we felt it important to assess the robustness of SCRI-f and SCRI-r to model misspecification as the imposed parametric working model in real-life applications would be at best approximately true.

For each analysis, we output the number of simulation replications that successfully converged (N) out of the 5,000 replications, the 95% CI coverage rate, the median bias of $\log(\widehat{RR})$, and the robust

standard error of $\log(\widehat{RR})$ which equals the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35.[14] Note that the robust standard error equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers. In addition, we output the median CI width and a relative efficiency measure defined as the ratio of the median CI width for the respective SCRI analysis divided by the median CI width for the corresponding SCRI-g with the same n_s and *RR* values. We bold the CI coverage rates that are below 94.4% and italicize those that are above 95.6%. With 5,000 simulation replications, a valid CI with a nominal coverage of 95% has a 95% chance of observing a coverage rate between 94.4% and 95.6%.

2. Results

We present in Tables 1-2 the results with $n_s = 20$ and $n_s = 100$ in scenario 1. In the absence of time-varying confounding, SCRI-g and SCRI-u are equivalent. We present in Tables 3-4 the results with $n_s = 20$ and $n_s = 100$ in scenario 2. We present in Tables 5-6 the results with $n_s = 20$ and $n_s = 100$ in scenario 3. The results for other n_s values show similar trend and thus are not presented. In all settings considered in Tables 1-6, the risk interval is 1-21 days. Results for the risk interval of 1-7 days are similar and thus not presented.

When n_s is small, which may occur with rare or very rare AEs, the performance of all SCRI analyses tends to be less stable, especially when RR is high and/or with between-site heterogeneity. In some simulated datasets, the logistic regression model in SCRI-f did not converge because most or all AEs occurred in the risk interval. CIs tended to be conservative likely because the variance estimates were large due to unbalanced distribution of cases between the risk and control intervals.

In all considered settings, SCRI-g had good performance as expected. In scenarios 2 and 3, SCRI-u yielded biased point estimates and too-narrow CIs since it failed to adjust for the time-varying baseline risks. As expected, the impact of the bias becomes more significant as n_s increases and the estimation of RR becomes more precise. While this may be counter-intuitive at first glance, an example may help elucidate this concept. For example, if the true RR RR_0 equals 1, then the RR estimator from SCRI-u is biased due to time-varying baseline risks and converges to 1.5 with an infinite sample. Suppose the RR estimator remains at 1.4 from two SCRI-u analyses with 20 and 100 AEs respectively, but the CI with $n_s = 20$ is [0.9, 2.0] and the CI with $n_s = 100$ is [1.3, 1.6]. Then we would not reject the null with $n_s = 20$ but would wrongly reject the null (Type I error) with $n_s = 100$. Essentially, with a biased estimator, the more precision we have the more likely we would reject the true RR value.

In all considered settings, SCRI-r had good performance despite the incorrectly specified parametric model for baseline risks, demonstrating the robustness of SCRI-r to model misspecification. The efficiency of SCRI-r was comparable to SCRI-g when n_b was at least as large as n_s. With a smaller n_b, SCRI-r had about 5%-15% efficiency loss compared to SCRI-g. The unnecessary adjustment of time-varying baseline risks in scenario 1 did not have any major impact on the efficiency of SCRI-r.

In most considered settings, when n_b was greater than or equal to $0.5n_s$ (i.e., the number of events used to estimate the external baseline risk function are at least 50% of the number of events in the main

analytic sample), SCRI-f had good performance on bias and efficiency and was robust to model misspecification in baseline risks estimation. In scenario 3 with a large between-site heterogeneity, a higher cut-off of $n_b \geq n_s$ or $n_b \geq 2n_s$ was needed. We suggest an ad-hoc cut-off of $n_b \geq n_s$, which is expected to hold in most, if not all, applications. This means that the number of events used to estimate the baseline risk function must be greater than or equal to the number of events in the main analytic cohort.

Table 1: Performance of the four SCRI-analyses in scenario 1 (single site, without time-varying baseline risks) with varying $\boldsymbol{RR}$ **and** $\boldsymbol{n_b, n_s = 20}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency[#]
1.0	N/A	SCRI-u	5000	*95.7*	0	0.0042	1.762	1.00
	$0.2n_s$	SCRI-f	4999	**85.6**	0.045	0.0083	1.799	1.97
		SCRI-r	5000	95.4	0.029	0.0073	1.973	1.74
	$0.5n_s$	SCRI-f	5000	94.8	0.009	0.0069	1.789	1.65
		SCRI-r	5000	*95.7*	0.003	0.0069	1.853	1.64
	n_s	SCRI-f	5000	*95.7*	0.012	0.0065	1.788	1.54
		SCRI-r	5000	*96*	0.012	0.0065	1.813	1.55
	$2n_s$	SCRI-f	5000	95.6	0.01	0.0065	1.789	1.54
		SCRI-r	5000	*95.7*	0.007	0.0065	1.8	1.55
3.0	N/A	SCRI-u	4990	*96*	0	0.0057	2.024	1.00
	$0.2n_s$	SCRI-f	4988	**85.9**	0.134	0.0099	2.056	1.76
		SCRI-r	5000	*96*	0.06	0.009	2.273	1.6
	$0.5n_s$	SCRI-f	4982	*95.8*	0.052	0.008	2.028	1.41
		SCRI-r	5000	*96.5*	0.048	0.008	2.094	1.41
	n_s	SCRI-f	4986	*97.5*	0.02	0.0071	2.025	1.25
		SCRI-r	5000	*97.4*	0.017	0.0072	2.053	1.26
	$2n_s$	SCRI-f	4980	*97.3*	0.022	0.0069	2.025	1.21
		SCRI-r	5000	*96.9*	0.02	0.007	2.039	1.23
5.0	N/A	SCRI-u	4854	*96.6*	0.125	0.0086	2.455	1.00
	$0.2n_s$	SCRI-f	4844	**88.9**	0.121	0.0107	2.458	1.24
		SCRI-r	5000	**94.3**	0.102	0.0099	2.562	1.15
	$0.5n_s$	SCRI-f	4858	*96.9*	0.064	0.0092	2.455	1.07
		SCRI-r	5000	94.4	0.073	0.0094	2.488	1.09
	n_s	SCRI-f	4852	*97.1*	0.072	0.0091	2.455	1.06
		SCRI-r	5000	94.4	0.082	0.0092	2.472	1.07
	$2n_s$	SCRI-f	4867	*97*	0.09	0.009	2.455	1.05
		SCRI-r	5000	94.5	0.097	0.0091	2.463	1.05

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the

baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 2: Performance of the four SCRI-analyses in scenario 1 (single site, without time-varying baseline risks) with varying $\boldsymbol{RR}$ **and** $\boldsymbol{n_b, n_s = 100}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-u	5000	**94.1**	0	0.0029	0.785	1.00
	$0.2n_s$	SCRI-f	5000	**92.2**	0.009	0.003	0.787	1.03
		SCRI-r	5000	94.4	0.004	0.0031	0.845	1.04
	$0.5n_s$	SCRI-f	5000	94.5	0.004	0.0029	0.787	0.97
		SCRI-r	5000	95.2	-0.002	0.0029	0.812	0.97
	n_s	SCRI-f	5000	94.7	0	0.0028	0.786	0.96
		SCRI-r	5000	95.1	-0.005	0.0028	0.799	0.96
	$2n_s$	SCRI-f	5000	95.2	0.003	0.0029	0.786	0.98
		SCRI-r	5000	95.2	-0.002	0.0029	0.792	0.98
3.0	N/A	SCRI-u	5000	95.3	0	0.0034	0.905	1.00
	$0.2n_s$	SCRI-f	5000	**93.7**	0.021	0.0036	0.906	1.07
		SCRI-r	5000	95.1	0.011	0.0036	0.962	1.06
	$0.5n_s$	SCRI-f	5000	95	0.001	0.0033	0.905	0.98
		SCRI-r	5000	*95.7*	-0.001	0.0033	0.926	0.98
	n_s	SCRI-f	5000	95	0.004	0.0032	0.905	0.96
		SCRI-r	5000	95.1	0	0.0032	0.916	0.96
	$2n_s$	SCRI-f	5000	94.8	0.008	0.0034	0.905	1.02
		SCRI-r	5000	94.8	0.003	0.0034	0.911	1.02
5.0	N/A	SCRI-u	5000	95.5	0.049	0.0038	1.068	1.00
	$0.2n_s$	SCRI-f	5000	**93.7**	0.027	0.0041	1.069	1.08
		SCRI-r	5000	95.1	0.021	0.0041	1.104	1.07
	$0.5n_s$	SCRI-f	5000	95.3	0.008	0.0038	1.044	0.99
		SCRI-r	5000	95.5	0.004	0.0038	1.068	1.00
	n_s	SCRI-f	5000	95.5	0.015	0.0038	1.044	1.00
		SCRI-r	5000	*95.7*	0.009	0.0038	1.057	1.00
	$2n_s$	SCRI-f	5000	94.9	0.005	0.0039	1.044	1.03
		SCRI-r	5000	94.9	0.001	0.0039	1.05	1.03

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 3: Performance of the four SCRI-analyses in scenario 2 (single site, with time-varying baseline risks) with varying $\boldsymbol{RR}$ **and** $\boldsymbol{n_b}$**,** $\boldsymbol{n_s = 20}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-g	5000	*96.1*	-0.014	0.0064	1.794	1.00
		SCRI-u	5000	94.5	-0.201	0.0064	1.789	1.00
	$0.2n_s$	SCRI-f	4999	**82.9**	0.118	0.009	1.835	1.02
		SCRI-r	5000	95.6	0.067	0.0076	2.066	1.15
	$0.5n_s$	SCRI-f	5000	**92.6**	0.046	0.007	1.796	1.00
		SCRI-r	5000	95.1	0.047	0.007	1.904	1.06
	n_s	SCRI-f	5000	95.2	0	0.0064	1.791	1.00
		SCRI-r	5000	*95.7*	0.016	0.0065	1.84	1.03
	$2n_s$	SCRI-f	5000	95.4	-0.008	0.0066	1.791	1.00
		SCRI-r	5000	95.6	0.004	0.0066	1.811	1.01
3.0	N/A	SCRI-g	4995	*96.6*	-0.013	0.0078	1.924	1.00
		SCRI-u	4995	**91.7**	-0.251	0.008	1.913	0.99
	$0.2n_s$	SCRI-f	4989	**81.8**	0.182	0.0105	2.032	1.06
		SCRI-r	5000	95.6	0.108	0.0085	2.278	1.18
	$0.5n_s$	SCRI-f	4992	94.4	0.075	0.0079	2.019	1.05
		SCRI-r	5000	*96.5*	0.074	0.0079	2.091	1.09
	n_s	SCRI-f	4993	*96.7*	0.054	0.0077	1.955	1.02
		SCRI-r	5000	*97.1*	0.062	0.0077	2.047	1.06
	$2n_s$	SCRI-f	4991	*96.7*	0.039	0.0076	1.937	1.01
		SCRI-r	5000	*96.7*	0.055	0.0076	1.987	1.03
5.0	N/A	SCRI-g	4925	*97.3*	-0.032	0.0073	2.2	1.00
		SCRI-u	4925	**93.0**	-0.223	0.0067	2.191	1.00
	$0.2n_s$	SCRI-f	4918	**83.8**	0.237	0.0122	2.452	1.11
		SCRI-r	5000	95.5	0.143	0.01	2.569	1.17
	$0.5n_s$	SCRI-f	4926	95.4	0.073	0.0096	2.211	1.01
		SCRI-r	5000	*95.8*	0.088	0.0097	2.353	1.07
	n_s	SCRI-f	4928	*96.9*	0.019	0.0084	2.203	1.00

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
		SCRI-r	5000	*96.1*	0.039	0.0089	2.255	1.02
	$2n_s$	SCRI-f	4927	*97.2*	-0.002	0.0079	2.202	1.00
		SCRI-r	5000	*96.1*	0.025	0.0084	2.225	1.01

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 4: Performance of the four SCRI-analyses in scenario 2 (single site, with time-varying baseline risks) with varying ***RR*** **and** $\boldsymbol{n_b}$ **,** $\boldsymbol{n_s = 100}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-g	5000	94.9	0.007	0.0029	0.791	1.00
		SCRI-u	5000	**86.6**	-0.16	0.003	0.787	0.99
	$0.2n_s$	SCRI-f	5000	**91.5**	0.014	0.0033	0.793	1.00
		SCRI-r	5000	94.8	0.021	0.0032	0.876	1.11
	$0.5n_s$	SCRI-f	5000	**94.3**	0	0.0029	0.791	1.00
		SCRI-r	5000	95	0.007	0.0029	0.827	1.05
	n_s	SCRI-f	5000	94.4	-0.003	0.003	0.791	1.00
		SCRI-r	5000	95	0.004	0.003	0.809	1.02
	$2n_s$	SCRI-f	5000	94.6	-0.002	0.0028	0.79	1.00
		SCRI-r	5000	94.9	0.005	0.0028	0.799	1.01
3.0	N/A	SCRI-g	5000	95.3	-0.001	0.0032	0.877	1.00
		SCRI-u	5000	**86.9**	-0.154	0.0031	0.873	1.00
	$0.2n_s$	SCRI-f	5000	**92.1**	0.012	0.0037	0.876	1.00
		SCRI-r	5000	95.3	0.016	0.0036	0.959	1.09
	$0.5n_s$	SCRI-f	5000	94.4	0.01	0.0033	0.875	1.00
		SCRI-r	5000	95.3	0.018	0.0033	0.912	1.04
	n_s	SCRI-f	5000	95	-0.002	0.0032	0.875	1.00
		SCRI-r	5000	95.6	0.006	0.0033	0.891	1.02
	$2n_s$	SCRI-f	5000	94.6	-0.002	0.0033	0.875	1.00
		SCRI-r	5000	95	0.005	0.0032	0.883	1.01

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
5.0	N/A	SCRI-g	5000	94.7	0.009	0.0038	1.003	1.00
		SCRI-u	5000	**90.6**	-0.16	0.0041	0.999	1.00
	$0.2n_s$	SCRI-f	5000	**92.9**	0.032	0.0041	1.004	1.00
		SCRI-r	5000	95.6	0.037	0.004	1.084	1.08
	$0.5n_s$	SCRI-f	5000	94.5	0.008	0.0038	1.002	1.00
		SCRI-r	5000	95.1	0.019	0.0038	1.036	1.03
	n_s	SCRI-f	5000	95	0.003	0.0038	1.002	1.00
		SCRI-r	5000	95.4	0.015	0.0037	1.017	1.01
	$2n_s$	SCRI-f	5000	95.3	-0.004	0.0036	1.001	1.00
		SCRI-r	5000	95.6	0.004	0.0036	1.008	1.01

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 5: Performance of the four SCRI-analyses in scenario 3 (2 sites, with time-varying baseline risks) with varying ***RR*** **and** $\boldsymbol{n_b}$ **,** $\boldsymbol{n_s = 20}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-g	4965	*97.7*	-0.077	0.0087	2.195	1.00
		SCRI-u	4965	*98.3*	-0.288	0.0093	2.191	1.00
	$0.2n_s$	SCRI-f	4351	**74**	-0.081	0.0152	2.391	1.09
		SCRI-r	4965	*96.9*	-0.004	0.0096	2.563	1.17
	$0.5n_s$	SCRI-f	4939	**91**	-0.052	0.0095	2.195	1.00
		SCRI-r	4953	*97*	-0.006	0.0086	2.289	1.04
	n_s	SCRI-f	4963	*95.8*	-0.058	0.0086	2.192	1.00
		SCRI-r	4963	*96.8*	-0.025	0.0085	2.234	1.02
	$2n_s$	SCRI-f	4967	*97.3*	-0.079	0.0084	2.192	1.00
		SCRI-r	4967	*97.1*	-0.044	0.0084	2.214	1.01
3.0	N/A	SCRI-g	5000	*96*	-0.019	0.0064	1.793	1.00
		SCRI-u	5000	94.7	-0.201	0.0064	1.789	1.00
	$0.2n_s$	SCRI-f	4669	**65**	0.076	0.0163	1.916	1.07
		SCRI-r	5000	95.2	0.054	0.0084	2.199	1.23
	$0.5n_s$	SCRI-f	4999	**87.5**	0.018	0.0081	1.81	1.01

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
		SCRI-r	5000	95.5	0.043	0.0073	1.952	1.09
	n_s	SCRI-f	5000	**93.1**	-0.024	0.007	1.793	1.00
		SCRI-r	5000	95.3	0.019	0.0068	1.871	1.04
	$2n_s$	SCRI-f	5000	95.4	-0.045	0.0066	1.79	1.00
		SCRI-r	5000	*96*	-0.002	0.0065	1.823	1.02
5.0	N/A	SCRI-g	5000	*96.1*	0.036	0.0065	1.798	1.00
		SCRI-u	5000	**92.6**	-0.105	0.0065	1.789	1.00
	$0.2n_s$	SCRI-f	4695	**64.3**	0.171	0.0177	1.938	1.08
		SCRI-r	5000	95.5	0.093	0.0088	2.235	1.24
	$0.5n_s$	SCRI-f	4990	**86.8**	0.035	0.0083	1.838	1.02
		SCRI-r	5000	*96.1*	0.045	0.0072	1.976	1.10
	n_s	SCRI-f	5000	**93.3**	0	0.0072	1.805	1.00
		SCRI-r	5000	*95.8*	0.034	0.007	1.887	1.05
	$2n_s$	SCRI-f	5000	94.9	-0.021	0.0068	1.795	1.00
		SCRI-r	5000	*96*	0.016	0.0068	1.838	1.02

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 6: Performance of the four SCRI-analyses in scenario 3 (2 sites, with time-varying baseline risks) with varying ***RR*** **and** $\boldsymbol{n_b}$ **,** $\boldsymbol{n_s = 100}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-g	5000	95.3	-0.005	0.0035	0.949	1.00
		SCRI-u	5000	**90.9**	-0.167	0.0037	0.946	1.00
	$0.2n_s$	SCRI-f	5000	**88.2**	-0.044	0.0043	0.953	1.00
		SCRI-r	5000	95	0.005	0.004	1.065	1.12
	$0.5n_s$	SCRI-f	5000	**93.3**	-0.047	0.0037	0.949	1.00
		SCRI-r	5000	95.1	-0.005	0.0037	0.999	1.05
	n_s	SCRI-f	5000	**94.2**	-0.045	0.0037	0.948	1.00
		SCRI-r	5000	95.1	-0.001	0.0036	0.972	1.02
	$2n_s$	SCRI-f	5000	94.8	-0.049	0.0036	0.947	1.00
		SCRI-r	5000	95.4	-0.009	0.0036	0.959	1.01
3.0	N/A	SCRI-g	5000	*95.7*	-0.003	0.0029	0.791	1.00
		SCRI-u	5000	**85.6**	-0.16	0.003	0.788	1.00

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
	$0.2n_s$	SCRI-f	5000	**85.5**	-0.037	0.0037	0.794	1.00
		SCRI-r	5000	94.7	0.006	0.0034	0.92	1.16
	$0.5n_s$	SCRI-f	5000	**91.8**	-0.031	0.0032	0.791	1.00
		SCRI-r	5000	95.4	0.007	0.0032	0.847	1.07
	n_s	SCRI-f	5000	**93.4**	-0.045	0.0029	0.79	1.00
		SCRI-r	5000	95	0	0.0029	0.819	1.04
	$2n_s$	SCRI-f	5000	95	-0.047	0.0028	0.789	1.00
		SCRI-r	5000	*95.7*	-0.006	0.0028	0.805	1.02
5.0	N/A	SCRI-g	5000	95	0.007	0.0029	0.8	1.00
		SCRI-u	5000	**89.2**	-0.147	0.003	0.797	1.00
	$0.2n_s$	SCRI-f	5000	**85.3**	-0.021	0.0038	0.804	1.00
		SCRI-r	5000	95.1	0.023	0.0034	0.931	1.16
	$0.5n_s$	SCRI-f	5000	**91.5**	-0.031	0.0033	0.8	1.00
		SCRI-r	5000	95	0.01	0.0032	0.857	1.07
	n_s	SCRI-f	5000	**94.1**	-0.04	0.0031	0.798	1.00
		SCRI-r	5000	95.1	0.005	0.003	0.828	1.03
	$2n_s$	SCRI-f	5000	**93.2**	-0.047	0.003	0.797	1.00
		SCRI-r	5000	94.8	-0.004	0.0029	0.812	1.01

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

B. THE SECONDARY SIMULATION STUDY

1. Study Design

We designed a secondary simulation study using empirical data from the MS-PRISM Influenza Vaccines and Febrile Seizures study[12] consisting of children ages 6 through 59 months in selected MS Data Partners. We assumed each child had a 50% probability of receiving the influenza vaccine. The vaccination rate turned out to be irrelevant because we fixed the number of AEs in both the SCRI and baseline samples. For vaccinated children, the distribution of vaccination time in age month was specified based on the observed proportions of influenza vaccine doses by each age month. After the vaccination time in age-month was simulated, the exact vaccination time was simulated uniformly within that age-month. For simplicity, we assumed in the simulation study all vaccinated children had continuous enrollment for 365 days post vaccination. For unvaccinated children, we simulated their

index date based on the same distribution and assumed they also had continuous enrollment for 365 days following the index date.

In the simulation study, we used the weekly baseline incidence rates obtained in the febrile seizures study by fitting quadratic B-spline bases with a single internal knot of 82 weeks to the study data.[12] Then the weekly rates were linearly extrapolated to obtain daily rates. Figure 4 presents the distribution of vaccination time and the average baseline AE rate per 1,000 person-days for each age-month 6 through 59.

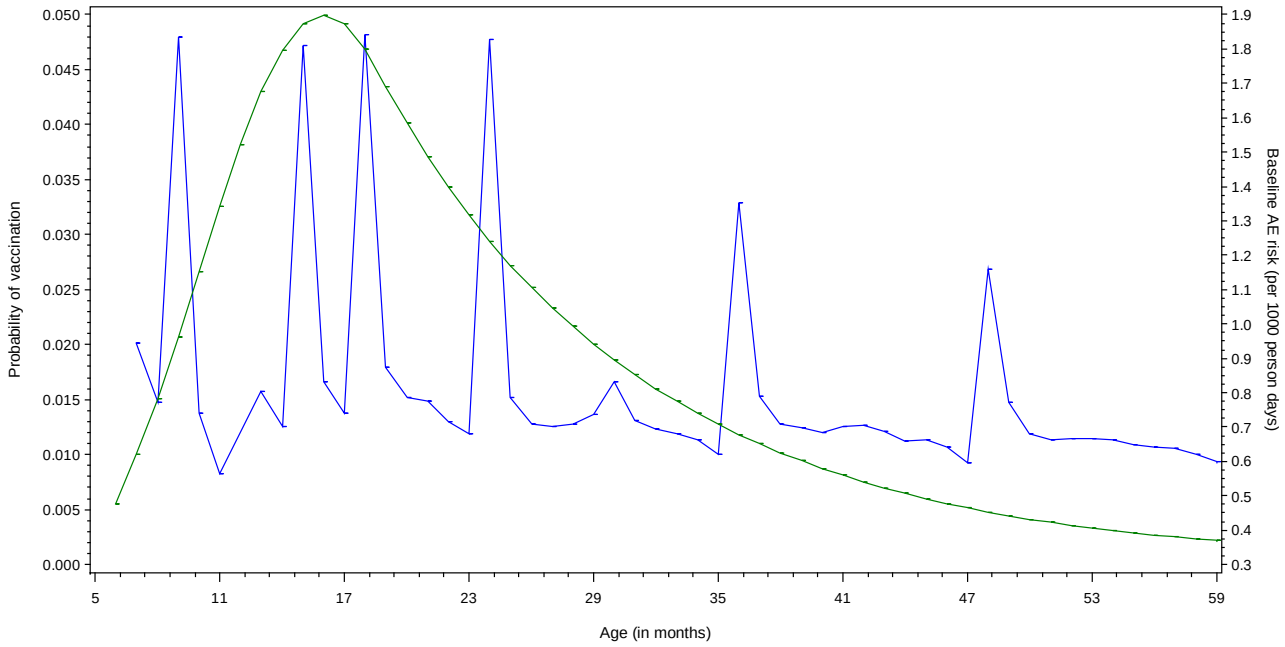

Figure 4: Probability of vaccination (blue) and average baseline AE risks (green), by age

For each vaccinated child, the AE risks for the risk interval were elevated from the baseline risks by the specified RR. We only counted the first AE for each child. The SCRI sample consists of children who were vaccinated and had an AE in either the risk or control interval. The baseline sample consists of all person-days for the unvaccinated children as well as the person-days outside the risk interval for the vaccinated children.

We varied the number of individuals in the SCRI sample, n_s, from 20 to 200; for each given n_s, we varied the number of AEs in the baseline sample, n_b, from as small as $0.2n_s$ to $2n_s$; for each given (n_s, n_b), we varied the RR to be 1.0, 3.0, or 5.0. While the baseline sample is typically larger than the SCRI sample, or it would not likely be used, for completeness it is important to evaluate both smaller and larger values. We considered a risk interval of 1-2 days post vaccination and a control interval of 15-21 days post vaccination. For each data setting, we conducted 5,000 simulation replications to assess the performance on bias, variance, and CI coverage.

For SCRI-f and SCRI-r, we used the linear and quadratic terms (age, age^2) to estimate varying baseline risks due to age. Same as in the primary simulation study, the imposed working model is expected to deviate from the true baseline risk function. However, we felt it is important to assess the robustness of SCRI-f and SCRI-r to model misspecification as the imposed parametric working model in real-life applications would be at best approximately true.

For each analysis, we output the number of simulation replications that successfully converged (N) out of the 5,000 replications, the 95% CI coverage rate, the median bias of $\log(\widehat{RR})$, the robust standard error of $\log(\widehat{RR})$ which equals the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35.[14] Note that the robust standard error equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers. In addition, we also output the median CI width and a relative efficiency measure which is defined as the ratio of the median CI width for the respective SCRI analysis divided by the median CI width for the corresponding SCRI-g with the same n_s and *RR* values. We bold the CI coverage rates that are below 94.4% and italicize those that are above 95.6%. With 5,000 simulation replications, a valid CI with a nominal coverage of 95% has a 95% chance of observing a coverage rate that is between 94.4% and 95.6%.

2. Results

We present in Tables 7-8 the results with $n_s = 20$ and $n_s = 100$. In this simulation study, the relative baseline incidence rates for the risk versus control intervals were much smaller compared to the primary simulation study in which the baseline risks varied dramatically over a short period of time. Therefore, the impact of time-varying baseline risks on SCRI analysis was much smaller. The unadjusted analysis SCRI-u had reasonable bias efficiency performance in the considered settings. SCRI-f had less than nominal level CI coverage probabilities in three settings with $n_b = 0.2n_s$, but otherwise performed very well. SCRI-r had good performance in all considered settings. The simulation results are consistent with the results from the actual MS-PRISM Influenza Vaccines and Febrile Seizures study[12] in which age adjustment had little impact on the RR estimates and CIs.

Table 7: Performance of the four SCRI-analyses in the secondary simulation study with varying ***RR*** **and** $\boldsymbol{n_b}$ **,** $\boldsymbol{n_s = 20}$

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency#
1.0	N/A	SCRI-g	4969	*97.8*	-0.111	0.0092	2.192	1.00
		SCRI-u	4969	*97.8*	-0.134	0.0093	2.191	1.00
	$0.2n_s$	SCRI-f	4969	**90.5**	-0.136	0.0092	2.193	1.00
		SCRI-r	4965	*97.4*	-0.094	0.0086	2.236	1.02
	$0.5n_s$	SCRI-f	4969	*96.7*	-0.1	0.0089	2.191	1.00
		SCRI-r	4968	*97.6*	-0.095	0.0088	2.195	1.00
	n_s	SCRI-f	4980	*97.6*	-0.108	0.009	2.191	1.00
		SCRI-r	4980	*97.6*	-0.107	0.009	2.193	1.00
	$2n_s$	SCRI-f	4958	*97.9*	-0.115	0.0092	2.191	1.00
		SCRI-r	4958	*97.8*	-0.114	0.0092	2.192	1.00
3.0	N/A	SCRI-g	5000	*96.2*	-0.03	0.0064	1.79	1.00
		SCRI-u	5000	*96.1*	-0.047	0.0064	1.789	1.00
	$0.2n_s$	SCRI-f	5000	**89.7**	-0.054	0.008	1.797	1.00
		SCRI-r	4996	*96.1*	-0.035	0.0071	1.879	1.05
	$0.5n_s$	SCRI-f	5000	95	-0.038	0.0067	1.789	1.00
		SCRI-r	5000	*95.7*	-0.036	0.0066	1.796	1.00
	n_s	SCRI-f	5000	*95.8*	-0.034	0.0064	1.789	1.00
		SCRI-r	5000	*96*	-0.036	0.0064	1.791	1.00
	$2n_s$	SCRI-f	5000	*96.2*	-0.036	0.0064	1.789	1.00
		SCRI-r	5000	*96.2*	-0.036	0.0064	1.79	1.00
5.0	N/A	SCRI-g	5000	*96*	0.051	0.0065	1.79	1.00
		SCRI-u	5000	*96*	0.049	0.0065	1.789	1.00
	$0.2n_s$	SCRI-f	5000	**90.4**	-0.035	0.0073	1.838	1.03
		SCRI-r	4987	*95.7*	-0.015	0.0069	1.896	1.06
	$0.5n_s$	SCRI-f	4999	95	0.02	0.0066	1.79	1.00
		SCRI-r	5000	95.4	0.02	0.0066	1.805	1.01
	n_s	SCRI-f	4999	*96.3*	0.031	0.0065	1.79	1.00
		SCRI-r	4999	*96.4*	0.029	0.0065	1.793	1.00
	$2n_s$	SCRI-f	5000	95.6	0.038	0.0065	1.789	1.00
		SCRI-r	5000	95.6	0.037	0.0065	1.791	1.00

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

[#]The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

Table 8: Performance of the four SCRI-analyses in the secondary simulation study with varying **RR and n_b , n_s=100**

RR	n_b	Analysis	N	95% CI coverage (%)	Median bias of $\log(\widehat{RR})$	Robust s.e. of $\log(\widehat{RR})^*$	Median CI width	Relative efficiency[#]
1.0	N/A	SCRI-g	5000	95.6	-0.004	0.0036	0.947	1.00
		SCRI-u	5000	95.6	-0.013	0.0037	0.946	1.00
	$0.2n_s$	SCRI-f	5000	94.8	-0.016	0.0035	0.946	1.00
		SCRI-r	5000	95.1	-0.016	0.0035	0.953	1.01
	$0.5n_s$	SCRI-f	5000	95.2	-0.015	0.0036	0.946	1.00
		SCRI-r	5000	95.2	-0.017	0.0036	0.949	1.00
	n_s	SCRI-f	5000	95.3	-0.012	0.0036	0.946	1.00
		SCRI-r	5000	95.4	-0.012	0.0036	0.947	1.00
	$2n_s$	SCRI-f	5000	95.1	-0.01	0.0036	0.946	1.00
		SCRI-r	5000	95.1	-0.011	0.0036	0.947	1.00
3.0	N/A	SCRI-g	5000	94.6	0.003	0.0028	0.788	1.00
		SCRI-u	5000	**94.1**	-0.006	0.0025	0.788	1.00
	$0.2n_s$	SCRI-f	5000	94.7	-0.008	0.0029	0.788	1.00
		SCRI-r	5000	95.2	-0.008	0.0029	0.796	1.01
	$0.5n_s$	SCRI-f	5000	95.3	-0.009	0.0029	0.788	1.00
		SCRI-r	5000	95.3	-0.01	0.0029	0.791	1.00
	n_s	SCRI-f	5000	94.4	-0.006	0.0029	0.788	1.00
		SCRI-r	5000	94.5	-0.007	0.0029	0.789	1.00
	$2n_s$	SCRI-f	5000	94.8	-0.004	0.0028	0.788	1.00
		SCRI-r	5000	94.8	-0.005	0.0028	0.789	1.00
5.0	N/A	SCRI-g	5000	*96.2*	0.009	0.003	0.797	1.00
		SCRI-u	5000	95.2	0.007	0.003	0.797	1.00
	$0.2n_s$	SCRI-f	5000	95.4	-0.006	0.0029	0.797	1.00
		SCRI-r	5000	95.7	-0.007	0.0029	0.805	1.01
	$0.5n_s$	SCRI-f	5000	95.4	-0.005	0.0029	0.797	1.00
		SCRI-r	5000	95.5	-0.006	0.0029	0.799	1.00
	n_s	SCRI-f	5000	95	-0.004	0.0029	0.797	1.00
		SCRI-r	5000	95	-0.007	0.0029	0.798	1.00
	$2n_s$	SCRI-f	5000	95	0.001	0.0029	0.797	1.00
		SCRI-r	5000	95	-0.002	0.0029	0.798	1.00

SCRI-u: the unadjusted SCRI analysis, SCRI-f: SCRI with fixed adjustment, SCRI-r: SCRI with random adjustment; n_s denotes the number of AEs in the SCRI sample, n_b denotes the number of AEs in the baseline sample, RR denotes the exposure-associated relative risk, N denotes the number of simulation replications that successfully converged.

*The robust s.e. of $\log(\widehat{RR})$ is calculated as the inter-quartile range of $\log(\widehat{RR})$ divided by 1.35, which equals the standard error if $\log(\widehat{RR})$ follows a normal distribution but is less sensitive to outliers.

#The relative efficiency is defined as the ratio of the median CI width for the specific SCRI analysis divided by the median CI width for the corresponding SCRI-g analysis with the same n_s and *RR*.

IV. DISCUSSION

We evaluated the performance of four SCRI analyses: SCRI-u, SCRI-g, SCRI-r, and SCRI-f in a variety of settings, including two vaccine-outcome examples that were previously evaluated in the MS-PRISM 2010-11 Influenza Vaccines and Febrile Seizures study[12] and the Rotavirus Vaccines and Intussusception study.[6] We also varied the number of AEs in the SCRI sample, the number of AEs in the baseline sample, RR, and risk interval length. We found that SCRI-f and SCRI-r have very good performance with respect to bias and efficiency, in most considered settings. Based on our simulation results, we anticipate that SCRI-f can be used as a good alternative to SCRI-r when the number of AEs in the baseline sample is at least the number of AEs in the SCRI sample ($n_b \geq n_s$) which is expected to hold in most, if not all, applications.

In the analyses for a SCRI design, the first step is to consider whether adjustment of time-varying baseline risks is needed. Our simulation results show that SCRI-r and SCRI-f with a quadratic adjustment have mild to little efficiency loss, depending on the value of n_b/n_s, in the absence of time-varying baseline risks. In contrast, SCRI-u may lead to severely biased point and interval estimates when baseline risks vary greatly between the risk and control intervals. Therefore, we suggest considering using either SCRI-r or SCRI-f whenever there is a potential concern about time-varying confounding.

If access to individual-level data for the baseline sample is available, SCRI-r in general is the preferred approach as it appropriately accounts for the random variation in the SCRI sample and baseline risk estimation. Otherwise, SCRI-f can be used instead as long as the baseline sample is reasonably large. Both SCRI-r and SCRI-f require that (i) the baseline sample be comparable to the SCRI sample regarding the baseline AE risks and (ii) the imposed working parametric model for baseline risks estimation be (approximately) correct. Condition (i) is achieved by selecting the appropriate data and condition (ii) is achieved by analyzing the selected data appropriately.

The selection of the baseline sample needs to consider the tradeoff between validity and sample size. The population from an external sample may differ in important characteristics such as age, race, or socio-economic status and may come from a different healthcare system with different exposure and/or AE diagnosis patterns. A historical sample may differ due to temporal trend in AE incidence rates or an evolving exposed population. We want to clarify that the robust performance of SCRI-f and SCRI-r to model misspecification in age adjustment does not imply that the methods are robust to systematic differences between the baseline and SCRI samples. An internal, concurrent sample is more likely to be valid than an external sample or a historical sample, although a restricted internal sample may have a smaller number of AEs. These limitations can be properly addressed, as illustrated by our simulation

study in which it was shown that the minimum number of AEs in the baseline sample needed for SCRI-f and SCRI-r to have reasonable, stable performance, is moderate (namely between $0.5n_s$ and n_s).

The choice of an appropriate working model for baseline risks estimation is beyond the "scope of" SCRI design and has been studied extensively.[15-18] In many settings, low-order polynomial structures (e.g., linear, quadratic, cubic) provide reasonable approximations to smooth mean functions,[15,16] which was confirmed in our simulation study by the robust performance of the SCRI-r and SCRI-f analyses. More advanced smoothing techniques such as B-splines[17,18] can be used if needed. We suggest conducting sensitivity analyses by using different parametric and semi-parametric working models. If the point estimates are similar, the parametric model may be preferred due to its simplicity and efficiency; otherwise the semi-parametric model may be preferred due to its robustness to model misspecification.

Extra caution needs to be taken to handle boundaries or sub-ranges with extremely low AE incidence. With very few AEs observed corresponding to these age values, both parametric and semi-parametric models can be problematic. Parametric models extrapolate information across the entire age range which may be particularly inappropriate in these settings. Semi-parametric models are much more flexible in functional forms but may encounter numeric issues due to a small number of AEs. We suggest conducting additional sensitivity analyses excluding the AEs from the SCRI sample with case onset dates in these problematic scenarios.

The SCRI design has been used in combination with a sequential analytic approach, the binomial maximized sequential probability ratio test (Binomial MaxSPRT),[19] in prospective surveillance studies in Vaccine Data Link Project[20] and MS-PRISM systems to sequentially monitor the AE risks following vaccine or drug exposure.[11,21,22] The fixed adjustment approach applies directly to the sequential version. We expect the sequential SCRI-f to have good performance as long as the baseline sample is reasonably large.

V. REFERENCES

1. Greene SK, Kulldorff M, Lewis EM, et al. Near Real-Time Surveillance for Influenza Vaccine Safety: Proof-of-Concept in the Vaccine Safety Datalink Project. American Journal of Epidemiology 2010;171:177-88.
2. Greene SK, Rett M, Weintraub ES, et al. Risk of Confirmed Guillain-Barre Syndrome Following Receipt of Monovalent Inactivated Influenza A (H1N1) and Seasonal Influenza Vaccines in the Vaccine Safety Datalink Project, 2009-2010. American Journal of Epidemiology 2012;175:1100-9.
3. Baker MA, Lieu TA, Li L, et al. A Vaccine Study Design Selection Framework for the Postlicensure Rapid Immunization Safety Monitoring Program. Am J Epidemiol 2015.
4. Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I, Smeeth L. Self-Controlled Case Series Analysis with Event-Dependent Observation Periods. Journal of the American Statistical Association 2011;ahead of print.
5. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Stat Med 2006;25:1768-97.
6. Yih WK, Lieu TA, Kulldorff M, et al. Intussusception Risk after Rotavirus Vaccination in US Infants. New England Journal of Medicine 2014;370:503-12.
7. France EK, Glanz JM, Xu S, et al. Safety of the trivalent inactivated influenza vaccine among children: a population-based study. Arch Pediatr Adolesc Med 2004;158:1031-6.
8. Kramarz P, DeStefano F, Gargiullo PM, et al. Does influenza vaccination exacerbate asthma? Analysis of a large cohort of children with asthma. Vaccine Safety Datalink Team. Arch Fam Med 2000;9:617-23.
9. O'Leary ST, Glanz JM, McClure DL, et al. The Risk of Immune Thrombocytopenic Purpura After Vaccination in Children and Adolescents. Pediatrics 2012;129:248-55.
10. Akinsanya-Beysolow I, Advisory Committee on Immunization Practices (ACIP), ACIP Child/Adolescent Immunization Work Group. Advisory Committee on Immunization Practices Recommended Immunization - Schedules for Persons Aged 0 Through 18 Years — United States, 2014. MMWR 2014;63:110-2.
11. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM, Group VSDRCAIW. 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;30:2024-31.
12. Kawai AT, Martin D, Kulldorff M, et al. Febrile Seizures After 2010-2011 Trivalent Inactivated Influenza Vaccine. Pediatrics 2015;136:e848-55.
13. Tate JE, Simonsen L, Viboud C, et al. Trends in intussusception hospitalizations among US infants, 1993-2004: Implications for monitoring the safety of the new rotavirus vaccination program. Pediatrics 2008;121:E1125-E32.
14. Casella G, Berger RL. Statistical Inference: Duxbury Pacific Grove, CA; 2002.
15. Rudin W. Functional analysis. New York, NY McGraw-Hill; 1973.
16. Rudin W. Principles of mathematical analysis. New York: McGraw-Hill; 1976.
17. De Boor C. A practical guide to splines. New York: New York : Springer-Verlag, c1978.; 1978.
18. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. Statistical Science 1996;11:89-102.
19. 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.
20. DeStefano F. The Vaccine Safety Datalink project. Pharmacoepidemiol Drug Saf 2001;10:403-6.
21. Greene SK, Kulldorff M, Lewis EM, et al. Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol 2010;171:177-88.

22. Lee GM, Greene SK, Weintraub ES, et al. H1N1 and Seasonal Influenza Vaccine Safety in the Vaccine Safety Datalink Project. Am J Prev Med 2011;41:121-8.