In this paper, we propose new methods for analyzing cases of vaccine adverse events spontaneously reported to a surveillance database. The methods use the self-controlled case series approach, extended in several ways with parametric and nonparametric assumptions to account for the specific features of the data (large amount of underreporting and variation of reporting with time since vaccination). This work was motivated by the documented risk of intussusception after RotaShield vaccination (Wyeth-Lederle Vaccines, Radnor, Pennsylvania) and used worldwide spontaneous reports of intussusception occurring after Rotarix vaccination (GlaxoSmithKline Biologics, Research Triangle Park, North Carolina) collected between January 2004 and February 2010. The estimated risk during the 3- to 7-day period after vaccination was approximately 5 times higher after dose 1 of Rotarix than after dose 2, which is similar to published findings on the same topic. We undertook a large simulation study to evaluate the performance of the method in different scenarios, including its robustness to different sample sizes and time-dependent reporting functions. The bias was generally small, the type I error rate was correctly controlled, and the power to detect a risk ratio of 4 was satisfactory, provided that the sample size was over 100. The proposed methods are an effective way to explore and quantify vaccine safety signals from spontaneous reports.

adverse events; intussusception; pharmacoepidemiology; pharmacovigilance; rotavirus; self-controlled case series; spontaneous reporting; vaccines

Abbreviation: SCCS, self-controlled case series.

Vaccination is one of the great advances in public health, one that has dramatically reduced the burden of a number of infectious diseases. Vaccines are delivered to healthy individuals, often to infants, so their safety is of paramount importance. Safety signals must be investigated as promptly as possible to invalidate or confirm them. The efficacy and safety of a vaccine are studied in preclinical and clinical trials before licensing (1); however, rare adverse effects will be observed only when a large population has been vaccinated. For instance, the first rotavirus vaccine, RotaShield (Wyeth-Lederle Vaccines, Radnor, Pennsylvania), initially marketed in the United States in 1998, was withdrawn in 1999 after an increased risk of intussusception was observed on days 3–7 after the first dose; the risk was on the order of 1 per 10,000 vaccinated infants (2). Spontaneous reporting is at the core of pharmacovigilance systems, some of which are vaccine-specific, like the Vaccine Adverse Event Reporting System in the United States. Spontaneous reports of adverse vaccine effects are used for hypothesis generation because underreporting is temporally related to vaccination (3–8). Current quantitative methods for spontaneous reports are either focused on a specific safety problem, using extra information such as the size of the exposed population (3–6), or unfocused and automated, using data-mining methods (7–15). The data-mining methods have been shown to be effective (16–23), yet detected signals need to be further assessed through more robust epidemiologic studies.

Pharmacoepidemiologic studies are typically case-control studies or retrospective cohort studies; however, case series models have been developed to estimate the increase in risk.
New SCCS Method for Analyzing Adverse Vaccine Events

of adverse events after vaccination (24–26). The self-controlled case series (SCCS) method uses only cases, with these cases acting as their own controls; hence, it adjusts implicitly for all time-independent confounders, while allowing explicit modeling of the variation of incidence with age. The method is derived from a cohort model that is then conditioned on the number of events experienced by each individual, whose observation period is partitioned into control and risk periods. The relative incidence is estimated by within-individual comparison of the incidences of events in different periods. The SCCS method has been successfully used to assess several vaccine adverse effects (2, 27–37). Several extensions exist (38–45), including a sequential setting for surveillance systems (46, 47). However, the latter is intended to be used for prospective monitoring of vaccine safety based on active surveillance. Indeed, as with other study designs, in order to produce unbiased relative incidence estimates, case collection for the SCCS method needs to be independent of vaccination; this condition is obviously not met with spontaneous reporting, since cases are reported because of their temporal association with vaccination.

In this paper, we extend the SCCS model to the analysis of spontaneous reports aiming at rapid evaluation of a risk. It is proposed not as an alternative to an epidemiologic study but as a tool for exploring a spontaneous reporting signal prior to a pharmacoepidemiologic study, if any. We were motivated by the worldwide spontaneous reports of intussusception after vaccination with a new rotavirus vaccine, the 2-dose monovalent Rotarix vaccine (GlaxoSmithKline Biologics, Research Triangle Park, North Carolina), licensed in 2006 in Europe and in 2008 in the United States. Before licensing, it had been tested in 75,000 infants. In the absence of vaccination, the rate of intussusception is known to vary with age during the first year of life. In the United States, it increases from less than 5 per 100,000 in the first 9 weeks of life to a maximum of 62 per 100,000 in weeks 26–29, and then decreases to 26 per 100,000 in week 52 (48). Rotarix vaccine is given in 2 doses, the first between 6 and 15 weeks of age and the second at least 4 weeks after the first dose and before week 24.

On the basis of the RotaShield experience, we compared the time distribution of cases after the first dose with the time distribution of cases after the second dose (49). Each case contributed only to the single dose period in which the event occurred. This analysis is detailed below, along with the underlying assumptions. We also provide an alternative SCCS extension that models the reporting process. Both approaches were applied to the Rotarix data and evaluated with a simulation study.

MATERIALS AND METHODS

Models

Standard SCCS approach. We describe the standard SCCS model in the context of Rotarix vaccination. The observation period is defined as days 0–30 after each vaccine dose, day 0 being the vaccination day. We selected this observation period to avoid any overlap between risk periods after the first and second doses, since the vaccine doses have to be separated by at least 30 days according to the recommended vaccine schedule. With the previous vaccine, RotaShield, an elevated risk of intussusception was observed during the first 2 weeks after administration of the first dose of vaccine, the risk being higher on days 3–7 (2). We therefore partitioned the 0- to 30-day risk period into 4 risk periods: 0–2, 3–7, 8–14, and 15–30 days after each dose, the last period being considered a reference period.

For each of the first 3 risk periods, the ratio of incidence during the risk period to incidence during the reference period was estimated. We call this ratio the relative incidence ratio, relative to the reference period.

A standard model for the incidence rate \( \mu_i(t,d) \) in the observation period after dose \( d \), without age effects, is

\[
\mu_i(t,d) = \exp\left( \alpha_{id} + \sum_{j=1}^{4} \beta_{dj} X_j(t) \right)
\]

for case \( i = 1, \ldots, n \), occurring at time \( t \) after dose \( d \), where \( X_j(t) \) is the indicator for risk period \( j \) (\( j = 1, 2, \ldots, 4 \)), period \( j = 4 \) is chosen as the reference period (\( \beta_{14} = \beta_{24} = 0 \)), and \( \alpha_{id} \) is an individual effect. The parameters \( \beta_{dj} \) denote the effect of the vaccine in period \( j \) contrasted with that in the reference period for the same dose \( d \).

The SCCS model assumes that the events occur according to a Poisson process with incidence rate \( \mu_i(t) \). The parameters are estimated by the maximum likelihood method. Considering only cases, each individual contribution to the likelihood is calculated as conditional on the number of events (see Web Appendix, available at http://aje.oxfordjournals.org/). As a result, the time-constant effects \( \alpha_{id} \) in model 1 cancel out after conditioning and are implicitly adjusted for.

We extend the SCCS model to analyze spontaneous reports. The main idea is that the probability of a case to be reported after dose \( d \), \( p_i(t,d) \), is a function of time since vaccination. The incidence rate of intussusception is \( \mu_i(t,d) \) at time \( t \) after dose \( d \), and the reported events occur according to a Poisson process with incidence rate \( \lambda_i(t,d) = p_i(t,d) \times \mu_i(t,d) \). Furthermore, for each individual, we condition on the total number of events after each dose, rather than on the total number of events within that person’s observation period.

We now introduce 2 sets of assumptions for \( p_i(t,d) \), corresponding to a nonparametric approach and a parametric approach, respectively. The nonparametric approach assumes that 1) the reporting probability depends only on time since vaccination and on dose number and 2) the ratio of the reporting probabilities between doses is constant. This allows estimation of the ratio of the relative incidences after dose 1 and after dose 2, but not of the relative incidences for each dose. This approach takes advantage of the multidose vaccine scheme and can only be used when the risks are expected to be different after each dose. The ratio of relative incidences has previously been used for vaccine comparison (45).

The second approach assumes that the reporting probability decreases exponentially with time since vaccination, which allows estimation of the relative incidences separately for each dose. The models are described briefly below and detailed in the Web Appendix. All computations are done using R, version 2.14.2 (R Foundation for Statistical Computing, Vienna, Austria (http://cran.r-project.org/)).
First extension of the SCCS model: nonparametric approach. Let \( d_i \) denote the dose at which the event occurs for individual \( i \). The observed sample consists of \( n \) cases of intussusception reported within 30 days after either dose 1 or dose 2 vaccination. We now write the incidence rate \( \mu_i(t, d_i) \) as

\[
\mu_i(t, d_i) = \exp \left( \alpha_i + \gamma I_{d_i=1} + \sum_{j=1}^{4} \left( \beta_j + \gamma_j I_{d_i=1} \right) X_j(t) \right),
\]

where \( j = 1, 2, 3, 4 \) codes for the 4 different periods, period \( j = 4 \) being the reference period (\( \beta_4 = 0, \gamma_4 = 0 \)); \( X_j(t) \) is the indicator for risk period \( j \), common to both doses; \( I_{d_i=1} \) is the indicator function of dose 1 (1 if \( d_i = 1 \) and 0 otherwise); and \( \alpha_i \) is an individual effect. The parameter \( \gamma \) is an overall dose effect, and the parameters \( \beta_j \) denote the effect of the vaccine in period \( j \) contrasted with the period of reference for dose 2; \( \exp(\beta_j) \) are the corresponding relative incidences. The parameters \( \gamma_j \) denote the interaction effects of dose on vaccine effect. As non-time-varying effects, the \( \alpha_i \) and \( \gamma \) cancel out after conditioning and are not estimable. Additional effects may be included, such as age at vaccination (categorized); in that case, main-effect parameters as well as interactions with risk period parameters are added in model 2, though only the interactions are estimable.

For the analysis of Rotarix data, we included age at vaccination in the model by considering 3 age groups: \( 1–3 \) months, 4 months, and \( \geq 5 \) months.

Because each case is reported with probability \( p_i(t, d_i) \), the observed process has the following incidence rate:

\[
\lambda_i(t, d_i) = p_i(t, d_i) \times \exp \left( \alpha_i + \gamma I_{d_i=1} + \sum_{j=1}^{4} \left( \beta_j + \gamma_j I_{d_i=1} \right) X_j(t) \right).
\]

When \( p_i(t, d_i) \) is a constant that is independent of \( t \) and of the dose, the parameters \( \beta_j \) and \( \gamma_j \) are estimated respectively by

\[
\exp(\hat{\beta}_j) = \frac{n_{2j}/e_j}{n_{24}/e_4},
\]

and

\[
\exp(\hat{\gamma}_j) = \frac{n_{1j}/n_{14}}{n_{2j}/n_{24}},
\]

where \( n_{dj} \) is the number of cases after vaccination dose \( d \) in risk period \( j \) and \( e_j \) is the number of days in period \( j \).

Let \( p_i(t, d_i) \), the time-dependent reporting probability after dose \( d \), be \( \pi_1(t) \) if case \( i \) occurred after dose 1 and \( \pi_2(t) \) if it occurred after dose 2. Under the assumption of a constant ratio for reporting probabilities between doses, one obtains

\[
\pi_1(t) = \kappa \pi_2(t),
\]

where \( \kappa \) is not estimable. Under the assumption stated in equation 6, \( \beta \) cannot be estimated, since \( \pi_2(t) \) is unknown.

Figure 1. Worldwide numbers of spontaneous reports of intussusception following Rotarix (GlaxoSmithKline Biologicals, Research Triangle Park, North Carolina) rotavirus vaccination, January 2004–February 2010.

Maximum likelihood estimation involves unknown and non-estimable quantities \( I_j \) in place of \( e_j \) in equation 4, where \( I_j \) is defined by

\[
I_j = \int_{Period_j} \pi_2(u) \, du,
\]

but \( \gamma_j \) is still estimated by means of equation 5, since those unknown quantities cancel out (see Web Appendix). The key assumption here is that \( \pi_2(t) \) and therefore \( I_j \) are the same for all individuals. Standard packages for conditional Poisson regression or conditional logistic regression can be used to fit model 3, using the logarithm of \( e_j \) as the offset as in a classical SCCS. Parameters \( \gamma_j \) are estimated, and their 95% confidence intervals are obtained. The null hypotheses \( \gamma_j = 0 \) are tested by considering exclusion of zero from the 95% confidence interval, or by the likelihood ratio test. If a null hypothesis is rejected with \( \hat{\gamma}_j > 0 \), it means that the relative incidence in period \( j \) as compared with period 4 is larger after dose 1 than after dose 2.

Second extension of the SCCS model: parametric approaches. In the first parametric model, the reporting probability \( p_i(t, d) \) is assumed to be an exponential function of time since dose \( d \):

\[
p_i(t, d) = \exp(\theta t).
\]

For this approach, dose-specific effects \( \beta_{dj} \) are directly estimated by maximum likelihood methods, though the estimation can no longer be done via a Poisson model (see Web Appendix).

Consistency between the parametric and nonparametric approaches is assessed by calculating the ratio between the relative incidences for each dose:

\[
\frac{\exp(\hat{\beta}_{dj})}{\exp(\hat{\beta}_{d2})},
\]

to be compared with \( \exp(\hat{\gamma}_j) \) from the nonparametric model. The 95% confidence interval is obtained for the \( \beta \)'s from the profile likelihood (see Web Appendix), and null hypotheses are tested using the confidence intervals.
Two extensions of model 7 are considered. The first is

\[ p_i(t, d) = \exp\left[\theta_1 I_1 + \theta_2 I_2 t\right], \]

where the reporting probability depends on the dose \( d \). The model can further be extended to

\[ p_i(t, d, u_i) = \exp\left[\theta_1 U_1 + \theta_2 U_2 + \theta_3 U_3 t\right], \]

where the reporting probability depends on the age at vaccination \( u_i \), with age-group indicators \( U_{im} \) defined by \( U_{im} = 1 \) if \( u_i \) belongs to the \( m \)th age group and 0 otherwise. Likelihood ratio tests are used to compare each extension with model 7 (i.e., models 8 and 9 are compared with model 7), since model 7 is nested in models 8 and 9.

As a sensitivity analysis, we also considered a hyperbolic reporting probability which has a longer tail than an exponential distribution with the same variants (dose or age dependency).

\[ p(t, d) = \frac{1}{(1 + t)^\omega}. \]

Because hyperbolic and exponential models were not nested, we used the Akaike Information Criterion to compare the fits (50).

**Materials**

*The Rotarix data.* A line listing of 370 worldwide reports of intussusception cases arising after Rotarix vaccination between January 2004 and February 2010 was made available to us by the French National Agency for the Safety of Medicines and Health Products (Agence Nationale de Sécurité du Médicament et des Produits de Santé, Saint Denis, France), which had obtained these data from GlaxoSmithKline. There were no patients with recurrent intussusception. Data on the following variables were recorded: sex, country, date of vaccination, date of intussusception, age at intussusception, and time between most recent dose and intussusception. The redundancy in the information allowed estimation of some missing data and identification of inconsistencies. When the data were inconsistent, priority was given to time between...
most recent dose and intussusception, to date of intussuscep-
tion, and to age, in that order. Despite this effort, 27% of the
cases had to be excluded because of incomplete data. In a
second step, cases of intussusception occurring more than 30
days after the most recent vaccination dose were excluded.
Finally, the working data set contained 151 cases, 111 occur-
ring after the first vaccine dose and 40 after the second dose.

Figure 1 details the data processing, and Table 1 and Figure 2
describe the data.

The nonparametric model and the parametric models were
fitted to these data.

Simulated data. A simulation study was performed, gen-
erating data according to the Rotarix situation, and the gen-
eration process was similar to the one presented by Kuhnert
et al. (42). Details on the data generation can be found in the
Web Appendix.

Each simulation scenario was repeated 1,000 times, with
data sets of 100, 200, or 500 reported cases. To evaluate the
type I error, simulations under the null hypothesis (i.e., with
all relative risks relative to period 4 equal to 1) were performed.
Other simulations, under the alternative hypothesis of a risk
multiplied by 4 only during the period 3–7 days after the
first dose, were performed to evaluate the power (Figure 3).

Three functions of time were investigated for the reporting
probability after dose 1:
• A constant reporting probability.
• A reporting probability decreasing with time according to
  a logistic function.
• A reporting probability decreasing with time correspond-
ing to a half-normal distribution.

These reporting probabilities are displayed in Figure 4. The
reporting probabilities after dose 1 and after dose 2 were
assumed to be proportional.

Lastly, errors in the recording of the dose number were con-
sidered, assuming that 30% of events recorded as occurring
after the second dose would be wrongly recorded as occur-
ring after the first dose. All parameters and chosen values used
in the simulation study are described in Table 2. The same
nonparametric model and the 3 parametric exponential models
as those used for the real Rotarix data set were fitted to the
simulated data.

RESULTS

Rotarix data

Results from the different models are summarized in Tables 3
and 4. Based on the nonparametric approach (Table 3), the

Table 2. Characteristics and Values Used in a Simulation Study of
Intussusception Following Rotavirus Vaccinationa

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, no.</td>
<td>100, 200, 500</td>
</tr>
<tr>
<td>Risk ratio for risk of intussusception after first dose during days 3–7 after vaccination versus days 15–30 after vaccination</td>
<td>1, 4</td>
</tr>
<tr>
<td>Reporting probability as a function of time</td>
<td>Constant, logistic, half-normal</td>
</tr>
<tr>
<td>Reporting ratio for dose 2 versus dose 1</td>
<td>1, 0.7</td>
</tr>
<tr>
<td>Misspecification rate (dose wrongly recorded as 1)</td>
<td>0, 0.3</td>
</tr>
</tbody>
</table>

a See text for details.

Table 3. Nonparametric Estimatessa of Intussusception Risk Parameters Among Worldwide Spontaneous Reports of Intussusception Following Rotarix® Rotavirus Vaccination, January 2004–February 2010

<table>
<thead>
<tr>
<th>Model</th>
<th>0–2 (j = 1)</th>
<th>3–7 (j = 2)</th>
<th>8–14 (j = 3)</th>
<th>15–30 (j = 4) (exp(4))</th>
<th>Degrees of Freedom</th>
<th>–2 Log Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>No age effect</td>
<td>1.13</td>
<td>0.38, 3.35</td>
<td>3.24*</td>
<td>0.64</td>
<td>1</td>
<td>360.10</td>
</tr>
<tr>
<td>.Age effect</td>
<td>1.56</td>
<td>0.45, 5.45</td>
<td>4.97*</td>
<td>0.42</td>
<td>12</td>
<td>349.66</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

* P < 0.05 (significantly different from 1).

a Each parameter exp(γ) represents the ratio (dose 1:dose 2) of risk ratios (period j/period 4).

b GlaxoSmithKline Biologics, Research Triangle Park, North Carolina.

c Reference period.
### Table 4. Parameter Estimates of Intussusception Risk Parameters Among Worldwide Spontaneous Reporting Systems Following Rotavirus Vaccination. January 2004–February 2010

<table>
<thead>
<tr>
<th>Days After Vaccination (Period)</th>
<th>Time-Dependent Reporting Probability Model (Model 7)</th>
<th>Time-Dependent Reporting Probability Model (Model 8)</th>
<th>Time- and Age-Dependent Reporting Probability Model (Model 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>1</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>2</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>3</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>4</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>5</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>6</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>7</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>8</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>9</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>10</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>11</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>12</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>13</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>14</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>15</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>16</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>17</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>18</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>19</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>20</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>21</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>22</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>23</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>24</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>25</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>26</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>27</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>28</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>29</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
<tr>
<td>30</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
<td>$\exp(\beta_{ij})$</td>
</tr>
</tbody>
</table>

#### Simulated data

Selected results from the simulation study are shown in Table 5 for simulations under the null hypothesis of no increase in risk after vaccination. The data presented are the mean values of the 1,000 estimates. With 1,000 simulations, the 95% fluctuation interval of the type I error was 3.65%–6.35% for a nominal 5% value, and the estimated type I error rates were almost always within these limits. The testing procedure was found to be conservative, since type I error rates were generally less than 5%, particularly with sample sizes of 100 or 200 and with the nonparametric approach. Almost all of the parameter estimates were below 0; the relative bias was never larger than 8.5%, and it decreased generally when the sample size increased.

Results from simulations with the same scenarios under the alternative hypothesis of a risk multiplied by 4 during the period 3–7 days after dose 1 are shown in Table 6. The power to detect departure from the null hypothesis was approximately 50% for a sample size of 100, 80% for a sample size of 200, and at least 98% for a sample size of 500, except for the parametric model with a dose-dependent reporting probability. Under this model, power was approximately 28% with a sample size of 100, but likelihood ratio tests comparing the 3 model fits almost never selected this model (data not shown). The estimates were often lower than the simulated value of 1.39, indicating a moderate negative bias. Coverage probabilities of the profile likelihood-based confidence intervals were never below 92% (data not shown).

When dose misspecification was introduced, the results did not deteriorate to the point of being unusable.
accounted for this variation. Our excess risk was estimated under 2 sets of hypotheses that explained by the variation of underreporting with time. This finding is similar to those reported by Patel et al. (37), who conducted an SCCS study and a case-control study in Brazil and Mexico. They estimated an incidence ratio of 5.3 in Mexico during the first week after the first dose as compared with the period beyond the third week, using the SCCS method, and obtained a much less clear picture for Brazil, for which they discussed the possible impact of different vaccine recommendations (37). In addition, our estimates were consistent between our 2 approaches.

The simulation study confirmed the good properties of the new methodology. We investigated the null hypothesis and a strong alternative hypothesis where the incidence ratio between dose 1 and dose 2 is equal to 4. The results showed good control of the type I error. Overall, incidence ratios were underestimated, possibly owing to uncontrolled age effects. We took advantage of the SCCS framework, which uses only cases for modeling the within-subject incidence ratio

<table>
<thead>
<tr>
<th>Sample Size, no.</th>
<th>Simulated Reporting Probabilitya</th>
<th>Nonparametric Estimation Method</th>
<th>Parametric Estimation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reporting Probability Is Dose- and Age-independent</td>
<td>Reporting Probability Is Dose-dependent</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>α, %</td>
</tr>
<tr>
<td>100</td>
<td>Constant</td>
<td>-0.06</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>0.02</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>-0.07</td>
<td>0.96</td>
</tr>
<tr>
<td>200</td>
<td>Constant</td>
<td>-0.09</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>-0.08</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>-0.04</td>
<td>0.69</td>
</tr>
<tr>
<td>500</td>
<td>Constant</td>
<td>-0.05</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>-0.07</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>-0.02</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Abbreviation: SE, standard error.

a Described in detail in the Web Appendix.

Tables 2 and 3 for corresponding results). Parameter estimates were more negatively biased, but type I error rates were still almost always within the interval 3.65%–6.35%, and power to detect departure from the null hypothesis was at least 70% for a sample size of 500.

**DISCUSSION**

Spontaneous reporting of adverse drug reactions constitutes a unique resource with which to identify drug safety signals for rare adverse events. The present work was motivated by the observation of an excess risk of intussusception after the first dose of Rotarix vaccine, which could not be completely explained by the variation of underreporting with time. This excess risk was estimated under 2 sets of hypotheses that accounted for this variation. Our findings are similar to those of Patel et al. (37), who conducted an SCCS study and a case-control study in Brazil and Mexico. They estimated an incidence ratio of 5.3 in Mexico during the first week after the first dose as compared with the period beyond the third week, using the SCCS method, and obtained a much less clear picture for Brazil, for which they discussed the possible impact of different vaccine recommendations (37). In addition, our estimates were consistent between our 2 approaches.

The simulation study confirmed the good properties of the new methodology. We investigated the null hypothesis and a strong alternative hypothesis where the incidence ratio between dose 1 and dose 2 is equal to 4. The results showed good control of the type I error. Overall, incidence ratios were underestimated, possibly owing to uncontrolled age effects. We took advantage of the SCCS framework, which uses only cases for modeling the within-subject incidence ratio

<table>
<thead>
<tr>
<th>Sample Size, no.</th>
<th>Simulated Reporting Probabilitya</th>
<th>Nonparametric Estimation Method</th>
<th>Parametric Estimation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reporting Probability Is Dose- and Age-independent</td>
<td>Reporting Probability Is Dose-dependent</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
<td>SE</td>
<td>Power, %</td>
</tr>
<tr>
<td>100</td>
<td>Constant</td>
<td>1.40</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>1.43</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>1.47</td>
<td>0.81</td>
</tr>
<tr>
<td>200</td>
<td>Constant</td>
<td>1.38</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>1.40</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>1.39</td>
<td>0.55</td>
</tr>
<tr>
<td>500</td>
<td>Constant</td>
<td>1.36</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Logistic</td>
<td>1.36</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Half-normal</td>
<td>1.38</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Abbreviation: SE, standard error.

a Described in detail in the Web Appendix.
and added extra hypotheses. As a result, the new SCCS analysis relies on some of the same modeling assumptions as the classical one, namely 1) a Poisson model for case occurrence and 2) the a priori choice of nonoverlapping risk/control periods. The first assumption is not particularly limiting, while the second one restrains the use of the proposed model to signals for which some a priori information is available; this was the case here because of the experience of an increased risk with the first antirotavirus vaccine RotaShield during days 3–7 after the first dose. Additional assumptions were required. Because the whole vaccine history of reported cases is not available in spontaneous reporting, the observation period had to be shorter than the time lag between vaccination doses—1 month here for the Rotarix data. Other hypotheses concerned the reporting process. If the vaccination schedule includes 2 doses, the nonparametric approach allows estimation of the relative incidence ratio between doses, under the assumption of proportional reporting processes after the 2 doses. This assumption does not seem unrealistic. If there is a single dose, the parametric approach can be used. However, the parametric assumptions are stronger. We adopted an exponential model for the reporting probability and introduced some flexibility by letting the exponential parameter depend on the dose or on the age groups. As a sensitivity analysis, we also used a hyperbolic model for the reporting probability, and findings were consistent with the 2 approaches. The present spontaneous report data did not allow for sorting out the effects of the vaccine and of underreporting. To further evaluate the reporting distribution as a function of time, one could explore the reporting probability of an event known to be unrelated to a given type of vaccine.

A limitation of our study was the quality of the data. The information on adverse events that occurred after vaccination was not collected in a standardized manner in the spontaneous reports made to vaccine manufacturing companies or regulatory agencies. Vaccination dates, event dates, and the patient’s age are often recorded imprecisely, and age is recorded in various units (days, weeks, months, years), with varying degrees of precision; the dose number is not always specified. The deletion of incomplete cases could have introduced a bias if completion was associated with the event; however, this seems unlikely. One could also envisage the use of some imputation method or expectation-maximization algorithm to handle missing data.

The present work was motivated by the experience with Rotarix 2-dose vaccination. In the case of a 3-dose vaccination schedule, the parametric approach can be used directly, but the nonparametric method must be adapted by selecting an appropriate reference dose.

In conclusion, we have proposed a statistical approach to the modeling of spontaneous reporting data that can be used in a timely way to explore and quantify a vaccine safety signal rapidly, before media attention (or any other event) induces perturbation in the probability of reporting.

ACKNOWLEDGMENTS

Author affiliations: Biostatistics Team, U1018, Centre for Research in Epidemiology and Population Health, Institut National de la Santé et de la Recherche Médicale (INSERM), Villejuif, France (Sylvie Escolano, Pascale Tubert-Bitter); Unité Mixte de Recherche Scientifique 1018, Université Paris-Sud, Villejuif, France (Sylvie Escolano, Pascale Tubert-Bitter); and Biostatistics and Epidemiology Service, Institut Gustave Roussy, Villejuif, France (Catherine Hill).

We thank Drs. Alexis Jacquet and Fanny Raguideau of the Agence Nationale de Sécurité du Médicament et des Produits de Santé (Saint Denis, France) and the Pharmacovigilance Department of GlaxoSmithKline for their valuable help.

Conflicts of interest: none declared.

REFERENCES


