American Journal of Epidemiology Advance Access published September 27, 2013

American Journal of Epidemiology

DOI: 10.1093/aje/kwt207

Practice of Epidemiology

Self-Controlled Case Series and Misclassification Bias Induced by Case Selection from Administrative Hospital Databases: Application to Febrile Convulsions in Pediatric Vaccine Pharmacoepidemiology

Catherine Quantin*, Eric Benzenine, Michel Velten, Frédéric Huet, C. Paddy Farrington, and Pascale Tubert-Bitter

*Correspondence to Prof. Catherine Quantin, Service de Biostatistique et Informatique Médicale, Centre Hospitalier Universitaire, BP 77908, 21079 Dijon Cedex, France (e-mail: catherine.quantin@chu-dijon.fr).

Initially submitted March 8, 2013; accepted for publication July 26, 2013.

Vaccine safety studies are increasingly conducted by using administrative health databases and self-controlled case series designs that are based on cases only. Often, several criteria are available to define the cases, which may yield different positive predictive values, as well as different sensitivities, and therefore different numbers of selected cases. The question then arises as to which is the best case definition. This article proposes new methodology to guide this choice based on the bias of the relative incidence and the power of the test. We apply this methodology in a validation study of 4 nested algorithms for identifying febrile convulsions from the administrative databases of 10 French hospitals. We used a sample of 695 children aged 1 month to 3 years who were hospitalized in 2008–2009 with at least 1 diagnosis code of febrile convulsions. The positive predictive values of the algorithms ranged from 81% to 98%, and their sensitivities were estimated to be 47%–99% in data from 1 large hospital. When applying our proposed methods, the algorithm we selected used a restricted diagnosis code and position on the discharge abstract. These criteria, which resulted in the selection of 502 cases with a positive predictive value of 95%, provided the best compromise between high power and low relative bias.

administrative data; bias; febrile convulsions; pharmacoepidemiology; positive predictive value; power; vaccines

Abbreviations: ICD-10, *International Classification of Diseases, Tenth Revision*; PMSI, Programme de Médicalisation des Systèmes d'Information; PPV, positive predictive value; SCCS, self-controlled case series.

Drug safety studies are increasingly based on information from large computerized health databases (1–11). Medical administrative databases represent low-cost and extensive sources of information on large populations. Hospital databases allow efficient counting of cases with events resulting in hospitalization. With linked hospital and prescription or health claim databases, association studies of adverse drug reactions may be undertaken. In this context, data accuracy is emerging as an important issue (12–24).

Because drug safety events are rare, commonly used designs for assessing the relationship between drug exposures and adverse events are case-control studies or studies based on cases only, notably the self-controlled case series (SCCS) (25, 26). The SCCS model was initially proposed and used to study vaccine adverse effects (4, 27–37) (see the recent review by Weldeselassie et al. (38)) but has since been applied to a wider range of drugs (39–49). One advantage of the SCCS model is that time-constant confounding factors are implicitly adjusted, this feature being crucial when drug exposure data are extracted from administrative databases in which only partial information on risk factors is available or in the presence of confounding by indication. Besides confounding, there are other potential sources of bias in epidemiologic studies, notably misclassification of the disease outcome. This has been extensively investigated for case-control studies (50–54) but not yet for SCCS studies.

In this article, we focus on case ascertainment from hospital databases in the context of SCCS studies of vaccine safety.

Often, several criteria are available to define the events, which, when evaluated against a "gold standard," may yield different positive predictive values (PPVs). These definitions may have different sensitivities and may yield different numbers of cases. The question then arises as to which event definition criterion is best. We propose new methodology to guide this choice. We consider the impact of false positives on the estimation of relative incidence in SCCS in terms of bias and power. Then, we report the results of a French validation study in which information recorded in administrative hospital databases was compared with that recorded in medical charts. We chose febrile convulsions among young children as the event of interest because they are known to be induced occasionally by some vaccines. Finally, we describe several case selection algorithms, assess their properties, and compare their merits.

MATERIALS AND METHODS

The SCCS model

For simplicity, we consider the relative incidence ρ of an acute adverse event such as febrile convulsions occurring in the presence or absence of recent vaccination, in which recent vaccination is defined by a single risk period of length e_1 (55). Additionally, it is assumed that all subjects are observed for the same period of time, and that they share the same constant background incidence. The observation period that is not part of the risk period is the control period, of length e_0. For all vaccinated individuals, the proportion of the observation period at risk is $r = e_1/(e_1 + e_0)$. For simplicity, we ignore age effects.

The SCCS analysis is conditional on the total number of events observed during the observation period for each individual. Suppose that $n_0 + n_1$ events arise in vaccines, where n_0 is the number of events occurring in the control period, and n_1 is the number of events occurring in the risk period. Maximum likelihood estimation of ρ involves only vaccinated individuals and, in the present scenario, takes the following simple form (Web Appendix 1, available at http://aje.oxfordjournals.org/):

$$\hat{\rho} = \frac{n_1 e_0}{n_0 e_1}.$$

Estimates are obtained under the assumption that case assessment is perfectly accurate, and maximum likelihood estimators are therefore asymptotically unbiased.

Misclassification bias

Now we assume that, in addition to n_0 and n_1 true events arising as before (with relative incidence ρ between risk and control periods), there are also m_0 and m_1 false positive events, which we will call pseudoevents, arising in the control period and the risk period, respectively. We shall assume that these pseudoevents arise with a relative incidence θ between the risk and control periods. So, if the pseudoevents are unrelated to the vaccine, $\theta = 1$.

Maximum likelihood applied to these data yields an asymptotically unbiased estimate

$$\hat{\tilde{\rho}} = \frac{(n_1 + m_1)e_0}{(n_0 + m_0)e_1}$$

of the parameter

$$\tilde{\rho} = \rho PPV_0 + \theta(1 - PPV_0), \quad (1)$$

where PPV_0 denotes the positive predictive value of events in the control period (Web Appendix 2). In particular, if

$$1 \leq \theta \leq \rho, \quad (2)$$

then

$$1 \leq \theta \leq \tilde{\rho} \leq \rho.$$

As a result, when there is no association between the vaccine and true events ($\rho = 1$), then the estimator is unbiased, because from equation 2, $\tilde{\rho} = 1$ as well. Equation 1 leads to the following expression for the relative bias in the relative incidence induced by the presence of false positive events:

$$\frac{\tilde{\rho} - \rho}{\rho} = \frac{(1 - PPV_0)(\theta - \rho)}{\rho}. \quad (3)$$

Often, equation 2 can be assumed to hold, in which case the inclusion of pseudoevents biases the estimated relative incidence away from its true value ρ and toward 1. When $\theta = \rho$, so that the vaccine is equally associated with the occurrence of pseudoevents and true events, then $\tilde{\rho} = \rho$, and ρ is correctly estimated. However, $\tilde{\rho}$ is biased even when $\theta = 1$, that is, when pseudoevents arise at the same rate in the risk and control periods. Assuming that $\theta = 1$ and that PPV_0 is known, one can correct the estimate of ρ by using the equation

$$\rho = \frac{\tilde{\rho} - (1 - PPV_0)}{PPV_0} \quad (4)$$

from equation 1 and applying it to $\hat{\tilde{\rho}}$ to get an estimate $\hat{\rho}$. It might be easier to obtain the overall PPV among all individuals, PPV_{ov} (Web Appendix 2). When r is small, as is often the case in vaccine studies, then PPV_0 and PPV_{ov} take similar values. Otherwise, one can correct the estimate of ρ, still assuming $\theta = 1$, based on PPV_{ov} (Web Appendix 2). Finally, corrected confidence intervals for ρ can be derived by means of the delta method or by the inverse transformation in equation 4 applied to the confidence bounds for $\tilde{\rho}$. With the delta method, one can additionally take into account the uncertainty in the PPV estimate (Web Appendix 3).

Tradeoffs between power and bias

Increasing the PPV to exclude pseudoevents may result in the exclusion of true events; therefore, in certain circumstances, higher power may be achieved with lower PPV. Thus, when several algorithms to select events are assessed in terms of

PPVs, one may consider the power of the test of no association in addition to the bias as a criterion to guide the choice of the algorithm that will be used to conduct the SCCS analysis.

In the presence of pseudoevents, the test that is performed is that of H_0: $\tilde{\rho} = 1$. Assuming $\theta = 1$, or more generally equation 2, $\rho = 1$ if and only if $\tilde{\rho} = 1$. For simplicity, we investigate the 1-sided test of H_0: $\tilde{\rho} = 1$ versus H_1: $\tilde{\rho} > 1$ (a 2-sided alternative hypothesis could also be considered). The power at significance level α with n total events (true events and pseudoevents) in vaccinees is

$$P(n, PPV_0, \rho, \theta) = 1 - \Phi\left(\frac{Z_\alpha\sqrt{r(1-r)} - \sqrt{n}(\tilde{\pi} - r)}{\sqrt{\tilde{\pi}(1-\tilde{\pi})}}\right), \quad (5)$$

where Φ is the cumulative distribution of the standard normal, Z_α is the $100(1-\alpha)$ percentile of the standard normal, and

$$\tilde{\pi} = \frac{\tilde{\rho}r}{\tilde{\rho}r + 1 - r}$$

is a function of ρ, PPV_0, and θ (Web Appendix 4). In PPV assessment studies, the total number of events (including both true events and pseudoevents), m, is usually available, as well as PPV_{ov}. Assuming a short risk period (so r is small), it is then possible to substitute mp for n in equation 5 (P being the vaccine coverage) and PPV_{ov} for PPV_0. Alternatively, PPV_0 can be expressed in terms of PPV_{ov} (Web Appendix 2). Thus, the power curve may be obtained for different values of ρ and θ.

It is important to note that higher powers do not necessarily correspond to larger n values or to smaller PPV_0. For example, suppose that $\theta = 1$, $\rho = 3$, and $r = 0.05$, and consider the 3 power values, 90%, 80%, and 60%. Equation 5 then makes it possible to derive the number of events in vaccinated individuals n to achieve these fixed powers as a function of PPV_0. Figure 1 displays 3 pairs (PPV_0, n) achieving these power values and ranked according to them as pair 1, pair 2, and pair 3. Ranking them according to the PPV_0 or to n would produce different ordering. In this example, a high PPV_0 of 90% requires 67 vaccinated cases to achieve 80% power (pair 2). A much less specific case definition, with PPV_0 of 55%, requires more cases ($n = 77$) to achieve a much smaller power (60%, pair 3), whereas a slightly lower PPV_0 (85%) requires two-thirds more cases ($n = 111$) to achieve the best of the 3 powers (90%, pair 1). Consequently, it is possible that, in certain settings, higher power may be achieved with a case definition with lower PPV. Thus, power criteria might lead to a different algorithm ranking from that obtained according to PPV_0 alone or n alone. We propose that both the relative bias (which increases as PPV_0 decreases) and the power be considered in selecting an algorithm.

We explore these ideas by using a validation study that aimed to estimate the accuracy of febrile convulsion event selection from a French hospital administrative database and by constructing optimized selection algorithms.

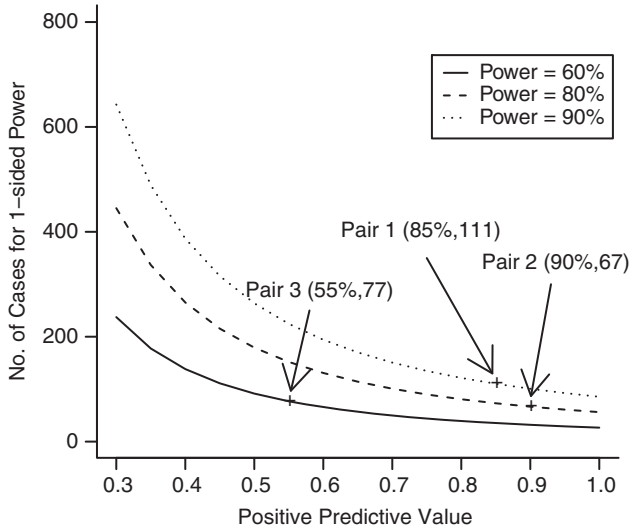

Figure 1. The number of vaccinated cases required to achieve a power of 60%, 80%, or 90% (1-sided test, 5% nominal type I error risk). The relative incidence of pseudoevents is $\theta = 1$, the relative incidence is $\rho = 3$, and the proportion of the observation period at risk is $r = 0.05$. The numbers in parentheses are the positive predictive values for events in the control period followed by the number of events required in vaccinated individuals.

Collection of administrative data in France

Inspired by the American diagnosis-related group model (56), the French established the collection of medical administrative data through the Programme de Médicalisation des Systèmes d'Information (PMSI) in 1991 (57) and extended it in 1997 to all French health care facilities (58). Initially designed to analyze hospital activity and contribute to the strategic development of facility plans, it has become an instrument of financial management. Since 2008, each hospital's budget (for all public and private hospitals) depends entirely on the medical activity recorded in this database (59), which compiles discharge abstracts for every admission. Information in these abstracts includes both administrative data (age; sex; postal code of residence; year, month, and type of admission; year, month, and type of discharge; and facility status) and medical data. Diagnoses identified during admission are coded according to the *International Classification of Diseases, Tenth Revision* (ICD-10). The condition involving the greatest use of resources during hospitalization is recorded as the main diagnosis, with other diseases listed as associated diagnoses (60). Administrative data collection rules are decided by the government and applied nationally. Each facility produces its own standardized data, which are then anonymized and compiled at the national level. These data have already proven to be useful in estimating the incidence or prevalence of cancers and some medical procedures in France (61–66). They are now linked to the national health insurance claims in a single database, the Système National d'Informations Inter Régime d'Assurance Maladie (67, 68), which offers new prospects for national pharmacoepidemiology studies (69).

The study population

A 2-step procedure was followed. First, to select the hospitals, anonymous administrative data from all hospitals in the 3 French administrative areas of Bas Rhin, Côte-d'Or, and Doubs were obtained from the national database. These included all hospitalizations of children between 1 month and 3 years of age with at least 1 diagnosis of febrile convulsions (ICD-10 code R56.0 or R56.8), whatever its position on the discharge abstract (principal diagnosis or associated diagnosis) who were discharged between January 1, 2008, and December 31, 2009. We kept only hospitals with at least 5 hospitalizations, resulting in the selection of 10 public hospitals, including 3 university hospitals (in Besançon, Dijon, and Strasbourg, France). In the second step, we constituted the study sample by extracting the hospitalizations from the "local" administrative database of each of these 10 hospitals to collect named patient data for medical chart review.

Algorithms for identifying febrile convulsions

Algorithms for identifying febrile convulsions were constructed from the available variables, including age, sex, length of stay, and diagnoses and procedures. We were interested in identifying cases for SCCS analyses and thus sought a good PPV. However, for power considerations, we also needed to select enough cases, so we measured the sensitivity of our algorithms. This was done by using local resources available in Dijon, France. The following 3 operating characteristics were then assessed: the PPV, sensitivity, and number of cases. The validation study involved the consultation of medical records and was approved by the national data protection authority (Commission Nationale de l'Informatique et des Libertés).

Estimation of the PPV

The PPV corresponds to the probability that hospitalization with a diagnosis code of febrile convulsions in the administrative database is indeed related to a febrile convulsion. The medical record was considered to be the gold standard. The events selected from the administrative data in each of the 10 hospitals were compared with the data from the matching medical records. This comparison was made by using a validation sheet and reading the following components of the medical record: discharge letters (to referring and primary care physicians), nursing records, hospital reports, and procedure reports (particularly for radiology). In case of any doubt regarding the interpretation of diagnoses, the opinion of an expert was sought.

The true positives were the hospitalizations for which febrile convulsions were identified both in the administrative abstracts and the corresponding medical records. False positives were febrile convulsions recorded in the administrative abstracts that were not identified as such in the patients' medical records.

Estimation of the sensitivity

Sensitivity is the probability that administrative data correctly identify hospitalizations for febrile convulsions. The registry of convulsion cases established by the pediatric emergency department of Dijon University Hospital (Dijon, France) was considered to be the gold standard. Therefore, the estimation of sensitivity was calculated from Dijon databases only. Inpatient data in this computerized registry were linked to the administrative database by using the following variables: first name, last name, and date of birth. The true positives were hospitalizations for which the febrile convulsion was identified both in the administrative database and in the Dijon registry. The false negatives were the hospitalizations listed in the Dijon registry of febrile convulsion cases but not reported in the administrative database.

Regression models to identify factors associated with false positives

A multivariable logistic regression model was used to estimate how the probability of a false positive depended on the patient's characteristics. The model was assessed by using data from each of the 10 hospital data sets for which the "true" status was known from the medical records. All events identified as positive were retrieved from the administrative database. Among these subjects, the binary response variable was assigned the value of 1 (false positive) or 0 (true positive). The independent variables included in this model were year of admission, season, hospital type (university or nonuniversity hospital) and administrative area, admission in an emergency department, length of hospital stay, age (considered in broad categories), diagnoses (code and type, main or associated) of febrile convulsions, previous neurological diseases (coded in the current or in a previous hospitalization), and procedures such as magnetic resonance imaging, tomodensitometry, or electroencephalography.

Selection of the algorithm

Different algorithms were proposed based on the variables selected by the regression model as significantly associated with a decrease in the rate of false positives. For all of these algorithms, PPV and sensitivity were estimated. We also derived the power functions and relative biases for each candidate algorithm under different scenarios. Finally, we selected an algorithm combining good power and low bias. The sensitivity analysis provided further insight into the performance of these algorithms.

RESULTS

Algorithms, PPV, and sensitivity

In 2008 and 2009, 695 hospitalizations of children between the ages of 1 month and 3 years with at least 1 diagnosis code of febrile convulsions (ICD-10 code R56.0 or R56.8) in any position on the discharge abstract occurred in the 10 hospitals. The corresponding PPV was 80.72%. Two-thirds of false positives were nonfebrile convulsions and one-third were not convulsions at all. The registry of convulsion cases established by the pediatric emergency department of Dijon University Hospital identified 137 hospitalizations for febrile convulsions during the same period, allowing us to estimate a sensitivity of 98.54% for this selection.

By using the multivariable logistic regression model, we selected 4 variables associated with an increase in the probability of false positives (Table 1). When the discharge abstract

Table 1. Logistic Regression Model to Identify Factors Associated With False Positive Febrile Convulsion Diagnoses in 10 Hospitals in the French Administrative Areas of Bas Rhin, Côte-d'Or, and Doubs, 2008–2009

Variable	OR	95% CI	*P* Value
Diagnosis of febrile convulsion[a]	30.7	16.9, 55.8	<0.0001
Admission into emergency department (no/yes)	25.0	6.6, 95.5	<0.0001
Febrile convulsion coded as main diagnosis (no/yes)	3.7	1.8, 7.6	<0.0001
Neurological investigation performed (no/yes)	4.1	2.1, 8.0	<0.0004

Abbreviations: CI, confidence interval; OR, odds ratio.
[a] Defined as *International Classification of Diseases, Tenth Revision*, code R56.8 or R56.0.

included ICD-10 code R56.8, the odds of being a false positive were about 30 times higher than when ICD-10 code R56.0 was used alone. When the diagnosis of febrile convulsions was not coded as the main diagnosis, the odds of being a false positive were approximately 4 times higher. When the child was not admitted to an emergency department, the odds of being a false positive were 25 times higher. When a procedure for neurological investigation was performed (magnetic resonance imaging, tomodensitometry, or electroencephalography), the odds of being a false positive were approximately 4 times higher.

Different algorithms were considered, involving the variables selected by the regression model. The first algorithm (the original selection) selects all hospitalizations with at least 1 diagnosis of febrile convulsions (ICD-10 code R56.0 or R56.8), whatever its position on the discharge abstract. The second algorithm relies only on the ICD-10 code R56.0 as a main diagnosis. The third algorithm uses this main diagnosis ICD-10 code R56.0 plus admission to an emergency department. The fourth algorithm adds a neurological investigation to the selection criteria of the third algorithm. For all of these algorithms, PPV and sensitivity were estimated (Table 2). As expected, the PPV increased as the sensitivity decreased. The decrease in sensitivity (of approximately 40%) is particularly noticeable between the third and fourth algorithms, although the PPV increased by only 2%.

Power and relative bias

We calculated the power function from equation 5 for all 4 algorithms. We used $P = 0.9$, the number of cases m from Table 2, and $r = 0.05$ or $r = 0.25$. We considered $\theta = 1$, $\theta = 1 + (\rho - 1) / 2$, which satisfies equation 2, and $\theta = 1.3$, which does not. Power functions are shown in Figure 2. PPVs and sample sizes (m) used in the power calculations are those obtained from the PPV assessment study in 10 hospitals from the 3 French administrative areas of Bas Rhin, Côte d'Or, and Doubs in 2008–2009.

The power is smaller for $r = 0.05$ than for $r = 0.25$. The power functions for algorithms 2 and 3 are always virtually equal. Algorithm 1 always shows the largest power, and algorithm 4 always shows the smallest power. For scenario $\theta = 1$, there is little difference among algorithms 1, 2, and 3. The differences appear for the 2 other scenarios. Finally, as expected, the type I error ($\rho = 1$) is controlled in scenarios $\theta = 1$ and $\theta = 1 + (\rho - 1) / 2$, where the hypothesis in equation 2 is met, but not in scenario $\theta = 1.3$.

We obtained the relative bias (equation 3) for each algorithm under the same scenarios for θ as for the power. Figure 3 indicates a clear hierarchy among the 4 algorithms, with algorithm 1 resulting in a markedly worse relative bias for all scenarios. The relative bias is less than 10% for the true value $\rho \leq 2$, which was the larger value that was considered and increases with ρ. PPVs used in the relative bias calculations are those obtained from the PPV assessment study in 10 hospitals in the 3 French administrative areas of Bas Rhin, Côte d'Or, and Doubs in 2008–2009.

Thus, algorithms 1–3 provide the best power; within these, algorithms 2 and 3 have the highest PPVs and hence the lowest bias. Algorithm 2 is the simplest of the 3, and is therefore our best choice.

DISCUSSION

Vaccine safety studies based on automated administrative data are cheaper and quicker to complete than studies using medical records. When appropriate, the choice of the SCCS

Table 2. Estimation of the Positive Predictive Value and Sensitivity of 4 Algorithms to Identify Febrile Convulsions in 10 Hospitals in the French Administrative Areas of Bas Rhin, Côte-d'Or, and Doubs, France, 2008–2009

Algorithm	No. of Cases	Sensitivity, %[a]	95% CI	No. of Cases	Positive Predictive Value[b]	95% CI
1	170	98.54	96.53, 99.99	695	80.72	77.79, 83.65
2	137	89.05	83.82, 94.28	502	95.02	93.12, 96.92
3	131	86.86	81.20, 92.52	490	96.33	94.67, 97.99
4	71	47.45	39.09, 55.81	229	98.25	96.55, 99.95

Abbreviation: CI, confidence interval.
[a] Sensitivity was computed from 137 cases in the Dijon University Hospital database. Registry of the pediatric emergency department was considered the "gold standard."
[b] Positive predictive value was computed from the 10 hospital databases. Medical records were considered the gold standard.

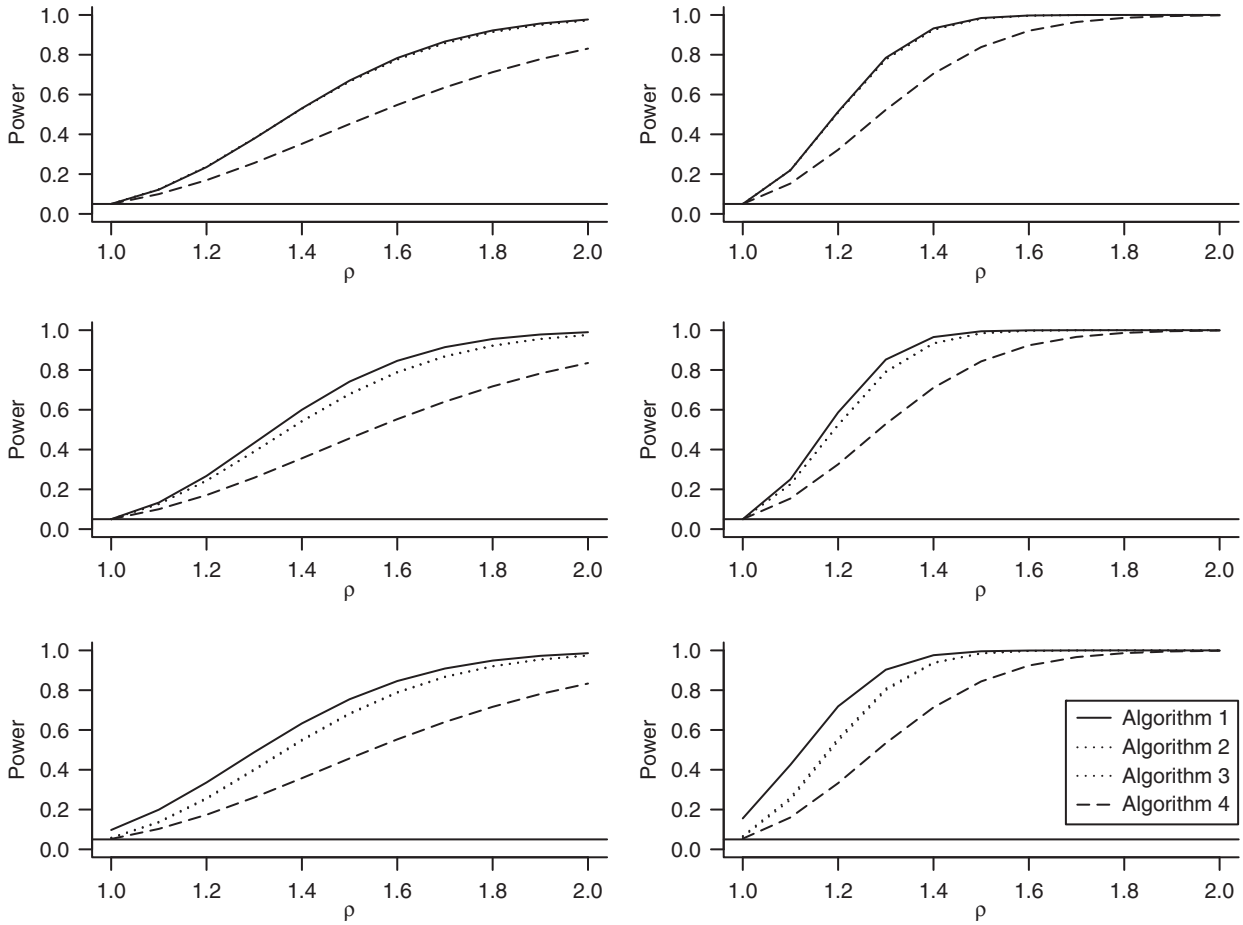

Figure 2. Power function of the self-controlled case series test (1-sided, 5% nominal type I error risk) by relative incidence ρ in detecting an association between vaccination and hospitalization for febrile convulsions when the cases include false positives. Case definition algorithm 1 (solid line), algorithms 2 and 3 (dotted line), and algorithm 4 (dashed line). Algorithms 2 and 3 are indiscernible. The relative incidence in pseudoevents is $\theta = 1$ in the top row, $\theta = 1 + (\rho - 1) / 2$ in the middle row, and $\theta = 1.3$ in the bottom row. The proportion of the observation period at risk is $r = 0.05$ in the left column and $r = 0.25$ in the right column. The proportion of vaccinated population is $P = 90\%$. The bottom horizontal straight line indicates the 5% level.

design contributes to this effectiveness. However, a preliminary question is the ability of hospital discharge databases to select true cases and the ability of the SCCS analysis to cope with false positives.

In this study, we used the French Système National d'Informations Inter Régime d'Assurance Maladie, which includes linked claim databases and hospital discharge (PMSI) databases at the national level, and explored the properties of case identification from the latter. In the future, SCCS studies could be conducted in this framework. We defined and compared algorithms to identify febrile convulsions from the PMSI, and we investigated the behavior of the SCCS according to the resulting PPVs.

Our results show that automated administrative hospital databases have the potential to identify cases of febrile convulsions in childhood with high PPV. Our estimated PPVs ranged from 81% to 98% according to the algorithm, which compares rather favorably with values obtained elsewhere (70, 71). We were able to explore factors that predict the occurrence of a false positive and found that nonadmission to the emergency department, a criterion also considered by Huang et al. (71), was a good predictor. This and the inclusion of ICD-10 diagnosis code R56.8 were the most strongly associated factors. We also found, to a lesser extent, that the absence of a neurological investigation was also predictive of being a false positive, as well as nonrestriction to the main diagnosis. These considerations led to the formulation of 4 nested candidate algorithms; the more restrictive the algorithm, the better it was in terms of PPV and the lower the sensitivity and the number of events.

We propose that the selection of the algorithm should be based on a compromise between bias and power, because including more events, even polluted with false positives, can increase the power. Algorithm 4 resulted in the lowest power, whereas the first 3 algorithms showed comparable power, with algorithms 2 and 3 indistinguishable under all scenarios.

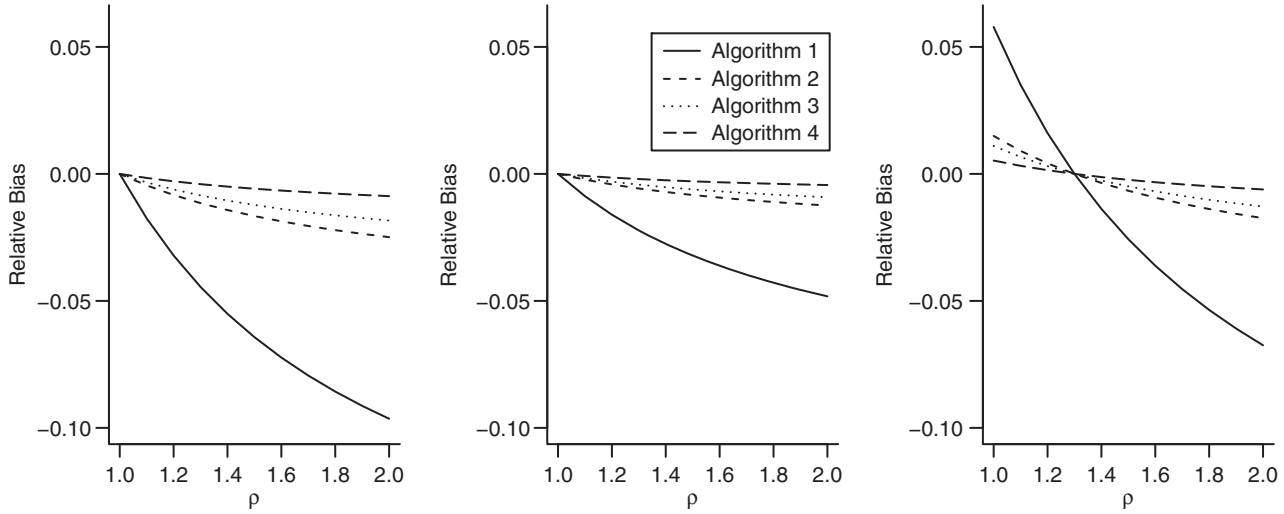

Figure 3. Relative bias by relative incidence ρ of hospitalization for febrile convulsions in children aged 1 month to 3 years, identified by the administrative hospital database when there are false positives in cases. The relative incidence in pseudoevents is $\theta = 1$ in the left graph, $\theta = 1 + (\rho - 1)/2$ in the middle graph, and $\theta = 1.3$ in the right graph.

Algorithm 1 produced the greatest relative bias. Overall, defining febrile convulsions on the basis of the main ICD-10 diagnosis code R56.0 (in algorithm 2) seemed to provide a good compromise in our setting.

Our power and bias criteria provide a novel framework for guiding the choice of algorithm, which is directly related to the aims of the investigation, namely estimation of the relative incidence and hypothesis testing. These quantities involve the PPV and the number of cases, but not (independently) the sensitivity, which, as it turns out, is not essential for this purpose, though it influences the PPV (72). In our study, we were able to estimate the sensitivity of the 4 algorithms. The results shed further light on their performance. The small gain in PPV from algorithm 4 is achieved at the cost of drastic reductions in the sensitivity and number of events selected.

The exact compromise to be struck between bias and power will depend on circumstances. Power is secondary when large numbers of cases are available. It may be more of an issue when the adverse event of interest is rare or not easily diagnosed or if its diagnosis relies on a severity threshold. These considerations arose in Farrington et al. (1), who used contrasting analyses based on febrile convulsions or aseptic meningitis (1,062 events, relative incidence = 1.51) and on aseptic meningitis alone (7 events, relative incidence = 38.1) to study the association between the measles, mumps, and rubella vaccine with the Urabe mumps strain and aseptic meningitis. In this study, not all febrile convulsions were caused by aseptic meningitis, but conversely, not all children with convulsions caused by aseptic meningitis underwent lumbar punctures.

In our theoretical calculations, we allowed for the possibility that pseudoevents might be associated with the event. If this association does not exist or is less than the association between true events and the vaccine (equation 2), the relative incidence will be conservatively estimated, and the type I error probability will be controlled even if pseudoevents are included. When an association between pseudoevents and the vaccine can be ruled out, it is possible, in principle, to correct the relative incidence estimate.

Our analysis has limitations. The validation study was limited to 10 public hospitals in 3 French geographical areas for the PPV assessment and to 1 public hospital for the sensitivity assessment, which raises the question of whether its results can be extrapolated to all areas and to all types of hospital. The theoretical study was restricted to a simple scheme with no age effect and 1 risk period, which made it possible to focus on the key issues and to formulate a framework to compare event definition algorithms. Overall, our findings give us confidence in the validity of SCCS analyses in the presence of false positive events.

ACKNOWLEDGMENTS

Department of Biostatistics and Medical Informatics, Teaching Hospital of Dijon, Dijon, France (Catherine Quantin, Eric Benzenine); Inserm U866, University of Burgundy, Dijon, France (Catherine Quantin); Faculty of Medicine, Department of Epidemiology and Public Health, University of Strasbourg, Strasbourg, France (Michel Velten); Department of Epidemiology and Biostatistics, Paul Strauss Comprehensive Cancer Center, Strasbourg, France (Michel Velten); Department of Pediatrics, University Hospital Dijon, Dijon, France (Frédéric Huet); Department of Mathematics and Statistics, The Open University, Milton Keynes, United Kingdom (C. Paddy Farrington); Biostatistics Team, Centre for Research in Epidemiology and Population Health, Inserm, Villejuif, France (Pascale

Tubert-Bitter); and University Paris-Sud, Villejuif, France (Pascale Tubert-Bitter)

This work was supported by grants from the Agence Nationale pour la Recherche, the Institut National de la Santé et de la Recherche Médicale (Inserm, Réseau de recherche clinique et en santé des populations), the Programme Hospitalier de Recherche Clinique, and by a Royal Society Wolfson Research merit award to C.P.F.

We thank the French National Health Insurance Fund for Salaried Workers (*Caisse nationale de l'assurance maladie des travailleurs salariés*), Centre Hospitalier Universitaire de Dijon, Centre Hospitalier de Beaune, Centre Hospitalier de Semur en Auxois, Centre Hospitalier Universitaire de Besançon, Centre Hospitalier de Pontarlier, Centre Hospitalier Universitaire de Strasbourg, Centre Hospitalier de Haguenau, Centre Hospitalier de Selestat, Centre Hospitalier de Saverne, and Centre Hospitalier de Wissembourg.

C.P.F. and P.T.-B. contributed equally to the work.

Conflict of interest: none declared.

References

1. Farrington P, Pugh S, Colville A, et al. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. *Lancet*. 1995;345(8949):567–569.
2. Chan KA, Andrade SE, Boles M, et al. Inhibitors of hydroxymethylglutaryl-coenzyme A reductase and risk of fracture among older women. *Lancet*. 2000;355(92222):2185–2188.
3. Ray WA, Daugherty JR, Griffin MR. Lipid-lowering agents and the risk of hip fracture in a Medicaid population. *Inj Prev*. 2002;8(4):276–279.
4. Cameron JC, Walsh D, Finlayson AR, et al. Oral polio vaccine and intussusception: a data linkage study using records for vaccination and hospitalization. *Am J Epidemiol*. 2006;163(6):528–533.
5. Verhamme K, Sturkenboom M. Study designs in paediatric pharmacoepidemiology. *Eur J Clin Pharmacol*. 2011;67(1 suppl):67S–74S.
6. Habel LA, Cooper WO, Sox CM, et al. ADHD medications and risk of serious cardiovascular events in young and middle-aged adults. *JAMA*. 2011;306(24):2673–2683.
7. Blin P, Lassalle R, Dureau-Pournin C, et al. Insulin glargine and risk of cancer: a cohort study in the French National Healthcare Insurance Database. *Diabetologia*. 2012;55(3):644–653.
8. Fagot JP, Blotiere PO, Ricordeau P, et al. Does insulin glargine increase the risk of cancer compared with other basal insulins? A French nationwide cohort study based on national administrative databases. *Diabetes Care*. 2013;36(2):294–301.
9. Neumann A, Weill A, Ricordeau P, et al. Pioglitazone and risk of bladder cancer among diabetic patients in France: a population-based cohort study. *Diabetologia*. 2012;55(7):1953–1962.
10. Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. *Eur J Clin Pharmacol*. 2012;68(2):123–129.
11. Moro ML, Nobilio L, Voci C, et al. A population based cohort study to assess the safety of pandemic influenza vaccine, Focetria in Emilia-Romagna region, Italy—part two. *Vaccine*. 2013;31(10):1438–1446.
12. Shapiro S. Automated record linkage: a response to the commentary and letters to the editor. *Clin Pharmacol Ther*. 1989;46(4):395–398.
13. Faich GA, Stadel BV. The future of automated record linkage for postmarketing surveillance: a response to Shapiro. *Clin Pharmacol Ther*. 1989;46(4):387–389.
14. Strom BL, Carson JL. Automated data bases used for pharmacoepidemiology research. *Clin Pharmacol Ther*. 1989;46(4):390–394.
15. Shapiro S. The role of automated record linkage in the postmarketing surveillance of drug safety: a critique. *Clin Pharmacol Ther*. 1989;46(4):371–386.
16. Mullooly JP. Misclassification model for person-time analysis of automated medical care databases. *Am J Epidemiol*. 1996;144(8):782–792.
17. Couris CM, Colin C, Rabilloud M, et al. Method of correction to assess the number of hospitalized incident breast cancer cases based on claims databases. *J Clin Epidemiol*. 2002;55(4):386–391.
18. Arbogast PG, Ray WA. Use of disease risk scores in pharmacoepidemiologic studies. *Stat Methods Med Res*. 2009;18(1):67–80.
19. Noize P, Bazin F, Dufouil C, et al. Comparison of health insurance claims and patient interviews in assessing drug use: data from the Three-City (3C) Study. *Pharmacoepidemiol Drug Saf*. 2009;18(4):310–319.
20. Hartzema AG, Racoosin JA, MaCurdy TE, et al. Utilizing Medicare claims data for real-time drug safety evaluations: Is it feasible?. *Pharmacoepidemiol Drug Saf*. 2011;20(7):684–688.
21. Ray WA. Improving automated database studies. *Epidemiology*. 2011;22(3):302–304.
22. Dart AB, Martens PJ, Sellers EA, et al. Validation of a pediatric diabetes case definition using administrative health data in Manitoba, Canada. *Diabetes Care*. 2011;34(4):898–903.
23. Kawai VK, Murray KT, Stein CM, et al. Validation of a computer case definition for sudden cardiac death in opioid users. *BMC Res Notes*. 2012;5:473.
24. Bernard MA, Benichou J, Blin P, et al. Use of health insurance claim patterns to identify patients using nonsteroidal anti-inflammatory drugs for rheumatoid arthritis. *Pharmacoepidemiol Drug Saf*. 2012;21(6):573–583.
25. Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*. 1995;51(1):228–235.
26. Whitaker HJ, Farrington CP, Spiessens B, et al. Tutorial in biostatistics: the self-controlled case series method. *Stat Med*. 2006;25(10):1768–1797.
27. Farrington P, Pugh S, Colville A, et al. A new method for active surveillance of adverse events from diphteria/tetanus/pertussis and measles/mumps/rubella vaccines. *Lancet*. 1995;345(8949):567–569.
28. Murphy TV, Gargiullo PM, Massoudi MS, et al. Intussusception among infants given an oral rotavirus vaccine. *N Engl J Med*. 2001;344(8):564–572.
29. Miller E, Waight P, Farrington CP, et al. Idiopathic thrombocytopenic purpura and MMR vaccine. *Arch Dis Child*. 2001;84(3):227–229.
30. Mutsch M, Zhou W, Rhodes P, et al. Use of the inactivated intranasal influenza vaccine and the risk of Bell's palsy in Switzerland. *N Engl J Med*. 2004;350(9):896–903.
31. Miller E, Andrews N, Stowe J, et al. Risks of convulsion and aseptic meningitis following measles-mumps-rubella vaccination in the United Kingdom. *Am J Epidemiol*. 2007;165(6):704–709.

32. France EK, Glanz J, Xu S, et al. Risk of immune thrombocytopenic purpura after measles-mumps-rubella immunization in children. *Pediatrics*. 2008;121(3):e687–e692.
33. Naleway AL, Belongia EA, Donahue JG, et al. Risk of immune hemolytic anemia in children following immunization. *Vaccine*. 2009;27(52):7394–7397.
34. Gwini SM, Coupland CA, Siriwardena AN. The effect of influenza vaccination on risk of acute myocardial infarction: self-controlled case-series study. *Vaccine*. 2011;29(6):1145–1149.
35. Patel MM, Lopez-Collada VR, Bulhoes MM, et al. Intussusception risk and health benefits of rotavirus vaccination in Mexico and Brazil. *N Engl J Med*. 2011; 364(24):2283–2292.
36. Sun Y, Christensen J, Hviid A, et al. Risk of febrile seizures and epilepsy after vaccination with diphtheria, tetanus, acellular pertussis, inactivated poliovirus, and Haemophilus influenzae type B. *JAMA*. 2012;307(8):823–831.
37. Arnheim-Dahlstrom L, Hallgren J, Weibull CE, et al. Risk of presentation to hospital with epileptic seizures after vaccination with monovalent AS03 adjuvanted pandemic A/H1N1 2009 influenza vaccine (Pandemrix): self controlled case series study. *BMJ*. 2012;345:e7594.
38. Weldeselassie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. *Epidemiol Infect*. 2011;139(12):1805–1817.
39. Hubbard R, Farrington P, Smith C, et al. Exposure to tricyclic and selective serotonin reuptake inhibitor antidepressants and the risk of hip fracture. *Am J Epidemiol*. 2003;158(1):77–84.
40. Hubbard R, Lewis S, West J, et al. Bupropion and the risk of sudden death: a self-controlled case-series analysis using the Health Improvement Network. *Thorax*. 2005;60(10):848–850.
41. Grosso A, Douglas I, Hingorani A, et al. Oral bisphosphonates and risk of atrial fibrillation and flutter in women: a self-controlled case-series safety analysis. *PLoS One*. 2009;4(3): e4720.
42. Douglas IJ, Evans SJ, Pocock S, et al. The risk of fractures associated with thiazolidinediones: a self-controlled case-series study. *PLoS Med*. 2009;6(9):e1000154.
43. Gibson JE, Hubbard RB, Smith CJ, et al. Use of self-controlled analytical techniques to assess the association between use of prescription medications and the risk of motor vehicle crashes. *Am J Epidemiol*. 2009;169(6):761–768.
44. Donaldson GC, Hurst JR, Smith CJ, et al. Increased risk of myocardial infarction and stroke following exacerbation of COPD. *Chest*. 2010;137(5):1091–1097.
45. Minassian C, D'Aiuto F, Hingorani AD, et al. Invasive dental treatment and risk for vascular events: a self-controlled case series. *Ann Intern Med*. 2010;153(8):499–506.
46. Pariente A, Fourrier-Reglat A, Ducruet T, et al. Antipsychotic use and myocardial infarction in older patients with treated dementia. *Arch Intern Med*. 2012;172(8):648–653.
47. Butt DA, Mamdani M, Austin PC, et al. The risk of hip fracture after initiating antihypertension drugs in the elderly. *Arch Intern Med*. 2012;172(22):1739–1744.
48. Maclure M, Fireman B, Nelson JC, et al. When should case-only designs be used for safety monitoring of medical products? *Pharmacoepidemiol Drug Saf*. 2012;21(1 suppl):50S–61S.
49. Nordmann S, Biard L, Ravaud P, et al. Case-only designs in pharmacoepidemiology: a systematic review. *PLoS One*. 2012;7(11):e49444.
50. Kopec JA, Esdaile JM. Bias in case-control studies. A review. *J Epidemiol Community Health*. 1990;44(3):179–186.
51. Brenner H, Savitz DA. The effects of sensitivity and specificity of case selection on validity, sample size, precision, and power in hospital-based case-control studies. *Am J Epidemiol*. 1990;132(1):181–192.
52. Chavance M, Dellatolas G, Lellouch J. Correlated nondifferential misclassifications of disease and exposure: application to a cross-sectional study of the relation between handedness and immune disorders. *Int J Epidemiol*. 1992; 21(3):537–546.
53. Kristensen P. Bias from nondifferential but dependent misclassification of exposure and outcome. *Epidemiology*. 1992;3(3):210–215.
54. Magder LS, Hughes JP. Logistic regression when the outcome is measured with uncertainty. *Am J Epidemiol*. 1997;146(2):195–203.
55. Farrington CP, Nash J, Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *Am J Epidemiol*. 1996;143(11):1165–1173.
56. Fetter RB, Shin Y, Freeman JL, et al. Case mix definition by diagnosis-related groups. *Med Care*. 1980;18(2 suppl):iii, 1S–53S.
57. French Law No. 94-748 of August 2, 1991. J.O. No. 179. Paris, p. 10255.
58. French Order of July 22, 1996. J.O. No. 173 of July 26, 1996, p. 11308.
59. French Law No. 2003-1199 of December 19, 2003. J.O. No. 293.
60. French Circular No. 119, 1989. Health Mo. Ministerial J.O. No. 90/2, pp. 29–71.
61. Couris CM, Polazzi S, Olive F, et al. Breast cancer incidence using administrative data: correction with sensitivity and specificity. *J Clinical Epidemiol*. 2009;62(6):660–666.
62. Quantin C, Benzenine E, Fassa M, et al. Evaluation of the interest of using discharge abstract databases to estimate breast cancer incidence in two French departments. *Stat J IAOS*. 2012;28:73–85.
63. Remontet L, Mitton N, Couris CM, et al. Is it possible to estimate the incidence of breast cancer from medico-administrative databases? *Eur J Epidemiol*. 2008;23(10):681–688.
64. Uhry Z, Colonna M, Remontet L, et al. Estimating infra-national and national thyroid cancer incidence in France from cancer registries data and national hospital discharge database. *Eur J Epidemiol*. 2007;22(9):607–614.
65. Quantin C, Benzenine E, Ferdynus C, et al. Advantages and limitations of using national administrative data on obstetric blood transfusions to estimate the frequency of obstetric hemorrhages. *J Public Health (Oxf)*. 2013;35(1):147–156.
66. Quantin C, Benzenine E, Hagi M, et al. Estimation of national colorectal-cancer incidence using claims databases. *J Cancer Epidemiol*. 2012;2012:298369.
67. Tuppin P, De Roquefeuil L, Weill A, et al. French national health insurance information system and the permanent beneficiaries sample. *Rev Epidemiol Sante Publique*. 2010; 58(4):286–290.
68. Goldberg M, Jougla E, Fassa M, et al. The French health information system. *Stat J IAOS*. 2012;28(1-2):31–41.
69. Martin-Latry K, Begaud B. Pharmacoepidemiological research using French reimbursement databases: Yes we can! *Pharmacoepidemiol Drug Saf*. 2010;19(3):256–265.
70. Barlow WE, Davis RL, Glasser JW, et al. The risk of seizures after receipt of whole-cell pertussis or measles, mumps, and rubella vaccine. *N Engl J Med*. 2001;345(9):656–661.
71. Huang WT, Gargiullo PM, Broder KR, et al. Lack of association between acellular pertussis vaccine and seizures in early childhood. *Pediatrics*. 2010;126(2): 263–269.
72. Zhou X, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. 2nd ed. Hoboken, NJ: Wiley; 2011.