Using existing data sources for assessment of vaccine safety:
a focus on methods

Het gebruik van bestaande gegevensbronnen ter beoordeling van de veiligheid van vaccins:
een focus op methodiek.

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof. dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board
The public defense shall be held

on 26 June 2019 at 15:30 hrs
by

Caitlin Nora Dodd
Born in Duluth, Minnesota, USA.

Erasmus University Rotterdam

Doctoral Committee

PROMOTER

Prof. dr. M.C.J.M. Sturkenboom

OTHER MEMBERS

Prof. S. Black

Prof. dr. E.P. van Puyenbroek

Prof. dr. M. Koopmans

COPROMOTERS

Dr. M.A.J. de Ridder

Dr. D.M. Weibel

Using existing data sources for assessment of vaccine safety:

A focus on methods

CAITLIN NORA DODD

Cover design by Caitlin Dodd

Layout by Caitlin Dodd

This work described in this thesis was carried out at the Department of Medical Informatics, Erasmus Medical Center, Rotterdam, The Netherlands

For Nora and Millie.

CONTENTS

CHAPTER 1. GENERAL INTRODUCTION

Using existing data-sources for assessment of vaccine safety: a focus on methods

Vaccines as a public health intervention

Vaccination is widely accepted as one of the foremost public health achievements of the last century, with the United States Centers for Disease Control reporting in 2014 that vaccines have prevented 322 million illnesses, 21 million hospitalizations, and 732,000 deaths over the lifespan of children born between 1984 and 2014 (1) . Vaccines confer this degree of protection only when a large enough proportion of the population is vaccinated to produce herd immunity, or protection of the unvaccinated by the vaccinated (2). Because vaccines serve as a primary prevention measure, they are principally administered to healthy individuals. This means that public expectations for measurement of the safety and effectiveness of vaccines may be more stringent in terms of timeliness and accuracy than those for drugs which treat illness.

The vaccine product life cycle

Vaccines are rigorously tested for safety and effectiveness in clinical trials before receiving market authorization (3). Because clinical trials are limited in size and scope, however, adverse events and unexpected benefits may go undetected in trials and arise only when the vaccine is administered to large populations. Chen et al, in their paper describing the then newly-initiated Vaccine Adverse Events Reporting System (VAERS), displayed the cycle of vaccine confidence in an often-cited figure (figure 1). The figure displays how increasing vaccination coverage is paired with decreasing incidence of the vaccine-preventable disease but also potentially with increasing reports of vaccine-associated (either causally or temporally) adverse events. This increased reporting of adverse events following vaccination can lead to decreased confidence in the vaccine and associated decreases in coverage and effectiveness. The Vaccine Confidence Project, led by the London School of Hygiene and Tropical Health, has found that concern about vaccine safety is the primary cause of vaccine hesitancy in Europe (4). It is of utmost importance to public health that any adverse reactions or deficiencies in efficacy are detected and addressed quickly and accurately in order to maintain public confidence in vaccines(5).

The phases of vaccine development, pharmacovigilance, and communication can be described using figure 1. Period one (Prevaccine) represents all phases of development and testing prior to market authorization of a vaccine while period two (Increasing Coverage) represents post-authorization vaccination of targeted populations. I focus in this thesis on time periods three (Loss of Confidence) and four (Resumption of Confidence) during which safety signals may be detected and verified, additional hypotheses regarding safety and efficacy may be tested, public health implications are assessed, and scientists, public health agencies, and healthcare providers communicate with the public about the benefits and risks of vaccines. Period five (Eradication), or complete elimination of the vaccine preventable disease, will not be addressed in this thesis.

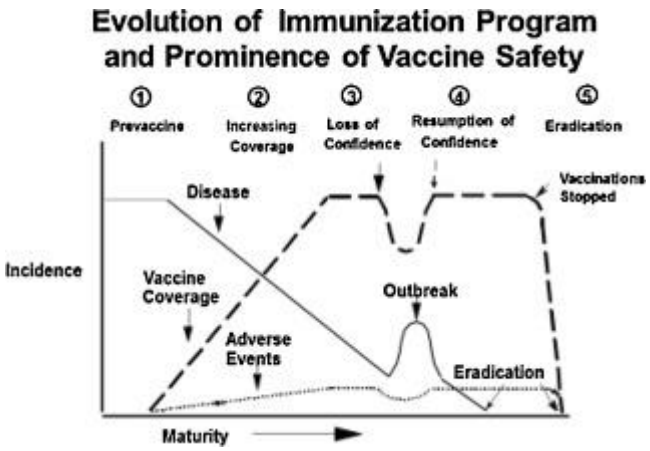


Figure 1. Potential stages in the evolution of immunization program, showing the dynamics of the interaction between vaccine coverage, disease incidence and vaccine adverse events, as the program matures from pre-vaccine to disease eradication (6, 7)

Methodological challenges of assessing vaccine safety

There are many methodological challenges inherent to the assessment of vaccine safety. First, adverse events following vaccinations are often rare (less than 1 per 1000 vaccine recipients) and/or non-specific (meaning that symptoms may be attributed to multiple syndromes), making them difficult to capture and retrieve in existing data sources that may have been assembled for other reasons (8). Rarity of some events can also lead to a limitation in statistical power, even in large databases, necessitating collaboration among data sources to obtain a sufficient number of cases for analysis (9). This type of collaboration can lead to challenges in and of itself, related to heterogeneity and pooling of estimates. Non-specificity of adverse events may mean absence of specific codes or definitions for the event. In addition, some reported adverse events may have long periods between onset of symptoms and diagnosis by a physician, complicating causality assessments (10).

Additional challenges are created by public insistence on vaccine safety and public awareness of potential risks associated with vaccination which may impact individual behavior. These behavioral changes may include increased or decreased healthcare seeking as well as increased or decreased motivation to report events associated with vaccination. These changes in behavior can lead to various forms of bias including ascertainment bias (meaning some subjects are more likely than others to be included in a study), recall bias (meaning that subjects inaccurately report exposures or symptoms), and effects such as the 'healthy vaccinee effect' in which subjects who have experienced an adverse event may delay or forego vaccination when ill (11-13). Additionally, vaccines are given on an age and sometimes seasonal schedule, meaning entanglement of these effects with the effects of the vaccine itself. For example, while Guillain-Barré Syndrome (GBS) has been repeatedly investigated as an adverse event following seasonal influenza vaccination, respiratory infections, which spike in the fall and winter when seasonal vaccines are also administered, are a known cause of GBS (14, 15). Studies which aim to elucidate the association must take into account this seasonally fluctuating rate of infection (16, 17). This association and the impact of seasonality is addressed in chapter 5 of this thesis using the example of pandemic H1N1 influenza vaccines. Finally, vaccination programs alter the host-pathogen relationship due to their scale via indirect effects such as shifts in the age at which peak incidence occurs, protection from sequelae following infection, and herd immunity. I attempt in this thesis to assess the scale and understand the impact of some of these challenges.

Assessment of vaccine safety using existing data-sources

While risks of vaccines may go undetected or prove difficult to estimate correctly in pre-licensure trials due to sample size and time limitations, other data sources are often used to detect and analyze risks and benefits post-licensure (18). These data sources include claims, hospital, and general practitioner databases, registries, birth cohorts, and spontaneous reporting systems. The advantages of using electronic health data are the scale of the population that is captured, the flexibility in design, and the fact that rapid retrospective analyses may be performed. Two pioneering projects which illustrate the utility of existing data sources for assessment of vaccine safety are the Vaccine Safety Datalink (VSD) and the Post-licensure Rapid Immunization Safety Monitoring (PRISM) program. The VSD, a collaboration founded in 1990 among healthcare organizations and the US Centers for Disease Control and Prevention, links medical records on over 8 million subjects from nine healthcare providers in order to perform rapid analysis of vaccine safety and to quickly deploy studies when a new safety signal emerges (19). Similarly, the PRISM program is the vaccine safety component of the US Food and Drug Administration's Sentinel system which was initiated in 2008 and links claims data on over 200 million Americans from national insurance providers in a distributed data model (20). This network is used to sequentially monitor vaccines and to address vaccine safety concerns rapidly. This thesis will focus specifically on spontaneous reporting databases, general practitioner databases, and collaboration of databases in multinational studies as well as simulated data to assess bias and heterogeneity in these types of data sources.

Signal detection in Spontaneous Reporting System Databases

Spontaneous Reporting System databases are those to which consumers, physicians, and others may report suspected adverse events following exposure to a drug or vaccine (21). These databases have the benefits of being relatively inexpensive to set up and maintain and may provide the opportunity for investigators and public health institutions to detect adverse events quickly. Additionally, the databases maintained by US agencies (the FDA Adverse Events Reporting System and the CDC and FDA Vaccine Adverse Event Reporting System) are publically available (22). However, they are also rife with shortcomings (23). Data are self-reported with no verification of exposures or diagnoses. Additionally, important data fields that could aid in more accurate assessment of risks, like time from exposure to onset of symptoms, outcome of the event, and vaccine manufacturer are often incomplete. Because these databases contain data only on suspected cases, no underlying population denominators exist, forcing researchers to use the database itself (i.e. all other reports) as the denominator in any calculations(24). Typically, when assessing spontaneous reports, the number of observed reports of the vaccine/AEFI association of interest is compared to the number expected in the absence of a causal association (i.e. observed vs. expected) using a 2x2 table such as the one below. In other words, the number of reports of the drug or vaccine together with the event of interest is compared to the number of reports of the drug-event combination we would expect if reports involving the drug or vaccine of interest were statistically independent from reports involving the event of interest (25).

Table 1. 2x2 table used in analysis of spontaneous reporting systems

	Experienced AEFI Y	Did Not Experience AEFI Y but other AEFI
Received Vaccine X	A	B
Received other Vaccine than X	C	D

These shortcomings mean that spontaneous reporting databases can be used quite efficiently to detect safety signals or to strengthen signals that arise elsewhere (i.e. a published case series) but is not suitable for hypothesis testing (26, 27).

Many methods have been proposed to diminish the impact of the well-recognized shortcomings of spontaneous reporting databases, some of which will be discussed in this thesis. The first which will be addressed is the Empirical Bayes Geometric Mean. While most signal detection algorithms (SDAs) use an expected count calculated from the counts in the cells of the table above, the Empirical Bayes Geometric Mean uses a Bayesian model of expected counts to shrink effect estimates toward the null, especially when counts of the drug-event combination of interest are low(25). We tested the performance of the EBGM in a drug-specific reporting database and a vaccine-specific reporting database. In the drug-specific database (FAERS), we sought to understand the impact of age stratification on the performance of SDAs in pediatrics while in the vaccine-specific database (VAERS), we compared performance of the EBGM with a new method which exploits time-to-onset distributions of vaccine-associated reports.

Since adverse reactions to vaccines or drugs may be specific to a certain age group, especially in pediatrics where ontogenic changes occur which may change susceptibility, stratification by age when conducting signal detection, has been recommended (28). Stratification is a method to assess confounding by age and to explore effect modification. The lack of stratification may either mask true signals or lead to false positives (29). Where confounding by age is suspected, the typical strategy is to stratify by age categories and subsequently pool stratum-specific estimates if no evidence of effect modification is seen. If differences in association are seen across strata, effect modification is present, and pooling should not be done. While stratification and subsequent adjustment by age has been advocated by some researchers (30), adjustment is routinely implemented in some Bayesian but not in frequentist SDAs. (25, 31, 32) Few studies have systematically addressed the impact of age stratification or adjustment and the results are contradictory (29, 33, 34). Another reason for confounding or effect modification is the presence of vaccine related reports. Vaccines are usually administered to healthy individuals, occur in early childhood (highly associated with age) and because of their intramuscular administration, do not have the same adverse events as drugs that are metabolized through the gastro-intestinal system. We therefore investigated how to deal with masking of positive associations in a spontaneous reporting database that comprises reports for therapeutics as well as vaccines.

Finally, Van Holle et al have developed a new method for detecting signals of disproportionate reporting in databases of vaccine reports using the reported time from administration of the vaccine to onset of the symptoms (35). In this methodology, three distributions are estimated: time from administration of the vaccine of interest to onset of the event of interest is (Vaccine AEFI), time from administration of the vaccine of interest to any other event (Other AEFI) and time from administration of any other vaccine to the event of interest (Other Vaccine). The latter two distributions (Other Vaccine or Other AEFI) are compared to the Vaccine/AEFI time to onset distribution using a Kolmogorov-Smirnov Test of the equality of probability distributions (See Figure 2) (36). In this thesis, we apply the time-to-onset method to the VAERS and EudraVigilance databases and compare its performance to an established SDA (See Chapter 2).

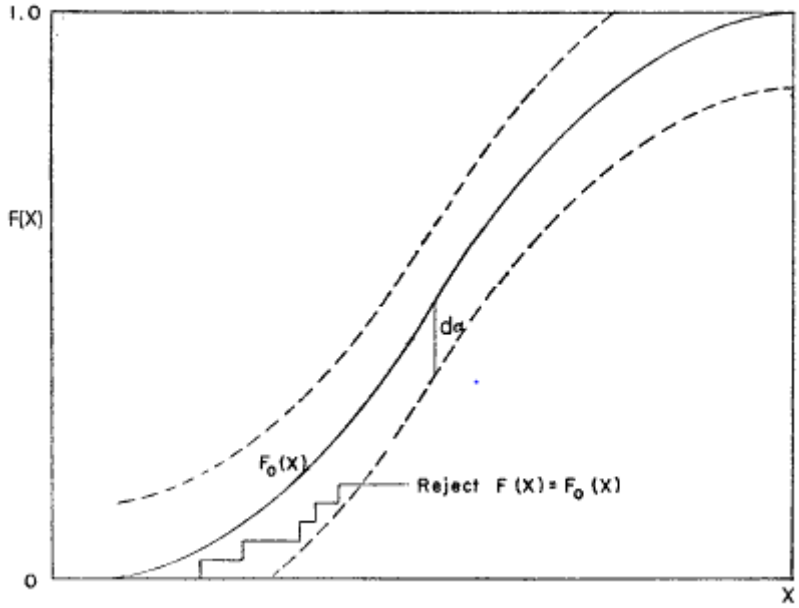


Figure 2. Two-sample Kolmogorov-Smirnov Test: The continuous curve represents the Vaccine/AEFI time-to-onset distribution, the dashed lines provide the critical distance from this distribution, and the step function represents the observed (Other Vaccine TTO or Other AEFI TTO) distribution. Reject unless the step function lies between the two dashed curves (37).

As depicted in figure 3, SDAs represent a first-pass screening after which biological plausibility and case series review must take place before proceeding to any further action. Each of the studies in this thesis using spontaneous reporting system data were conducted within the Global Research in Pediatrics (GRiP) Network of Excellence (38).

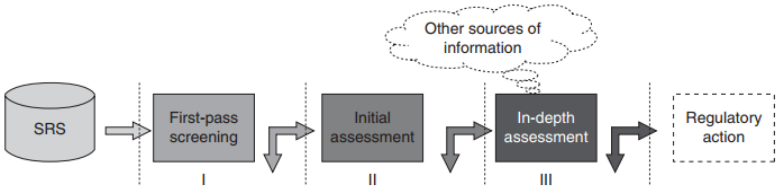


Figure 3. Schematic overview of the signal detection process. SRS = spontaneous reporting system (39)

Whenever new safety issues are detected, next steps are rapid assessment of impact and plausibility and subsequently full in-depth evaluation.

Rapid assessment & Sequential Methods

In order to arrest the potential process of public loss of confidence when safety issues occur for vaccines, these safety concerns must be addressed quickly and accurately. Several methods have been proposed to do so, including running of modular programs in large data networks such as the VSD and PRISM which are maintained (amongst other goals) to allow for rapid assessment at the population level using ecological methods and sequential methods (19, 40).

One approach to rapid detection of an adverse event is to quantify drug or vaccine effects sequentially within a population based electronic data source. This was developed and is still applied in the VSD. Specifically, Lieu et al proposed in 2007 that sequential surveillance was needed to detect safety concerns early and applied their newly developed method, maximized sequential probability ratio testing (maxSPRT), to the test case of meningococcal conjugate vaccine using VSD data (41). When using sequential methods to assess a drug or vaccine, evidence in a data source or sources is accumulated and tested sequentially until a predefined threshold is met. Sequential methods have been applied for vaccine safety primarily using variations on the sequential probability ratio test such as the MaxSPRT and conditional MaxSPRT to monitor the safety of numerous vaccines in the VSD (42-45) and more recently in the FDA PRISM system (46).

Alternatively, ecological methods make use of data on the population as a whole rather than individual exposure and outcome data. They are generally easier to implement than other designs because individual-level privacy is not a concern and data at the population level is often easily available. For use in the study of vaccinations, ecological methods can be used to compare the rate of an event – either a vaccine-preventable disease or an adverse event – in periods before and after a targeted vaccination campaign, licensure of a new vaccine, or change in vaccination schedule. Ecological designs such as these are complicated by the fact that changes over time can be driven by other factors such as changes in policy or diagnostic practices, growing awareness in the population, or the background incidence of an event. Also ecological studies cannot deal with confounding factors, but can serve as a way for rapid assessment. In this thesis we looked at the ability of ecological methods to assess risk of narcolepsy following exposure to different pandemic influenza vaccines on a global scale using fluctuating background incidence of the event in the Systematic Observational Method for Narcolepsy and Influenza Immunization Assessment (SOMNIA) project (See Chapter 3). Ecological methods were also used in the VAESCO project in Europe to monitor changes in narcolepsy rates using a distributed model (47)**.**

Single database studies to study vaccine effects

Single database studies were the most prevalent in the study of post-licensure vaccine effects and have therefore been well-reviewed in the literature (48-51) . Electronic healthcare records and specifically general practitioner databases are powerful in that they contain data typically on large numbers of subjects often over a long period of time. An advantage of limiting a study to a single database is that the database has a consistent structure and generally makes use of one coding system. If the database represents a geographical area (as is usually the case), certain assumptions about the underlying population can be made. General practitioner databases that span a long period of time may also offer the possibility to study effects in periods before and after introduction of a vaccine, outbreak of an infectious disease, or change in legislation.

While single database studies have long been the primary approach for assessing risks of vaccines, they are often limited in size and scope. Individual database systems have proven to be underpowered to detect rare adverse events, especially for subpopulations, reinforcing the need for collaborative studies (52). The withdrawal of Vioxx in 2004 following detection of an increased risk of heart attack in long-term users prompted the US FDA to set up the Sentinel system in accordance with a US Congress mandate (53, 54). The 2009 H1N1 influenza pandemic required active surveillance for rare adverse events following worldwide vaccination. The FDA PRISM program, in a proof-of-concept study, system successfully linked data from health plans and immunization registries to monitor vaccine safety, significantly reducing the time to data generation (55). Similarly, Izurieta et al have called for a collaborative monitoring system for vaccines, especially those scheduled to be introduced in low and middle income countries in the near future (56). Collaborative vaccine safety studies will be addressed in Chapter 5 of this thesis.

Collaborative studies to assess vaccine safety

In the field of vaccine safety, collaborative studies allow researchers to gain power to study a rare outcome or to investigate differences in vaccine types and schedules which vary by region or country. In preparation for the 2009 H1N1 pandemic, the Vaccine Adverse Event Surveillance and Communication network (VAESCO) was formed to conduct collaborative vaccine safety studies in Europe and following the pandemic, the Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) consortium was initiated to monitor the benefits and risks of licensed vaccines (57, 58). As pandemic preparedness has become a priority for governments, collaborative studies allow public health institutions to understand infectious diseases before they spread globally and to study the implications of decisions made either at the policy-maker or individual level before implementing vaccination or surveillance programs. Finally, as many vaccines currently in development are targeted for pathogens prevalent in the developing world, collaboration to build capacity for data collection and sharing in those countries which will introduce new vaccines becomes crucial (56, 59).

While collaborative studies are essential, they also present methodological challenges. First, data sharing across borders and even institutions is often restricted due to privacy concerns, meaning that new tools and processes which ensure anonymity must be developed to enable collaboration. In recent years, many of these tools and processes have been developed within European and global projects. For example, within the Exploring and Understanding Adverse Drug Reactions by Integrative Mining of Clinical Records and Biomedical Data (EU-ADR) project, purpose-built Java software was developed to read and process common input files locally at each data-contributing site, producing encrypted data which can be analyzed centrally (60). Within the aforementioned SOMNIA, VAESCO, and ADVANCE projects, common protocols were used across databases. Within ADVANCE, coding systems were harmonized using a newly developed tool which exploits the Unified Medical Language System to link user-defined concepts to codes in multiple coding systems (61). Data sources vary widely from institution to institution and country to country meaning that harmonization of event and exposure coding is often the longest and most challenging phase of a collaborative study. Finally, methodological challenges when combining data must be addressed. Heterogeneity in data sources and database-specific estimates must be assessed before data can be pooled.

In this thesis, collaborative studies are addressed in chapter 5 with a study assessing the risk of Guillain-Barré Syndrome following pandemic H1N1 vaccines in The Global H1N1 GBS Consortium (62), a case-control study of narcolepsy following pandemic H1N1 vaccines in the SOMNIA study (63), and a proof-of-concept study assessing the capacity for a hospital-based vaccine safety active surveillance system in the WHO Global Vaccine Safety Multi-Country Collaboration (59).

Heterogeneity in collaborative studies

As described above, there are many advantages to conducting collaborative studies of vaccine effects. However, there are significant challenges, one of which is heterogeneity in data sources. Data sources which may be utilized in collaborative studies may have their own data capture methods, coding practices, and data structure. In order for data to be shared, these factors must be taken into account. In recent years, many collaborative groups have worked to develop solutions to these problems. Heterogeneity was encountered in each of the collaborative projects addressed in this thesis: SOMNIA, ADVANCE, The Global H1N1 GBS Consortium, and the WHO Global Vaccine Safety-Multi Country Collaboration (GVS-MCC). Types of heterogeneity include heterogeneity in data sources and coding systems, differences in vaccine schedules and vaccine types utilized in populations. These collaborations all required pooling either heterogeneous data from different data structures or heterogeneous database-specific results.

Many groups have addressed these issues of heterogeneity and developed a variety of solutions. In the EU-ADR project, as previously mentioned, purpose-built software was designed to process common input files. Within the VAESCO project, a data entry tool (Chameleon) was developed to allow researchers to enter data into a common report form, after which the data is anonymized and transmitted to a data processing center, maintaining the original data locally and allowing for common input files (64). This same software was used in the SOMNIA and WHO GVS-MCC projects for data sharing and case validation. The Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium (PROTECT) project alternatively made use of common data specifications implemented by each database rather than common input files or data entry software (65). In the US FDA PRISM project, researchers have tested algorithms for outcome detection and the effect of misclassification on time to detection of safety signals in heterogeneous databases within the network (66) . Many projects use a distributed data model with common input files to reduce heterogeneity among databases and allow for rapid launch of studies using common analysis scripts and/or a remote research environment (RRE) within which to securely share data and scripts. Among these are PRISM, VSD and the Observational Medical Outcomes Partnership (OMOP) (9, 67). A remote research environment (Octopus) that was built to allow for datashring in the ARITMO project (9) by Erasmus Medical Center has been re-used in the studies within SOMNIA, GRiP, ADVANCE, GVS-MCC, and The Global H1N1 GBS Consortium included in this thesis.

In the collaborative studies included in this thesis, challenges were encountered in defining events in a harmonized manner. In order to address this challenge, we make use in chapter 5 of a new method for extracting events from existing data sources using combinations of disease, laboratory, and treatment codes. Typically, events of interest are extracted from data sources using a list of pre-defined diagnostic codes which may be more sensitive or specific in some databases as compared to others. The component analysis method attempts to address this shortcoming by using the combination of codes from multiple domains that maximizes the positive predictive value in each data source. Within the EMIF project, Roberto et al developed a strategy to combine codes or 'components' from different data domains to improve identification of events in databases (68). This methodology, the 'component algorithm strategy' is used in chapter 5 of this thesis.

Collaborative studies also often have, as one of their objectives, the comparison of effects across data sources. This becomes difficult or impossible if exposures are recorded in a different manner in different databases – or not recorded at all. In a multi-database study of vaccine effects, exposure data may be obtained from patient records or, depending upon the health care system and database structure, it is may be necessary to obtain exposure data from secondary sources such as a registry or individual vaccination cards. These differences in availability and quality of vaccine exposure data invariably lead to heterogeneity. In the studies included in this thesis, we make use of common data entry tools to minimize the impact of heterogeneous exposure data. More specifically, exposures of interest may be unknown or missing. In chapter five of this thesis, we report a methodological study conducted within ADVANCE which tests four methods for estimating incidence of an adverse event following vaccination in a multi-database study when data on vaccination is incomplete or missing for a subset of databases.

Bias in studies of vaccine effects

As in all epidemiological studies, bias and confounding must be taken into account in studies of vaccine effects. Bias can take the form of reporting (i.e. subjects or their physicians are more or less likely to report an event or exposure due to other underlying factors), ascertainment (i.e. events and exposures become more or less easy to identify in the data due to underlying factors), recall (i.e. subjects or physicians recall exposures or events differentially based upon other factors) and selection (i.e. subjects are more or less likely to be included/excluded in a study due to factors such as healthcare seeking behavior, socioeconomic status, or underlying health status) (69). Confounding must also be taken into account. These sources of bias and confounding are typically addressed

through matching, stratification, control for covariates, and use of case-only designs. In database studies, variations on propensity score methodologies have become prevalent in recent years while in vaccine safety, the self-controlled case series (SCCS) and variations on this method are dominant for control of bias and confounding (70, 71). However, these measures and designs cannot completely remove the effects of bias. In order to address the impact of two specific forms of bias: ascertainment and recall, in chapter 5 we present a simulation study in which behavior of subjects was modified to mimic these two sources of bias and the impact of these behavior changes on effect estimates was measured.

Aims and Outline of this thesis

In **Chapter Two,** *Methods for signal detection in spontaneous reporting system databases*: The thesis begins with assessment of the impact of stratification by age in the FDA FAERS database using a pediatric-specific drug-adverse event reference set (72). We then proceed to assess how the presence of vaccine exposures in a mixed vaccine/therapeutic database affects performance of SDAs in pediatric age strata within the EudraVigilance database using the same pediatric drug reference set. Finally, we use a vaccine-adverse event reference set to assess the suitability of standard SDAs for reports associated with vaccines in the US VAERS and EudraVigilance databases by comparing EBGM to the vaccine-specific time-to-onset method(35).

In **Chapter Three**: *Methods for rapid signal assessment*, we describe use of population-level incidence data to assess the association between Pandemic H1N1 vaccines and narcolepsy globally. Because the vaccination campaign of the 2009 pandemic was targeted, we were able to calculate incidence of narcolepsy before and after the vaccination campaign. We discuss suitability of population-based ecological methods for rapid assessment of vaccine safety concerns.

In **Chapter Four:** *Single database studies for study of vaccine effects,* we use the United Kingdom The Health Information Network (THIN) database to study the potential association between pandemic H1N1-containing vaccines and facial paralysis. We then use the same database to study the effectiveness of measles-containing vaccines by exploiting the reduced MMR uptake and subsequent measles cases in the UK following the since disproved claim of a MMR-autism link.

In **Chapter Five:** *Collaborative studies to assess vaccine effects,* we again focus on pandemic influenza vaccine in studies of the association of pH1N1 vaccine and Guillain-Barré Syndrome (GBS) and subsequently of pH1N1 vaccine and narcolepsy. These studies address many of the issues inherent to collaborative studies – namely power considerations, heterogeneity, and pooling – and approach them in different ways. The GBS study uses a case-only approach as proof of concept for the WHO that sentinel based studies are possible on a global scale. The narcolepsy global study went much further and used a case-control study to be conducted within each site with one common protocol, data collection methods and use of a remote research environment for pooling. While these two studies include only developed countries, the final study of this chapter is a proof-of-concept study, based on sentinel sites, which aimed to evaluate known adverse events in order to prove the suitability of the collaborative framework as well as the data sources for conduct of vaccine safety studies in the future. This study employed a common protocol, data collection tool and secure sharing of data for pooling.

In **Chapter Six:** *Methods for dealing with heterogeneity and bias*, we look further into some of the sources of heterogeneity and bias acknowledged in the other chapters of this thesis. In the first manuscript, conducted in response to the VAESCO narcolepsy studies and in conjunction with the SOMNIA study, data was simulated to mimic hypothesized reporting patterns of narcolepsy following pH1N1 vaccine which may have changed due to public awareness of the purported association (47, 73). We then analyzed the simulated data to understand how

such changes in reporting may bias effect estimates. Second, within ADVANCE, we investigate a new method for better identifying events in electronic health records by combining codes and records from various domains such as diagnosis, prescribing, and laboratory tests. For our test case, we use pertussis, a vaccine-preventable disease. Finally, within ADVANCE and reusing data from the first ADVANCE proof-of-concept study, we test four methods to obtain estimates of adverse events following vaccine exposure when data on exposure is incomplete or missing.

Chapter	Title	Project	Data Source(s)	Purpose-built Tools Used
2.1	Drug safety monitoring in children: Performance of signal detection algorithms and impact of age stratification	GRiP	FAERS	RRE
2.2	Masking by vaccines in pediatric signal detection: A study in the EudraVigilance database	GRiP	EudraVigilance	Vacc-O, RRE
2.3	Signal detection in VAERS and Eudravigilance using disproportionality analysis and time to onset	GRiP	VAERS	GRiP Common Data Model, RRE
3.1	Incidence rates of narcolepsy diagnoses in Taiwan, Canada, and Europe: methods for assessment of potential safety issues on a population level in the SOMNIA study	SOMNIA	Global data sources using common protocol and harmonized event definitions	Jerboa, RRE
4.1	Bell's palsy and influenza(H1N1)pdm09 containing vaccines: a self-controlled case series	N/A	THIN	N/A
4.2	The impact and longevity of measles-associated immune suppression: a population-based matched cohort study	N/A	THIN	N/A
5.1	International Collaboration to Assess the Risk of Guillain Barré Syndrome Following Influenza A (H1N1) 2009 Monovalent Vaccines	The Global H1N1 GBS Consortium	Global data sources using common protocol and harmonized event definitions	Common protocol
5.2	Narcolepsy and adjuvanted pandemic influenza A (H1N1) 2009 vaccines: a global investigation	SOMNIA	Global data sources using common protocol, common data entry tools, and harmonized event definitions	Chameleon, Jerboa, RRE
5.3	Enhancing global vaccine pharmacovigilance: Proof-of-concept study on aseptic meningitis and immune thrombocytopenic purpura following measles-mumps containing vaccination	WHO GVS-MCC	Global data sources using common protocol, common data entry tools, and harmonized event definitions	Chameleon, Jerboa, RRE
6.1	Pandemic Influenza vaccine & Narcolepsy: Simulations on the potential impact of bias	SOMNIA	Simulated data	
6.2	Heterogeneity in disease misclassification: the component	ADVANCE	European ADVANCE partners using	Codemapper, RRE

	analysis		harmonized code lists, common input files, and common analysis scripts	
6.3	Title: Estimating incidence of adverse events following vaccination in observational databases with incomplete exposure data	ADVANCE	European ADVANCE partners using harmonized code lists, common input files, and common analysis scripts	Codemapper, Advance common data model, Common protocol, RRE

References

1. Whitney CG, Zhou F, Singleton J, Schuchat A. Benefits from immunization during the vaccines for children program era - United States, 1994-2013. MMWR Morbidity and mortality weekly report. 2014;63(16):352-5. Epub 2014/04/25.

2. Rothman KJ. Epidemiology : an introduction. New York, N.Y.: Oxford University Press; 2002. viii, 223 p. p.

3. GUIDELINE ON CLINICAL EVALUATION OF NEW VACCINES EMEA/CHMP/VWP/164653/2005 (2006).

4. Heidi Larson EK. ECDC TECHNICAL REPORT: Rapid literature review on motivating hesitant population groups in Europe to vaccinate. Stockholm: European Centers for Disease Prevention and Control, 2015 October 2015. Report No.

5. Larson HJ, Cooper LZ, Eskola J, Katz SL, Ratzan S. Addressing the vaccine confidence gap. Lancet. 2011;378(9790):526-35. Epub 2011/06/15.

6. Chen RT, Rastogi SC, Mullen JR, Hayes SW, Cochi SL, Donlon JA, et al. The Vaccine Adverse Event Reporting System (VAERS). Vaccine. 1994;12(6):542-50. Epub 1994/05/01.

7. Chen RT, Shimabukuro TT, Martin DB, Zuber PL, Weibel DM, Sturkenboom M. Enhancing Vaccine Safety Capacity Globally: A Lifecycle Perspective. American journal of preventive medicine. 2015;49(6 Suppl 4):S364-76. Epub 2015/11/23.

8. Adverse Effects of Vaccines: Evidence and Causality. Washington, DC: Institute of Medicine, 2012.

9. Trifiro G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? Journal of internal medicine. 2014;275(6):551-61. Epub 2014/03/19.

10. Stratton KR, Howe CJ, Johnston RB, Jr. Adverse events associated with childhood vaccines other than pertussis and rubella. Summary of a report from the Institute of Medicine. Jama. 1994;271(20):1602-5. Epub 1994/05/25.

11. Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. International journal of epidemiology. 2006;35(2):337-44. Epub 2005/12/22.

12. Davis RL, Marcuse E, Black S, Shinefield H, Givens B, Schwalbe J, et al. MMR2 immunization at 4 to 5 years and 10 to 12 years of age: a comparison of adverse clinical events after immunization in the Vaccine Safety Datalink project. The Vaccine Safety Datalink Team. Pediatrics. 1997;100(5):767-71. Epub 1997/11/14.

13. Fine PE, Chen RT. Confounding in studies of adverse reactions to vaccines. American journal of epidemiology. 1992;136(2):121-35. Epub 1992/07/15.

14. Hurwitz ES, Schonberger LB, Nelson DB, Holman RC. Guillain-Barré syndrome and the 1978-1979 influenza vaccine. The New England journal of medicine. 1981;304(26):1557-61. Epub 1981/06/25.

15. Haber P, DeStefano F, Angulo FJ, Iskander J, Shadomy SV, Weintraub E, et al. Guillain-Barré syndrome following influenza vaccination. Jama. 2004;292(20):2478-81. Epub 2004/11/25.

16. Stowe J, Andrews N, Wise L, Miller E. Investigation of the temporal association of Guillain-Barré syndrome with influenza vaccine and influenzalike illness using the United Kingdom General Practice Research Database. American journal of epidemiology. 2009;169(3):382-8. Epub 2008/11/27.

17. Dieleman J, Romio S, Johansen K, Weibel D, Bonhoeffer J, Sturkenboom M. Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe. BMJ. 2011;343:d3908. Epub 2011/07/14.

18. Wijnans L, de Bie S, Dieleman J, Bonhoeffer J, Sturkenboom M. Safety of pandemic H1N1 vaccines in children and adolescents. Vaccine. 2011;29(43):7559-71. Epub 2011/08/09.

19. Chen RT, Glasser JW, Rhodes PH, Davis RL, Barlow WE, Thompson RS, et al. Vaccine Safety Datalink project: a new tool for improving vaccine safety monitoring in the United States. The Vaccine Safety Datalink Team. Pediatrics. 1997;99(6):765-73. Epub 1997/06/01.

20. Nguyen M, Ball R, Midthun K, Lieu TA. The Food and Drug Administration's Post-Licensure Rapid Immunization Safety Monitoring program: strengthening the federal vaccine safety enterprise. Pharmacoepidemiology and drug safety. 2012;21 Suppl 1:291-7. Epub 2012/01/25.

21. Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. Drug safety. 2002;25(6):381-92. Epub 2002/06/20.

22. Zhou W, Pool V, Iskander JK, English-Bullard R, Ball R, Wise RP, et al. Surveillance for safety after immunization: Vaccine Adverse Event Reporting System (VAERS)--United States, 1991-2001. MMWR Surveill Summ. 2003;52(1):1-24. Epub 2003/06/27.

23. Stephenson WP, Hauben M. Data mining for signals in spontaneous reporting databases: proceed with caution. Pharmacoepidemiology and drug safety. 2007;16(4):359-65. Epub 2006/10/05.

24. van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiology and drug safety. 2002;11(1):3-10. Epub 2002/05/10.

25. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. The American Statistician. 1999;53(3):177-90.

26. Goldman SA. Limitations and strengths of spontaneous reports data. Clinical therapeutics. 1998;20 Suppl C:C40-4. Epub 1999/01/23.

27. Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiology and drug safety. 2009;18(6):427-36. Epub 2009/04/10.

28. Star K, Edwards IR. Pharmacovigilance for children's sake. Drug Safety. 2014;37:91-8.

29. Woo EJ, Ball R, Burwen DR, Braun MM. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug Safety. 2008;31:667-74.

30. Orre R, Lansner, A., Bate, A., & Lindquist, M. Bayesian neural networks with confidence estimations applied to data mining. Computational Statistics & Data Analysis,. 2000;34(4):473-93.

31. Almenoff JS, LaCroix KK, Yuen NA, Fram D, DuMouchel W. Comparative performance of two quantitative safety signalling methods. Drug Safety. 2006;29(10):875-87.

32. Norén GN, Bate A, Orre R, Edwards IR. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. Statistics in Medicine. 2006;25:3740-57.

33. Almenoff JS, Pattishall EN, Gibbs TG, DuMouchel W, Evans SJW, Yuen N. Novel statistical tools for monitoring the safety of marketed drugs. Clinical Pharmacology & Therapeutics. 2007;82(2):157-66.

34. Hopstadius J, Norén GN, Bate A, Edwards IR. Impact of stratification on adverse drug reaction surveillance. Drug Safety. 2008;31(11):1035-48.

35. Van Holle L, Zeinoun Z, Bauchau V, Verstraeten T. Using time-to-onset for detecting safety signals in spontaneous reports of adverse events following immunization: a proof of concept study. Pharmacoepidemiology and drug safety. 2012;21(6):603-10. Epub 2012/03/03.

36. Smirnov NV. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. Bull Math Univ Moscou 1939;2(2).

37. Massey FJ. The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association. 1951;46(253):68–78.

38. GRiP. Global Research in Paediatrics–Network of Excellence (GRiP) 2011 [cited 2017 30 August 2017]; Available from: http://grip-network.org/.

39. Hauben M, Noren GN. A decade of data mining and still counting. Drug safety. 2010;33(7):527-34. Epub 2010/06/18.

40. The Sentinel Initiative: National Strategy for Monitoring Medical Product Safety. Programs OoCP; 2008.

41. Lieu TA, Kulldorff M, Davis RL, Lewis EM, Weintraub E, Yih K, et al. Real-time vaccine safety surveillance for the early detection of adverse events. Medical care. 2007;45(10 Supl 2):S89-95. Epub 2007/10/25.

42. Brown JS, Kulldorff M, Chan KA, Davis RL, Graham D, Pettus PT, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. Pharmacoepidemiology and drug safety. 2007;16(12):1275-84. Epub 2007/10/24.

43. Li L, Kulldorff M. A conditional maximized sequential probability ratio test for pharmacovigilance. Statistics in medicine. 2010;29(2):284-95. Epub 2009/11/27.

44. Lee GM, Greene SK, Weintraub ES, Baggs J, Kulldorff M, Fireman BH, et al. H1N1 and seasonal influenza vaccine safety in the vaccine safety datalink project. American journal of preventive medicine. 2011;41(2):121-8. Epub 2011/07/20.

45. Gee J, Naleway A, Shui I, Baggs J, Yin R, Li R, et al. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the Vaccine Safety Datalink. Vaccine. 2011;29(46):8279-84. Epub 2011/09/13.

46. Yih WK, Kulldorff M, Sandhu SK, Zichittella L, Maro JC, Cole DV, et al. Prospective influenza vaccine safety surveillance using fresh data in the Sentinel System. Pharmacoepidemiology and drug safety. 2016;25(5):481-92. Epub 2015/11/18.

47. Wijnans L, Lecomte C, de Vries C, Weibel D, Sammon C, Hviid A, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine. 2013;31(8):1246-54. Epub 2012/12/19.

48. Maglione MA, Das L, Raaen L, Smith A, Chari R, Newberry S, et al. Safety of Vaccines Used for Routine Immunization of US Children: A Systematic Review. Pediatrics. 2014;134(2):325-37.

49. Osokogu OU, Dukanovic J, Ferrajolo C, Dodd C, Pacurariu AC, Bramer WM, et al. Pharmacoepidemiological safety studies in children: a systematic review. Pharmacoepidemiology and drug safety. 2016;25(8):861-70. Epub 2016/06/04.

50. Garcia Rodriguez LA, Perez Gutthann S. Use of the UK General Practice Research Database for pharmacoepidemiology. British journal of clinical pharmacology. 1998;45(5):419-25. Epub 1998/06/27.

51. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages. Nature clinical practice Rheumatology. 2007;3(12):725-32. Epub 2007/11/27.

52. Salmon DA, Proschan M, Forshee R, Gargiullo P, Bleser W, Burwen DR, et al. Association between Guillain-Barré syndrome and influenza A (H1N1) 2009 monovalent inactivated vaccines in the USA: a meta-analysis. Lancet. 2013;381(9876):1461-8. Epub 2013/03/19.

53. Sibbald B. Rofecoxib (Vioxx) voluntarily withdrawn from market. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2004;171(9):1027-8. Epub 2004/10/27.

54. Ball R, Robb M, Anderson SA, Dal Pan G. The FDA's sentinel initiative--A comprehensive approach to medical product surveillance. Clinical pharmacology and therapeutics. 2016;99(3):265-8. Epub 2015/12/17.

55. Salmon D, Yih WK, Lee G, Rosofsky R, Brown J, Vannice K, et al. Success of program linking data sources to monitor H1N1 vaccine safety points to potential for even broader safety surveillance. Health Aff (Millwood). 2012;31(11):2518-27. Epub 2012/11/07.

56. Izurieta HS, Zuber P, Bonhoeffer J, Chen RT, Sankohg O, Laserson KF, et al. Roadmap for the international collaborative epidemiologic monitoring of safety and effectiveness of new high priority vaccines. Vaccine. 2013;31(35):3623-7. Epub 2013/05/28.

57. CONSORTIUM A. [cited 2017 28 August 2017]; Available from: http://www.advance-vaccines.eu/.

58. Destefano F, and C. Vellozzi. . Facilitators Report, Lessons learned exercise, ECDC Vaccine Adverse Event Surveillance and Communication:(VAESCO II) project. Stockholm: (2009).

59. Perez-Vilar S, Weibel D, Sturkenboom M, Black S, Maure C, Castro JL, et al. Enhancing global vaccine pharmacovigilance: Proof-of-concept study on aseptic meningitis and immune thrombocytopenic purpura following measles-mumps containing vaccination. Vaccine. 2017. Epub 2017/06/01.

60. Coloma PM, Schuemie MJ, Trifiro G, Gini R, Herings R, Hippisley-Cox J, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiology and drug safety. 2011;20(1):1-11. Epub 2010/12/25.

61. Becker BFH, Avillach P, Romio S, van Mulligen EM, Weibel D, Sturkenboom M, et al. CodeMapper: semiautomatic coding of case definitions. A contribution from the ADVANCE project. Pharmacoepidemiology and drug safety. 2017;26(8):998-1005. Epub 2017/06/29.

62. Dodd CN, Romio SA, Black S, Vellozzi C, Andrews N, Sturkenboom M, et al. International collaboration to assess the risk of Guillain Barre Syndrome following Influenza A (H1N1) 2009 monovalent vaccines. Vaccine. 2013;31(40):4448-58. Epub 2013/06/19.

63. Black S, Miriam Sturkenboom, Jan Bonhoeffer, Frank Destefano, Tom Shimabukuro, Caitlin Dodd, Maria De Ridder, Daniel Weibel, Angela Gentile, Noberto Giglio, Salahddin Mahmud. The Risk of Narcolepsy Following Receipt of Adjuvanted Pandemic 2009 The Subtype of Influenza A Virus (H1N1) Vaccines: Results of the SOMNIA Global Collaborative Study. Open Forum Infectious Diseases. 2016;3(suppl 1.):792.

64. Coloma PM, Valkhoff VE, Mazzaglia G, Nielsson MS, Pedersen L, Molokhia M, et al. Identification of acute myocardial infarction from electronic healthcare records using different disease coding systems: a validation study in three European countries. BMJ open. 2013;3(6). Epub 2013/06/26.

65. Klungel OH, Kurz X, de Groot MC, Schlienger RG, Tcherny-Lessenot S, Grimaldi L, et al. Multi-centre, multi-database studies with common protocols: lessons learnt from the IMI PROTECT project. Pharmacoepidemiology and drug safety. 2016;25 Suppl 1:156-65. Epub 2016/04/03.

66. Maro JC, Brown JS, Dal Pan GJ, Kulldorff M. Minimizing signal detection time in postmarket sequential analysis: balancing positive predictive value and sensitivity. Pharmacoepidemiology and drug safety. 2014;23(8):839-48. Epub 2014/04/05.

67. Reisinger SJ, Ryan PB, O'Hara DJ, Powell GE, Painter JL, Pattishall EN, et al. Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases. Journal of the American Medical Informatics Association : JAMIA. 2010;17(6):652-62. Epub 2010/10/22.

68. Roberto G, Leal I, Sattar N, Loomis AK, Avillach P, Egger P, et al. Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. PloS one. 2016;11(8):e0160648. Epub 2016/09/01.

69. Rothman KJ, Sander Greenland, and Timothy L. Lash. Modern epidemiology Lippincott Williams & Wilkins; 2008.

70. Sturmer T, Joshi M, Glynn RJ, Avorn J, Rothman KJ, Schneeweiss S. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. Journal of clinical epidemiology. 2006;59(5):437-47. Epub 2006/04/25.

71. Weldeselassie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. Epidemiology and infection. 2011;139(12):1805-17. Epub 2011/08/19.

72. Osokogu OU, Fregonese F, Ferrajolo C, Verhamme K, de Bie S, t Jong G, et al. Pediatric drug safety signal detection: a new drug-event reference set for performance testing of data-mining methods and systems. Drug safety. 2015;38(2):207-17. Epub 2015/02/11.

73. Black S, et al. . The Risk of Narcolepsy Following Receipt of Adjuvanted Pandemic 2009 The Subtype of Influenza A Virus (H1N1) Vaccines: Results of the SOMNIA Global Collaborative Study. Open Forum Infectious Diseases: Oxford University Press; 2016.

CHAPTER 2. METHODS FOR SIGNAL DETECTION IN SPONTANEOUS REPORTING SYSTEM DATABASES

2.1 Drug safety monitoring in children: Performance of signal detection algorithms and impact of age stratification

Osemeke U Osokogu*

Caitlin Dodd*

Alexandra Pacurariu

Florentia Kaguelidou

Daniel Weibel

Miriam C J M Sturkenboom

*Authors have contributed equally to the manuscript

Drug safety 39.9 (2016): 873-881.

Abstract

Introduction: Spontaneous reports of suspected adverse drug reactions (ADRs) can be analysed to yield additional drug safety evidence for the pediatric population. Signal detection algorithms (SDAs) are required however the performance of SDAs in the pediatric population specifically is unknown. We tested the performance of two SDAs on pediatric data from the US FDA Adverse Event Reporting System (FAERS) and investigated the impact of age stratification and age adjustment on SDAs' performance.

Methods: We tested the performance of two established SDAs: Proportional Reporting Ratio (PRR) and Empirical Bayes Geometric Mean (EBGM) on a pediatric dataset from FAERS (2004 to 2012). We compared SDAs' performance to a published pediatric-specific reference set, by calculating diagnostic-test related statistics including the area under the Receiver Operating Characteristics curve (AUC). Impact of age stratification and age-adjustment SDAs' performance was assessed. Age adjustment was performed by pooling (Mantel-Hanszel) stratum-specific estimates.

Results: A total of 115,674 pediatric reports (patients aged 0-18 years) comprising 893,587 drug-event combinations (DECs) were analysed. Crude values of the AUC were similar for both SDAs: 0.731 (PRR) and 0.745 (EBGM). Stratification unmasked four DECs, for example 'ibuprofen and thrombocytopenia'. Age-adjustment did not improve performance.

Conclusion: The performance of the two tested SDAs was similar in the pediatric population. Age adjustment does not improve performance and is therefore not recommended to be performed routinely. Stratification can reveal new associations, therefore is recommended when either drug use is age-specific or when an age-specific risk is suspected.

Introduction

Spontaneous reports of suspected adverse drug reactions (ADRs) can yield important information regarding the safety of drugs [1]. Usually, such reports are screened for emerging safety issues by applying statistical methods called signal detection algorithms (SDAs). Current SDAs compare the reporting rate of a drug-event combination (DEC) of interest with the expected count calculated from the overall reporting rate of that reaction in the entire database [1, 2]. Although SDAs are routinely applied to reports pertaining to the general population, the performance of SDAs in the pediatric population specifically has not been investigated to date. Compared to adults, the pattern of drug use and occurrence of ADRs in pediatrics may differ [3-5] since the latter population comprises a heterogeneous group of subjects at various stages of development with age-dependent organ maturation and hormonal changes [6]. Several studies investigating ADR reporting in children identified different reporting patterns in this population compared to adults [3, 5, 7, 8]. Since ADRs may be age–specific, adjustment for age seems to be a logical step when investigating pediatric ADRs and has been advocated by some researchers [4]. The major aim of stratification is verification of confounding and effect modification which otherwise may mask true signals [9]. Confounding by age can be dealt with by stratifying for age categories and pooling stratum-specific estimates. However if age specific estimates differ (in case of effect modification) pooling/adjustment should not be done, but instead, a verification of each individual stratum. While stratification has been investigated by some researchers [10], adjustment is routinely implemented in some Bayesian but not in frequentist SDAs [11-13] Few studies have systematically addressed the impact of age stratification or adjustment and the results are contradictory [9, 14, 15].

Within the context of the Global Research in Pediatrics (GRiP) – Network of excellence [16], we aimed to evaluate the performance of two well-established SDAs in the pediatric population and determine if age stratification or adjustment impacts signal detection in this population.

Methods

Data source

Data was retrieved from the publicly available version of the US FDA Adverse Event Reporting System (FAERS), which comprises spontaneous reports of suspected ADRs submitted by manufacturers, healthcare professionals and patients. FAERS is one of the largest repositories of spontaneous reports in the world [17, 18]. In this study, we analyzed reports received from the first quarter of 2004 through the third quarter of 2012.

For performance analysis, only reports of ADRs occurring in children and adolescents (<18 years of age) were retained. The ADRs in FAERS are coded according to the Medical Dictionary for Regulatory Activities (MedDRA®)[19].

To improve the quality of the dataset, we excluded reports with missing age, the main variable in our study. Also, reports with reported age equal to zero and with a MedDRA® preferred term indicating prenatal exposure were removed, as these imply *in-utero* drug exposure and were therefore not relevant for our study. We minimized the number of duplicates (i.e. the same report submitted by different reporters) by applying an algorithm based on case identifier, report identifier, drug and event names. For multiple reports (i.e. the same report is reported at a later time, with additional and updated information) [20], the most recent (and most updated) report was retained for analysis.

As drug names included in FAERS are not standardized, a harmonization procedure was implemented. Briefly, this consisted of removing superfluous characters and applying a generalized edit distance matching algorithm [21] to

map free text drug names to synonyms and finally to the corresponding active substance and World Health Organization-Anatomic Therapeutic Chemical (WHO-ATC) code.

In this study, only those drugs reported as primary or secondary suspect in the FAERS database were retained for analysis. Analysis was performed at Drug-Event Combination (DEC) level, meaning that within each report, every suspect drug was combined with all reported ADRs. Thus, one report may comprise more than one DEC.

Signal detection algorithms (SDAs)

We tested two well established SDAs which are routinely used by various national and international regulatory and/or research institutions for signal detection: the proportional reporting ratio (PRR) [2] and the empirical Bayes geometric mean (EBGM) [13] (see Table 1). We also tested count of reports, as a positive control. In order to define a signal of disproportionate reporting (SDR) [22, 23], we selected thresholds that are currently applied in routine practice. We applied the SDAs at the end of the study period, when the maximum number of reports had accrued.

Table 1 Signal detection algorithms and corresponding thresholds applied

Signal Detection Algorithm	Applied Threshold[a]	Institution where the method and the respective threshold is currently used
Number of reports	n ≥5	*NA*
PRR	PRR lower bound 95% CI ≥ 1 & n ≥ 5 reports	European Medicines Agency
EBGM	EB05 CI≥ 1.8 and n≥ 3 reports & EBGM ≥2.5	Medicines and Healthcare products Regulatory Agency (MHRA)

PRR= Proportional reporting ratio; EBGM= Empirical Bayes Geometric Mean; CI=confidence interval; NA= Not available; EB05= Lower bound of the 95% confidence interval

[a] Thresholds were obtained from Candore et al [23]

Performance assessment measures

The performance of the SDAs was assessed by calculating diagnostic-test related statistics, namely specificity and sensitivity, positive predictive value (PPV), and negative predictive value (NPV) [24][25]. Sensitivity is the ability of the method to correctly identify true signals while specificity is the ability to correctly exclude false signals. PPV and NPV are posterior probabilities, describing how many of the signals classified as positive or negative are indeed correctly classified [24, 25].

Since diagnostic-test related statistics are dependent on the threshold choice, their individual comparison has only limited, albeit practical value. Therefore, we also estimated the area under the curve (AUC) of receiver operating characteristics (ROC) in order to compare the performance of the SDAs [32]; the AUC incorporates both sensitivity and specificity across all the possible values for a certain SDA. Calculation of AUCs was conducted by varying only the point estimate of each SDA and did not take into account the other components of the SDA.

For the purpose of performance evaluation, a previously constructed pediatric-specific GRIP reference set of positive and negative drug-event associations was used. It consists of 37 positive and 90 negative DECs and includes drugs that are administered to children and events that are regarded as important for this population. The positive DECs are those that were confirmed to occur based on evidence from Summary of Product Characteristics (SmPC) and the published literature, while the negative DECs are those that could not be confirmed at the time of literature review by neither the SmPC nor the published literature. For a full description of the reference set, see Osokogu et al [26].

Stratification and adjustment for age

The impact of age stratification and adjustment on the performance of the SDAs was investigated. First, we checked for possible effect modification across age strata, by stratifying the data according to age categories defined by the International Conference on Harmonization (ICH) [27] and calculating stratum-specific measures for each SDA.

Secondly, we calculated age-adjusted estimates for PRR and EBGM by combining the stratum-specific estimates in an overall measure [28]. The performance of each SDA was reassessed after adjustment.

Statistical analysis

Differences in the performance (AUC) of each SDA, crude versus age-adjusted and crude versus count of reports (positive control) were tested using paired chi-squared tests. Stratum-specific contingency tables were tested for homogeneity using the Breslow Day Tarone test [29]. The Mantel-Haenszel approach was used for pooling and calculating age-adjusted estimates [28]. The lower bound of the EBGM 95% confidence interval (EBGM05) was calculated using the EB05 for each stratum and then computing a Mantel-Haenszel average based upon Zeinoun [30]. Statistical significance was defined by p value < 0.05.

Analysis was performed using SAS software version 9.2. Graphs were made in SAS software version 9.2 and R version 3.1.3.

Results

Descriptive analysis

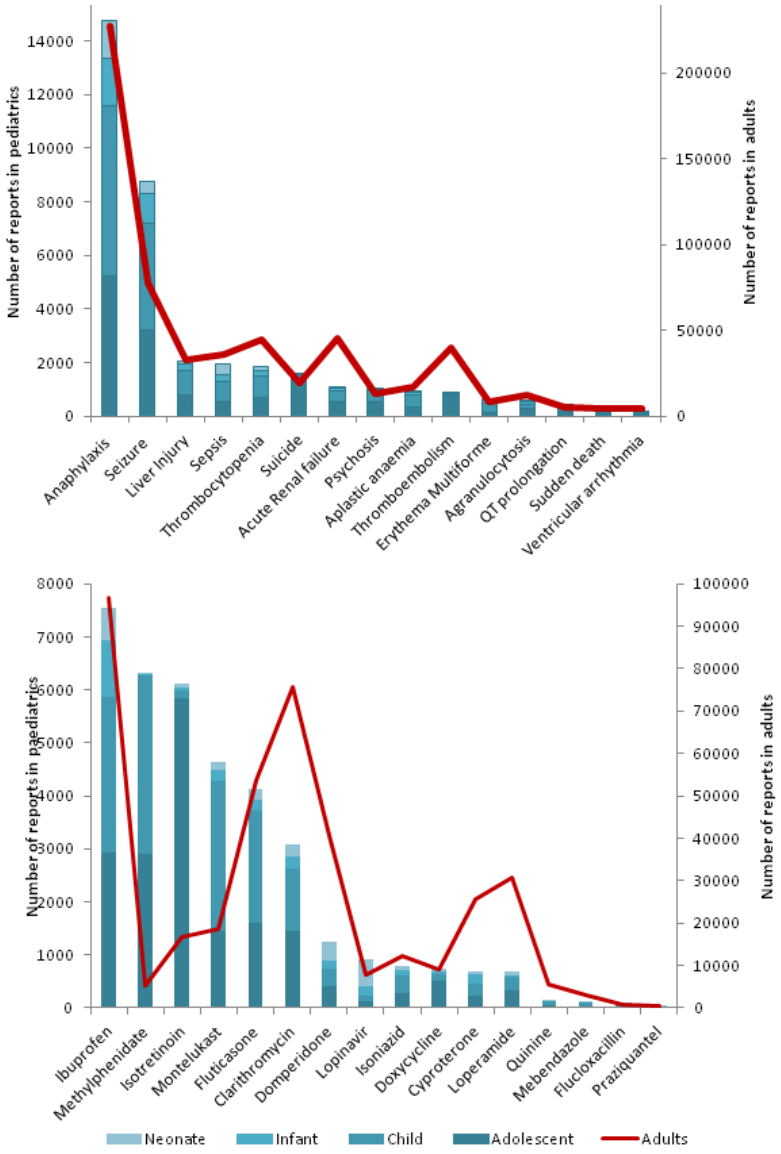
For the study period (first quarter of 2004 through the third quarter of 2012), a total of 4,285,088 reports were retrieved from FAERS. After eliminating duplicates (n=43,125), removal of adult reports (n=2,686,530) and reports with missing age (n=1,419,524) or age equal to zero with a MedDRA® preferred term indicating prenatal exposure (n=20,235), 115,674 reports corresponding to 893,587 individual DECs were retained for analysis of pediatric spontaneous reports (see Table 2).

Table 2 Description of pediatric reports by age categories

Age group	**Number of reports, n (%)**
Neonates: 0-27 days	5,091 (4.40%)
Infants: 28 days-23 months	12,566 (10.86%)
Children: 2-11 years	49,982 (43.21%)
Adolescents: 12-17 years	48,035 (41.53 %)
Total	**115,674 (100%)**

The total number of pediatric reports that included the investigated drugs and ADRs from the reference set can be observed in Fig. 1, which also shows data regarding adults (for comparison purposes). The number of children exposed to the drugs of interest, for whom any of the investigated ADRs was reported, varied from 26 patients (for praziquantel) to 7,535 patients (for ibuprofen) with a median of 781 patients exposed across all drugs. The number of events of interest in FAERS ranged from 164 reports (ventricular arrhythmia) to 14,777 (anaphylaxis), with a median of 1,004 reports across all events. For a more detailed description of reports counts please refer to *Electronic Supplementary material Table 1.*

Fig. 1 Count of reports in pediatric and adult population for the investigated ADRs and drugs, cumulatively for the period Q1 2004 -Q3 2012[a]



a -Number of reports in children is represented by bars and plotted on the left axis, while the number of reports in adults is represented by the red line and plotted on the right axis; Reports with missing age or age=0 were excluded. Only reports mentioning any of the drugs or events in the reference set were considered.

Overall performance of SDAs

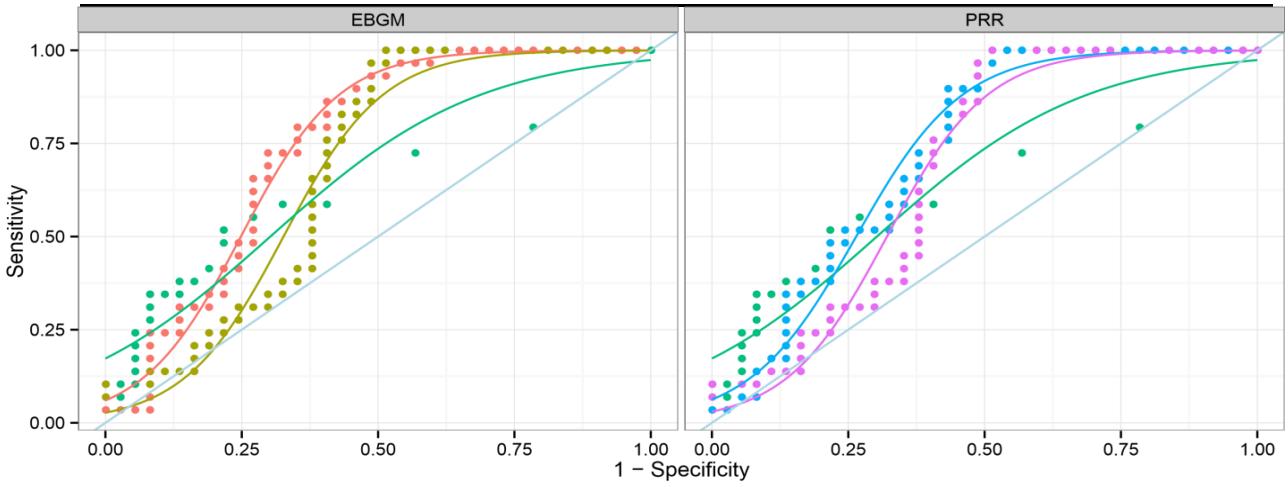
Both SDAs showed high specificity and low sensitivity. They both had similar specificity values (PRR:83.8% and EBGM:91.9%), while sensitivity was lower for EBGM than for PRR (17.2% vs. 37.9%). The NPV and PPV were similar for both SDAs. When we applied the threshold-independent (AUC-based) approach, the tested SDAs showed similar performance in the pediatric population although the AUC value for EBGM (0.745) was slightly higher than for PRR (0.731). None of the SDAs performed better than the simple report count (AUC=0.634, *p-values: PRR=0.27 and EBGM=0.14)*

Stratification and adjustment for age and its impact on performance

Upon calculating SDA values per age stratum and testing for heterogeneity across strata, we observed effect modification for some associations. Some false negatives (positive DECs which failed to be highlighted as signals when analyzing data pertaining to the entire pediatric population) were unmasked in some strata. Four DECs were unmasked in total: ibuprofen-thrombocytopenia and isoniazid-seizure (by PRR) and clarithromycin-erythema multiforme and ibuprofen-erythema multiforme (by EBGM). Conversely, 'ibuprofen-acute liver injury', also a positive DEC, was highlighted when we analyzed data pertaining to the entire pediatric population but after stratifying, it became clear that this DEC was highlighted only in older children (adolescents), and not highlighted in younger children (see Fig. 3). For an overview of SDA values across age strata and results of heterogeneity tests please refer to the *Electronic Supplementary material figures 1A and 1B.*

Fig. 2 Performance of signal detection algorithms within the entire pediatric population

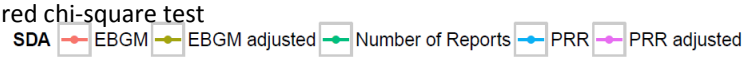
SDA	Sensitivity	Specificity	PPV	NPV	AUC	p-value[b]
Number of reports	58.62	67.57	58.62	67.57	0.634	*reference*
PRR	37.93	83.78	64.71	63.27	0.731	0.266
EBGM	17.24	91.89	62.50	58.62	0.745	0.144
After age adjustment [a]						*(reference-crude PRR/EBGM)*
PRR	34.48	86.49	66.67	62.75	0.688	0.267
EBGM	10.34	97.30	75.00	58.06	0.683	0.216



SDA-signal detection algorithm; PRR= Proportional reporting ratio; EBGM= Empirical Bayes Geometric Mean; AUC=area under the curve; PPV=positive predictive value; NPV-negative predictive value

[a] adjusted PRR/ROR values calculated by combining the individual estimates from each age stratum into one measure according to the Mantel-Haenszel approach.

[b] paired chi-square test



We evaluated the performance of the methods within individual age strata (see Table 3). On average, performance of the SDAs was lower within age strata compared to the entire pediatric population and performance improved with increasing stratum size. For infants and neonates, the performance was very low, not better than chance (*p-value* > 0.5 for both SDAs). The adolescent group exhibited the best performance which was similar to the overall performance.

Table 3 Performance of signal detection algorithms across age strata

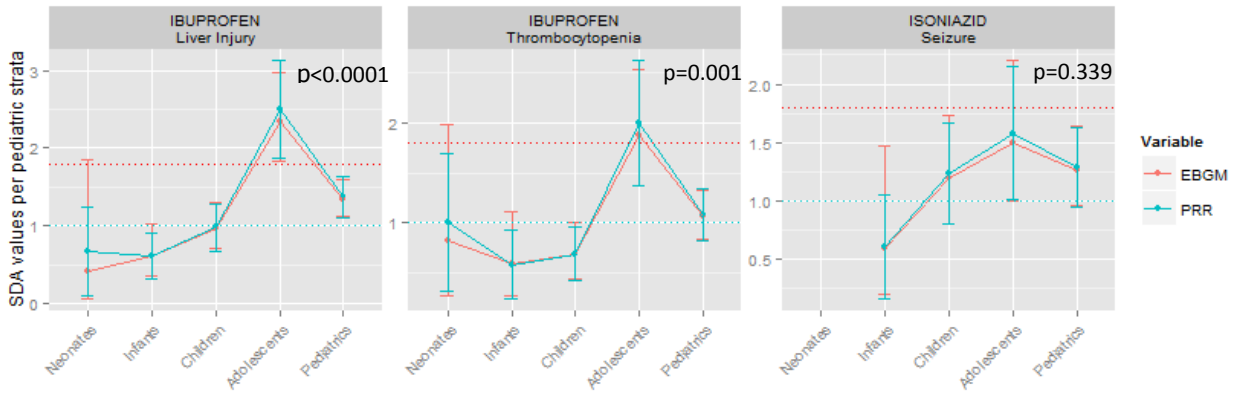
Age groups	**Signal Detection Algorithms**	**Size of the age stratum (number of reports)**	**AUC**
Neonates		5,091	
	Number of Reports		0.625
	EBGM		0.600
	PRR		0.65
Infants		12,566	
	Number of Reports		0.667
	EBGM		0.548
	PRR		0.554
Children		49,982	
	Number of Reports		0.654
	EBGM		0.698
	PRR		0.649
Adolescents		48,035	
	Number of Reports		0.698
	EBGM		0.771
	PRR		0.718
Entire pediatric population		115,674	

	Number of Reports	0.634
	EBGM	0.746
	PRR	0.733

PRR= Proportional reporting ratio; EBGM= Empirical Bayes Geometric Mean; AUC=area under the curve

After adjusting for age by pooling the stratum-specific estimates, the performance of the SDAs decreased, although not significantly (see Fig. 2; crude vs. adjusted AUC for PRR 0.731 vs. 0.688, *p-value = 0.267;* crude vs. adjusted AUC for EBGM 0.745 vs. 0.683, *p-value = 0.216*).

Fig. 3 Variation of PRR and EBGM estimates across pediatric specific strata –selected examples



p-values were calculated with Breslow Day Tarone test for homogeneity

Discussion

In this study, we have demonstrated that age stratification for detection of drug safety signals in children may unmask some signals that do not appear in neither crude nor adjusted analysis. Adjustment for age does not improve performance of the PRR and EBGM.

For the investigated events, similar reporting patterns were observed for children and adults while the investigated drugs appeared to have different reporting patterns (see Figure 1). Different drug-related reporting patterns in children vs adults were previously reported [5].Consequently, reported drug-event associations for children may differ from adults [3, 5], underlining the need for pediatric-specific approaches to signal detection especially when we consider that even within the pediatric population, reported drugs may vary by age group [3, 31].

Overall, the PRR and EBGM showed good performance although results were slightly lower than results reported on other (not pediatric-specific) reference sets [32, 33]. The similarity in performance between PRR and EBGM is in accordance with recent results from the PROTECT project [23]. The fact that the performance (based on AUC) of PRR and EBGM was not statistically significantly better than simple report count may be due to the lack of power. Within age strata, performance seemed to correlate with stratum size: the poorest results were observed for infants and neonates (the smaller groups), slightly improving for children while the best performance was observed for adolescents, the age stratum with the highest number of tested DECs. Decrease in power due to fewer reports and therefore DECs may account for this observation. The fact that we used lower bounds of confidence intervals for signaling instead of point estimates might have exacerbated the influence of sample size on the results, since smaller strata will have higher variability. In neonates and infants for whom expected counts were

difficult to calculate because of few reports, we observed that simple report counts performed similar or even better than the SDAs and might be an alternative to commonly used SDAs. The fact that simple report count performed better than SDAs may have been because the reference set comprised known DECs (which in turn may have influenced reporting) rather than emerging safety issues, a hypothesis proposed by Noren et al [34].

Inspection of SDA values across child specific strata (age-stratification) revealed some heterogeneity in estimates pointing to some effect modification. For example, 'ibuprofen-thrombocytopenia', was found as a signal in the adolescents' group but not detected in the entire pediatric population or the younger age categories. This suggests that age-specific SDA calculations are sometimes needed, rather than age-adjusted SDA estimates. The age-adjusted estimates did not improve performance; in fact even PPV unexpectedly decreased. Simulation studies have shown that when adjusted for strata, Bayesian methods such as EBGM tend to be underestimated when there are sparse strata [15]; this was also the case in our study. Previous studies in adults show contradictory results, with some showing a beneficial effect [9] while others did not [15]. The reason for our finding is not entirely clear; a possible explanation is that age is not a strong confounder for the investigated DECs. Also, the method of weighting (Mantel-Haenszel approach) may have played a role since more weight was assigned to age groups with more reports (adolescents and children). This may have masked signals occurring in age groups with fewer reports.

The limitations of data mining in FAERS include those inherent to spontaneous reporting databases: underreporting, lack of denominator data and control group, biases in reporting, as well as missing and poor quality data [35]. Missing information regarding age substantially reduced the study sample size since we could not determine whether these reports described patients aged less than 18 years old. While these biases are well acknowledged and have a definite impact, they cannot be completely avoided. Compared to adults, there are fewer reports and different reporting patterns for children [3, 36, 37] which may complicate signal detection in the pediatric population.

Evaluating performance of SDAs is a constant challenge due to lack of standard methodologies, imperfect reference standards and uncertainty regarding the best thresholds (See supplementary material for measures of performance using alternative thresholds). Some of the drugs and events in the reference set are specific to one age group within pediatrics and this is obvious in Fig. 1, even though the reference set was designed to be relevant for the entire pediatric population. We acknowledge that the reference set used, although specifically constructed for this purpose, does not include all the ADRs that are highly specific for pediatrics. This highlights the need for pediatric-specific approaches to signal detection; accounting for not just the entire pediatric population but also the different age strata within pediatrics. Still, the reference set captures various drug use and ADRs patterns [38] and is currently the only available pediatric-specific reference set. The thresholds applied to define a signal were obtained from previous publications and other cut-off points may generate better results; further research on pediatric-specific thresholds should be encouraged.

Conclusion

Our study revealed that age adjustment did not improve performance of the SDAs. However, stratification revealed some variation in SDAs' values across strata (effect modification) and inspection of stratum-specific estimates might sometimes yield useful information during routine surveillance.

Compliance with ethical standards

Funding: The Global Research in Pediatrics–Network of Excellence is funded under the European Union's Seventh Framework Program (FP7/2007–2013) for research, technological development and demonstration under grant agreement number 261060. Funding for this study was also received from the 'Priority Medicines Kinderen project

ZONMW: EVIPED: Novel methods to assess and compare drug effects in pediatrics' (Grant agreement number 113201007). The funders had no role whatsoever in designing and conducting the study, the collection and management of data, and preparation, review, or approval of the manuscript.

References

1. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiology and Drug Safety. 2009 2009;18:427–36.

2. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiology and Drug Safety. 2001 10/2001;10:483-6.

3. Blake KV, Zaccaria C, Domergue F, La Mache E, Saint-Raymond A, Hidalgo-Simon A. Comparison between paediatric and adult suspected adverse drug reactions reported to the European medicines agency: implications for pharmacovigilance. Paediatric Drugs. 2014 Aug 2014;16:309-19.

4. Star K, Edwards IR. Pharmacovigilance for children's sake. Drug Safety. 2014 Feb 2014;37:91-8.

5. Star K, Norén GN, Nordin K, Edwards IR. Suspected adverse drug reactions reported for children worldwide: an exploratory study using VigiBase. Drug Safety. 2011 May 1, 2011;34:415-28.

6. Kearns GL, Abdel-Rahman SM, Alander SW, Blowey DL, Leeder JS, Kauffman RE. Developmental pharmacology - Drug disposition, action, and therapy in infants and children. New England Journal of Medicine. 2003 2003;349:1157-67.

7. Aagaard L, Weber CB, Hansen EH. Adverse drug reactions in the paediatric population in Denmark: a retrospective analysis of reports made to the Danish Medicines Agency from 1998 to 2007. Drug Safety. 2010 Apr 1, 2010;33:327-39.

8. Kimland E, Rane A, Ufer M, Panagiotidis G. Paediatric adverse drug reactions reported in Sweden from 1987 to 2001. Pharmacoepidemiology and Drug Safety. 2005 Jul 2005;14:493-9.

9. Woo EJ, Ball R, Burwen DR, Braun MM. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug Safety. 2008 2008;31:667-74.

10. Orre R, Lansner, A., Bate, A., & Lindquist, M. Bayesian neural networks with confidence estimations applied to data mining. Computational Statistics & Data Analysis,. 2000 2000;34(4):473-93.

11. Almenoff JS, LaCroix KK, Yuen NA, Fram D, DuMouchel W. Comparative performance of two quantitative safety signalling methods. Drug Safety. 2006;29(10):875-87.

12. Norén GN, Bate A, Orre R, Edwards IR. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. Statistics in Medicine. 2006 Nov 15, 2006;25:3740-57.

13. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. The American Statistician. 1999;53(3):177-90.

14. Almenoff JS, Pattishall EN, Gibbs TG, DuMouchel W, Evans SJW, Yuen N. Novel statistical tools for monitoring the safety of marketed drugs. Clinical Pharmacology & Therapeutics. 2007;82(2):157-66.

15. Hopstadius J, Norén GN, Bate A, Edwards IR. Impact of stratification on adverse drug reaction surveillance. Drug Safety. 2008;31(11):1035-48.

16. Global Research in Paediatrics. Network of excellence. Global Research in Paediatrics. Network of excellence,. [cited 2016 28th February]; Available from: http://www.grip-network.org/index.php/cms/en/home

17. Moore TJ, Furberg CD. The Safety Risks of Innovation: The FDA's Expedited Drug Development Pathway. JAMA. 2012 2012-09-05;308:869.

18. Weiss-Smith S, Deshpande G, Chung S, Gogolak V. The FDA drug safety surveillance program: adverse event reporting trends. Archives of Internal Medicine. 2011 Mar 28, 2011;171:591-3.

19. MedDRA MSSO. Introductory Guide MedDRA Version 15.1. 2012 [cited 2016 28th February]; Available from: http://www.meddra.org/sites/default/files/guidance/file/intguide_15_1_English_0.pdf

20. Poluzzi E, Piccinni C, Raschi E, De Ponti F. Data Mining Techniques in Pharmacovigilance: Analysis of the Publicly Accessible FDA Adverse Event Reporting System (AERS): INTECH Open Access Publisher; 2012.

21. SAS. The power to know. SAS(R) 9.2 Language Reference: Dictionary, Fourth Edition. COMPGED Function. Returns the generalized edit distance between two strings. Computing the Generalized Edit Distance. [cited; Available from: http://support.sas.com/documentation/cdl/en/lrdict/64316/HTML/default/viewer.htm#a002206133.htm

22. Hauben M, Aronson JK. Defining 'signal'and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Safety. 2009;32(2):99-110.

23. Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et al. Comparison of Statistical Signal Detection Methods Within and Across Spontaneous Reporting Databases. Drug Safety. 2015 2015.

24. Zhou X-H, McClish DK, Obuchowski NA. Statistical methods in diagnostic medicine: John Wiley & Sons; 2009.

25. Strom BL KS, Hennessy S,. Pharmacoepidemiology. 5th ed. Chichester, West Sussex, UK: Wiley-Blackwell; 2012.

26. Osokogu OU, Fregonese F, Ferrajolo C, Verhamme K, de Bie S, Catapano M, et al. Pediatric Drug Safety Signal Detection: A New Drug–Event Reference Set for Performance Testing of Data-Mining Methods and Systems. Drug Safety. 2015;38(2):207-17.

27. INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE ICH HARMONISED TRIPARTITE GUIDELINE. CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS IN THE PEDIATRIC POPULATION E11. Current Step 4 version 2000 [cited 2016 28th February]; Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E11/Step4/E11_Guideline.pdf

28. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute. 1959 Apr 1959;22:719-48.

29. Breslow NE, Day NE. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC scientific publications. 1980 1980:5-338.

30. Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. Human vaccines. 2009 Sep;5(9):599-607.

31. de Bie S, Ferrajolo C, Straus SMJM, Verhamme KMC, Bonhoeffer J, Wong ICK, et al. Pediatric Drug Safety Surveillance in FDA-AERS: A Description of Adverse Events from GRiP Project. PloS One. 2015 2015;10:e0130399.

32. van Puijenbroek EP, Bate A, Leufkens HGM, Lindquist M, Orre R, Egberts ACG. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiology and Drug Safety. 2002 2002;11:3–10.

33. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of Pharmacovigilance Signal-Detection Algorithms for the FDA Adverse Event Reporting System. Clin Pharmacol Ther. 2013 2013;93:539-46.

34. Norén GN, Caster O, Juhlin K, Lindquist M. Zoo or savannah? Choice of training ground for evidence-based pharmacovigilance. Drug Safety. 2014 Sep 2014;37:655-9.

35. Almenoff J, Tonning JM, Gould AL, Szarfman A, Hauben M, Ouellet-Hellstrom R, et al. Perspectives on the use of data mining in pharmacovigilance. Drug Safety. 2005;28(11):981-1007.

36. Aagaard L, Strandell J, Melskens L, Petersen PSG, Holme Hansen E. Global patterns of adverse drug reactions over a decade: analyses of spontaneous reports to VigiBase™. Drug Safety. 2012 Dec 1, 2012;35:1171-82.

37. Morales-Olivas FJ, Martínez-Mir I, Ferrer JM, Rubio E, Palop V. Adverse drug reactions in children reported by means of the yellow card in Spain. Journal of Clinical Epidemiology. 2000 Oct 2000;53:1076-80.

38. Impicciatore P, Choonara I, Clarkson A, Provasi D, Pandolfini C, Bonati M. Incidence of adverse drug reactions in paediatric in/out-patients: a systematic review and meta-analysis of prospective studies. British Journal of Clinical Pharmacology. 2001 Jul 2001;52:77-83.

Supplementary Table 1 Counts of reports DECs from reference set in pediatrics

True Positive Associations [a]					
Drug	**Event**	**A**	**B**	**C**	**D**[b]
Clarithromycin	Erythema Multiforme	11	778	540	113446
Clarithromycin	Liver Injury	18	771	2216	111770
Clarithromycin	Psychosis	9	780	1323	112663
Clarithromycin	QT prolongation	6	783	511	113475
Clarithromycin	Sudden death	1	788	276	113710
Clarithromycin	Thrombocytopenia	11	778	1689	112297
Clarithromycin	Ventricular arrhythmia	3	786	243	113743
Domperidone	Sudden death	1	78	276	114420
Doxycycline	Erythema Multiforme	1	208	550	114016
Doxycycline	Thrombocytopenia	3	206	1697	112869
Flucloxacillin	Liver Injury	3	14	2231	112527
Ibuprofen	Acute Renal failure	247	3015	1262	110251
Ibuprofen	Anaphylaxis	547	2715	14296	97217
Ibuprofen	Erythema Multiforme	36	3226	515	110998
Ibuprofen	Liver Injury	86	3176	2148	109365
Ibuprofen	Thrombocytopenia	52	3210	1648	109865
Isoniazid	Liver Injury	13	344	2221	112197
Isoniazid	Psychosis	3	354	1329	113089
Isoniazid	Seizure	37	320	9198	105220
Isotretinoin	Psychosis	99	3130	1233	110313
Isotretinoin	Suicide	197	3032	2023	109523
Lopinavir	Liver Injury	5	202	2229	112339

Mebendazole	Liver Injury	2	56	2232	112485
Montelukast	Psychosis	95	2316	1237	111127
Montelukast	Suicide	130	2281	2090	110274
Quinine	Agranulocytosis	1	96	528	114150
Quinine	Liver Injury	2	95	2232	112446
Quinine	Thrombocytopenia	1	96	1699	112979
Quinine	Ventricular arrhythmia	1	96	245	114433

True Negative Associations [a]

Drug	**Event**	**A**	**B**	**C**	**D**[b]
Clarithromycin	Suicide	3	786	2217	111769
Clarithromycin	Thromboembolism	1	788	697	113289
Cyproterone	Anaphylaxis	2	17	14841	99915
Domperidone	Acute Renal failure	2	77	1507	113189
Domperidone	Agranulocytosis	2	77	527	114169
Domperidone	Anaphylaxis	5	74	14838	99858
Domperidone	Aplastic anemia	1	78	994	113702
Domperidone	Liver Injury	2	77	2232	112464
Domperidone	Sepsis	2	77	2012	112684
Domperidone	Suicide	3	76	2217	112479
Domperidone	Thrombocytopenia	4	75	1696	113000
Doxycycline	QT prolongation	1	208	516	114050
Doxycycline	Suicide	5	204	2215	112351
Doxycycline	Thromboembolism	3	206	695	113871
Fluticasone	Acute Renal failure	2	1703	1507	111563
Fluticasone	Liver Injury	3	1702	2231	110839

Fluticasone	QT prolongation	1	1704	516	112554
Fluticasone	Sepsis	3	1702	2011	111059
Fluticasone	Thrombocytopenia	6	1699	1694	111376
Ibuprofen	Sepsis	79	3183	1935	109578
Isoniazid	Sepsis	1	356	2013	112405
Isoniazid	Ventricular arrhythmia	6	351	240	114178
Isotretinoin	Anaphylaxis	230	2999	14613	96933
Isotretinoin	QT prolongation	2	3227	515	111031
Isotretinoin	Sepsis	28	3201	1986	109560
Loperamide	Acute Renal failure	2	166	1507	113100
Loperamide	Liver Injury	1	167	2233	112374
Loperamide	Sepsis	3	165	2011	112596
Loperamide	Suicide	21	147	2199	112408
Loperamide	Thrombocytopenia	11	157	1689	112918
Mebendazole	Anaphylaxis	13	45	14830	99887
Montelukast	Anaphylaxis	142	2269	14701	97663
Montelukast	Aplastic anemia	4	2407	991	111373
Montelukast	QT prolongation	1	2410	516	111848
Montelukast	Sepsis	4	2407	2010	110354
Montelukast	Thrombocytopenia	15	2396	1685	110679
Montelukast	Thromboembolism	1	2410	697	111667

[a] Reference set associations with no reports in pediatrics not presented

[b] A, B, C, and D represent the following cell counts:

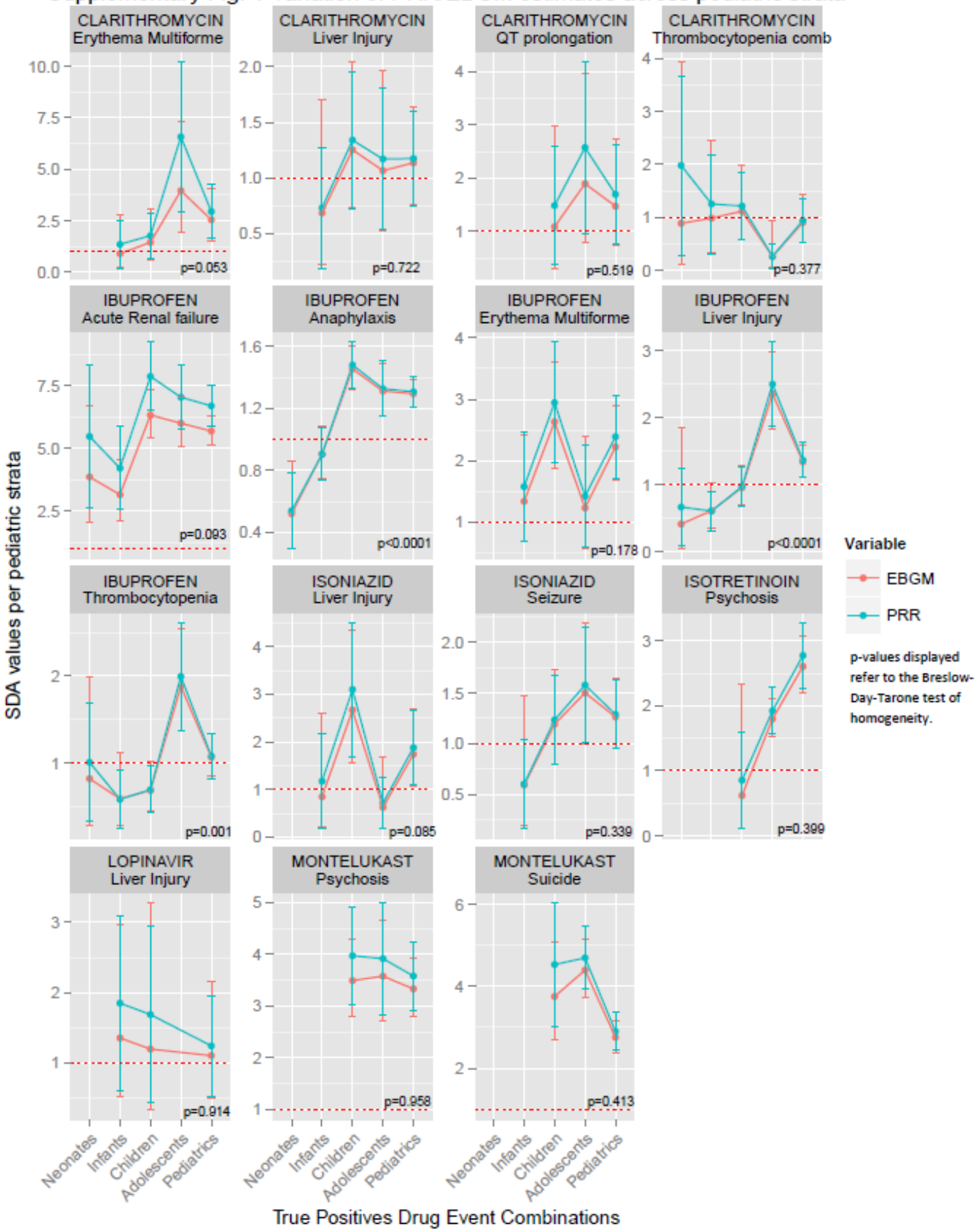
A = Reports related to the drug of interest and the event of interest

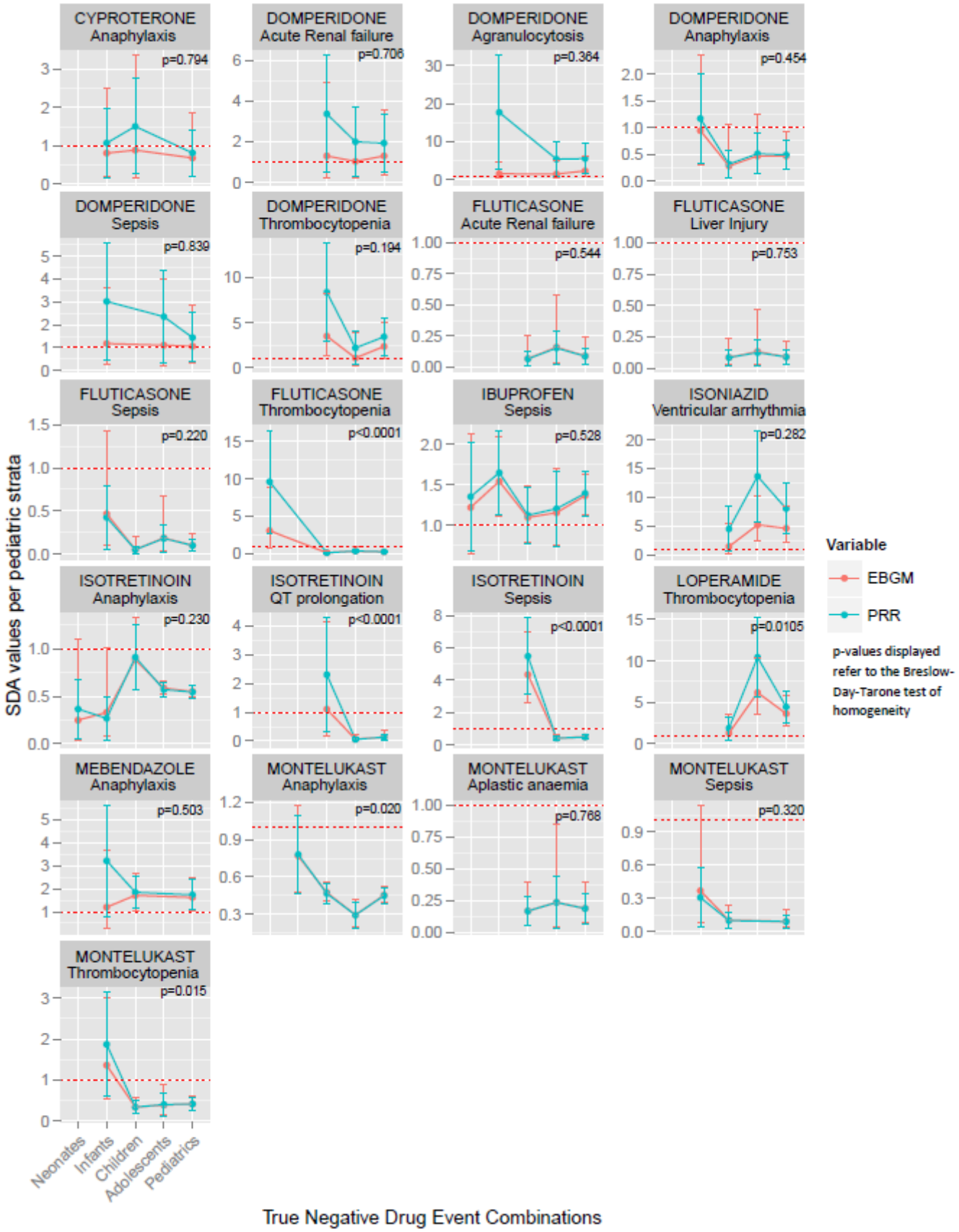
B = Reports of the exposure of interest associated with a different event

C = Reports of the event of interest associated with a different exposure

D = Reports related to exposures and events other than those of interest

Supplementary Fig. 1 Variation of PRR/EBGM estimates across pediatric strata




CYPROTERONE
Anaphylaxis
p=0.794
DOMPERIDONE
Acute Renal failure
p=0.706
DOMPERIDONE
Agranulocytosis
p=0.364
DOMPERIDONE
Anaphylaxis
p=0.454
DOMPERIDONE
Sepsis
p=0.839
DOMPERIDONE
Thrombocytopenia
p=0.194
FLUTICASONE
Acute Renal failure
p=0.544
FLUTICASONE
Liver Injury
p=0.753
FLUTICASONE
Sepsis
p=0.220
FLUTICASONE
Thrombocytopenia
p<0.0001
IBUPROFEN
Sepsis
p=0.528
ISONIAZID
Ventricular arrhythmia
p=0.282
ISOTRETINOIN
Anaphylaxis
p=0.230
ISOTRETINOIN
QT prolongation
p<0.0001
ISOTRETINOIN
Sepsis
p<0.0001
LOPERAMIDE
Thrombocytopenia
p=0.0105
MEBENDAZOLE
Anaphylaxis
p=0.503
MONTELUKAST
Anaphylaxis
p=0.020
MONTELUKAST
Aplastic anaemia
p=0.768
MONTELUKAST
Sepsis
p=0.320
MONTELUKAST
Thrombocytopenia
p=0.015
SDA values per pediatric strata
Neonates
Infants
Children
Adolescents
Pediatrics
True Negative Drug Event Combinations
Variable
EBGM
PRR
p-values displayed refer to the Breslow-Day-Tarone test of homogeneity

2.2 MASKING BY VACCINES IN PEDIATRIC SIGNAL DETECTION: A STUDY IN THE EUDRAVIGILANCE DATABASE

Caitlin Dodd

Alexandra Pacurariu

Osemeke U Osokogu

Daniel Weibel

Carmen Ferrajolo

Dang H Vo

Benedikt Becker

Jan A Kors

Miriam C J M Sturnkenboom

Pharmacoepidemiology and drug safety 27.11 (2018): 1249-1256.

Abstract

Post-marketing drug safety surveillance relies upon measures of disproportionate reporting in spontaneous reporting systems. It has been hypothesized that products or events reported frequently may 'mask' signals.

We analyzed the masking effect of vaccines in pediatrics in the EudraVigilance database by conducting disproportionality analysis in the full database (containing vaccine exposures) and in a restricted set (excluding vaccine exposures). We measured performance of the reporting odds ratio (ROR) in both data sets using a pediatric-specific drug reference set and in the absence of a reference set. We assessed masking effects across age groups and conducted a classification tree (CART) analysis.

Removal of vaccines decreased the ROR values both in negative and positive controls. Exceptions were drug-event combinations including outcomes frequent in vaccine reports. When restricted to positive control associations, removal of vaccine-related events resulted in increased ROR values for events commonly reported following vaccination. For events rarely associated with vaccination, ROR values decreased for all age groups, especially infants. Analysis in the absence of a reference set showed decrease in ROR following vaccine removal and CART revealed that change in ROR with vaccine removal depended upon age and proportion of reports including a vaccine.

Removal of vaccines for signal detection in a pediatric population has an impact on ROR, dependent upon the reporting frequency of the event of interest in combination with vaccines. We recommend stratification by age and removal of vaccine exposures if the investigated ADRs include those typically reported in association with vaccines for the age strata.

Introduction

The paucity of clinical studies in the pediatric population is well known, resulting in lack of direct evidence regarding drug safety in this vulnerable population (1). Consequently, legislation has been introduced in the last decade, aimed at stimulating pediatric-specific drug research (2,3). The Global Research in Pediatrics (GRiP) Network of Excellence (2) was conceptualized with one of the aims being to improve global pediatric pharmacoepidemiology and pharmacovigilance.

Post-marketing drug safety surveillance relies to a large extent on spontaneous reporting systems (SRSs). To automatically estimate if a specific drug-event combination (DEC) is a potential signal, statistical disproportionality analysis is usually applied (4). Such analysis consists of computing an observed to expected event ratio, obtained by comparing a specific DEC with the background reporting rate for that event in relation to all other drugs. SRSs lack real, population-based denominator data or an appropriate control group, and are subject to various biases in reporting (5). The denominator in this case is represented by the background reporting rate in the database. For disproportionality calculation, the background reporting rate in the database has a large impact on the results and it is hypothesized that products or events which are present in large numbers due to frequent reporting may induce a masking effect for certain associations (6).

A masking effect might occur when the signal for a given DEC is suppressed by the presence of other known DECs that are overrepresented in the reporting database, thereby increasing the threshold for detection (7,8). The extent and impact of masking on the detection of new signals of public health relevance is not fully understood; some studies suggest that its effect is limited (9,10) while others have demonstrated that removal of masking can lead to discovery of new relevant signals (7). Maignen et al. created an algorithm for detection and quantification of the masking effect, which is a ratio of the value of the signal-detection algorithm (SDA) in the presence of the hypothesized masking agent(s) divided by the value of the SDA after the removal of the hypothesized masking agent(s) (8). Masking was shown to occur more frequently in small databases (11), compared to larger databases (9,11). The reason for that might be that larger databases have a larger and more diverse denominator, which seems to be more robust to masking. Additionally, masking is reported to have a larger impact on events that are rarely reported compared to frequent events (8).

The highest masking effect is induced by products for which the reaction (event) is known, or extensively reported (8). Since vaccines are frequently administered in the pediatric population and vaccine-related reports constitute up to 50% of the pediatric reports within spontaneous reporting databases (11-13), it is plausible that in the younger population, vaccine-related reports might induce a masking effect for safety signals arising from use of small molecules and non-vaccine biologicals (non-vaccines). On the other hand, it is widely accepted that vaccines are often administered to healthy subjects unlike non-vaccines, possibly making the set of adverse events reported following vaccines distinct enough from those of non-vaccine medicines to avoid a masking effect (14).

In this study we aimed to achieve the following: first, to analyze the masking effect of vaccines in the pediatric population in an European SRS (EudraVigilance) database; second, to investigate any differential masking effect across different pediatric age groups; and third, to compare our findings to those reported by others who have investigated masking by vaccines.

Methods

Study setting and data source

The study was conducted within EudraVigilance, the European spontaneous reporting database maintained by the European Medicines Agency, which contained approximately 6.2 million spontaneous reports of suspected adverse drug reactions (ADRs) at the end of 2015 (15). Only reports of ADRs pertaining to the pediatric population (<18 years of age) during the study period of 2002-2015 were requested through an academic agreement for the GRiP project. Subsequently, a subset of anonymized data fields was provided. For classification of the pediatric population in age groups, we applied the International Conference on Harmonization (ICH) categories with the exception that neonates (<1 month) and infants (1-23 month(s)) were pooled into one category due to the sparseness of reports, resulting in the following categories: neonates and infants (0-23 months), children (24 months-11 years) and adolescents (12-17 years) (16).

Drug and event mapping

All events in EudraVigilance are coded according to the Medical Dictionary for Regulatory Activities (MedDRA®) (17). Drug and vaccine exposures are coded according to the EudraVigilance medicinal product dictionary (XEVMPD or Article 57 database) but in the dataset that we obtained through the academic agreement most drugs had no corresponding codes. Therefore we developed, evaluated, and applied an algorithm to automatically map drug entries in EudraVigilance into World Health Organization-Anatomic Therapeutic Chemical (WHO-ATC) codes. First, we constructed a dictionary of WHO-ATC codes and related drug terms (including trade names and ingredients) based on the Unified Medical Language System (UMLS), version 2016AA (18). Secondly, we applied the Solr search engine (http://lucene.apache.org/solr/) to match EudraVigilance entries to the dictionary.

Reports of vaccine exposures in EudraVigilance were also mapped to WHO-ATC codes using the Solr engine, but with a slightly different approach that was developed as part of the vaccine ontology development in the ADVANCE (Accelerated Development of Vaccine benefit-risk Collaboration in Europe) project (http://www.advance-vaccines.eu/). In EudraVigilance, combination vaccines may be reported with all antigens in one row, or with each antigen in a separate row. If reported separately, a combination vaccine against measles, mumps and rubella, for example, would be assigned to J07BD (measles vaccine), J07BE (mumps vaccine), and J07BJ (rubella vaccine), instead of to J07BD72 (measles-mumps-rubella vaccine). For this reason, we developed an algorithm to group separately recorded antigens within reports according to existing combination vaccines, leaving jointly recorded antigens unchanged, resulting in the lowest possible number of unique exposures per report.

De-duplication

We minimized the number of duplicate reports by applying an algorithm based on report identifier, drug, and event names. For multiple reports (i.e., the same report is reported at a later time, with additional and updated information), the most recent (and most updated) report based upon date of the report was retained for analysis. All reports indicating antenatal exposure (preferred terms 'Maternal exposure during pregnancy', 'Foetal exposure during pregnancy', and 'Exposure via father') were excluded.

One report usually contains more than one suspect drug and more than one reaction. All suspect drugs and vaccines were considered in combination with all reported ADRs and analyzed at the drug/vaccine event combination (DEC/VEC) level.

Analysis

For assessing the association between drug and outcome, we applied a signal detection algorithm - the reporting odds ratio (ROR), using all DECs in the database as the comparator and defined as ad/cb (Table 1). Multiple studies have found little substantive difference among performance of various SDAs (4, 18, 19); therefore for simplicity of

interpretation, we limited our investigations to the ROR. We applied thresholds that are currently applied in routine regulatory signal detection practice (19), which is lower 95% limit of the ROR above 1 and the number of reports of the DEC of interest (cell a) at least 3.

Table 1. Contingency table for a drug or vaccine and adverse event pair in spontaneous reporting data[a]

	Experienced adverse event of interest	Experienced another adverse event	Total
Exposed to drug or vaccine of interest	a	b	(a+b)
Exposed to other drug or vaccine	c	d	(c+d)
Total	(a+c)	(b+d)	(a+b+c+d)

[a]ROR = ab/cd

The impact of masking of associations in non-vaccine reports due to vaccine related reports was assessed via two approaches. First, we assessed the change in performance of the ROR, when applied to the full (vaccine and drug DECs retained) versus restricted (vaccine DECs removed) setting. The analysis was also stratified by age group. For this approach, we used the GRiP pediatric-specific drug-event reference set for non-vaccines which comprises 256 unique DECs, 37 of which are classified as positive control pairs, 90 as negative control pairs, and the remainder unclassifiable (20). We calculated standard performance metrics: area under the curve (AUC) (calculated by varying the threshold for the ROR value), sensitivity, specificity, positive predictive value, and negative predictive value at the predefined threshold of ROR > 1 and number of reports ≥ 3 in the full and restricted settings. Since the AUC assesses performance based upon the entire reference set and may potentially hide different patterns across DECs, we also looked at the change per DEC for the positive and negative controls. Since the amount of confounding will be best observed in DECs where we would expect to observe an association (positive controls) we subsequently calculated the percent change in ROR between the full and restricted settings for the DECs that were indicated as positive controls in the GRiP reference set for non-vaccines, per age stratum.

The second performance assessment was conducted independent of any reference set. We calculated SDA values for all DECs occurring in relation to the top-15 most frequently reported pediatric drugs (non-vaccines) in each pediatric age group. The selection of top-15 drugs was based on convenience and to increase the sample size. The concordance between the ROR values in the full and restricted setting (only non-vaccines) was evaluated by plotting all ROR values against each other in each age group. A slope equal to one would indicate equivalence of the ROR values between the different settings (including vs. excluding vaccine-related reports). A lack of concordance, as defined by deviation from the line of identity, would indicate that removal of vaccine-related reports impacts the ROR value and thereby potential signal detection. This non-concordance was tested by fitting a regression line to the plotted values and assessing whether the 95% confidence interval for its slope included one.

In order to understand which types of associations are mostly affected by masking, we conducted a classification tree analysis on all non-vaccine DECs in the full data set for which an ROR value could be calculated. Classification and regression trees (CART) is a non-parametric method for determining class membership based upon a set of variables. Briefly, the algorithm attempts to predict a target variable by splitting the data recursively according to categories of the predictor variables and repeating until no further gain in group 'purity' can be achieved or until a user-specified stopping rule is reached. We used algorithms which minimized entropy (or maximized node homogeneity) while growing the classification tree and balanced cost (misclassification) with complexity (tree size) while 'pruning' or reducing the tree (20). In this case, the classification tree algorithm was used to split all DECs into the classes 'increased ROR', 'decreased ROR' or 'no change'. An increase in ROR was defined as >10% increase in

ROR value and a decrease was defined as >10% decrease (21). The selected predictor variables were age group and proportion of DECs within the age group containing the event of interest for which the event was reported in association with a vaccine. We also conducted a regression analysis of the impact of vaccine removal by regressing the change in ROR (ROR restricted - ROR full) on age strata, proportion of DECs for the event of interest reported in association with a vaccine, and the interaction of these terms.

Results

A total of 448,364 pediatric reports describing 1,115,324 events and 794,713 exposures were initially retained. Seventy-one reports pertaining to adults or with missing age and 10,589 events indicating exposure *in utero* were excluded. Following removal of duplicate DECs and DECs with missing exposure data (N=121,205), we analyzed 366,062 reports comprising 1,177,375 DECs involving 479,595 suspect drugs/vaccines (Table 2).

Table 2. Description of all ADR reports in pediatrics in EudraVigilance by age categories

Age group	Number of DECs, n (%) non-vaccines & biologicals	Number of DECs, n (%) related to vaccines
Infants: 0 days-23 months	402,817 (34.21%)	208,658 (61.19%)
Children: 2-11 years	406,136 (34.49%)	72,271 (21.19%)
Adolescents: 12-17 years	368,422 (31.29%)	60,064 (17.61%)
Total	**1,177,375**	**340,993**

In figure 1 we report the events mentioned in the GRiP drug reference set, comparing their frequency in the vaccine vs. non-vaccine drug reports. In infants, most events were commonly reported in association with vaccines except suicide and QT prolongation. Anaphylaxis, seizure and sudden death were mentioned more commonly in vaccine reports in infants unlike in children and adolescents for whom the same events were mentioned more commonly in non-vaccine drug reports.

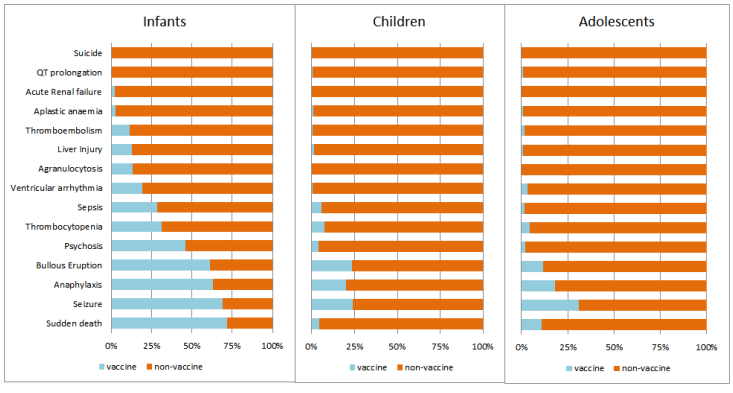


Figure 1 Percentage of each event from reference set reported in combination with vaccines (orange) or nonvaccines (blue) by age strata

Events like acute renal failure, agranulocytosis, aplastic anemia, liver injury, psychosis, QT prolongation, suicide, and thromboembolism occur in less than 5% of the vaccine-related reports in children and adolescents (Figure 1).

For the entire pediatric age range, we could calculate ROR estimates for 28 of 37 positive controls and 34 of 90 negative controls (see Supplementary table 1). The AUC for the full dataset was 0.887 and when we excluded vaccine reports, the AUC was reduced slightly to 0.881. Results of performance assessment using the full reference set can be found in the supplementary material (Supplementary table 2). In comparison of ROR values from the full and restricted data sets for reference set DECs, removal of vaccines from the dataset generally decreased the ROR values both in negative controls as well as positive DECs - with the exception of those DECs that included outcomes that were also relatively frequently reported in vaccine-related reports such as seizure, anaphylaxis, or erythema multiforme (bullous eruption) (Figure 2). For these events the ROR increased upon removal of the vaccines.

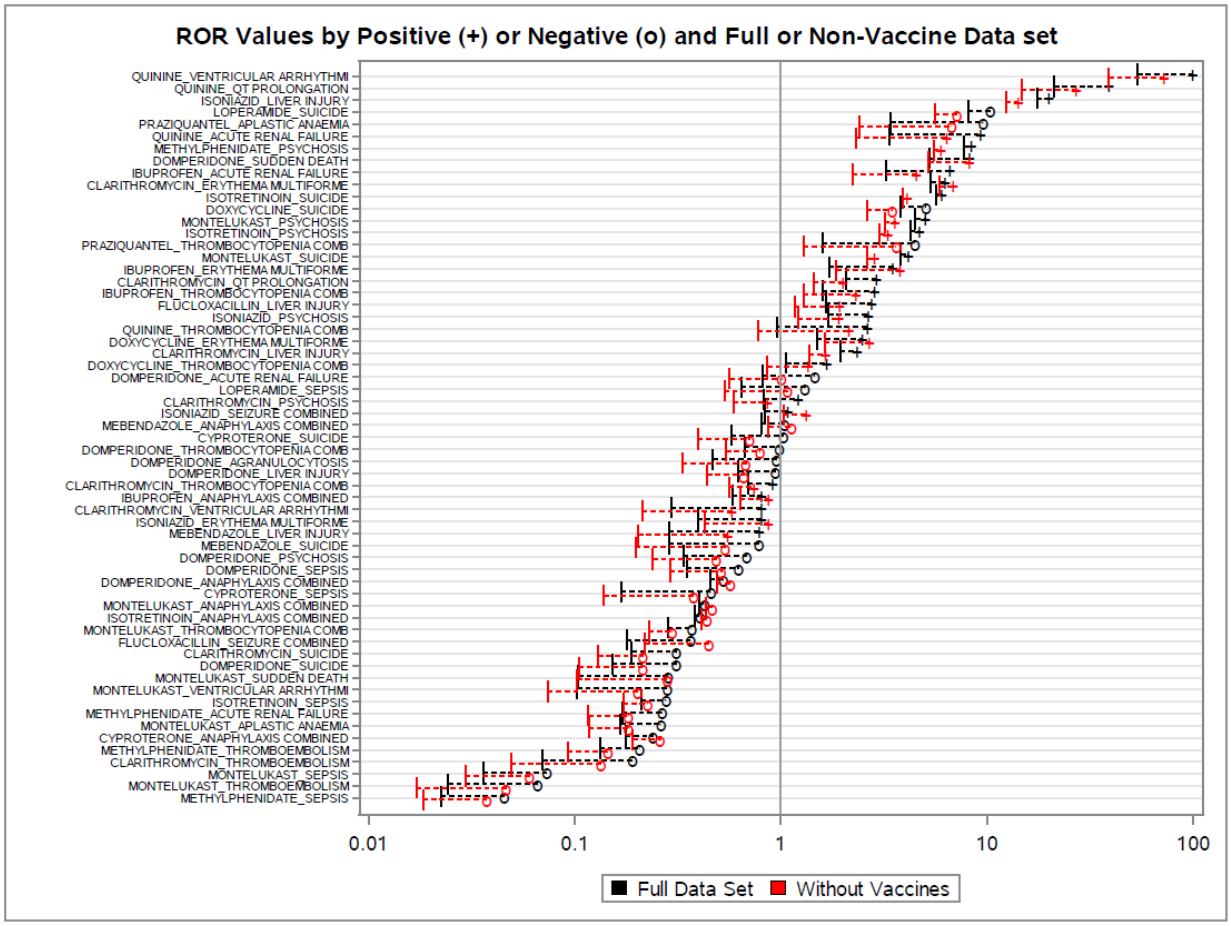


Figure 2 Reporting odds ratio values for drug-event combinations in the Global Research in Pediatrics reference set estimated from the full EudraVigilance pediatric data set (vaccine and nonvaccine) or after exclusion of vaccine-related reports. Symbols represent reporting odds ratio values while vertical bars represent lower 95% confidence bounds

When the analyses were restricted to the positive DECs and stratified by age, we observed that removal of vaccine-related events resulted in an increase in the ROR values for DECs related to seizure, anaphylaxis, or bullous eruption and particularly in infants (Figure 3). For events rarely associated with vaccination such as

thrombocytopenia, liver injury, QT prolongation, and renal failure, the ROR value decreased for all age groups with the most pronounced decrease in infants (Figure 3).

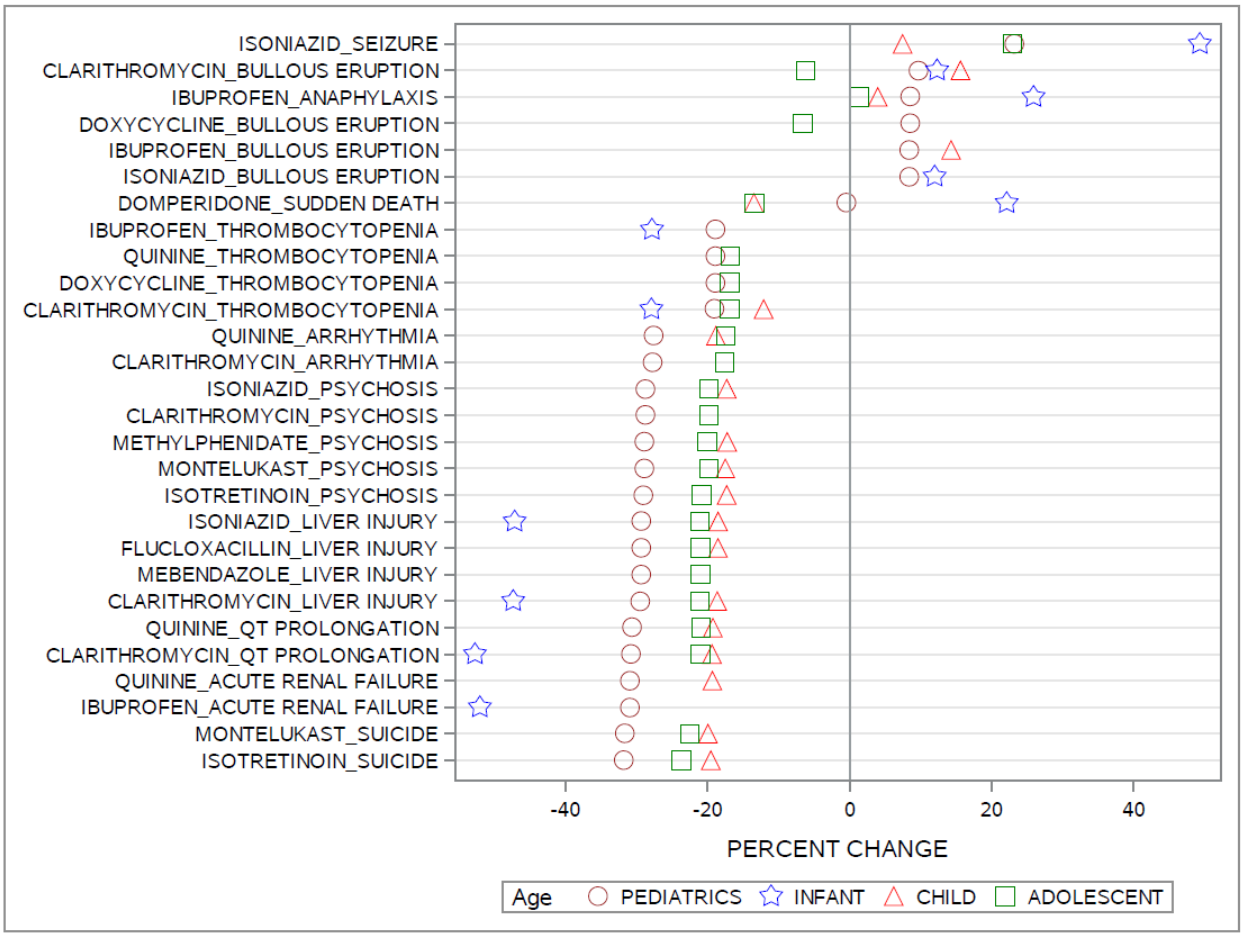


Figure 3 Percent change in reporting odds ratio values for drug-event combinations from the Global Research in Pediatrics reference set that were indicated as positive controls, after removal of vaccine-related reports from the EudraVigilance data set, stratified by age: Pediatrics (0 to <18), infant (0 to <2), child (2 to <12), adolescent (12 to <18)

Analyzing all events related to the top-15 used drugs, showed that in general, vaccine removal decreased the ROR values for DECs (see Figure 4). Regression slopes for infants (0-2 years) and adolescents (12-18) were lower than those in children (2-12 years of age). Regression slopes and their corresponding 95% confidence intervals by age strata are displayed in table 3.

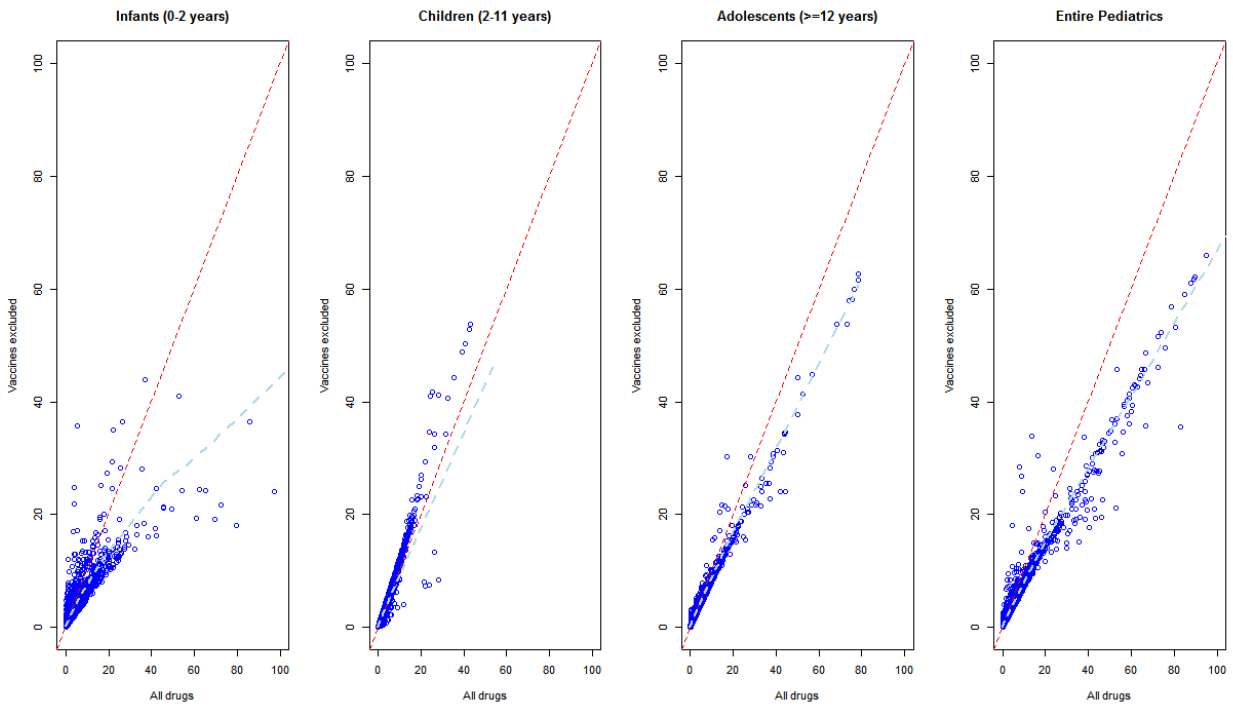


Figure 4 Reporting odds ratio (ROR) values for any event reported for the 15 most frequently reported drugs in EudraVigilance per age category, upon exclusion of vaccine-related reports. The x-axis represents ROR values in the full (vaccine and nonvaccine) database while the y-axis represents ROR values for the same drug-event combinations following removal of vaccine reports

Table 3. Regression slopes of reporting odds ratio (ROR) calculated using full (drug + vaccine) vs nonvaccine (drug) data sets in all drug-event combinations reported for top 15 drugs[a]

	Neonates & Infants	Children	Adolescents	Total Pediatrics
Regression Slope	0.472	0.816	0.788	0.696
(95% CI)	(0.471, 0.474)	(0.814, 0.818)	(0.787, 0.789)	(0.695, 0.697)

[a]Regression Model: ROR (non-vaccine) = a + b*ROR (vaccine)

In the classification-tree analysis of all DECs in the database, we found that increase in ROR estimate with removal of vaccines was dependent upon age group and the proportion of reports for each event that included a vaccine exposure. For infants and adolescents, this proportion was quite high, at 58%; the ROR values were predicted to remain unchanged if this proportion was between 48 and 57% and to decrease if 47% or less. For children, the proportion of reports for each event which included a vaccine exposure was lower at 27%; the ROR values were predicted to remain unchanged if this proportion was between 12 and 27%; and to decrease following vaccine removal if this proportion was less than 11%. See supplementary material for further details. In regression analysis, we found a significant interaction between the proportion of DECs for the event of interest reported in association with a vaccine and age strata on the change in ROR between full and restricted data sets. Removal of vaccine exposures tended to reduce ROR estimates for infants and adolescents when considering all events reported in the database. However, as shown in other analyses presented here, this effect does not hold for those events commonly reported in association with vaccines.

Discussion

In this study we demonstrated that in a spontaneous reporting database that contains vaccines and non-vaccine related adverse reactions, exclusion of vaccines has an important impact on the measures of association in the pediatric population. The direction of the impact differs by age and the frequency of reporting of the event after vaccine exposure. For events mostly reported in non-vaccines, the ROR decreased upon removal of vaccines, whereas it increased for those events that are frequently reported after vaccination. The change was most pronounced in infants (0-2 years), the age group with most vaccine-related reports. Beyond this major finding we observed several other important issues: first, events that were listed in the reference set were reported at very different frequencies, depending on age group, and varied between vaccine and non-vaccine reports. Many events were seldom reported for vaccine-related reports, whereas others had high rates (especially in infants) in vaccine-related reports. In infants, for example, only suicide, which should be non-existent in infants (we assume these reports are errors), and QT prolongation, which is usually not diagnosed in that age range, were not reported in association with vaccines. This imbalance and "association" of exposure and outcome is the pre-condition for confounding/effect modification which we were investigating. Secondly, for the majority of negative controls, we indeed could not calculate a ROR, because there were no reports in children in EudraVigilance for that combination, thereby validating the reference set. For some negative controls however, we actually observed a significant association (lower bound of the ROR 95% confidence interval > 1 and number of reports ≥ 3), indicating a signal. This suggests that the classification of DECs based on the literature may have been wrong or may have changed due to newly occurring evidence. Thirdly, the masking/confounding did not have the same impact when we investigated the change in ROR. For the positive controls specifically, the ROR decreased when vaccines were excluded and for events that were not mentioned in vaccine-related reports. The ROR increased (masking effect) following removal of vaccine reports, when it concerned events that frequently occurred with vaccines. The change in ROR was highest in infants, which is expected since most vaccine-related reports concern infants. True positive associations involving seizure, bullous eruption and anaphylaxis were generally signals whether vaccine exposures were present in the data set or not, although with a lower ROR value in the full data set. While this does not indicate masking, it does indicate risk for lower ROR values when vaccine exposures are retained in the data set.

In the common 2x2 table upon which most SDAs are based, removal of vaccine-associated reports will not alter cell a (reports of the DEC of interest) or b (reports of the drug of interest in association with other events). However, cells c (reports of the event of interest with other drugs) and d (reports of other events with other drugs) will decrease. If this decrease is more extreme in cell d than in cell c, the effect will be a decrease in the SDA value. This is exactly what we observed: for most associations, removal of vaccine-associated DECs reduces the number of reports in cell d – vaccines in this case – with other events without simultaneously reducing the number of reports of the event of interest in association with other drugs. This suggests that, in general, the adverse events reported in association with vaccines are different than those reported in association with the non-vaccines included in our reference set or with the most commonly reported non-vaccines in pediatrics. This is supported by the effect we see in the ROR values for the reference set. The only associations for which the SDA values increase with removal of vaccines are those including events frequently reported following vaccination: bullous eruption, seizure, and anaphylaxis (see Supplementary table 1). Blake et al. have reported the preponderance of vaccine-associated reports in pediatrics while Juhlin et al. have reported the considerable impact of some vaccine-adverse event pairs in signal detection, specifically common adverse events like fever (2,22). This effect is most pronounced in infants and adolescents who are the target groups for most routine childhood vaccinations.

Regarding the analysis without a reference set, we found that removal of vaccine exposures typically decreases the value of SDAs, especially in infants. In the decision-tree analysis, we found that the impact of vaccine removal on ROR values was dependent upon age group and the proportion of reports for an event which included a vaccine exposure. Since drug-event associations for events that occur frequently after vaccination may be underestimated for non-vaccine drugs we recommend that vaccines be removed from the data set when the event of interest is a known adverse event following vaccination.

Other studies that specifically investigated masking by vaccine-related reports had similar results. De Bie et al., after analyzing a large international spontaneous reports database (Vigibase), found that vaccines have a large and mathematically predictable impact on signal detection in the pediatric population, also after stratification by age (12,23). Specifically, they found that in analysis of non-vaccine DECs, when the non-vaccine proportion ratio (defined as the ratio between the proportion of non-vaccine-related pairs in cell c and the proportion of non-vaccine-related pairs in cell d in the full database) was < 1, additional ADRs were detected after restriction to non-vaccine DECs. Similar to the present study, they found that the masking effect was most pronounced in infants. Another study performed on a smaller company-owned vaccine-specific database (9) found a rather modest masking effect. While the findings of this study may be generalizable to any database containing reports of vaccines and non-vaccines in children, because other databases vary in size, completeness and type of reports contained, generalizability cannot be assumed. Our study is limited in its focus on only one frequentist SDA. Performance of a Bayesian SDA may differ from that of the ROR but due to the preponderance of studies demonstrating little difference in performance among SDAs (4,22,23) and in the interest of ease of interpretation, we chose to limit our focus to the ROR. Method-wise there were other methods proposed in the literature to evaluate masking, for example the masking ratio developed and validated by Maignen (8, 24). This method, when conducted at the DEC level, is conceptually similar to the percent change in ROR which we have reported..

In conclusion, removal of vaccines prior to performing signal detection in a pediatric population has an impact on ROR, dependent upon the frequency with which the event of interest is reported in combination with vaccines. We recommend removal of vaccine exposures only if the investigated ADRs include those typically reported in association with vaccines, such as anaphylaxis, fever, and convulsions. Because the impact of masking differed by age group due both to frequency of vaccination and diversity of adverse events experienced in each age group, we additionally recommend stratification by age. Since we could not assess if the gain in specificity offsets the loss in sensitivity, we recommend further evaluation using a larger pediatric-tailored reference set.

References

1. Sturkenboom MCJM, Verhamme KMC, Nicolosi A, Murray ML, Neubert A, Caudri D, et al. Drug use in children: cohort study in three European countries. BMJ. 2008 Nov 24;337:a2245.

2. GRiP. Global Research in Paediatrics – Network of Excellence (GRiP). Available from: http://grip-network.org/,

3. Blake KV, Saint-Raymond A, Zaccaria C, Domergue F, Pelle B, Slattery J. Enhanced Paediatric Pharmacovigilance at the European Medicines Agency: A Novel Query Applied to Adverse Drug Reaction Reports. Paediatr Drugs. 2016 Feb;18(1):55–63.

4. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf. 2009;18(6):427–436.

5. CIOMS Working Group. Practical aspects of signal detection in pharmacovigilance . Geneva; 2010.

6. Almenoff J, Tonning JM, Gould AL, Szarfman A, Hauben M, Ouellet-Hellstrom R, et al. Perspectives on the use of data mining in pharmaco-vigilance. Drug Saf Int J Med Toxicol Drug Exp. 2005;28(11):981–1007.

7. Pariente A, Didailler M, Avillach P, Miremont-Salamé G, Fourrier-Reglat A, Haramburu F, et al. A potential competition bias in the detection of safety signals from spontaneous reporting databases. Pharmacoepidemiol Drug Saf. 2010 Nov;19(11):1166–71.

8. Maignen F, Hauben M, Hung E, Holle LV, Dogne JM. A conceptual approach to the masking effect of measures of disproportionality.Pharmacoepidemiol Drug Saf. 2014 Feb;23(2):208-17. doi: 10.1002/pds.3530. Epub 2013 Nov 15

9. Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. Hum Vaccin. 2009 Sep;5(9):599-607. Epub 2009 Sep 7.

10. Wang H, Hochberg AM, Pearson RK, Hauben M. An experimental investigation of masking in the US FDA adverse event reporting system database. Drug Saf Int J Med Toxicol Drug Exp. 2010 Dec 1;33(12):1117–33.

11. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiol Drug Saf. 2003 Nov;12(7):559–74.

12. Aagaard L, Weber CB, Hansen EH. Adverse drug reactions in the paediatric population in Denmark: a retrospective analysis of reports made to the Danish Medicines Agency from 1998 to 2007. Drug Saf. 2010 Apr 1;33(4):327–39.

13. de Bie S, Verhamme KMC, Straus SMJM, Stricker BHC, Sturkenboom MCJM. Vaccine-Based Subgroup Analysis in VigiBase. Drug Saf. 2012 Apr;35(4):335–46.

14. CIOMS. Definitions and Applications of Terms for Vaccine Pharmacovigilance. 2012.

15. Annual Report on EudraVigilance 2016 [cited 2016 Oct 9]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Report/2016/03/WC500203705.pdf

16. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. Clinical investigation of medicinal products in the pediatric population E11, CPMP/ICH/2711/99, July 2012;

17. MedDRA Data retrieval and presentations-points to consider [Internet]. Available from: https://www.meddra.org/sites/default/files/guidance/file/9610-1610_datretptc_r3.6_sep2013.pdf

18. Bodenreider O. Using UMLS semantics for classification purposes. Proc AMIA Symp. 2000;86–90.

19. Hauben M, Aronson JK. Defining “signal” and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Saf Int J Med Toxicol Drug Exp. 2009;32(2):99–110.

20. Osokogu O, Fregonese F, Ferrajolo C, Verhamme K, de Bie S, Jong G ’t, et al. Pediatric Drug Safety Signal Detection: A New Drug–Event Reference Set for Performance Testing of Data-Mining Methods and Systems. Drug Saf. 2015 Feb 1;38(2):207–17.

21. Juhlin K, Ye X, Star K, Norén GN. Outlier removal to uncover patterns in adverse drug reaction surveillance - a simple unmasking strategy. Pharmacoepidemiol Drug Saf. 2013 Oct;22(10):1119–29.

22. Schuemie MJ, Coloma PM, Straatman H, Herings RMC, Trifirò G, Matthews JN, et al. Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods. Med Care. 2012 Oct;50(10):890–7.

23. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. Clin Pharmacol Ther. 2013 Jun;93(6):539–46.

24. Maignen F, Hauben M, Dogné J-M. A mathematical framework to quantify the masking effect associated with the confidence intervals of measures of disproportionality. Therapeutic Advances in Drug Safety. 2017;8(7):231-244. doi:10.1177/2042098617704143.

Supplementary Material

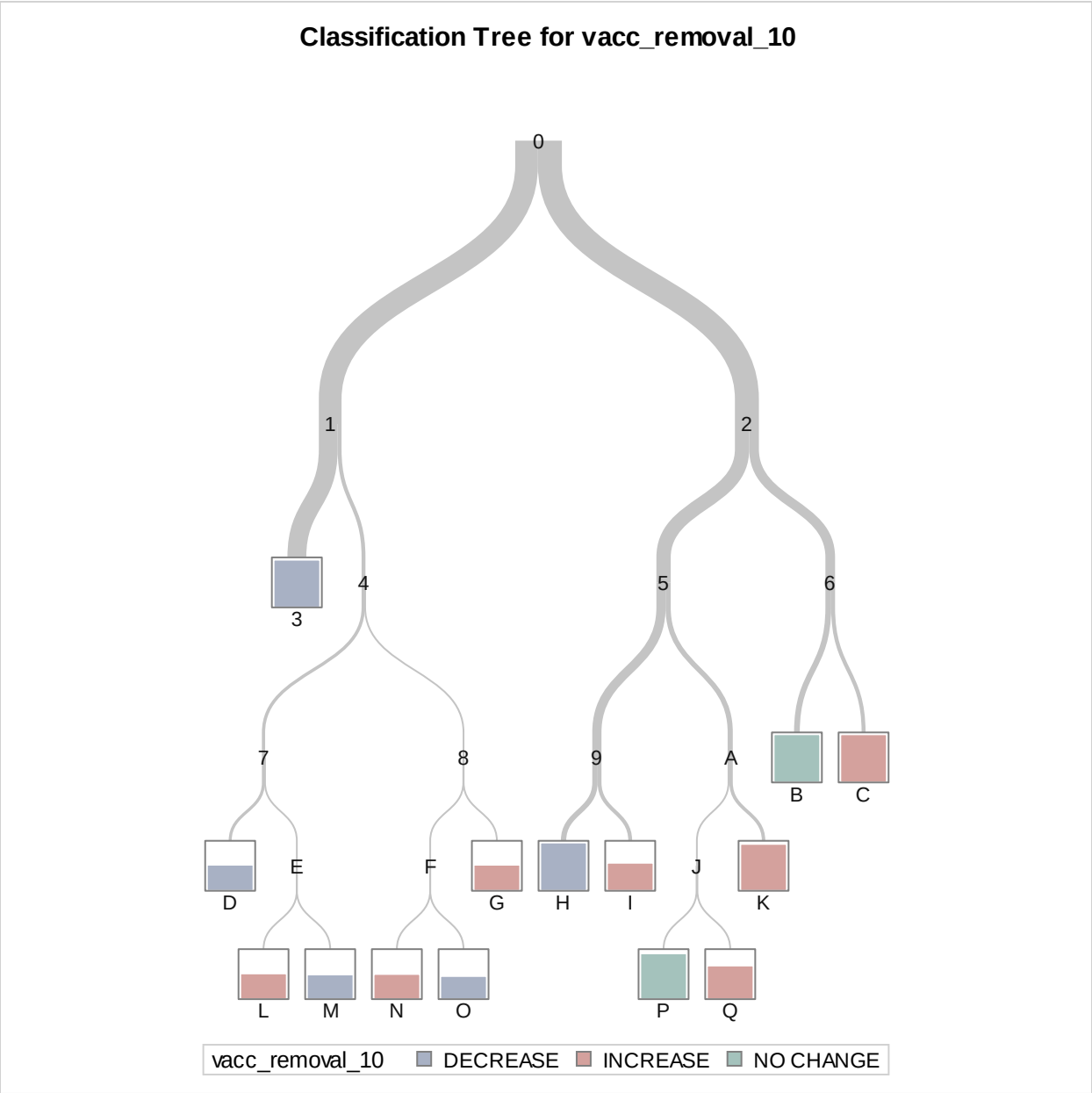
Supplementary Table 1. Reporting odds ratio values in EUDRAVIGILACE in population 0-18 years of age for GRiP reference set associations (true positive = green cells, true negative = red cells)

		BULLOUS ERUPTION	APLASTIC ANEMIA	AGRANULOCYTOSIS	THROMBOCYTOPENIA	PSYCHOSIS	SUICIDE	VENTRICULAR ARRHYTHMIA	SUDDEN DEATH	QT PROLONGATION	THROMBOEMBOLISM	ANAPHYLAXIS	SEIZURE	ACUTE KIDNEY INJURY	LIVER INJURY	SEPSIS	SIDS
FLUCLOXACILLIN	ROR	5.77		4.15	2.45							2.09	0.30	6.39	2.75		
	ROR NO VACCINES	6.27		2.95	1.99							2.27	0.45	4.42	1.94		
CLARITHROMYCIN	ROR	6.26	0.60	1.21	0.92	1.21	0.31	0.80		2.91	0.19	2.18	0.45	1.23	2.35	0.21	
	ROR NO VACCINES	6.87	0.42	0.86	0.75	0.86	0.21	0.58		2.01	0.13	2.38	0.56	0.85	1.66	0.17	
DOXYCYCLINE	ROR	2.49	5.09	2.27	1.67	1.63	5.07					0.75	0.61		1.12		
	ROR NO VACCINES	2.70	3.58	1.61	1.36	1.16	3.49					0.82	0.75		0.79		
LOPINAVIR	ROR				3.87				22.04			0.56				2.83	
	ROR NO VACCINES				3.14				21.85			0.61				2.34	
ISONIAZID	ROR	0.80	0.92	1.48	3.13	2.68	1.96					0.69	1.09	1.51	20.10		
	ROR NO VACCINES	0.87	0.65	1.05	2.54	1.91	1.35					0.75	1.34	1.04	14.21		
PRAZIQUANTEL	ROR		9.61		4.48							2.69					
	ROR NO VACCINES		6.76		3.63							2.92					
MEBENDAZOLE	ROR				0.70		0.79		8.21			1.04	0.21		0.79		
	ROR NO VACCINES				0.57		0.54		8.14			1.13	0.26		0.56		
QUININE	ROR				2.65			99.87		39.26		0.38	2.60	9.30			
	ROR NO VACCINES				2.15			72.41		27.23		0.41	3.20	6.43			
FLUTICASONE	ROR	1.97										0.39	1.33				
	ROR NO VACCINES	2.13										0.42	1.64				
MONTELUKAST	ROR	0.92	0.26		0.37	5.07	4.18	0.28	0.28		0.07	0.43	0.63		1.03	0.07	
	ROR NO VACCINES	1.00	0.18		0.30	3.61	2.86	0.20	0.28		0.05	0.47	0.77		0.73	0.06	
ISOTRETINOIN	ROR	0.40	0.34	0.84	0.88	4.69	6.06	0.16			1.43	0.40	0.25		3.13	0.28	
	ROR NO VACCINES	0.43	0.23	0.59	0.71	3.33	4.13	0.12			1.00	0.44	0.30		2.19	0.23	
LOPERAMIDE	ROR						10.41					0.92	0.54	0.25		1.31	
	ROR NO VACCINES						7.16					0.99	0.66	0.17		1.08	
DOMPERIDONE	ROR			0.95	0.98	0.68	0.31		8.26	2.55		0.53	1.82	1.46	0.94	0.63	
	ROR NO VACCINES			0.68	0.80	0.49	0.21		8.23	1.76		0.57	2.24	1.01	0.66	0.52	
METHYLPHENIDATE	ROR	0.14	0.63	1.01	1.23	8.39	2.57	2.95	3.38	2.62	0.21	0.25	0.74	0.27	1.12	0.05	
	ROR NO VACCINES	0.15	0.44	0.71	0.99	5.97	1.76	2.12	3.41	1.80	0.14	0.27	0.92	0.18	0.79	0.04	
IBUPROFEN	ROR	3.52			2.87		7.91					0.81	0.28	6.61			
	ROR NO VACCINES	3.82			2.32		5.44					0.88	0.35	4.57			
CYPROTERON	ROR	0.57					1.03					0.24	0.37		0.69	0.46	
	ROR NO VACCINES	0.62					0.71					0.26	0.46		0.48	0.38	

Supplementary Table 2. Overall performance of ROR signal detection algorithm using the GRiP drug reference set on data in EUDRAVIGILACE in population 0-18 years of age

Age group	Set	AUC	SENSITIVITY	SPECIFICITY
0-18 years	Vaccine & non-vaccine	0.887	0.671	0.939
	Non-vaccine only	0.881	0.607	0.939
0-<2 years	Vaccine & non-vaccine	0.861	0.583	0.889
	Non-vaccine only	0.898	0.500	1.000
2-<12 years	Vaccine & non-vaccine	0.856	0.421	0.913
	Non-vaccine only	0.860	0.421	0.957
12-<18 years	Vaccine & non-vaccine	0.815	0.609	0.917
	Non-vaccine only	0.817	0.565	0.917

SDA=signal detection algorithm; Non-vaccine=data set with all vaccine-adverse event pairs removed; Sensitivity and Specificity calculated using reporting odds ratio (ROR) threshold of lower confidence interval of ROR > 1 and number of reports ≥ 3



Supplementary figure 1. Classification tree for vaccine removal. RORs for reference set associations following vaccine removal classified as DECREASE (> -10% change in ROR), INCREASE (> +10% change in ROR), NO CHANGE (-10% > change in ROR < +10%)

Node Information								
		Training Data						
ID	Path	Count	DECREASE		INCREASE		NO CHANGE	
3 (DECREASE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc < 0.109476	93741	0.9082		0.0670		0.0249	
	AGE = CHILD,INFANTS	77868	0.9887	*	0.0000		0.0113	
B (NO CHANGE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = CHILD	39884	0.0006		0.4062		0.5932	
	prop_with_vacc < 0.268714	23658	0.0010		0.0003		0.9987	*
C (INCREASE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = CHILD	39884	0.0006		0.4062		0.5932	
	prop_with_vacc >= 0.268714	16226	0.0000		0.9979	*	0.0021	
H (DECREASE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = ADOL,INFANTS	59134	0.4760		0.4226		0.1013	
	prop_with_vacc < 0.477714	37907	0.7183		0.2428		0.0389	
	AGE = INFANTS	22063	0.9998	*	0.0000		0.0002	
K (INCREASE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = ADOL,INFANTS	59134	0.4760		0.4226		0.1013	
	prop_with_vacc >= 0.477714	21227	0.0433		0.7439		0.2128	
	prop_with_vacc >= 0.577238	14086	0.0210		0.9719	*	0.0071	
P (NO CHANGE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = ADOL,INFANTS	59134	0.4760		0.4226		0.1013	
	prop_with_vacc >= 0.477714	21227	0.0433		0.7439		0.2128	
	prop_with_vacc < 0.577238	7141	0.0872		0.2941		0.6187	
	AGE = INFANTS	4438	0.0023		0.0489		0.9489	*
Q (INCREASE)	Root Node	193E3	0.5878		0.2463		0.1659	
	prop_with_vacc >= 0.109476	99018	0.2845		0.4160		0.2995	
	AGE = ADOL,INFANTS	59134	0.4760		0.4226		0.1013	
	prop_with_vacc >= 0.477714	21227	0.0433		0.7439		0.2128	
	prop_with_vacc < 0.577238	7141	0.0872		0.2941		0.6187	
	AGE = ADOL	2703	0.2268		0.6966	*	0.0766	
* Selected target level								

2.3 PEDIATRIC VACCINE SAFETY SIGNAL DETECTION IN VAERS AND EUDRAVIGILANCE USING DISPROPORTIONALITY ANALYSIS, TIME TO ONSET, AND THEIR COMBINATION

Caitlin Dodd

Lionel van Holle

Maria de Ridder

Daniel Weibel

Miriam C J M Sturkenboom

Abstract

Background

Adverse events following immunization (AEFIs) may be monitored in spontaneous reporting systems using signal detection algorithms which compare the frequency of reports of a specific vaccine-AEFI combination to the frequency of all other reports in the database. Alternatively, a recently developed method uses distribution tests of reported time-to-onset (TTO) to detect vaccine safety signals with unexpected reported temporal relationship.

Objectives

To compare performance of the Empirical Bayes Geometric Mean (EBGM), the TTO method, and their combination in pediatric vaccine reports in the databases (DBs) of the US Vaccine Adverse Event Reporting System (VAERS) and EU EudraVigilance (EV) as part of the Global Research in Pediatrics project.

Methods

Following EV and VAERS conversion to a common data model, EBGM and TTO methods were applied to all pediatric vaccine reports in VAERS, EV, and their union. Performance of each method was assessed using a previously published vaccine reference set comprising 18 true positive and 113 true negative associations to calculate area under the receiver operating characteristic curve (AUC). The method of Pepe & Thompson (2000) was used to determine the linear combination of EBGM and TTO which maximized the AUC in the EV database, VAERS database, and their union (EV + VAERS).

Results

VAERS contained 1.56 million pediatric vaccine-AEFI combinations, 7% of which were missing TTO data. EV contained 228,181 pediatric vaccine-AEFI combinations, 54% of which were missing TTO data. In VAERS, the TTO method AUCs was 0.86 while EBGM AUC was 0.77. Performance in EV was similar across methods with TTO and EBGM AUCs of 0.83 and 0.86, respectively. In the union of VAERS and EV the TTO AUC was 0.86 while the EBGM AUC was 0.80. The linear combination of TTO and EBGM increased AUC to 0.90 in VAERS, 0.94 in EV, and 0.92 in the union (EV + VAERS) database.

Conclusion

The EBGM method performed better than the TTO method in EV while TTO performed better than EBGM in VAERS and the union of the two databases. Unifying pediatric reports from the EV and VAERS databases did not improve performance over the best performing method in each DB alone. The combination of EBGM and TTO methods improved performance over either method alone in both databases and their union. When time to onset data is available, the TTO method is recommended. Linearly combining methods to make use of the most available data may improve performance over each method alone while unifying spontaneous reporting databases to conduct signal detection does not offer obvious benefit.

Introduction

Vaccines have been demonstrated to be very effective and have prevented more deaths than any other medical intervention (1). Administration of vaccines differs from that of other medicines in that vaccines are mostly administered to healthy rather than diseased persons and, for most vaccines, to large segments of the population, usually in childhood. While vaccines are tested for quality, efficacy and safety in clinical trials, the sample size and duration of follow-up is generally limited. Additionally, assessment of concomitant administration of other vaccines is not studied in clinical trials. Consequentially, additional safety concerns must be monitored in a post-marketing setting.

The traditional way to identify potential safety issues is through the analysis of spontaneously reported case reports that are submitted to systems to which physicians, consumers, and others can report suspected adverse events following immunization (AEFIs). Databases of this type contain only reports of suspected adverse reactions and lack data on the number of exposed persons. The absence of the number of exposed persons has led to the development of signal detection algorithms (SDAs) based on the reported cases only. These algorithms assess disproportionality of reporting by comparing the frequency with which a specific vaccine-AEFI combination is reported to all other vaccine-AEFI combinations in the database.

While SDAs developed for detection of safety signals associated with drugs have been extensively studied, little attention has been paid in the literature to the appropriateness of applying these methods to vaccine exposures, or to the development of vaccine-specific SDAs (2-4). One notable exception to this is the time-to-onset (TTO) method developed by Van Holle et al., which uses reported time from vaccination to onset of symptoms to detect safety signals (5).

The performance of an SDA is often tested by determining the area under the curve (AUC) of a receiver operator curve (ROC), which is a common practice in testing the performance of diagnostic tests. In the field of diagnostic testing much research has been done regarding the combination of test results to increase predictive accuracy (6-8). In contrast, while SDAs have been extensively tested and compared, little work to date has been done on the combination of SDA estimates to increase performance. With the exception of the 2014 study by Van Holle et al. which employs logistic regression to combine SDA estimates, the work which has been done is limited to combining signals from different data sources, such as spontaneous reporting systems and observational healthcare data (9-11). We were interested in the potential to increase performance of signal detection by combining different SDA methodologies within the same data source as well as in unifying multiple data sources.

As part of the Global Research in Pediatrics (GriP) project, which was funded by the European Commission Seventh Framework Programme, we used the public versions of the United States Food and Drug Administration (FDA) Vaccine Adverse Event Reporting System (VAERS) and a subset of the European Medicines Association (EMA) EudraVigilance (EV) system which was obtained through an academic agreement. We developed a common data model (CDM), which means that the structure and variable names of the databases are harmonized . This allows similar analysis and pooling. Both the EV and VAERS database were converted to this CDM (Supplementary material). We used the data to address three research questions: first, to compare the performance of the Empirical Bayes Geometric Mean (EBGM), a commonly used SDA, to the TTO method. Second, to investigate whether performance of signal detection changes when the VAERS and EV databases were united as compared to keeping them single. Third, to derive a linear combination of EBGM and TTO thresholds to investigate whether this would alter performance.

Methods

Data sources

VAERS database

The United States FDA together with the US Centers for Disease Control maintain the Vaccine Adverse Event Reporting System (VAERS) database (12). VAERS was set up in 1990 in response to The National Childhood Vaccine Injury Act of 1986 which required physicians to report suspected AEFIs (13). The data contained in VAERS is publically available for download and is updated monthly. Reports for serious events occurring outside the United States must also be reported to VAERS by vaccine manufacturers (14). From its inception until 2007, AEFIs in VAERS were coded using the FDA's Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) system but have since been coded using the Medical Dictionary for Regulatory Activities (MedDRA) coding system. All events reported prior to 2007 have been converted from COSTART to the MedDRA coding system. Vaccine exposures are coded using a VAERS-specific coding system with separate values for vaccine type which groups vaccines containing the same antigens together regardless of manufacturer, additives, or year with additional variables such as manufacturer, dose and route. As of December 2017, VAERS contained 541,018 reports, 43% of which pertained to events occurring in children, defined as those with non-missing age less than 18 years.

EudraVigilance database

The European Medicines Agency (EMA) maintains the EudraVigilance (EV) database since December 2001 with data dating back to 1995. Like VAERS, EV was set up in response to legislation requiring reporting of all serious adverse drug and vaccine reactions in Europe. Marketing authorization holders and sponsors of clinical trials must submit suspected adverse events occurring after exposure to products licensed in the EU - regardless of the location of exposure or event - to the EMA. The database has since been expanded in 2015 to include non-serious reactions; in 2017 reporting of non-serious reactions was made mandatory and the database was also expanded to include literature reports. EV data can be queried online with standard methods, for research a dataset can be requested from EMA (15). For the current study, we obtained individual case safety reports for spontaneous reports following drug and vaccine exposures only in which the subject was identified as pediatric based upon reported or calculated age less than 18 or reported age group of neonate, infant, child, or adolescent. For the current study, only events reported in association with vaccine exposures are considered. As of December 2017, EudraVigilance contained 12.45 million safety reports, 34% of which originated in Europe and 66% of which originated elsewhere (16). The EV data obtained for the Global Research in Pediatrics (GriP) project, which extends through the end of 2016, contains 0.5 million pediatric safety reports.

Common data model

Within the GRiP project, we developed a common data model for spontaneous reporting systems in order to combine data and compare systems. We transformed the publicly available VAERS database as well as the GriP pediatric subset of the EV database into this common data model (see appendix 1)

Vaccines in VAERS are coded using a standardized dictionary developed for the VAERS database while all drugs and vaccines within EV are reported using the non-proprietary substance name (17, 18).

Using an algorithm described previously (19), reports of vaccine exposures in EV were mapped to WHO-ATC codes using the Solr search engine (http://lucene.apache.org/solr/). The approach was part of the vaccine ontology development in the ADVANCE (Accelerated Development of Vaccine benefit-risk Collaboration in Europe) project (ADVANCE Deliverable 5.5: https://goo.gl/hkvAV4) (20). In EV, combination vaccines may be reported with all

antigens in one row, or with each antigen in a separate row. If reported separately, a combination vaccine against measles, mumps and rubella, for example, would be assigned to J07BD (measles vaccine), J07BE (mumps vaccine), and J07BJ (rubella vaccine), instead of to J07BD72 (measles-mumps-rubella vaccine). For this reason, we developed an algorithm to group separately recorded antigens within the same case report (defined by report id and date) according to existing combination vaccines, leaving jointly recorded antigens unchanged, resulting in the lowest possible number of unique exposures per report.

Following conversion to the CDM, the VAERS and EV datasets were separately de-duplicated using a combination of report id, substance name, vaccine type, age, sex, outcome, event, country and year of the event. When duplicates were detected, the record with the lower proportion of missing data was retained. Using the common data model (see supplementary material), we were able to combine the de-duplicated pediatric subset of the VAERS database with the pediatric vaccine-related reports from EudraVigilance. Because both databases contain reports for events occurring outside of their jurisdiction, the resulting data set was de-duplicated again using the same combination of variables with the exception of report id. In a sensitivity analysis, the impact of deduplication was assessed by performing a subset of analyses on the union of EV and VAERS without deduplication.

Reference Set

In order to test the performance of SDAs in spontaneous databases, we used a reference set that was developed in the GRIP project. This reference set contains 182 vaccine-AEFI pairs, 18 of which are classified as positive controls (true relation between vaccine and AEFI), 113 as negative controls (no known relation between the vaccine and AEFI), and 51 as unclassifiable (21). Mappings from AEFIs to MedDRA preferred terms can be found in supplementary material (Supplementary material). The number of reports of each reference set vaccine-AEFI pair was described by database and age group as defined by the International Conference on Harmonization (ICH), with the exception that neonates and infants were combined due to the paucity of vaccine exposures in neonates. Resulting age groups were infants (0 to < 2 years), children (2 to < 12 years), and adolescents (12 to < 18 years) (22).

Signal Detection Algorithms

The EBGM is a well-established SDA originally developed by DuMouchel for use in the US FDA Adverse Event Reporting System (FAERS) (23). Briefly, the method uses a Bayesian framework in which ratios of observed to expected counts are assumed to be drawn from a prior distribution which is the mixture of two gamma distributions. This method serves to shrink estimates toward the null, especially when observed or expected counts are small (24). We calculated EBGM values for each vaccine-AEFI combination and categorized each reference set association as a signal based upon a predefined threshold. The threshold used by the Medicines and Healthcare products Regulatory Agency (MHRA) for declaring a vaccine-AEFI pair a signal is: EBGM lower 90% credibility interval bound (EB05) ≥ 1.8, number of reports ≥ 3, and EBGM ≥ 2.5 (25). Under the null assumption of no association between a and AEFI and no confounding or bias we would expect an EBGM of 1. In a sensitivity analysis, we calculated age-stratified EGBM values using the Mantel-Haenszel approach for weighting of each ICH age strata and subsequent pooling (26).

Van Holle et al have developed an alternative vaccine safety signal detection method. This method makes use of the reported time from exposure to onset of symptoms (TTO) (5). Briefly, for each vaccine/AEFI combination of interest, three distributions are constructed using the recorded times in the reports in the database: 1. Time to onset from the vaccine of interest to the AEFI of interest (Vaccine/AEFI), 2. Time to onset from the vaccine of interest to any AEFI other than that of interest (Vaccine/Other events), and 3. Time to onset from any other vaccine to the AEFI of interest (Other vaccines/AEFI). Subsequently, distributions 2 (Vaccine/Other events) and 3 (Other vaccines/AEFI) are compared to distribution 1 (Vaccine/AEFI) using the Kolmogorov-Smirnov (KS) two sample test.

The KS statistic tests whether two distributions come from the same underlying distribution (27). In the TTO method, if p-values from both KS-tests (Vaccine/AEFI vs. Vaccine/Other events and Vaccine/AEFI vs. Other vaccines /AEFI) are below a predefined threshold, the association of interest is determined to be a signal. We calculated both KS-test p-values for each vaccine-AEFI combination and categorized each reference set association as a signal based upon the threshold of a p-value < 0.05 for both KS tests. Under the null assumption of no association between that vaccine and AEFI, the reported time-to-onset distribution is assumed to be similar to what is observed for that event following other vaccines. That assumption is tested by one of the aforementioned KS test (Vaccine/AEFI vs. Other vaccines/AEFI). The other KS test (Vaccine/AEFI vs. Vaccine/Other events) serves to increase specificity and therefore has less predictive value (9)

Performance of the EBGM and TTO methods was assessed in EV and VAERS separately and in the union of the VAERS and EV databases using the reference set as 'gold standard'. First, detection of true positives at the MHRA recommended threshold for EBGM (>2.5, EBGM05 ≥ 1.8, number of reports ≥ 3) and the p-value < 0.05 threshold of the TTO method was assessed. Subsequently, performance of both the EBGM and TTO test was measured via area under the receiver operating characteristic (ROC) curve. True positives were the 18 associations in the reference set listed as positive controls; true negatives were the 113 listed as negative controls, while the remaining 51 associations with conflicting evidence or absence of evidence were excluded from performance assessments. For vaccine-AEFI combinations in the reference set without any observations in the database, a null value of 0 for EBGM and 1 for TTO p-values was imputed prior to AUC calculation. For the EBGM based ROC analysis, the threshold of the EBGM test statistic was varied from 0 to infinity. For the ROC the number of reports and lower 90% credibility interval bound of the EBGM test statistic (which are used by MHRA) were not considered. For the ROC analysis of the TTO method, p-value thresholds for both KS tests were varied simultaneously and identically from 0 to 1.

To find the optimal combination of the EBGM test statistic and the two TTO KS p-values, we applied the methodology of Pepe and Thompson (19). This method uses a distribution-free rank-based approach to find the linear combination of test results (in our study: the log-transformed EBGM results and log-transformed p-values of the two KS tests in TTO) which maximizes the area under the ROC curve (28). The three test results were submitted to the maximization algorithm sequentially, and all six possible sequences of three tests were analyzed. The sequence yielding the maximum AUC was retained.

Results

The public VAERS database that was downloaded on 31 March 2017 contained 258,241 reports comprising 2,184,765 vaccine-AEFI combinations reported in children 0 to < 18 years of age. De -duplication removed 158,048 (7.2%) vaccine-AEFI combinations from VAERS. The EV database contained 448,364 reports related to pediatrics (0- < 18 years) and 77,679 of those related to vaccines. These 77,679 reports comprised 325,148 vaccine-AEFI combinations. De-duplication within EV removed 33,239 duplicate (10.2%) vaccine-AEFI combinations. Following concatenation of the de-duplicated pediatric VAERS and EV vaccine data, we obtained 2,318,627 vaccine-AEFI combinations. 533,783 (23.0%) additional duplicate vaccine-AEFI combinations were removed, leading to a final analysis data set containing 1,784,844 vaccine-AEFI combinations (Table 1). For our sensitivity analysis of data without de-duplication, we concatenated the EV and VAERS data sets without any deduplication, leading to an analysis data set with 2,509,913 vaccine-AEFI combinations. Data on time-to-onset was missing in less than 10% of VAERS reports but in over half of EV reports, meaning that either exposure date, onset date, or both was missing. Missingness was lowest in infants and highest in adolescents (Table 1).

Table 1. Number of vaccine/AEFI combination and missingness of time-to-onset by database and age

Database	Age Strata	Number of Vaccine/AEFI combinations (%)	Number of Missing Time-to-onset values	% Missing Time-to-onset
VAERS	Infants	848,365 (54.5%)	58,173	6.9%
	Children	437,082 (28.1%)	33,059	7.6%
	Adolescents	271,216 (17.4%)	23,465	8.7%
	Total	1,556,663	114,697	7.4%
EudraVigilance	Infants	139,922 (61.3%)	69,811	49.9%
	Children	53,122 (23.3%)	28,914	54.4 %
	Adolescents	35,137 (15.4%)	25,095	71.4%
	Total	228,181	123,820	54.3%
VAERS + EV	Infants	988,287 (55.4%)	127,984	12.2%
	Children	490,204 (27.5%)	61,973	14.0%
	Adolescents	306,353 (17.2%)	48,560	16.2%
	Total	1,784,844	238,517	13.4%

Each of the events in the GriP positive and negative reference set were reported at least once in both VAERS and EV, with anaphylaxis, HHE and seizure being most commonly reported in both data sources while disseminated tuberculosis was reported least frequently (Table 2). Most events were most frequently reported for the age group 0-23 months (infants), except for arthritis which was most frequently reported in adolescents. Across both databases and all age groups, 82-86 percent of events reported were events not included in the reference set (Table 2). Naturally, reports were mostly related to routine childhood vaccines and seasonal vaccines while reports related to BCG vaccine (a travel vaccine or risk group vaccine) were less frequent (Table 3). The percentage of reports of vaccines not in the reference set was higher in EV (25%) than in VAERS (14%). Adolescents had a higher percentage of non-reference set vaccine reports, the majority of which (79%, data not shown) were related to HPV vaccines (Table 3).The number of reports for all reference set associations by database can be found in supplementary material (Supplementary Table 1).

Table 2. Number of reports for events in the GriP reference set of positive and negative controls in association with any vaccine in VAERS and EUDRAVIGILANCE by age group (infants: 0-2 years, children 2-11, adolescents: 12-< 18)

	VAERS				EudraVigilance				Overall
	Infants	Children	Adolescents	Total Pediatrics	Infants	Children	Adolescents	Total Pediatrics	Total
ANAPHYLAXIS	73,560	56,324	21,251	151,135	12,180	4,558	2,405	19,143	170,278
ARTHRITIS	3,074	2,665	3,231	8,970	474	431	454	1,359	10,329
BELLS PALSY	46	47	104	197	16	22	36	74	271
DIABETES MELLITUS	239	219	192	650	65	56	42	163	813
DISSEMINATED OKA VZV	1,074	653	134	1,861	14	100	5	119	1,980
DISSEMINATED TB	1	0	1	2	5	3	1	9	11
ENCEPHALITIS	1,027	468	414	1,909	228	160	113	501	2,410
GUILLAIN BARRÉ SYNDROME	512	411	570	1,493	46	65	74	185	1,678
HHE	23,142	6,105	4,243	33,490	6,759	882	666	8,307	41,797

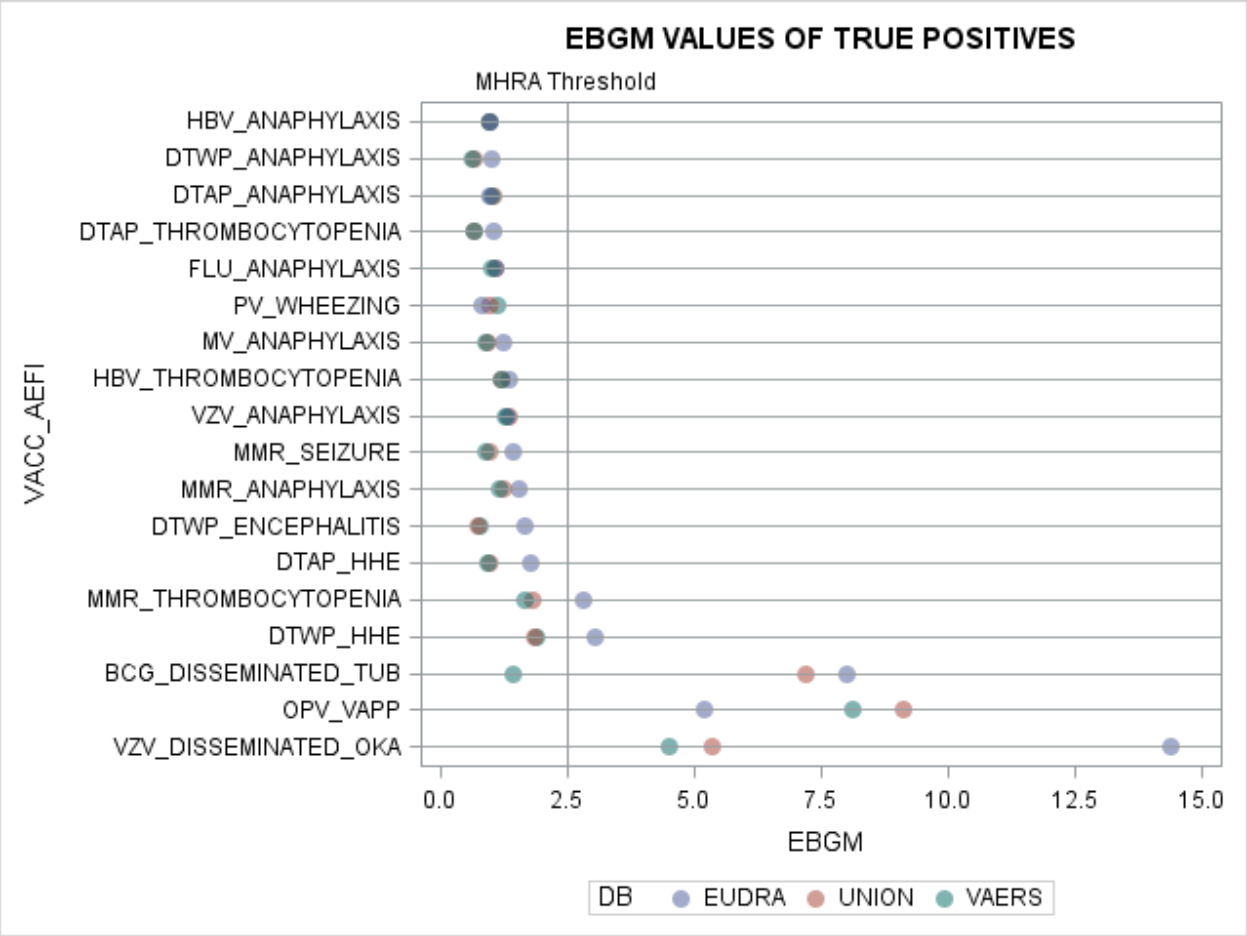
INTUSSUSCEPTION	3,448	23	3	3,474	1150	19	2	1,171	4,645
SEIZURE	29,398	7,898	6,709	44,005	3,832	1,173	894	5,899	49,904
THROMBOCYTOPENIA	2,824	779	485	4,088	451	160	66	677	4,765
VAPP	98	18	4	120	6	6	0	12	132
WHEEZING, R. AIRWAY	1,431	1,555	471	3,457	134	104	34	272	3,729
Events Not in GriP Reference Set	708,491 (83.5%)	359,917 (82.3%)	233,404 (86.1%)	1,301,812 (83.6%)	114,562 (81.9%)	45,383 (85.4%)	30,345 (86.4%)	190,290 (83.4%)	1,492,102 (83.6%)
TOTAL	848,365	437,082	271,216	1,556,663	139,922	53,122	35,137	228,181	1,784,844

Table 3. Number of reports for vaccines in the GriP reference set that we reported in association with any event for VAERS and EUDRAVIGILANCE by age group (infants: 0-2 years, children 2-11, adolescents: 12-< 18)

	VAERS				**EudraVigilance**				**Overall**
	Infants	**Children**	**Adolescents**	**Total Pediatrics**	**Infants**	**Children**	**Adolescents**	**Total Pediatrics**	**Total**
BCG	1,360	123	141	1,624	796	61	121	978	2,602
DTAP	123,217	83,248	26,239	232,704	12,302	3,049	971	16,322	249,026
DTWP	35,154	13,634	175	48,963	647	128	17	792	49,755
FLU	25,335	53,449	24,799	103,583	2,772	5,261	1,664	9,697	113,280
HAV	22,011	26,090	20,198	68,299	461	1,289	1210	2,960	71,259
HBV	66,701	20,582	22,275	109,558	14,543	3,277	2,104	19,924	129,482
HIB	159,902	14,988	636	175,526	7,981	3,812	7,676	19,469	194,995
MMR	66,623	57,340	12,564	136,527	11,180	6,226	940	18,346	154,873
MV	5,631	13,360	42,258	61,249	1,967	2,160	1,801	5,928	67,177
OPV	35,631	19,544	958	56,133	695	280	6	981	57,114
PV	134,175	20,776	3,678	158,629	42,907	5,713	272	48,892	207,521
RV	60,914	518	22	61,454	14,409	363	11	14,783	76,237
VZV	48,427	56,602	19,279	124,308	2,962	7,232	1,394	11,588	135,896
Vaccines Not in GriP Reference Set	63,284 (7.5%)	56,828 (13.0%)	97,994 (36.1%)	218,106 (14.0%)	26,300 (18.8%)	14,271 (26.9%)	16,950 (48.2%)	57,521 (25.2%)	275,627 (15.4%)
Total	848,365	437,082	271,216	1,556,663	139,922	53,122	35,137	228,181	1,784,844

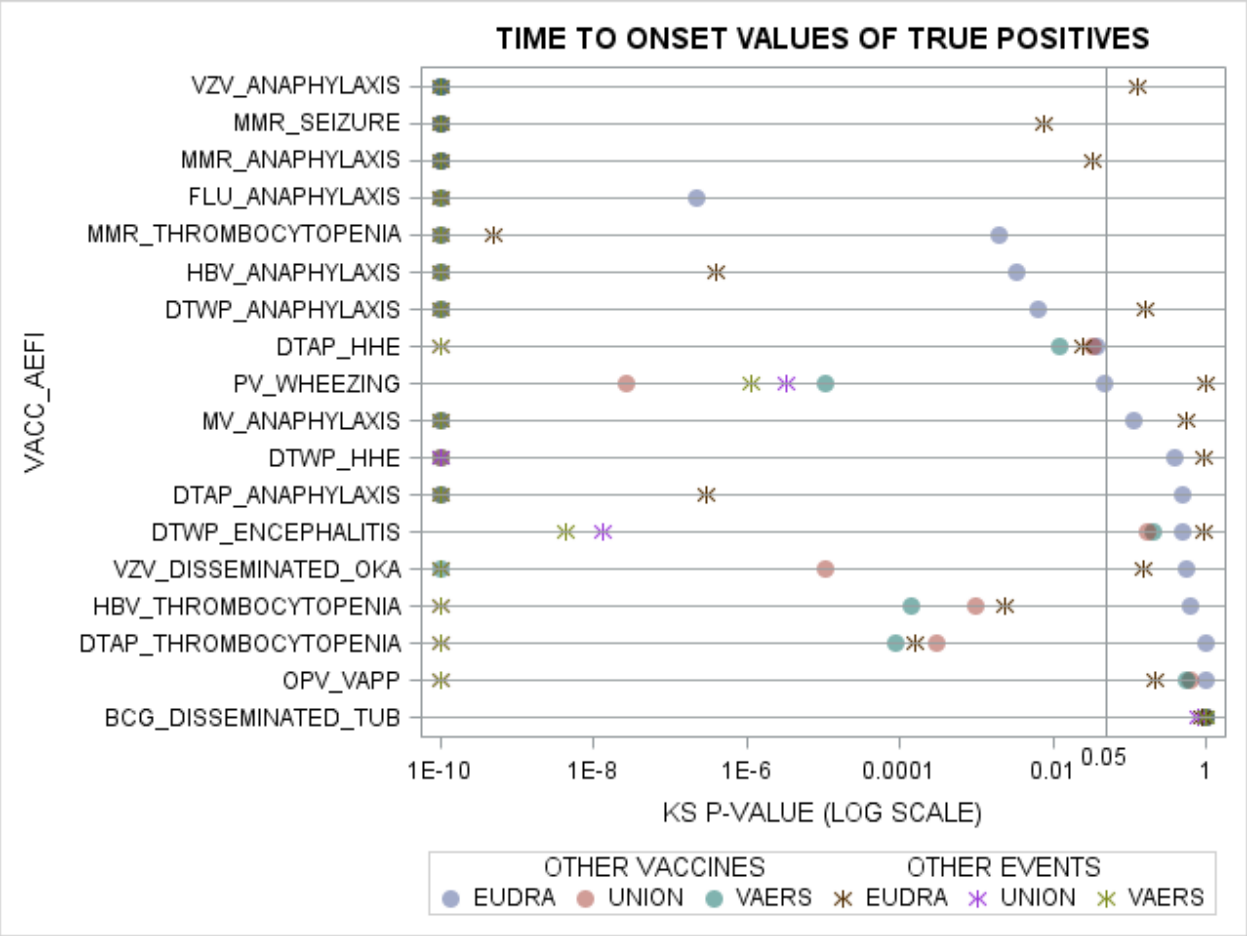
In VAERS, only 2 (11%) of the positive controls met MHRA criteria (EBGM >2.5, EB05 > 1.8, N ≥ 3) (Figure 1) while 111 (97%) of the negative controls did not meet MHRA criteria (Supplementary Figure 1). Using TTO and the standard p-value threshold of 0.05 for testing similarity in distributions, 15 (83%) of the positive controls were classified to be a signal (Figure 2) and 83 (72%) of the negative controls were had at least one p-value above the 0.05 threshold (Supplementary Figure 2).

Figure 1. EBGM values for true positive associations in VAERS



In EV, EBGM classified 5 (28%) of the positive associations as signals according to MHRA criteria (Figure 1) and 113 (98%) did not meet the criteria (Supplementary Figure 1). The TTO method flagged 6 (33%) of the true positive associations as signals (Figure 2) and 111 (97%) of true negatives had at least one p-value above the 0.05 threshold (Supplementary Figure 2).

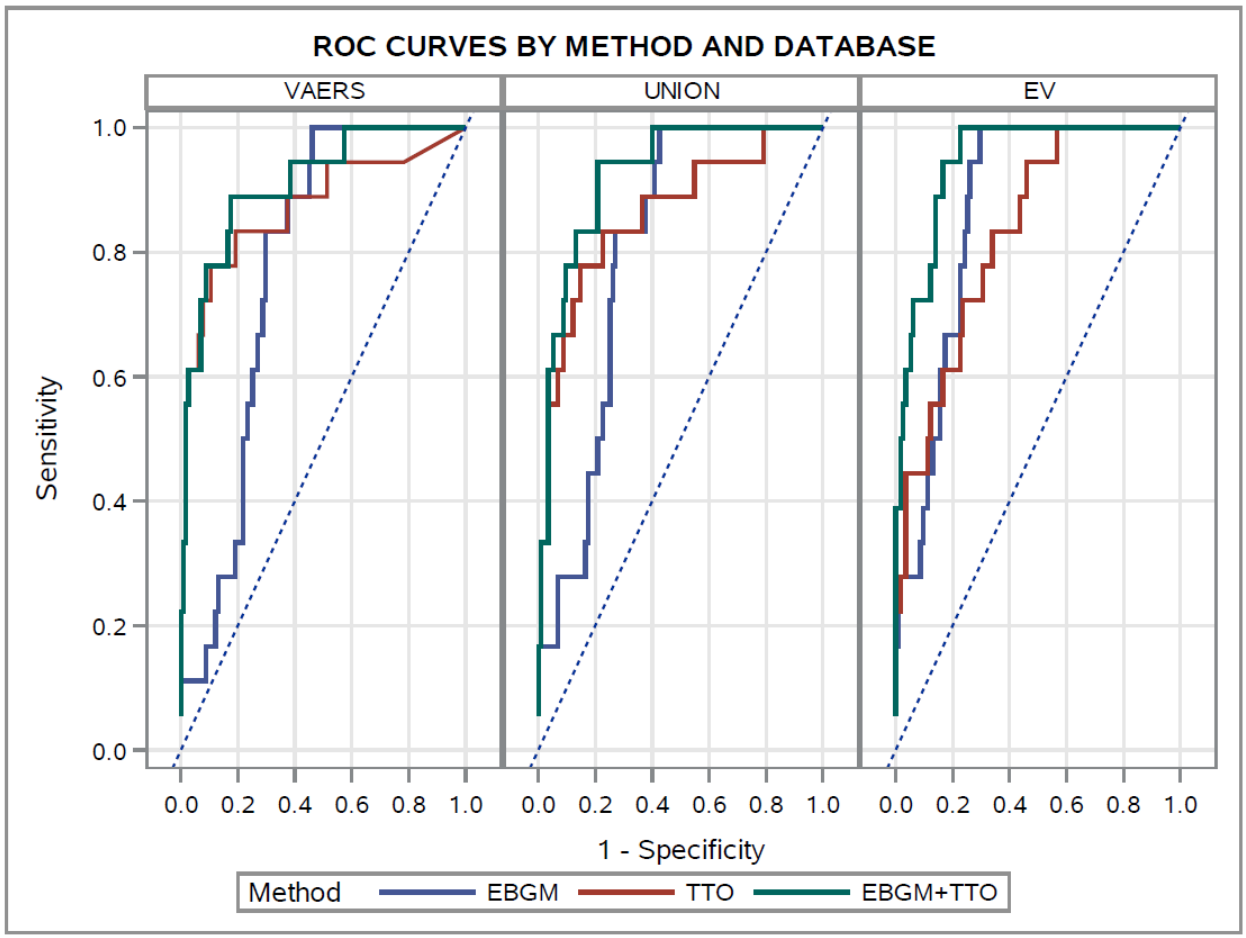
Figure 2. Time-to-onset Kolmogorov-Smirnov test p-values for true positive associations in VAERS



In the union of VAERS and EV, the EBGM method with MHRA criteria detected 3 (17%) of positive controls and 112 (97%) of the negative controls did not meet MHRA criteria for EBGM (Supplementary Figure 1). The TTO method correctly detected 15 (83%) of the true positives while 87 (76%) of true negatives had at least one p-value above the 0.05 threshold (Supplementary Figure 2).

In VAERS, the performance of the EBGM method was lower than that of the TTO method with AUCs (95% CI) of 77.2 (68.7, 85.7) and 86.4 (75.3, 97.5), respectively. Combining the EBGM and TTO methods led to an increased AUC of 90.9 (83.2, 98.6). In EV, EBGM performed slightly better than TTO with AUCs of 86.4 (79.9, 93.0) and 82.9 (73.6, 92.1), while the combination of EBGM and TTO markedly improved performance, producing an AUC of 94.4 (90.4, 98.5) (Figure 3, Table 4).

Figure 3. ROC Curves of the signal detection methods in pediatrics by database



TTO = Time To Onset, EBGM = Empirical Bayes Geometric Mean

In the union of the EV and VAERS databases, performance of the EBGM method was lower than that of TTO, with AUCs of 80.1 (71.9, 88.3) and 85.7 (75..3, 96.1). The combination of methods in the unified database improved performance over either method alone, producing an AUC of 92.3 (86.8, 97.9) (Figure 3, Table 4).

In sensitivity analyses retaining duplicate reports, the performance of EBGM improved slightly in both databases and their union. The performance of the TTO method improved in EV and the unified database with retention of duplicate reports but marginally decreased in VAERS (Table 4).

In EBGM analysis by age strata, performance in each DB and their union was similar for adolescents (AUC EV: 0.83, VAERS: 0.80, Union: 0.83), children (AUC EV: 0.83, VAERS: 0.80, Union: 0.87), and infants (AUC EV: 0.87, VAERS: 0.74, Union: 0.77). Area under the curve for age-adjusted EBGM in EV, VAERS, and their Union were 0.83, 0.76, and 0.79, respectively (Table 4). In TTO analysis stratified by age, performance in each DB and their union was similar for adolescents (AUC EV: 0.78, VAERS: 0.81, Union: 0.82), children (AUC EV: 0.86, VAERS: 0.88, Union: 0.91), and infants (AUC EV: 0.78, VAERS: 0.83, Union: 0.85) (Table 4).

Table 4. Performance of different signal detection algorithms expressed by area under the curve (AUC) of the receiver operating curve (ROC) in the different datasets

	EBGM			TTO		EBGM + TTO	
Database	Primary analysis (with deduplicaiton)	Sensitivity analysis 1 (without deduplication)	Sensitivity analysis 2 (with age-adjustment)	Primary analysis (with deduplicaiton)	Sensitivity (without deduplication)	Primary (with deduplicaiton)	**Equation for linear combination of EBGM and TTO thresholds**
VAERS	77.2 (68.7, 85.7)	78.4 (70.1, 86.6)	76.3 (67.2, 85.3)	86.4(75.3, 97.5)	85.2 (74.0, 96.4)	90.9 (83.2, 98.6)	= EBGM – 0.27*ln(TTO_Other_Event) -- 0.33* ln(TTO_Other_Vaccine)
EudraVigilance	86.4 (79.9, 93.0)	86.7 (80.2, 93.2)	82.8 (75.3, 90.2)	82.9 (73.6, 92.1)	83.3 (73.7, 93.0)	94.4 (90.4, 98.5)	= EBGM – 0.04*ln(TTO_Other_Event) – 0.09*ln(TTO_Other_Vaccine)
Union (VAERS + EudraVigilance)	80.1 (71.9, 88.3)	81.4 (73.4, 89.4)	79.3 (70.9, 87.8)	85.7 (75.3, 96.1)	85.8 (75.4, 96.2)	92.3 (86.8, 97.9)	= EBGM – 0.09*ln(TTO_Other_Event) – 0.08 *ln(TTO_Other_Vaccine)

Discussion

This study aimed to compare the performance of the EBGM and TTO methods separately and in combination on the Eudravigilance and VAERS databases focusing on pediatric vaccine reports separately and pooled (unified). We have several key findings. First of all we noted that there is a significant duplication of reports between EV and VAERS(23%). Of the reports in EV, 35% originate from the European Union and 65% originate from the rest of the world while relating to products licensed in Europe (29). Second, we showed based on the GriP reference set as gold standard that the TTO method performed better than the EBGM in VAERS whereas EBGM performed better than the TTO method in Eudravigilance. When Eudravigilance pediatric vaccine reports and VAERS pediatric reports were combined, the TTO method performed better than EBGM.

Lower performance of the TTO method in EV is likely related to the higher percentage of missing values in that database. As fewer reports can be used by the TTO method, it results into a lower power to flag time-to-onset distributions. In EV, time to onset is not requested from the reporter but is calculated from the reported reaction start and exposure dates only if these are both provided in a valid date format. In VAERS, vaccination date and onset date are required fields from which CDC calculates the interval prior to making data available.

Better performance of the EBGM method in EV may be due to the fact that there is greater diversity of events in the vaccine-specific subset of the EV database (3,774 unique events/228,181 reports in EV vs. 6,641 unique events/1,556,663 reports in VAERS) Because adverse events following both drug and vaccine exposures are reported to EV, sometimes within the same report, there is a larger number of distinct non-reference set events present in the EV database than in VAERS, which is vaccine-specific.

Unification of the two databases did not improve performance over the better performing method in each database alone. This seems to indicate that performance is driven not by the increased power provided by a larger database but by the selection of the best method for a database. While the size of the database increased, the absence of performance improvement in adding EV to VAERS could potentially be due to the fact that the pediatric EV data represents only 14.7% of the size of the pediatric subset of VAERS with TTO data of poorer quality.

The linear combination of methods improved performance over either method alone, particularly in EV where time-to-onset data was frequently missing.

There are several limitations to this study. First, because of absence of other reference sets in pediatrics we used the reference set published by Pernus et al. as gold standard; this reference set includes only 18 positive controls. The positive controls were identified as such because they displayed obvious signs of causality including a strong temporal relationship (30). Because of a proven causality, reporting may be aligned to this which could potentially make the TTO signal detection look artificially more performant than it will be in prospective signal detection setting. It is to be noted though, that a similar performance assessment was done for the GSK Vaccines spontaneous report data using events included in the summary of product characteristics for GSK vaccines as positive controls and it provided similar conclusions with higher performance for the TTO signal detection compared to the EBGM (11(30)).

Second, the reference set contains associations which are not applicable to the total pediatric age group. Positive controls associated with MMR vaccine, for example, are not likely to be detected in adolescents who will only receive MMR vaccine in unusual situations such as a catch-up vaccination. Age stratification, however, served to decrease overall performance of the EBGM method. Previous studies have shown that age stratification may serve to decrease performance for Bayesian methods such as EBGM when some strata are sparse, as was the adolescent age strata in our study (31, 32).

Conclusion

In databases in which time to onset data is available, the TTO method, as opposed to EBGM, should be used for vaccine signal detection. Databases such as EudraVigilance with a large proportion of missing times to onset benefit from the linear combination of EBGM and TTO methods. The pooling of VAERS and EudraVigilance data did not lead to improved performance for any of the tested methods, since it was dominated by VAERS data and time to onset was often missing in Eudravigilance. We suggest that spontaneous reporting systems rigorously collect time to onset data , especially for reports of AEFIs.

References

1. Plotkin S. History of vaccination. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(34):12283-7. Epub 2014/08/20.

2. van Holle L, Bauchau V. Signal detection on spontaneous reports of adverse events following immunisation: a comparison of the performance of a disproportionality-based algorithm and a time-to-onset-based algorithm. Pharmacoepidemiology and drug safety. 2014;23(2):178-85. Epub 2013/09/17.

3. Van Holle L, Bauchau V. Optimization of a quantitative signal detection algorithm for spontaneous reports of adverse events post immunization. Pharmacoepidemiology and drug safety. 2013;22(5):477-87. Epub 2012/12/21.

4. Walker AM. Signal detection for vaccine side effects that have not been specified in advance. Pharmacoepidemiology and drug safety. 2010;19(3):311-7. Epub 2009/12/17.

5. Van Holle L, Zeinoun Z, Bauchau V, Verstraeten T. Using time-to-onset for detecting safety signals in spontaneous reports of adverse events following immunization: a proof of concept study. Pharmacoepidemiology and drug safety. 2012;21(6):603-10. Epub 2012/03/03.

6. Huang X, Qin G, Fang Y. Optimal Combinations of Diagnostic Tests Based on AUC. Biometrics. 2011;67(2):568-76. Epub 2010/06/22.

7. Kang L, Liu A, Tian L. Linear combination methods to improve diagnostic/prognostic accuracy on future observations. Statistical methods in medical research. 2016;25(4):1359-80. Epub 2013/04/18.

8. Pfeiffer RM, Bur E. A model free approach to combining biomarkers. Biometrical journal Biometrische Zeitschrift. 2008;50(4):558-70. Epub 2008/07/30.

9. Van Holle L, Bauchau V. Use of logistic regression to combine two causality criteria for signal detection in vaccine spontaneous report data. Drug safety. 2014;37(12):1047-57. Epub 2014/11/15.

10. Li Y, Ryan PB, Wei Y, Friedman C. A Method to Combine Signals from Spontaneous Reporting Systems and Observational Healthcare Data to Detect Adverse Drug Reactions. Drug safety. 2015;38(10):895-908. Epub 2015/07/15.

11. Harpaz R, Vilar S, Dumouchel W, Salmasian H, Haerian K, Shah NH, et al. Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions. Journal of the American Medical Informatics Association : JAMIA. 2013;20(3):413-9. Epub 2012/11/03.

12. Chen RT, Rastogi SC, Mullen JR, Hayes SW, Cochi SL, Donlon JA, et al. The Vaccine Adverse Event Reporting System (VAERS). Vaccine. 1994;12(6):542-50. Epub 1994/05/01.

13. Varricchio F. The vaccine adverse event reporting system. Journal of toxicology Clinical toxicology. 1998;36(7):765-8. Epub 1998/12/29.

14. Braun M. Vaccine adverse event reporting system (VAERS): usefulness and limitations. . Johns Hopkins Bloomberg School of Public Health; 2017; Available from: http://www.vaccinesafety.edu/VAERS-Braun.htm.

15. European Medicines Agency policy on access to EudraVigilance data for medicinal products for human use, (2016).

16. 2017 Annual Report on EudraVigilance for the European Parliament, the Council and the Commission. London: European Medicines Agency, 2018 15 March 2018. Report No.

17. United States Department of Health and Human Services CfDCaP, Food and Drug Administration. VAERS DATA USE GUIDE. 2017 [cited 2017]; Available from: https://vaers.hhs.gov/docs/VAERSDataUseGuide_June2017.pdf.

18. Hans-Georg Wagner OS, Timothy Buxton. The European System for Monitoring Drug Safety, EudraVigilance (Version 1.0) 2005. Available from: http://www.lexjansen.com/phuse/2005/rc/rc07.pdf.

19. Dodd CNP, A.; Osokogu, O.U.; Weibel, D,; Ferrajolo, C.; Vo, D.H.; Becker, B.; Kors, J.A.; Sturkenboom, M. . Masking by vaccines in pediatric drug safety signal detection in the EudraVigilance database. Pharmacoepidemiology and drug safety. 2018.

20. Becker BS, Miriam; Kors, Jan; Merino, Elisa Martin; Roberto, Giuseppe. Ontology for the Integration and Extraction of Vaccine-related Information in Europe: a proof of concept. ADVANCE Consortium, 2017 September 29, 2017. Report No.

21. Brauchli Pernus Y, Nan C, Verstraeten T, Pedenko M, Osokogu OU, Weibel D, et al. Reference set for performance testing of pediatric vaccine safety signal detection methods and systems. Vaccine. 2016;34(51):6626-33. Epub 2015/10/27.

22. Use. ICOHOTRFROPFH. Clinical Investigation Of Medicinal Products In The Pediatric Population E11. 2000.

23. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. The American Statistician. 1999;53(3):177-90.

24. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiology and drug safety. 2003;12(7):559-74. Epub 2003/10/16.

25. Candore G, Juhlin K, Manlik K, Thakrar B, Quarcoo N, Seabroke S, et al. Comparison of statistical signal detection methods within and across spontaneous reporting databases. Drug safety. 2015;38(6):577-87. Epub 2015/04/23.

26. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22(4):719-48. Epub 1959/04/01.

27. Smirnov NV. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. Bull Math Univ Moscou 1939;2(2).

28. Pepe MS, Thompson ML. Combining diagnostic test results to increase accuracy. Biostatistics. 2000;1(2):123-40. Epub 2003/08/23.

29. Agency EM. 2015 Annual Report on EudraVigilance for the European Parliament, the Council and the Commission. European Medicines Agency, Division IaHMP; 2016.

30. Norén GN, Caster O, Juhlin K, Lindquist M. Zoo or savannah? Choice of training ground for evidence-based pharmacovigilance. Drug safety. 2014;37:655-9.

31. Hopstadius J, Norén GN, Bate A, Edwards IR. Impact of stratification on adverse drug reaction surveillance. Drug safety. 2008;31(11):1035-48.

32. Osokogu OU, Dodd C, Pacurariu A, Kaguelidou F, Weibel D, Sturkenboom MC. Drug Safety Monitoring in Children: Performance of Signal Detection Algorithms and Impact of Age Stratification. Drug safety. 2016;39(9):873-81. Epub 2016/06/04.

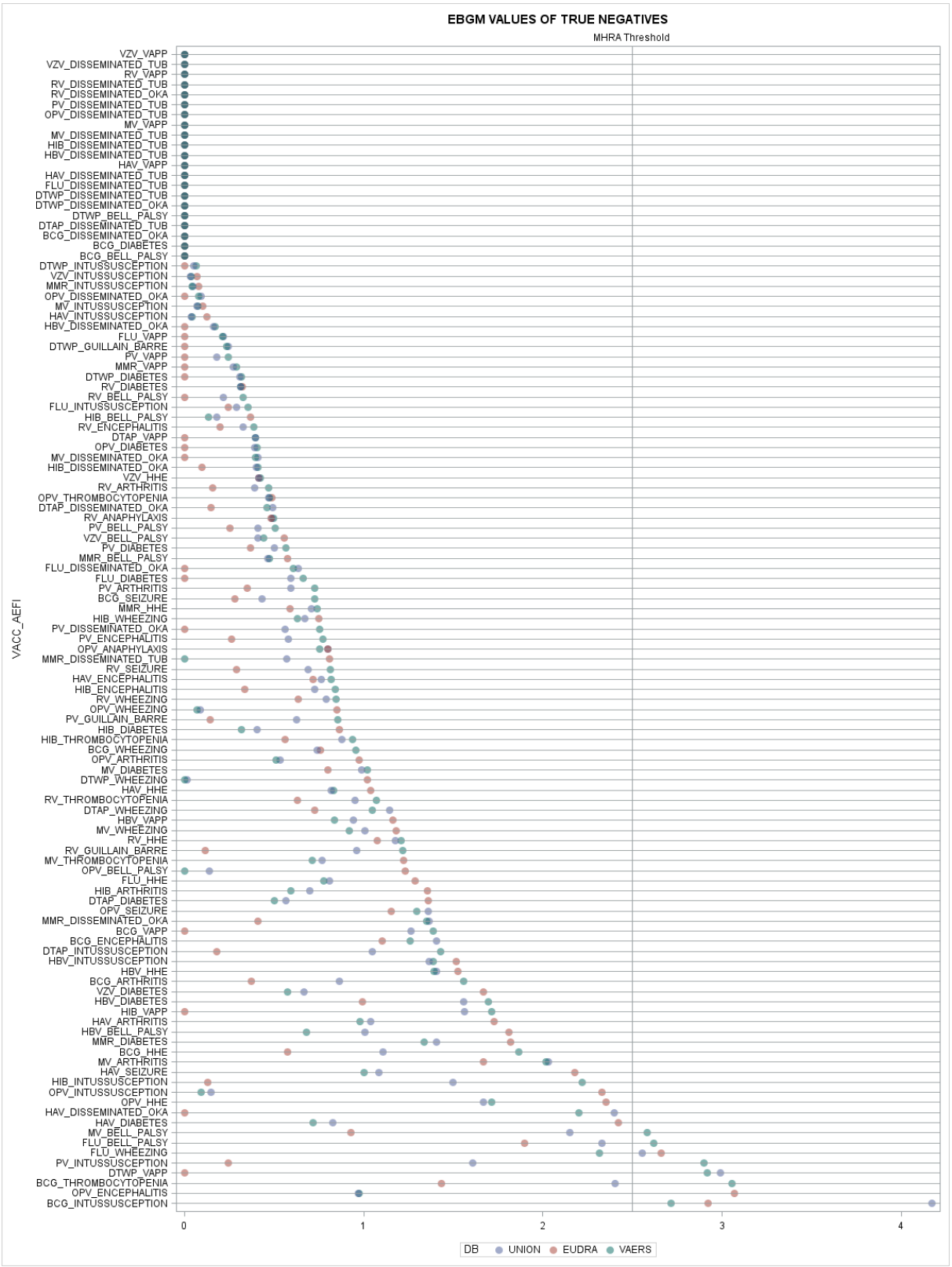
Supplemenatry Material

Supplementary Table 1. Counts of GRiP reference set associations reported in the pediatric subsets of EudraVigilance and VAERS

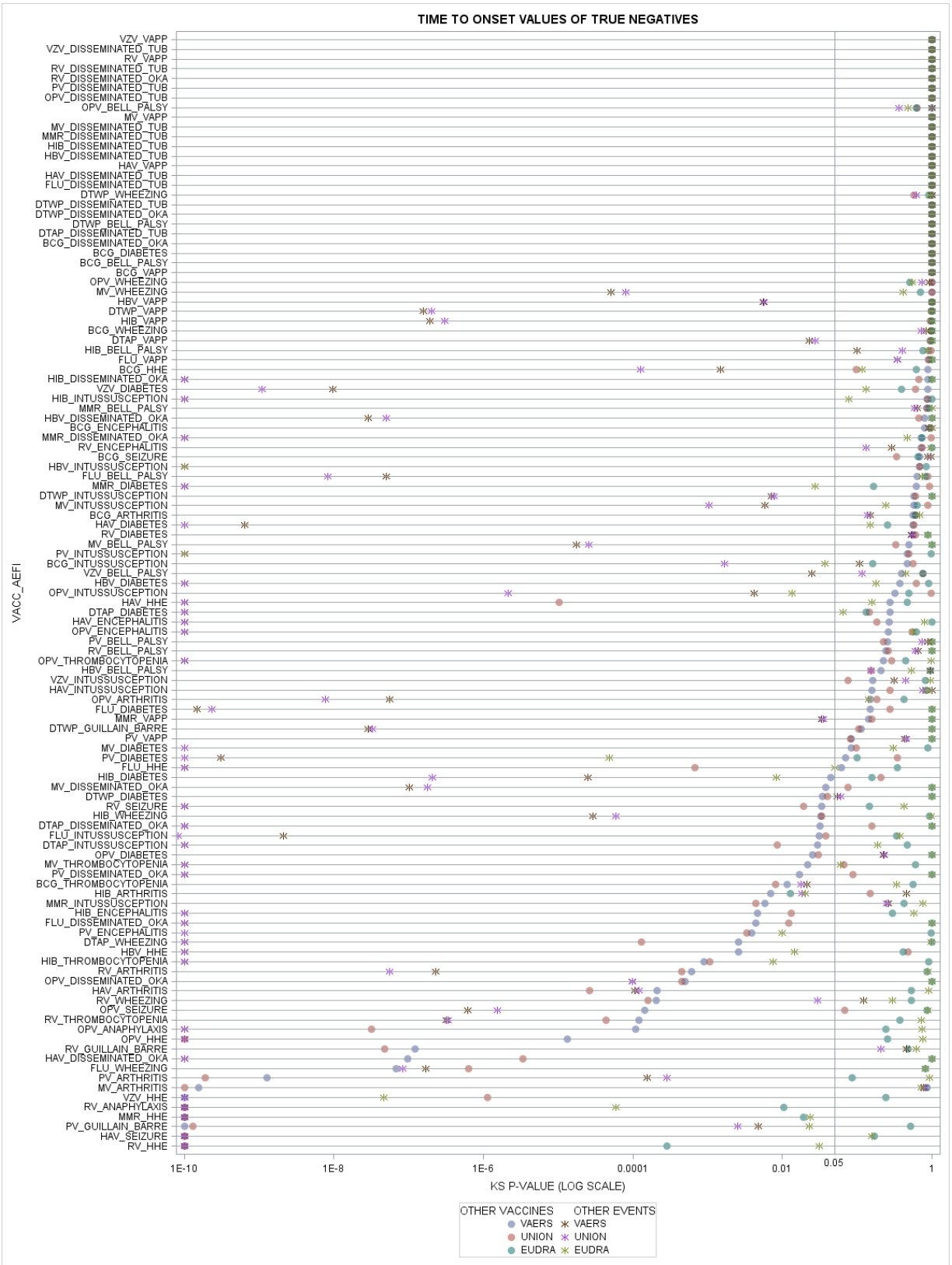
		ANAPHYLAXIS	ARTHRITIS	BELLS PALSY	CONVULSIONS	DISSEMINATED BCG-ITIS	DISSEMINATED OKA VZV	ENCEPHALITIS	GBS	HHE	IDDM	INTUSSUSCEPTION	THROMBOCYTOPENIA	VAPP	WHEEZING/REACTIVE AIRWAY DISEASE
BCG	EUDRAVIGILANCE	36	2	0	6	8	0	3	0	14	0	17	5	0	1
	VAERS	33	8	0	13	1	0	2	1	25	0	6	8	1	2
DTAP	EUDRAVIGILANCE	999	63	8	500	0	0	39	7	641		13	45	0	12
	VAERS	18469	1148	17	5438	0	140	183	158	3769	60	592	353	7	572
DTPW	EUDRAVIGILANCE	36	2	0	34	0	0	3	0	40	0	0	3	0	1
	VAERS	2814	145	0	2107	0	0	60	13	1922	9	6	34	16	0
HAV	EUDRAVIGILANCE	194	26	1	119	0	0	4	6	61	6	1	7	0	53
	VAERS	6785	391	13	1610	0	189	73	76	950	24	4	120	0	212
HBV	EUDRAVIGILANCE	1191	103	13	381	0	0	47	18	666	13	137	70	2	19
	VAERS	7850	794	11	2680	0	22	171	102	2490	90	251	278	7	192
HIB	EUDRAVIGILANCE	1544	196	3	387	0	1	18	7	503	15	15	39	0	20
	VAERS	11489	611	3	5601	0	89	193	95	4181	27	660	361	24	247
INFLUENZA	EUDRAVIGILANCE	503	77	6	278	0	0	31	33	212	0	8	29	0	22
	VAERS	7454	536	40	2278	0	72	146	211	1242	31	57	122	1	494
MMR	EUDRAVIGILANCE	2070	143	4	624	1	5	71	11	275	26	7	159	0	22
	VAERS	14013	855	11	3313	0	266	244	77	1946	104	11	569	3	340
MV	EUDRAVIGILANCE	419	50	2	141	0	0	21	17	118	3	2	18	0	7
	VAERS	3600	627	23	1359	0	26	75	133	942	27	6	83	0	109
OPV	EUDRAVIGILANCE	44	5	1	21	0	0	8	5	47	0	11	1	5	1
	VAERS	3820	198	0	2033	0	6	85	22	1910	13	10	69	50	10
PV	EUDRAVIGILANCE	4040	121	5	1089	0	0	34	7	1571	15	71	92	0	52
	VAERS	9396	567	10	3544	0	126	135	110	2389	37	656	393	2	338
RV	EUDRAVIGILANCE	476	13	0	92	0	0	6	1	373	3	793	26	0	10
	VAERS	1811	132	2	922	0	0	24	57	972	7	854	113	0	91
VZV	EUDRAVIGILANCE	915	47	2		0	110	50	6	101	13	3	43	3	18
	VAERS	14973	449	10	2029	0	868	123	50	1086	43	9	287	0	378

	Positive Control
	Negative Control
	Unclassifiable

Supplementary Figure 1. EBGM values of true negative associations in EudraVigilance, VAERS, and their union.



Supplementary Figure 2. TTO values of true negative associations in EudraVigilance, VAERS, and their union



Supplementary Material. GRiP Common Data Model

GRiP COMMON DATA MODEL

Reference document

Authors: Alexandra Pacurariu, Caitlin Dodd, Florentia Kaguelidou, Marius Gheorghe

Table definitions

Report table

ID_REPORT (character) – Primary Key

FOLLOW_UP (numeric) – the version of the report (1 = initial, 2, 3…)

REPORTER (character) – qualification of the reporter – extracted as reported in each database. (Data will be further classified as: MD= physician; PH+ pharmacist; OT= other health professional; LW= lawyer; CN=consumer)

DATE (date) – the date of registration of the report in the database (DDMMYYYY)

COUNTRY (character) – the country, region or state of origin of the report, this is not the country where the event occurred.

TYPE_SERIOUSNESS(character) – the type of seriousness of the report based on the WHO categories(Hospitalization or prolongation of existing hospitalization; Life-threatening; Death; Significant or persistent disability/incapacity; Congenital anomalies; other relevant conditions). More than one criterion may be present per report (to be put in same variable: later choose the most relevant)

Drug table

ID_DRUG (numeric) – Primary Key

ID_REPORT (character) – Foreign Key (FK) from **Report table**

NAME (character)- international nonproprietary name (when possible) = active substance name

MANUFACTURER (character) – this information will be extracted only for vaccines

ATC (character) - code assigned to an active substance

DOSE_AMOUNT (character) – the quantity of active substance per intake = drug dose per intake

DOSE_UNIT (character) – the unit of the drug dose per intake

DOSE_FREQ (character) – the frequency of drug administration

CUMULATIVE_DOSE (numeric) – the quantity of active substance until first event

CUMULATIVE_DOSE_UNIT (character) – the unit of the drug cumulative dose

RECHALLENGE (character) – this variable is defined as follows: positive= event reoccurred when drug therapy was restarted; negative= event did not reoccur when drug therapy was restarted; unknown; does not apply

DECHALLENGE (character) – this variable is defined as follows: positive= event abated when drug therapy stopped; negative= event did not abate when drug therapy stopped; unknown; does not apply

ROUTE (character) – the route of administration (to be classified as: topical, enteral and parenteral)

DOSE_NB (numeric) –current number of administrations at the occurrence of event. This information will be extracted only for vaccines.

LOT_NUMBER (numeric) – this information will be extracted only for vaccines.

ROLE (character) – drug's reported role in the event (will be classified as suspect, concomitant, interacting)

Indication table

ID_REPORT (character) – FK from the **Drug table**

ID_DRUG (numeric) – FK from the **Drug table**

IND_DESC (character) – MedDRA preferred term (PT) describing the indication for the use of the drug

IND_CODE (numeric) – MedDRA code corresponding to the PT for drug indication

Event table

ID_REPORT (character) – Foreign Key (FK) from the **Report table**

FOLLOW_UP (numeric) – FK from the **Report table**

DATE (date) – the date of occurrence of the eventOUTCOME (character) - reported outcome of the (Fully recovered/resolved; Recovering/resolving; Not recovered/not resolved; Recovered/resolved with sequelae; Caused death; Unknown)

PT_DESC (character) – preferred term (PT) of the MedDRA terminology describing the reported adverse event

PT_CODE (numeric) –MedDRA code corresponding to the PT term

SOC_DESC (character) – system organ class (SOC) of the MedDRA terminology of the reported adverse event

SOC_CODE (numeric) –MedDRA code corresponding to the SOC

Therapy table

ID_REPORT (character) – Foreign Key (FK) from the **Report table**

ID_DRUG (numeric) – Foreign Key from **Drug table**

START (date) – the date the therapy begins

END (date) – the date the therapy ends

DURATION (numeric) – the length of the therapy in days

START_UNTIL_EVENT (numeric) – difference between START **(Therapy)** and DATE (**Event)** (where missing, to be calculated)

END_UNTIL_EVENT (numeric) – difference between END **(Therapy)** and DATE (**Event)** (where missing, to be calculated)(allowed to be negative)

Demographics table (a.k.a Patient table)

ID_REPORT (character) - FK from the **Report table**

FOLLOW_UP (numeric) – the version of the report (if not present: to be calculated)

CALCULATED_AGE – age at occurrence of the eventalready calculated by the database

UNIT_CALCULATED_AGE – unit of calculated age

REPORTED_AGE – age at occurrence of the event provided in the report

UNIT_REPORTED_AGE – unit of reported age

AGE_GROUP (character) – as provided by the database (variable to be created as: newborn <= 27; infant/toddler = 28d – 2y; child = 2y+1d – 11y; adolescent = 12y - <18y; unknown)

AGE (numeric) – in months, either from 'calculated_age' or 'reported_age'; if both provided, 'calculated age' is to be kept

SEX(character) – choice between FEMALE, MALE, UNKNOWN

MAPPING Eudravigilance

Report table

ID_REPORT – EV_LOCAL_NUMBER (grip_cases)

FOLLOW_UP – create from EV_LOCAL_NUMBER (grip_cases) and MESSAGEGATEWAYDATE (grip_cases) (sorted dates and based on those figure out initial = first date and so on)

REPORTER – QUALIFICATION_TXT (grip_cases)

DATE – MESSAGEGATEWAYDATE (grip_cases)

COUNTRY –to be calculated from REPORTERSUBREGION(grip_cases)

TYPE_SERIOUSNESS – N/A – can be assumed that all is to be considered serious

Drug table

ID_DRUG -FK_DRUG_SUBSTANCE or FK_DRUG_PRODUCT (grip_drugs)- to be chosen

ID_REPORT – EV_LOCAL_NUMBER (grip_drugs)

NAME - ACTIVESUBSTANCENAME_REC (grip_drugs).

MANUFACTURER –only for vaccines N/A

ATC – ATCCODE (grip_drugs) – mostly empty – Eric mapping subsequently

DOSE_AMOUNT – DRUGSTRUCTUREDDOSAGENUMB (grip_drugs)

DOSE_FREQ– DRUGSEPARATEDOSAGENUMB+ DRUGINTERVALDOSAGEUNITNUMB + DRUGINTERVALDOSAGEDEFINITION_TXT (grip_drugs)

DOSE_UNIT – DRUGSTRUCTUREDOSAGEUNIT_TXT (grip_drugs)

CUMULATIVE_DOSE – DRUGCUMULATIVEDOSAGENUMB (grip_drugs)

CUMULATIVE_DOSE_UNIT – DRUGCUMULATIVEDOSAGEUNIT_TXT (grip_drugs)

RECHALLENGE – DRUGRECURREADMINISTRATION_TXT (grip_drugs)

DECHALLENGE – N/A -combination between ACTIONDRUG_TXT (grip_drugs) and REACTIONOUTCOME_TXT (grip_reactions)

ROUTE – DRUGADMINISTRATIONROUTE_TXT(grip_drugs) and if DRUGPARADMINISTRATION_TXT has value then fill “Transplacental”

DOSE_NB – N/A

LOT_NUMBER – N/A

ROLE – DRUGCHARACTERIZATION_TXT(grip_drugs)

Indication table

ID_REPORT – FK from the **Drug table**

ID_DRUG – FK from the**Drug table**

IND_CODE – N/A

IND_DESC – N/A

Event table

ID_REPORT – EV_LOCAL_NUMBER (grip_reactions)

FOLLOW_UP – N/A - to be calculated from MESSAGEGATEWAYDATE and ID_REPORT – sorted dates

DATE–N/A

OUTCOME – reactionoutcome_txt (GRIP_REACTIONS)

PT_CODE – pt_code (GRIP_REACTION)

PT_DESC– N/A - will be taken from MedDRA based on the code

SOC_CODE– N/A - will be completed by biosemantics based on the code

SOC_DESC – N/A - will be taken from MedDRA based on the code

Therapy table

ID_DRUG – FK **Drugtable**

START – N/A

END – N/A

DURATION – DRUGTREATMENTDURATION_CALC(grip_drugs)

START_UNTIL_EVENT – DRUGSTARTPERIOD(grip_drugs)

END_UNTIL_EVENT – DRUGLASTPERIOD(grip_drugs)

Demographics table (a.k.a Patient table)

ID_REPORT - EV_LOCAL_NUMBER (grip_cases)

FOLLOW_UP – N/A

AGE – AGEREACTION_CALC_MIN + PATIENTONSETAGE+PATIENTONSETAGEUNIT_TXT(grip_patients) – to be converted into months

AGE_GROUP – PATIENTAGEGROUP_TXT(grip_patients)

SEX – PATIENTSEX_Txt(grip_patients)

MAPPING FAERS

Report table

ID_REPORT- ISR (demographic file)

FOLLOW_UP – Create using ISR and FDA_DT (demographic file)

DATE – FDA_DT (demographic file)

REPORTER – OCCP_COD (demographic file)

COUNTRY –REPORTER_COUNTRY (demographic file)

REPORTER_COUNTRY is only available starting from 2005Q3

SERIOUSNESS – OUTC_COD (outcome file)

Drug table

ID_DRUG – own generated identifier

ID_REPORT –ISR (drug file)

ISR is used up till (and including) 2012Q3

NAME – DRUGNAME (drug file)

MANUFACTURER – N/A

ATC – N/A, to be added by Erik

DOSE_AMOUNT – DOSE_AMT (drug file)

DOSE_UNIT – DOSE_UNIT (drug file)

DOSE_FREQ – DOSE_FREQ (drug file)

RECHALLENGE - RECHAL (drug file)

DECHALLENGE - DECHAL (drug file)

ROUTE – ROUTE (drugs file)

DOSE_NB – N/A

LOT_NUMBER – LOT_NUM(drug file)

ROLE – ROLE_COD (drug file)

Indication table

ID_REPORT – FK from the **Drug table**

ID_DRUG – FKfrom the**Drug table**

IND_CODE – N/A (to be mapped)

IND_DESC – INDI_PT (indication file)

Event table

ID_REPORT –ISR (reaction file)

FOLLOW_UP N/A

DATE –EVENT_DT (demographicfile)

OUTCOME –N/A

PT_CODE – N/A

PT_DESC –PT (reaction file)

SOC_CODE –N/A

SOC_DESC – N/A

Therapy table

ID_DRUG – FK **Drugtable**

START –START_DT (therapy file) – sometimes only year and month

END –END_DT (therapy file) – most of the times missing

DURATION – DUR + DUR_COD (therapy table)

START_UNTIL_EVENT – N/A

END_UNTIL_EVENT – N/

Demographics table (a.k.a Patient table)

ID_REPORT –ISR (demographicfile)

FOLLOW_UP –CREATE FROM ISR AND FDA_DT (demographic file)

CALCULATED_AGE – N/A

UNIT_CALCULATED_AGE – N/A

REPORTED_AGE – AGE (demographic file)age at occurrence of the event provided in the report

UNIT_REPORTED_AGE – GE_COD (demographic file)

AGE_GROUP – N/A

AGE – N/A

SEX – GNDR_COD (demographic file)

MAPPING VAERS

Report table

ID_REPORT - VAERS_ID (vaersdata, vaerssymptoms, vaersvax)

FOLLOW_UP – Create from VAERS_ID + RECVDATE (vaersdata)

DATE – RECVDATE (vaersdata)

REPORTER – N/A

COUNTRY – STATE (vaersdata)

TYPE_SERIOUSNESS – DIED (vaersdata) + L_THREAT (vaersdata) + ER_VISIT (vaersdata) + HOSPITAL (vaersdata) + X_STAY (vaersdata) + DISABLE (vaersdata) + congenital anomalies that have to be extracted from the vaerssymptoms via PT

Drug table

ID_DRUG – own generated identifier

ID_REPORT – VAERS_ID (vaersdata)

NAME – VAX_NAM (vaersvax)

MANUFACTURER – VAX_MANU (vaersvax)

ATC – N/A

DOSE_AMOUNT – N/A

DOSE_FREQ – N/A

DOSE_UNIT – N/A

CUMULATIVE_DOSE – N/A

CUMULATIVE_DOSE_UNIT – N /A

RECHALLENGE - N/A

DECHALLENGE - N/A

ROUTE – VAX_ROUTE (vaersvax)

DOSE_NB – VAX_DOSE (vaersvax)

LOT_NUMBER – VAX_LOT (vaersvax)

ROLE – N/A

Indication table

ID_REPORT – FK from the **Drug table**

ID_DRUG – FK from the**Drug table**

IND_CODE –N/A

IND_DESC – N/A

Event table

ID_REPORT – VAERS_ID (one of the following vaersdata, vaerssymptoms, vaersvax)

FOLLOW_UP – Create from VAERS_ID + recvdate (vaersdata)

DATE – ONSET_DATE (vaersdata)

OUTCOME –DIED, RECOVD (vaersdata), if no info in DIED or RECOVD the value is 'unknown')

PT_CODE – add PT_CODE from symptom1-symptom5 (vaerssymptoms)

PT_DESC – SYMPTOM1 – SYMPTOM5 (vaerssymptoms)

SOC_CODE –add from Symptom (vaerssymptoms)

SOC_DESC – add from Symptom (vaerssymptoms)

Therapy table

ID_DRUG – FK from the Drug Table

START – N/A

END – N/A

DURATION – N/A

START_UNTIL_EVENT – NUMDAYS (vaersdata) – days from vaccination to onset

END_UNTIL_EVENT – N/A

Demographics table (a.k.a Patient table)

ID_REPORT - VAERS_ID (one of the following vaersdata, vaerssymptoms, vaersvax)

FOLLOW_UP – Create from VAERS_ID + recvdate (vaersdata)

AGE – AGE_YRS (vaersdata) + CAGE_YR (if calculated age is missing (CAGE), AGE_YRS is to be kept)+ CAGE_MO (both CAGE_YRS and CAGE_MO necessary to create Calculated age variables) – to be converted into months

AGE_GROUP –N/A – to be created from AGE

GENDER – SEX (vaersdata)

CHAPTER 3. METHODS FOR RAPID ASSESSMENT

3.1 Incidence rates of narcolepsy diagnoses in Taiwan, Canada, and Europe: methods for assessment of potential safety issues on a population level in the SOMNIA study

Caitlin Dodd

Maria de Ridder

Wan-Ting Huang

Daniel Weibel

Maria Giner-Soriano

Silvia Perez-Vilar

Javier Diez-Domingo

Lawrence W. Svenson

Salahddin M Mahmud

Bruce Carleton

Monica Naus

Jeffrey C Kwong

Brian J Murray

Lisen Arnheim-Dahlstrom

Lars Pedersen

Rosa Morros

Franciso Javier Puertas

Steven Black

Miriam C J M Sturkenboom

*PloS one*13.10 (2018): e0204799.

Abstract

Vaccine safety signals require investigation, which may be done rapidly at the population level using ecological studies, before embarking on hypothesis-testing studies. Incidence rates were used to assess a signal of narcolepsy following AS03-adjuvanted monovalent pandemic H1N1 (pH1N1) influenza vaccination among children and adolescents in Sweden and Finland in 2010. We explored the utility of ecological data to assess incidence of narcolepsy following exposure to pandemic H1N1 virus or vaccination in 10 sites that used different vaccines, adjuvants, and had varying vaccine coverage.

We calculated incidence rates of diagnosed narcolepsy for periods defined by influenza virus circulation and vaccination campaign dates, and used Poisson regression to estimate incidence rate ratios (IRRs) comparing the periods during which wild-type virus circulated and after the start of vaccination campaigns vs. the period prior to pH1N1 virus circulation. We used electronic health care data from Sweden, Denmark, the United Kingdom, Canada (3 provinces), Taiwan, Netherlands, and Spain (2 regions) from 2003 to 2013. We investigated interactions between age group and adjuvant in European sites and conducted a simulation study to investigate how vaccine coverage, age, and the interval from onset to diagnosis may impact the ability to detect safety signals.

Incidence rates of narcolepsy varied by age, continent, and period. Only in Taiwan and Sweden were significant time-period-by-age-group interactions observed. Associations were found for children in Taiwan (following pH1N1 virus circulation) and Sweden (following vaccination). Simulations showed that the individual-level relative risk of narcolepsy was underestimated using ecological methods comparing post- vs. pre-vaccination periods; this effect was attenuated with higher vaccine coverage and a shorter interval from disease onset to diagnosis.

Ecological methods can be useful for vaccine safety assessment but the results are influenced by diagnostic delay and vaccine coverage. Because ecological methods assess risk at the population level, these methods should be treated as signal-generating methods and drawing conclusions regarding individual-level risk should be avoided.

Introduction

In August 2010, a safety signal of narcolepsy following AS03-adjuvanted pdm(09)H1N1 influenza vaccine Pandemrix® was reported in Finland and Sweden among children and adolescents [1]. Other rapid risk assessment studies conducted in the European Union (EU) did not show changes in incidence rates of narcolepsy diagnoses, except in Finland, Sweden, and Norway [2], all countries that achieved high coverage rates with Pandemrix. Subsequent hypothesis-testing studies showed associations; these had high within- and between-study variation [3]. In China, where vaccine coverage was very low, a 3-fold increase in narcolepsy onset was reported following the peak of the pandemic [4].

Narcolepsy is a rare disease with a long interval from onset of symptoms to diagnosis, especially in adults. Several possible explanations for the purported pdm(09)H1N1 and narcolepsy link have been proposed but none confirmed. Hypotheses range from a causal effect of the AS03 adjuvant, the manufacturing process, presence of nucleoproteins in Pandemrix, and molecular mimicry, to awareness and assessment biases, and residual confounding [5-10]. Based upon simulation studies conducted by Wijnans et al., in the absence of a causal association but in the presence of accelerated diagnosis due to awareness, we would expect to see an increased incidence of narcolepsy diagnosis following awareness of the association followed by a decrease, even to levels below the background incidence, due to depletion of cases[8]. This effect may be particularly important in conditions with a long delay to diagnosis such as in narcolepsy where the delay in diagnosis from initial symptoms can be 10-20 years [11].

The SOMNIA (Systematic Observational Method for Narcolepsy and Influenza Immunization Assessment) study was funded by the US Centers for Disease Control and Prevention (CDC) and used information from countries that used different types of adjuvanted pandemic influenza vaccines to assess whether the pdm(09)H1N1 influenza vaccine and specifically the MF59 and AS03 adjuvants were associated with narcolepsy.

One of the goals of SOMNIA was to assess patterns of incidence rates of narcolepsy in multiple geographic areas and to understand changes in incidence rates of narcolepsy diagnoses before, during, and after the pdm(09)H1N1 influenza pandemic by using electronic health care data, which may be rapidly available. In this paper, we explore whether assessment of safety signals based on ecological methods and population-based electronic health care data are suitable for vaccine safety risk assessment, by exploiting the heterogeneity in vaccine coverage, types of vaccines, and vaccination programs across countries. We assess what strength of signals can be detected using population-level data collected before and after a hypothetical targeted vaccination campaign.

Ecological studies can be defined as those that measure exposure and outcomes at the group level rather than at the individual level [12, 13]. In such a study, groups are defined by a naturally occurring difference in space or time such as a change in the vaccination schedule [14] or the beginning and end of a targeted vaccination campaign [15].

This study may serve as an example of the utility of ecological methods to assess vaccine safety signals, particularly regarding events with long onset-to-diagnosis intervals.

Materials and Methods

Narcolepsy diagnosis incidence rates were evaluated in ten sites representing seven countries spanning three continents (Taiwan (TW), Canada (CA) [Manitoba, Alberta, and British Columbia], The Netherlands (NL), The United Kingdom (UK), Sweden (SE), Denmark (DK), and Spain (ES) [Valencia and Cataluña]) using population-based electronic healthcare databases originating from general practitioners (GPs) (UK, ES, NL) or claims/record linkage databases (SE, DK, TW [16-18], and CA) (Supplementary Table 1).

Study population and follow-up

For data sources in which individual linkage can be made between population and diagnoses (all sites except Sweden and British Columbia, Canada), the study population comprised all individuals registered within each of the databases during the study period. Observation time began on the date of first registration of an individual in the database, the start of the study period (January 2003), or the start date of data collection for the database, whichever was the latest and ended on the date of death, the date registration was terminated, the end of data collection, or the end of the study period (December 2013), whichever was the earliest. Sweden and British Columbia, Canada used census data to calculate person-time denominators. We used a harmonized approach in which databases locally extracted their data into simple input files in a common format that could be locally analyzed and aggregated using SAS or JAVA-based software [2, 19].

Case identification and validation

Cases were persons with a new diagnosis of narcolepsy with or without cataplexy. Validation of the diagnostic codes using patient discharge letters and medical records was conducted in the GP databases in the Netherlands and Valencia, Spain. For these two sites, only validated cases were used in the analysis. The other sites used algorithms combining diagnostic codes for narcolepsy with claims for multiple sleep latency tests (MSLTs) to reduce the false positive rate. The same method was used at each site over the entire time period. No further validation was done in other sites (Supplementary Table 1).

Analysis

To investigate the purported narcolepsy-pandemic vaccine effect, incidence rates of narcolepsy diagnosis were calculated by calendar year and month and also categorized into three periods based on specific circulation/vaccination periods in each country: 1) pre-pandemic (from January 2003 until the start of the period of pH1N1 circulation); 2) during pH1N1 wild-type virus circulation until the start of the country's pH1N1 vaccination campaign; and 3) from the start of the pH1N1 influenza vaccination campaign through the end of the study (Supplementary Table 1). Pandemic H1N1 virus circulation was defined as the period during which weekly influenza test positivity for pH1N1 infection exceeded 10%. Dynamic age groups were categorized as <5 years, 5-19 years, 20-59 years, and ≥60 years at the time of diagnosis. These age groups were motivated by differences in diagnosis for each age group, and particularly the challenges of differential diagnosis in young children and the elderly [20, 21]. Incidence rates of narcolepsy diagnoses were calculated by dividing the number of narcolepsy cases by the accumulated person-time. Ninety-five percent confidence intervals (CIs) were calculated assuming a negative binomial distribution. Following confirmation of homogeneity in incidence rates among databases within the same country, further analyses were conducted at the level of the country rather than the site.

Within each country, we estimated incidence rate ratios (IRRs) and 95% CIs for each time period using Poisson regression, with the pre-circulation period as a reference. We included terms for age strata, time periods, and an age*time period interaction using time periods as defined by pH1N1 circulation and vaccination campaign dates.

We conducted additional analyses restricted to European countries to estimate the impact of vaccine coverage and adjuvant among children and adolescents and separately among adults. For this analysis, a composite variable summarizing vaccine coverage classified as low (<20%) or high (≥20%) and adjuvant (MF59 or AS03) was created, and incidence in the period after vaccination had started was compared to the pre-pH1N1 circulation period. Because the composite adjuvant/coverage variable was collinear with database and country, neither database nor country was included in the European model.

Simulation

To better understand the utility of ecological methods for assessing vaccine safety signals and whether an association in one age group may be masked in a population-level analysis, we conducted statistical simulations population of 10,000 subjects aged 0 to 100 years. Baseline narcolepsy incidence in these subjects was simulated based on reported estimates of incidence by decade of age [22]. We then varied vaccine coverage and the individual-level relative risk of the association between vaccination and narcolepsy. Using vaccine coverage ranging from 1 to 99 percent, part of the population was given a vaccination date in the period from October to December 2009. Pre-vaccination campaign observation time was held constant at 2466 days (6.75 years). Total observation time was varied between 2739 days (7.5 years) to allow for 6 months of observation time following the simulated 90-day vaccination campaign, and 3625 days (10 years) to allow for 3 years post-vaccination and to mimic the current study period of 2003-2013. Relative risk of narcolepsy onset was varied from 0.5 to 10 in the six months following vaccination in subjects aged <20 years and held constant at 1 for subjects aged ≥20 years. . The median time from onset to diagnosis was initially set at 4 years for adults and at 1.5 years for children (aged ≤18 years) based on SOMNIA data (not shown). The effect of the length of the interval from onset to diagnosis was also tested by varying a scale parameter, with 0 removing the interval (i.e. immediate diagnosis following onset), 0.5 halving the interval, and 1 retaining the full simulated interval (see table 1 for description of simulation parameters). For each set of simulation parameters, 500 replications were run. The population-level incidence rate ratio for the period following the vaccination campaign vs. the period prior to the vaccination campaign was estimated using Poisson regression and compared to the simulated individual-level relative risk associated with vaccination. The estimated median incidence rate ratio from the 500 replications, as well as differences by age group and overall between each median IRR and the simulated relative risk for that age group were calculated and plotted against vaccine coverage. Using the median IRRs estimated from the simulated data, percent bias was calculated by subtracting the simulated relative risk from the estimated IRR and dividing by the simulated IRR.

Table 1. Simulation Parameters

Parameter	Definition	Levels
Scale	Reduction of onset to diagnosis interval	0 (Immediate diagnosis following onset), 0.5 (Halving of the onset-to-diagnosis interval), 1 (Full onset-to-diagnosis interval is retained)
Vaccine Coverage	Probability of vaccination	.01, .05, .10, .25, .50, .75, .90, .95, .99
Relative Risk	Relative risk of narcolepsy onset in first 6 months after vaccination	0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Observation Length	Total length of observation time following start of observation in January 2003	2739 days (pre-vaccination campaign time + 90 days vaccination campaign + 6-month risk period), 3625 days (10 years)

Calibration

A model was constructed using the simulation input parameters and results to predict individual-level relative risk among those aged 5-19 years, based on median estimated IRR, vaccine coverage, scale parameter, and the interaction of these terms (Model: Simulated individual-level true RR = Estimated IRR + Simulated vaccine coverage + Simulated onset-to-diagnosis scale + Interactions). Calibration was restricted to the 5-19 year age group as this group was the source of the safety signal and the only subjects for whom an increased risk was simulated. This prediction model was then applied to results obtained from the observed data along, using known values for vaccine coverage and diagnostic delay (Supplementary Table 1), in order to calculate the underlying relative risks in

5- to 19-year olds, that would have been necessary to produce the IRRs we found in the SOMNIA study in the absence of other sources of bias.

This study was conducted under the principles of the Helsinki declaration and each site was responsible for obtaining appropriate ethical approvals. The overall study was also approved by the central institutional review board for this study at Cincinnati Children's Hospital, Cincinnati, Ohio, USA.

Results

Observed Incidence

Incidence rates of narcolepsy diagnoses ranged from 0.22 to 1.52 per 100,000 person-years by site (table 2). Incidence rates in databases within the same country (Canada and Spain) were similar so for further analysis country-specific data were pooled.

Table 2: Crude Incidence Rates by Site

Site	Period	Events	Person-years	IR
EU				
Denmark	2003-2013	269	17,850,129	1.50 (1.33-1.69)
United Kingdom	2003-2013	467	42,897,721	1.09 (0.99-1.19)
ES, Valencia (validated)	2009-2013	46	20,458,082	0.22 (0.17-0.28)
ES, Cataluña	2007-2013	240	34,861,809	0.69 (0.50-0.78)
Sweden	2003-2013	1536	102,027,209	1.52 (1.43-1.59)
The Netherlands (validated)	2003-2013	14	2,879,712	0.49 (0.29-0.76)
North America				
CA, British Columbia	2003-2013	278	47,857,684	0.58 (0.32-0.64)
CA, Alberta	2003-2013	427	51,885,946	0.82 (0.74-0.90)
CA, Manitoba	2003-2010	42	6,335,257	0.66 (0.50-0.86)
Asia				
Taiwan	2003-2012	472	161,407,503	0.29 (0.27-0.32)

Abbreviations: IR (Incidence Rate), EU (European Union), ES (Spain), CA (Canada)

Due to very low rates observed among the very young (<5 years) and the elderly (≥60 years) in age-stratified analysis and known differences in diagnosis which precluded collapsing of strata, these age groups were not included in further stratified analyses. In Figure 1, IRs are shown stratified by age group and time period. In investigation of age group and time period and the interaction of these factors, IRRs were significantly elevated in both age groups in Taiwan (where MF59-adjuvanted vaccine coverage was 59% for those <19 years and 11% for those >19 years) in the period during circulation of wild-type virus prior to vaccination. For those aged 5-19 years, the IRR was 2.50 (95%CI 1.46, 4.28), and for those aged 20-59 years, the IRR was 2.23 (95%CI 1.26, 3.94). This continued in the period after the vaccination campaign had started, with IRR 1.60 (95%CI 1.20, 2.13) for those aged 5-19 years and IRR 2.13 (95%CI 1.62, 2.79) for those aged 20-59 years (Table 3). In Sweden, where AS03-adjuvanted Pandemrix vaccine coverage was 60%, in the period after vaccination incidence rates among those aged 5-19 years and 20-59 years were elevated [IRR=9.01 (95%CI 6.89, 11.80) and IRR=1.69 (95%CI 1.46, 1.95)], respectively (Figure 1, Table 3). None of the other countries showed significant time-period-by-age-group interactions.

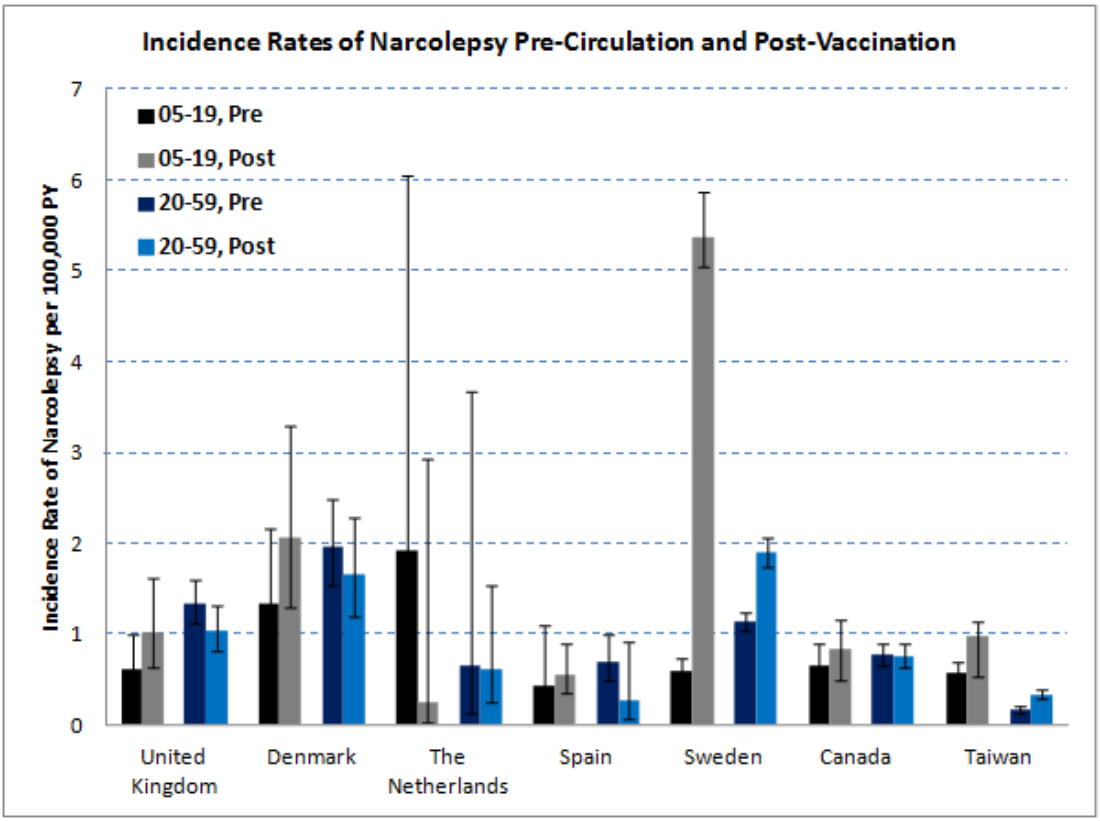


Figure 1. Incidence rates of narcolepsy pre-circulation and post-vaccination in children (05-19 yrs) and adults (20-59 yrs)

In the analysis restricted to Europe and including a vaccine coverage/adjuvant composite variable, IRRs were elevated in the period following start of vaccination in the high-coverage AS03 (Sweden) and low-coverage AS03 groups for children and adolescents (Supplementary Table 3). In adults, an elevated incidence in the period following vaccination was detected in the AS03 high-coverage group, which was limited to Sweden. In this analysis, no changes in the incidence of narcolepsy in the post-vaccination period were seen in sites using MF59-adjuvanted vaccine, all of which had low coverage (Supplementary Table 3).

Table 3: IRs and IRRs by country, age and period

Site	Age	Period*	Cases	Person years	IR	IRR†	95%CI
United Kingdom	5-19	Pre-Circulation	26	4247239	0.61	Ref	--
		Circulation	0	229303	0.00	NA	--
		Vaccination & Post	28	2752486	1.02	1.66	0.97, 2.83
	20-59	Pre-Circulation	183	13782669	1.33	Ref	--
		Circulation	9	744620	1.21	0.91	0.47, 1.78
		Vaccination & Post	90	8706262	1.03	0.78	0.61, 1.00
Denmark	5-19	Pre-Circulation	26	1941950	1.34	Ref	--
		Circulation	3	160562	1.87	1.40	0.42, 4.61
		Vaccination & Post	28	1352428	2.07	1.55	0.91, 2.64
	20-59	Pre-Circulation	103	5258884	1.96	Ref	--
		Circulation	8	416864	1.92	0.98	0.48, 2.01

		Vaccination & Post	58	3492758	1.66	0.85	0.62, 1.17
The Netherlands	5-19	Pre-Circulation	2	103950	1.92	Ref	
		Circulation	0	29453	0.00	NA	
		Vaccination & Post	1	394895	0.25	0.13	0.01, 1.45
	20-59	Pre-Circulation	2	306773	0.65	Ref	
		Circulation	1	87315	1.15	1.76	0.16, 19.37
		Vaccination & Post	7	1144346	0.61	0.94	0.20, 4.52
Spain	5-19	Pre-Circulation	7	1617473	0.43	Ref	
		Circulation	4	1488771	0.27	0.62	0.18, 1.13
		Vaccination & Post	26	4715178	0.55	1.27	0.55, 2.94
	20-59	Pre-Circulation	48	6847254	0.70	Ref	--
		Circulation	33	610444	0.54	0.77	0.50, 1.20
		Vaccination & Post	125	18915104	0.27	0.94	0.68, 1.31
Sweden	5-19	Pre-Circulation	62	10381883	0.60	Ref	--
		Circulation	1	819877	0.12	0.20	0.03, 1.47
		Vaccination & Post	369	6854603	5.38	9.01	6.89, 11.80
	20-59	Pre-Circulation	338	29823712	1.13	Ref	--
		Circulation	26	2418238	1.08	0.95	0.64, 1.41
		Vaccination & Post	401	20992445	1.91	1.69	1.46, 1.95
Canada	5-19	Pre-Circulation	67	10107116	0.66	Ref	--
		Circulation	6	1261204	0.48	0.72	0.31, 1.70
		Vaccination & Post	53	6378494	0.83	1.25	0.87, 1.80
	20-59	Pre-Circulation	265	34413993	0.77	Ref	--
		Circulation	36	4574717	0.79	1.02	0.72, 1.45
		Vaccination & Post	182	24228401	0.75	0.98	0.81, 1.18
Taiwan	5-19	Pre-Circulation	81	13985353	0.58	Ref	--
		Circulation	16	1103680	1.45	2.50	1.46, 4.28
		Vaccination & Post	110	11867183	0.93	1.60	1.20, 2.13
	20-59	Pre-Circulation	78	46806947	0.17	Ref	--
		Circulation	14	3768896	0.37	2.23	1.26, 3.94
		Vaccination & Post	158	44542437	0.35	2.13	1.62, 2.79

*Periods are as follows: Pre-Circulation = January 2003-the beginning of wild-type H1N1 circulation (defined per country); Circulation = Period from the beginning of wild-type H1N1 circulation until the start of the vaccination campaign (defined per country); Vaccination & Post = Period from the beginning of the vaccination campaign through December 2013.

† IRR comparing the period to the pre-circulation period, within the age group

Simulation

The simulation study showed that in an analysis such as the one described above, the true RR is consistently underestimated when it is greater than one and overestimated when less than one (Figure 2). Underestimation of true relative risks greater than one is attenuated as vaccination coverage increases but remains about 6%-26% underestimated even with vaccination coverage as high as 99% with no delay from onset to diagnosis and a 10 year observation period(Figure 3). As the interval from onset to diagnosis decreases, the estimated relative risk approaches the true relative risk; performance is improved if the observation time captures only the period of increased risk (data not shown). When the time from onset to diagnosis was set equal to empirical estimates (4 years for adults and at 1.5 years for children), no increased risk was detected for any set of simulation parameters

(data not shown). Stratification by age group was effective in elucidating the group that was the source of the increased risk.

Calibration

When using a model derived from the results of the simulation study to predict which should have been the true (calibrated) IRR to produce the estimates found in the SOMNIA study, these true IRRs were always considerably higher than the estimates. Notably, according to the model, the required individual-level relative risk underlying a 9-fold IRR which was found in 5-19 year olds in Sweden should have been 36.04 (95% CI: 27.79, 46.90) (table 4). .

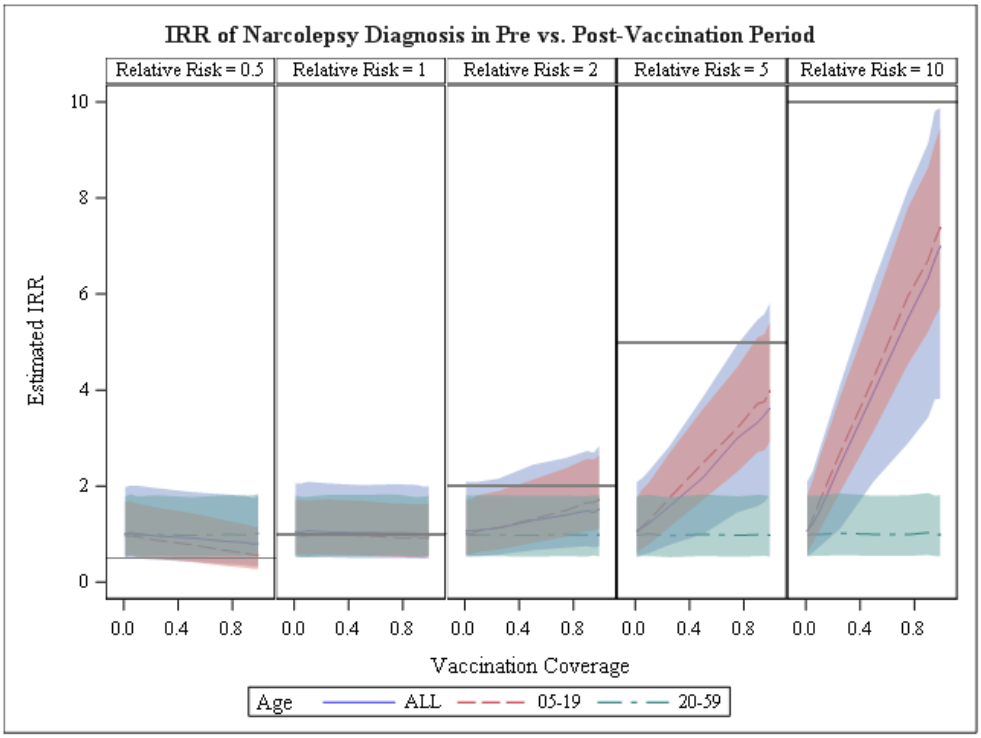


Figure 2: IRR estimates in simulated data, immediate diagnosis (scale = 0). Population-level incidence rate ratio estimated from simulated data with observation time equal to 3625 days and true individual-level relative risk equal to .05, 1, 2, 5, or 10 (columns). Gray horizontal reference lines represent the true simulated individual-level relative risk of narcolepsy diagnosis. The scale parameter is set equal to zero (meaning immediate diagnosis following onset of symptoms). Vaccination coverage increases within each column along the x-axis. Colored lines represent age group-specific IRRs as noted in the legend and colored bands represent associated 95% confidence intervals.

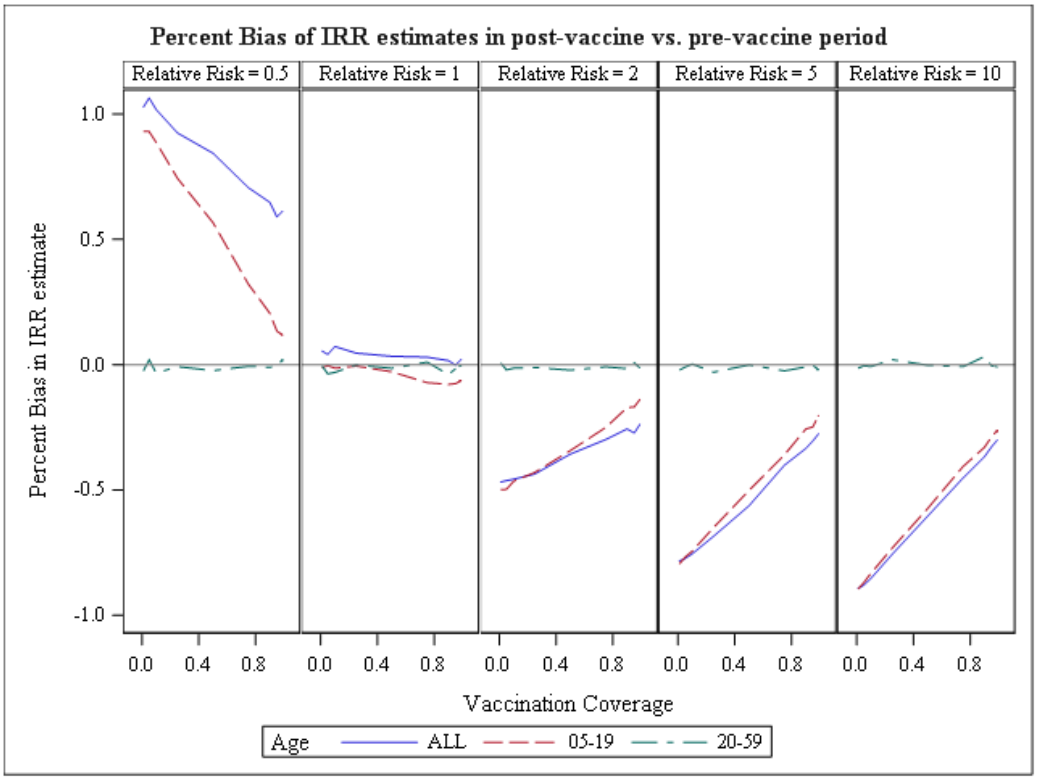


Figure 3: Percent Bias of IRR estimates in simulated data, immediate diagnosis (scale = 0). Population-level percent bias estimated from simulated data with observation time equal to 3625 days and true individual-level relative risk equal to .05, 1, 2, 5, or 10 (columns). Gray horizontal reference lines represent the absence of bias (the estimated IRR = the simulated RR). The scale parameter is set equal to zero (meaning immediate diagnosis following onset of symptoms). Vaccination coverage increases within each column along the x-axis. Colored lines represent age group-specific bias as noted in the legend.

Table 4: IRRs calibrated using model derived from simulated data

Country	Age Group	Observed Ecological IRR	95% CI	Gamma Scale	Vaccine Coverage	Simulation-predicted individual-level RR	95% CI
UK	5-19	1.66	0.97, 2.83	1	.05	7.76	4.95, 12.53
Denmark	5-19	1.55	0.91, 2.64	1	.05	7.31	4.71, 11.76
Netherlands	5-19	0.13	0.01, 1.45	1	.1	1.64	1.16, 6.84
Spain	5-19	1.27	0.55, 2.94	1	.05	6.17	3.24, 12.98
Sweden	5-19	9.01	6.89, 11.80	0.5	.12	36.04	27.79, 46.90
Canada	5-19	1.25	0.87, 1.80	1	.35	5.87	4.62, 7.68
Taiwan	5-19	1.60	1.20, 2.13	0	.65	2.75	1.75, 4.07

Discussion

Evaluation of incidence rates on a population level can be done relatively quickly in countries/regions with accessible population-based electronic health care databases. This is useful for assessing potential vaccine safety signals. In order to calculate rates quickly in a standardized manner, harmonization of data into simple input files in

a common format allowed for the pooling and sharing of data across three continents. The method was capable of identifying the signal in Sweden in 5- to 19-year olds.

In the analysis by country, elevated rates of narcolepsy were only detected in Taiwan during wild-type virus circulation through the period following vaccination with MF59-adjuvanted and non-adjuvanted vaccines, and in Sweden following vaccination with AS03-adjuvanted vaccines. The finding in Taiwan may be due to circulation of wild-type influenza virus prior to the start of the vaccination campaign [17]. This is consistent with the finding of a 3-fold increase in narcolepsy onset in China following the peak of the pH1N1 pandemic in a population with very low vaccine coverage [4]. Taiwan vaccinated children aged <1 year with MF59-adjuvanted vaccine and adults and school children with mainly non-adjuvanted vaccine. In Sweden, where the signal of a narcolepsy safety concern was originally detected [23] and where patients diagnosed with narcolepsy are being compensated [7], rates were much higher than in the other countries. This could be due to differential reporting due to increased awareness of the putative association, a true causal effect in this population with this vaccine, or some combination of these factors. As shown in simulations, reduction in the time from onset to diagnosis due to awareness of an association can lead to artificial inflations in risk estimates [8]. In Canadian provinces, with around 40% vaccine coverage of a different AS03-adjuvanted vaccine (Arepanrix^{TM}), no effect was seen in any of the age groups or periods.This study, which is by necessity observational, has several limitations. Data were collected according to a shared protocol but using locally derived algorithms, which may have led to differences in sensitivity and specificity. Case validation in some sites revealed low specificity of the original extraction, which may be the case in other sites as well. In our analysis by adjuvant and vaccine coverage, high coverage with AS03-containing vaccine was only present in Sweden, making Sweden and this adjuvant/coverage group collinear. This makes it impossible to determine whether we are seeing the effects of the vaccine itself or of the reporting and detection patterns in each country. Additionally, the manufacturing process of Arepanrix^{TM} differed from that of Pandemrix^{TM}, leading to vaccines containing different quantities of influenza virus components[24]. The potential effects of these differences in manufacturing cannot be differentiated from adjuvant specific-effects or from other country-specific effects using an ecological design such as the one presented here. Similarly, the countries in which MF59 was used were the same countries in which case validation was conducted. This limits comparability between these countries and others and, therefore, between MF59-containing vaccines and other pandemic vaccines.

Differences in case ascertainment could also have impacted our estimates. For example, it has been noted elsewhere that the safety signal originated in Sweden and that this, together with compensation for cases, may have impacted diagnosis patterns [6, 7]. Additionally, due to the healthcare system in Taiwan, children complaining of excessive daytime sleepiness are seen by a specialist quickly, making the interval from onset to diagnosis for these children shorter. A median time from symptom onset to MSLT referral of 60 days has been reported for pediatric narcolepsy cases in Taiwan[25]. While it is not possible to rule out a causal association, it is important to note that these factors undoubtedly contributed to the estimates obtained in this study. Differences in the prevalence of the underlying risk allele for narcolepsy, which has been reported to vary widely by country, may affect the incidence at the population level but is unlikely to have affected relative risk estimates [26].

Ecological methods, when applied to assessment of a signal association with a targeted vaccination campaign and a disease with a potentially long interval from onset to diagnosis, can provide an unbiased estimate of vaccine-associated risk in a very limited set of circumstances. Obtaining an unbiased estimate in the absence of an association is possible even with very low vaccine coverage and a long onset to diagnosis interval. However, in the presence of a true vaccine-associated risk, all estimates will be biased toward one; this bias is reduced when cases are detected quickly and vaccination coverage is high. Based upon simulations, the estimates we obtained in the current study appear to be underestimations of true relative risks greater than one. However, the simulations did

not take into account the possibility of increased reporting due to an awareness of the association, which has been shown in previous simulations to inflate risk estimates [8].

The predicted underlying individual-level relative risk obtained using models derived from simulated data, given the low vaccine coverage attained in most sites, are remarkably high. It is very unlikely that the true relative risk in Sweden, for example, is 36-fold. Calibrated estimates for The United Kingdom, however, are only slightly lower than those reported by Miller et al in their study of narcolepsy following Pandemrix™ vaccination in children aged 4-18 [27]. In general, these predicted relative risks are not in line with results found in the case-control study conducted within SOMNIA, in which no increased risk following pH1N1 vaccines was detected [28]. It is important to note that while the case-control portion of the SOMNIA study did not include any pediatric cases exposed to Pandemrix™, the case-coverage sub-study of pediatric cases in The Netherlands did not find an association with Pandemrix™ [28]. These inconsistencies may be an illustration of the ecological fallacy, namely that associations detected at the population level may not be causal at the individual level [29]. In fact, as coverage in our simulations approached 100%, our population-level analysis also approached an individual-level analysis with accurate exposure data for all subject, explaining why increased vaccine coverage in simulations leads to more accurate estimates of the simulated relative risk.

Previous simulations have shown that reduction in the time from onset to diagnosis following awareness of an association increase risk estimates [8]. Our simulation did not take into account factors that may have changed over the course of the study period such as awareness of narcolepsy and of the pH1N1-narcolepsy association as well as changes in diagnostic and coding practices. Each of these likely contributed to the IRR estimates we obtained. What our simulations do show is that in the absence of factors that increase case detection in the post-exposure period, detection of increased population-level risk of a disease with a long onset to diagnosis interval using ecological methods requires an extreme underlying individual-level risk.

The ecological approach fails to detect any increased risk unless the time from onset to diagnosis is short and both coverage and the true relative risk are high. Because of this, we recommend that population-level methods be used in assessment of outcomes with a delay from onset to diagnosis only to generate hypotheses or to strengthen signals when population-level exposure is high. Analysis of the full population when increased risk is only present in one age stratum performs as well as stratified analysis in terms of the magnitude of risk detected, but fails to identify the source of the increased risk. Therefore, if increased risk is suspected in a subset of the population, analyses should be stratified.

Conclusions

Ecological methods can be useful in assessment of vaccine safety but it is important for investigators to understand the impacts of masking by strata not at risk, patterns of onset and diagnosis, and vaccine coverage. What appears to be an estimate of no effect could be valid or, as shown in our simulations, could be an underestimation.

Acknowledgements

The authors would like to thank Tom Shimabukuro and Frank DeStefano for oversight and facilitation of the SOMNIA project.

References

1. Nohynek H, Jokinen J, Partinen M, Vaarala O, Kirjavainen T, Sundman J, et al. AS03 adjuvanted AH1N1 vaccine associated with an abrupt increase in the incidence of childhood narcolepsy in Finland. PloS one. 2012;7(3):e33536. Epub 2012/04/04. doi: 10.1371/journal.pone.0033536. PubMed PMID: 22470453; PubMed Central PMCID: PMC3314666.

2. Wijnans L, Lecomte C, de Vries C, Weibel D, Sammon C, Hviid A, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine. 2013;31(8):1246-54. Epub 2012/12/19. doi: 10.1016/j.vaccine.2012.12.015. PubMed PMID: 23246544.

3. Control ECfDPa. Narcolepsy in association with pandemic influenza vaccination (a multi-country European epidemiological investigation). Stockholm2012.

4. Han F, Lin L, Warby SC, Faraco J, Li J, Dong SX, et al. Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. Annals of neurology. 2011;70(3):410-7. Epub 2011/08/26. doi: 10.1002/ana.22587. PubMed PMID: 21866560.

5. Bollaerts K, Shinde V, Dos Santos G, Ferreira G, Bauchau V, Cohet C, et al. Application of Probabilistic Multiple-Bias Analyses to a Cohort- and a Case-Control Study on the Association between Pandemrix and Narcolepsy. PloS one. 2016;11(2):e0149289. Epub 2016/02/24. doi: 10.1371/journal.pone.0149289. PubMed PMID: 26901063; PubMed Central PMCID: PMC4762678.

6. Verstraeten T, Cohet C, Dos Santos G, Ferreira GL, Bollaerts K, Bauchau V, et al. Pandemrix and narcolepsy: A critical appraisal of the observational studies. Human vaccines & immunotherapeutics. 2016;12(1):187-93. Epub 2015/09/18. doi: 10.1080/21645515.2015.1068486. PubMed PMID: 26379011; PubMed Central PMCID: PMC4962758.

7. Sturkenboom MC. The narcolepsy-pandemic influenza story: can the truth ever be unraveled? Vaccine. 2015;33 Suppl 2:B6-B13. Epub 2015/05/30. doi: 10.1016/j.vaccine.2015.03.026. PubMed PMID: 26022571.

8. Wijnans L, Dodd C, de Ridder M, Romio S, Weibel D, Overeem S, et al. Pandemic influenza vaccine & narcolepsy: simulations on the potential impact of bias. Expert review of vaccines. 2016;15(5):573-84. Epub 2016/03/12. doi: 10.1586/14760584.2016.1164045. PubMed PMID: 26967200.

9. Ahmed SS, Schur PH, MacDonald NE, Steinman L. Narcolepsy, 2009 A(H1N1) pandemic influenza, and pandemic influenza vaccinations: what is known and unknown about the neurological disorder, the role for autoimmunity, and vaccine adjuvants. Journal of autoimmunity. 2014;50:1-11. Epub 2014/02/25. doi: 10.1016/j.jaut.2014.01.033. PubMed PMID: 24559657.

10. Ahmed SS, Volkmuth W, Duca J, Corti L, Pallaoro M, Pezzicoli A, et al. Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. Science translational medicine. 2015;7(294):294ra105. Epub 2015/07/03. doi: 10.1126/scitranslmed.aab2354. PubMed PMID: 26136476.

11. Taddei RN, Werth E, Poryazova R, Baumann CR, Valko PO. Diagnostic delay in narcolepsy type 1: combining the patients' and the doctors' perspectives. Journal of sleep research. 2016;25(6):709-15. Epub 2016/05/07. doi: 10.1111/jsr.12420. PubMed PMID: 27149919.

12. Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods. Vaccine. 2004;22(15-16):2064-70. Epub 2004/05/04. doi: 10.1016/j.vaccine.2004.01.017. PubMed PMID: 15121324.

13. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Annual review of public health. 1995;16:61-81. Epub 1995/01/01. doi: 10.1146/annurev.pu.16.050195.000425. PubMed PMID: 7639884.

14. Shields WD, Nielsen C, Buch D, Jacobsen V, Christenson P, Zachau-Christiansen B, et al. Relationship of pertussis immunization to the onset of neurologic disorders: a retrospective epidemiologic study. The Journal of pediatrics. 1988;113(5):801-5. Epub 1988/11/01. PubMed PMID: 3263484.

15. da Silveira CM, Salisbury DM, de Quadros CA. Measles vaccination and Guillain-Barre syndrome. Lancet. 1997;349(9044):14-6. Epub 1997/01/04. PubMed PMID: 8988116.

16. Hsing AW, Ioannidis JP. Nationwide Population Science: Lessons From the Taiwan National Health Insurance Research Database. JAMA internal medicine. 2015;175(9):1527-9. Epub 2015/07/21. doi: 10.1001/jamainternmed.2015.3540. PubMed PMID: 26192815.

17. Chuang JH, Huang AS, Huang WT, Liu MT, Chou JH, Chang FY, et al. Nationwide surveillance of influenza during the pandemic (2009-10) and post-pandemic (2010-11) periods in Taiwan. PloS one. 2012;7(4):e36120. Epub 2012/05/01. doi: 10.1371/journal.pone.0036120. PubMed PMID: 22545158; PubMed Central PMCID: PMC3335813.

18. Huang WT, Chen WW, Yang HW, Chen WC, Chao YN, Huang YW, et al. Design of a robust infrastructure to monitor the safety of the pandemic A(H1N1) 2009 vaccination program in Taiwan. Vaccine. 2010;28(44):7161-6. Epub 2010/09/02. doi: 10.1016/j.vaccine.2010.08.069. PubMed PMID: 20804804.

19. Trifiro G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? Journal of internal medicine. 2014;275(6):551-61. Epub 2014/03/19. doi: 10.1111/joim.12159. PubMed PMID: 24635221.

20. Hood BM, Harbord MG. Paediatric narcolepsy: complexities of diagnosis. Journal of paediatrics and child health. 2002;38(6):618-21. Epub 2002/11/02. PubMed PMID: 12410881.

21. Chakravorty SS, Rye DB. Narcolepsy in the older adult: epidemiology, diagnosis and management. Drugs & aging. 2003;20(5):361-76. Epub 2003/04/17. PubMed PMID: 12696996.

22. Silber MH, Krahn LE, Olson EJ, Pankratz VS. The epidemiology of narcolepsy in Olmsted County, Minnesota: a population-based study. Sleep. 2002;25(2):197-202. Epub 2002/03/21. PubMed PMID: 11902429.

23. MP. S. Occurrence of narcolepsy with cataplexy among children and adolescents in relation to the H1N1 pandemic and Pandemrix vaccinations–results of a case inventory study by the MPA in Sweden during 2009–2010. 2011.

24. Vaarala O, Vuorela A, Partinen M, Baumann M, Freitag TL, Meri S, et al. Antigenic differences between AS03 adjuvanted influenza A (H1N1) pandemic vaccines: implications for pandemrix-associated narcolepsy risk. PloS one. 2014;9(12):e114361. Epub 2014/12/17. doi: 10.1371/journal.pone.0114361. PubMed PMID: 25501681; PubMed Central PMCID: PMC4266499.

25. Huang W-TH, Chung-Yao; Chen, Hsi-Chung; Lee, Hsin-Chun; Lin, Hui-Chen; Hsieh, Cheng-Fang; Wu, Meng-Ni; Yang, Chin-Hui. Case-control Analysis of Narcolepsy after 2009 H1N1 Pandemic Vaccination in Taiwan. ICPE; 25 January 2017; Montreal, Canada2017.

26. Ahmed SS, Steinman L. Mechanistic insights into influenza vaccine-associated narcolepsy. Human vaccines & immunotherapeutics. 2016;12(12):3196-201. Epub 2016/04/01. doi: 10.1080/21645515.2016.1171439. PubMed PMID: 27031682; PubMed Central PMCID: PMC5215445.

27. Miller E, Andrews N, Stellitano L, Stowe J, Winstone AM, Shneerson J, et al. Risk of narcolepsy in children and young people receiving AS03 adjuvanted pandemic A/H1N1 2009 influenza vaccine: retrospective analysis. BMJ. 2013;346:f794. Epub 2013/02/28. doi: 10.1136/bmj.f794. PubMed PMID: 23444425.

28. Steven Black MS, Maria de Ridder, Caitlin Dodd, Daniel Weibel, Jan Bonhoeffer, Tom Shimabukuro, Frank Destefano, Angela Gentil, Norberto Giglio. The SOMNIA Study: Assessment of the Risk of Narcolepsy following Adjuvanted 2009 H1N1 Pandemic Vaccines IDWeek; October 27, 2016; New Orleans, LA2016.

29. Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. International journal of epidemiology. 1989;18(1):269-74. Epub 1989/03/01. PubMed PMID: 2656561.

Supplementary Material

Supplementary Table 1: Characteristics of the Databases in this study

Data site & source	Type of data*	Algorithm used	H1N1 virus circulation (weeks)	Vaccination Coverage by age group	Adjuvant Used by age group
UK THIN[1]	Linked Medical Records	Read codes F27.00, F270.00, F271.00, F27z.00	2009: 26-52	6 mo-5yr (20%) 5-18 yr (4%) >65yr (35%)	05-19: AS03 (Pandemrix) 20-59: AS03 (Pandemrix)
NL IPCI[2]	Linked Medical Records	Free text narcolepsy & MSLT. Followed by manual review	2009: 30-50	< 5yr (75%) Risk groups (70%)	05-19: MF59 20-59: MF59
DK AARHUS[3]	Population -based registry	ICD-10 code G47.4 (primary and secondary) diagnosis	2009: 29-45	>18yr (20%)	05-19: AS03 (Pandemrix) 20-59: AS03
SIDIAP (Spain, Catalunya)[4]	Linked Medical Records	ICD-10 code G47.4 diagnosis	2009: 31-50	< 18yr (1%) all population (3.5%)	05-19: MF59 20-59: AS03 (Pandemrix)
FISABIO (Spain, Valencia)[5]	Linked Medical Records	ICD-9CM codes 347.* with Manual validation	2009: 31-50	6mo-14yr, risk groups (11%) 15-59yr, risk groups (13%) >60yr, risk groups (28%) Pregnant women (9%) Healthcare workers (30%)	05-19: MF59 20-59: AS03 (Pandemrix)
Sweden[6]	Medical Record diagnoses + Census Population	ICD-10 code G47.4 diagnosis	2009: 30-50	< 18 yr (12%) ≥ 18 yr (13%)	05-19: AS03 (Pandemrix) 20-59: AS03 (Pandemrix)
Taiwan[7]	Linked Medical Records	ICD9-CM codes 347.* with MSLT procedure	2009: 30-52	6mo-18yr (59%) ≥ 19yr (11%)	05-19: MF59 20-59: MF59
Canada, Alberta[8]	Linked Medical Records	ICD9-CM codes 347.* with MSLT procedure	2009: 19-27	≥12 yr (37%)	05-19: AS03 (Arepanrix) 20-59: AS03

Canada, Manitoba[9]	Linked Medical Records	ICD9-CM codes 347.* with MSLT procedure	2009: 19-27	≥12 yr (37%)	05-19: AS03 (Arepanrix) 20-59: AS03
Canada, British Columbia[10]	Medical Record diagnoses + Census Population	ICD9-CM codes 347.* with MSLT procedure	2009: 19-27	<10yr (46%) 10-18yr (32%) 19-39yr (33%) 40-64yr (45%) >64 (58%)	05-19: AS03 (Arepanrix) 20-59: AS03

* Linked Medical Records = Population based medical records (GP and specialist diagnoses), directly linked; Population-based registry = Population based registries (emergency room, in and out patient diagnoses); Medical Record diagnoses + Census Population = In and outpatient diagnoses, case counts and population counts (census);

1. http://www.epic-uk.org/our-data/our-data.shtml
2. http://www.erasmusmc.nl/med_informatica/research/555688/?lang=en#
3. http://www.kea.au.dk/en/ResearchRegistries.html
4. http://www.sidiap.org/index.php/en
5. http://fisabio.san.gva.es/en/fisabio;jsessionid=AFE38E9ACF0A380A692A9739E88F2FF4
6. http://www.socialstyrelsen.se/english
7. http://www.mohw.gov.tw/CHT/DOS/DM1.aspx?f_list_no=812 (Chinese)
8. http://www.health.alberta.ca/documents/Research-Health-Datasets.pdf
9. http://umanitoba.ca/faculties/health_sciences/medicine/units/chs/departmental_units/mchp/resources/repository/index.html
10. https://www.popdata.bc.ca/data

Supplementary Table 2: incidence rates by continent, country, age and period

Continent	Site	Age	Period*	Cases	Person years	IR	IRR†	95%CI
Europe	United Kingdom	0-4	Pre-Circulation	0	1,438,688	0.00	Ref	--
			Circulation	0	82033	0.00	NA	--
			Vaccination & Post	1	986145	0.10	NA	--
		5-19	Pre-Circulation	26	4247239	0.61	Ref	--
			Circulation	0	229303	0.00	NA	--
			Vaccination & Post	28	2752486	1.02	1.66	0.97, 2.83
		20-59	Pre-Circulation	183	13782669	1.33	Ref	--
			Circulation	9	744620	1.21	0.91	0.47, 1.78
			Vaccination & Post	90	8706262	1.03	0.78	0.61, 1.00
		60+	Pre-Circulation	84	5738919	1.46	Ref	--
			Circulation	8	331682	2.41	1.65	0.80, 3.40
			Vaccination & Post	38	3857675	0.99	0.67	0.46, 0.99
	Denmark	0-4	Pre-Circulation	0	625074	0.00	Ref	--
			Circulation	0	51527	0.00	NA	--
			Vaccination & Post	2	426447	0.47	NA	--
		5-19	Pre-Circulation	26	1941950	1.34	Ref	--
			Circulation	3	160562	1.87	1.40	0.42, 4.61
			Vaccination & Post	28	1352428	2.07	1.55	0.91, 2.64
		20-59	Pre-Circulation	103	5258884	1.96	Ref	--
			Circulation	8	416864	1.92	0.98	0.48, 2.01
			Vaccination & Post	58	3492758	1.66	0.85	0.62, 1.17

		60+	Pre-Circulation	25	2189220	1.14	Ref	--
			Circulation	3	195788	1.53	1.34	0.41, 4.44
			Vaccination & Post	13	1738626	0.75	0.66	0.34, 1.28
	The Netherlands	0-4	Pre-Circulation	0	41272	0.00	Ref	--
			Circulation	0	10442	0.00	NA	--
			Vaccination & Post	0	129291	0.00	NA	--
		5-19	Pre-Circulation	2	103950	1.92	Ref	
			Circulation	0	29453	0.00	NA	
			Vaccination & Post	1	394895	0.25	0.13	0.01, 1.45
		20-59	Pre-Circulation	2	306773	0.65	Ref	
			Circulation	1	87315	1.15	1.76	0.16, 19.37
			Vaccination & Post	7	1144346	0.61	0.94	0.20, 4.52
		60+	Pre-Circulation	0	111122	0.00	Ref	--
			Circulation	0	35025	0.00	NA	--
			Vaccination & Post	1	488348	0.20	NA	--
	Spain	0-4	Pre-Circulation	0	755198	0.00	Ref	--
			Circulation	0	601869	0.00	NA	--
			Vaccination & Post	0	1647061	0.00	NA	--
		5-19	Pre-Circulation	7	1617473	0.43	Ref	
			Circulation	4	1488771	0.27	0.62	0.18, 1.13
			Vaccination & Post	26	4715178	0.55	1.27	0.55, 2.94
		20-59	Pre-Circulation	48	6847254	0.70	Ref	--
			Circulation	33	610444	0.54	0.77	0.50, 1.20
			Vaccination & Post	125	18915104	0.27	0.94	0.68, 1.31
		60+	Pre-Circulation	7	2728360	0.26	Ref	--
			Circulation	11	2372937	0.46	1.82	0.70, 4.76
			Vaccination & Post	25	7526241	0.33	1.29	0.56, 2.99
	Sweden	0-4	Pre-Circulation	13	3166715	0.41	Ref	--
			Circulation	4	274111	1.46	3.56	1.16, 10.90
			Vaccination & Post	7	2418815	0.29	0.71	0.28, 1.77
		5-19	Pre-Circulation	62	10381883	0.60	Ref	--
			Circulation	1	819877	0.12	0.20	0.03, 1.47
			Vaccination & Post	369	6854603	5.38	9.01	6.89, 11.80
		20-59	Pre-Circulation	338	29823712	1.13	Ref	--
			Circulation	26	2418238	1.08	0.95	0.64, 1.41
			Vaccination & Post	401	20992445	1.91	1.69	1.46, 1.95
		60+	Pre-Circulation	200	13550206	1.48	Ref	--
			Circulation	12	1158116	1.04	0.70	0.39, 1.26
			Vaccination & Post	103	10168490	1.01	0.69	0.54, 0.87
North America	Canada	0-4	Pre-Circulation	1	2669395	0.04	Ref	--
			Circulation	0	346294	0.00	NA	--
			Vaccination & Post	3	2021834	0.15	3.96	0.41, 38.08
		5-19	Pre-Circulation	67	10107116	0.66	Ref	--
			Circulation	6	1261204	0.48	0.72	0.31, 1.70
			Vaccination & Post	53	6378494	0.83	1.25	0.87, 1.80
		20-59	Pre-Circulation	265	34413993	0.77	Ref	--
			Circulation	36	4574717	0.79	1.02	0.72, 1.45
			Vaccination & Post	182	24228401	0.75	0.98	0.81, 1.18
		60+	Pre-Circulation	77	10259739	0.75	Ref	--
			Circulation	8	1481882	0.54	0.72	0.35, 1.49
			Vaccination & Post	49	8335817	0.59	0.78	0.55, 1.12
Asia	Taiwan	0-4	Pre-Circulation	0	3647009	0.00	Ref	--
			Circulation	0	256532	0.00	NA	--
			Vaccination & Post	0	2526902	0.00	NA	--
		5-19	Pre-Circulation	81	13985353	0.58	Ref	--
			Circulation	16	1103680	1.45	2.50	1.46, 4.28

			Vaccination & Post	110	11867183	0.93	1.60	1.20, 2.13
		20-59	Pre-Circulation	78	46806947	0.17	Ref	--
			Circulation	14	3768896	0.37	2.23	1.26, 3.94
			Vaccination & Post	158	44542437	0.35	2.13	1.62, 2.79
		60+	Pre-Circulation	8	17512300	0.05	Ref	--
			Circulation	‡	‡	0.08	1.72	0.22, 13.76
			Vaccination & Post	‡	‡	0.04	0.93	0.32, 2.68

*Periods are as follows: Pre-Circulation = January 2003-the beginning of wild-type H1N1 circulation (defined per country); Circulation = Period from the beginning of wild-type H1N1 circulation until the start of the vaccination campaign (defined per country); Vaccination & Post = Period from the beginning of the vaccination campaign through December 2013.

† IRR comparing the period to the pre-circulation period, within the age group

‡The count has been suppressed either because (1) the observed number of events is very small (n ≤ 2) and not appropriate for publication; or (2) it could be used to calculate the number in a cell that has been suppressed.

Supplementary Table 3: IRRs Post-Vaccination vs. Pre-Circulation in categories of Coverage and Adjuvant

Age Group	Adjuvant	Coverage	Countries	IRR	95% CI
Children (5-19 years)	MF59	Low	Netherlands Spain	1.01	0.45, 2.25
	AS03	Low	Denmark United Kingdom	1.62	1.09, 2.42
	AS03	High	Sweden	9.01	6.78, 11.99
Adults (20-59 years)	MF59	Low	Netherlands	0.94	0.18, 4.97
	AS03	Low	Spain Denmark United Kingdom	0.68	0.57, 0.81
	AS03	High	Sweden	1.69	1.45, 1.97

CHAPTER 4. SINGLE DATABASE STUDIES FOR THE ASSESSMENT OF VACCINE EFFECTS

4.1. BELL'S PALSY AND INFLUENZA(H1N1)PDM09 CONTAINING VACCINES: A SELF-CONTROLLED CASE SERIES

Leonoor Wijnans

Caitlin Dodd

Daniel Weibel

Miriam C J M Sturkenboom

PloS one 12.5 (2017): e0175539.

Abstract

Background

An association between AS03 adjuvanted pandemic influenza vaccine and the occurrence of Bell's palsy was found in a population based cohort study in Stockholm, Sweden. To evaluate this association in a different population, we conducted a self-controlled case series in a primary health care database, THIN, in the United Kingdom. The aim of this study was to determine whether there was an increased risk of Bell's palsy following vaccination with any influenza vaccine containing A/California/7/2009 (H1N1)-like viral strains. Secondly, we investigated whether risks were different following pandemic influenza A(H1N1)pdm09 vaccines and seasonal influenza vaccines containing the influenza A(H1N1)pdm09 strain.

Methods

The study population comprised all incident Bell's palsy cases between 1 June 2009 and 30 June 2013 identified in THIN. We determined the relative incidence (RI) of Bell's palsy during the 6 weeks following vaccination with either pandemic or seasonal influenza vaccine. All analyses were adjusted for seasonality and confounding variables.

Results

We found an incidence rate of Bell's palsy of 38.7 per 100,000 person years. Both acute respiratory infection (ARI) consultations and pregnancy were found to be confounders. When adjusted for seasonality, ARI consultations and pregnancies, the RI during the 42 days after vaccination with an influenza vaccine was 0.85 (95% CI: 0.72– 1.01). The RI was similar during the 42 days following seasonal vaccine (0.96, 95%CI: 0.82-1.13) or pandemic vaccine (0.73, 95%CI: 0.47-1.12).

Conclusion

We found no evidence for an increased incidence of Bell's palsy following seasonal influenza vaccination overall, nor for monovalent pandemic influenza vaccine in 2009.

Introduction

Bell's palsy is an idiopathic peripheral-nerve palsy affecting the cranial nerve and the most common cause for facial paralysis (1) with an incidence between 15 to 50 cases per 100,000 people per year (1-3). It is characterized by acute onset, unilateral facial paralysis, numbness or pain around the ear, reduction in taste and hypersensitivity to sounds. The diagnosis is made after excluding other possible causes for facial paralysis, including congenital, genetic and acquired causes. Standard diagnostic criteria are not available (4). Bell's palsy resolves spontaneously without treatment in most patients within 6 months. Some patients experience long-term sequelae with incomplete return of facial motor function and synkinesis (1). The cause of Bell's palsy is unknown. Inflammation is thought to play an important role in the aetiology of Bell's palsy (1) and an auto-immune aetiology has also been suggested (5). Known risk factors for Bell's palsy include diabetes, a weakened immune system and pregnancy (1, 6).

Bell's palsy has been associated with influenza vaccines in the past (7-12). A large population-based study in the UK did not detect a relationship between inactivated influenza vaccines and Bell's palsy (13), nor did a recent study in the US in children (14). Due to the earlier associations and the unknown aetiology, Bell's palsy remains an adverse event of interest following influenza vaccination.

Following the 2009/2010 influenza A(H1N1) pandemic, an association with Bell's palsy was found with an AS03 adjuvanted pandemic influenza vaccine, Pandemrix, in a population based cohort study in Stockholm, Sweden with a hazard ratio (HR) of 1.25, 95% CI 1.06 to 1.48 (15). The risk was highest during the first 6 weeks following vaccination (HR: 1.60, 1.25 to 2.05) and particularly present in those vaccinated early in the campaign (HR: 1.74, 95% CI 1.16 to 2.59), which were those with more (severe) underlying co-morbidity. Similarly, a signal was detected for monovalent pandemic influenza vaccines used in the Vaccine Safety Datalink (VSD) Project in the US in adults over the age of 25 years with a relative risk of 1.6 (16). This last signal was not verified in a case centred analysis which found and odds ratio of 1.21 (95% CI: 0.93 – 1.57). Finally, a signal of an increased risk of Bell's palsy during the 42 days after vaccination with pandemic (H1N1) 2009 vaccine was detected in Taiwan (17).

In order to evaluate the potential association of Bell's palsy following influenza A(H1N1)pdm09 vaccination in a different population, we conducted a self-controlled case series study. The aim of this study was to determine whether there was an increased risk of Bell's palsy following vaccination with any influenza vaccine containing A/California/7/2009 (H1N1)-like viral strains. Secondly, we looked whether risks were different following pandemic influenza A(H1N1)pdm09 vaccines and seasonal influenza vaccines containing the influenza A(H1N1)pdm09 strain.

Methods

We used a self-controlled case series (SCCS) (18, 19) design in The Health Improvement Network (THIN) database. THIN includes data from 562 general practices across the UK and the population covered by THIN is representative of the UK population. The data in THIN have been validated for pharmacoepidemiology studies (20, 21).

Study population, study period and outcome

The study population comprised all incident Bell's palsy cases between 1 June 2009 and 30 June 2013 identified in THIN, from a total population in this time period of nearly 6 million. A Bell's palsy case was defined as a person who had a consultation with a READ diagnosis code for Bell's palsy (F310.00). Multiple cases per person were allowed. If diagnosis dates were more than 6 months apart, they constituted two separate cases. Considering the relatively high predictive value of over 75% of READ diagnosis codes for Bell's palsy (13) no validation on identified cases was performed.

Exposures

Influenza vaccination was identified through relevant codes and recorded by year and vaccine type (seasonal or pandemic), including pandemic influenza vaccination and seasonal influenza vaccinations for the years 2009/2010, 2010/2011, 2011/2012 and 2012/2013. In the UK both Pandemrix and Celvapan were used during the 2009-2010 influenza A(H1N1) pandemic, and information on brand was retrieved if available. Moreover, during the 2009-2010 season, persons could have received both a seasonal vaccine and a pandemic influenza vaccine. In theory these could have been given on the same day or close together making it difficult to attribute the risk to either. Considering the study by Stowe *et al* (13) no increased risk was expected for the seasonal vaccine, therefore this was disregarded in the primary analysis.

Because each person serves as his or her own control, stable confounders such as gender, genetics, socio-economic status, and underlying disease are controlled for. Covariates that were considered as potential confounders were calendar time, occurrence of acute respiratory infections (ARI), influenza diagnoses, and pregnancy. Considering the short observation period, no age effect was expected. ARI episodes and influenza diagnoses were identified by relevant READ codes (provided in annex table). Consultations for ARI or influenza occurring within 28 days of a previous consultation were excluded as likely related to the same episode. The risk window for ARI and influenza was 0 to 7 days following the date of infection. Pregnancies were identified by the date of delivery. The risk period was the 270 days (9 months) before the date of delivery.

Analysis

We used means and standard deviations to describe continuous variables. For categorical variables, we used counts and percentages. We calculated the incidence rate of codes for Bell's palsy using all person time in the database within the study period and similarly determined vaccination rates per season using all subjects in the database.

All descriptive statistics were compared between vaccinated cases and unvaccinated cases using t-tests for continuous variables and chi-squared tests for categorical variables. Associations between pregnancy, ARI consultations and influenza diagnoses and Bell's palsy and influenza vaccination were determined. We determined the relative incidence (RI) of Bell's palsy during the 6 weeks following vaccination with either pandemic or seasonal influenza vaccine using a conditional Poisson regression conducted on Bell's palsy cases only. The risk period of interest was from days one to 42 following vaccination (D1 to D42), as this was the period with the highest risk found by Bardage et al (15). As vaccination could be delayed following an episode of Bell's palsy, the 14 days prior to vaccination were treated as a separate risk period in the analysis. The day of vaccination (D0) was also regarded as a separate risk period. Relative incidence of Bell's palsy associated with pregnancy, influenza diagnosis, and ARI were estimated using the SCCS method (18). If these univariate associations were significant, pregnancy, influenza, and ARI would be included as additional exposures in the primary analysis. All analyses were adjusted for calendar time by quarter.

Relative incidences were calculated separately for pandemic and for seasonal influenza vaccines, and for each season (vaccination period). All person time was included in analysis of risk following any pH1N1-containing vaccine exposure while person time was limited to October 1 to 61 days following the last administration in separate analyses of pandemic H1N1 vaccine and seasonal vaccines. As less than 0.1% of vaccinated cases received Celvapan during the 2009/2010 pandemic (22), we consider the findings with pandemic vaccines in our study to be applicable to Pandemrix. Age and sex specific relative incidences of Bell's palsy within 6 weeks of influenza vaccination were calculated.

To further account for the risk of deferral of vaccination after receiving a diagnosis of Bell´s palsy, we performed sensitivity analyses in which only the observation time after vaccination was considered.

All analyses were conducted using SAS 9.2.

Results

We identified 6381 Bell's palsy cases in 6288 persons. Of these, 6198 persons had one code for Bell's palsy, 87 persons had two and three persons had three consultations with a Bell's palsy code during the study period. The incidence rate was 38.7 per 100,000 person years (Table 1)

Table 1. Main characteristics of THIN population overall and by case status

	Non-Case n=5,726,368		Case n=6,288		Total n=5,732,656	
Demographics						
Female (n (%))	2,913,751	(50.88)	3,194	(50.80)	2,917,343	(50.88)
Follow Up Time in years (SD)	3.13	(1.58)	3.72	(1.27)	3.13	(1.58)
Mean age in years (SD)	37.10	(23.32)	45.00	(20.20)	37.11	(23.32)
Age (n (%))						
<45 yrs	3,646,538	(63.68)	3,217	(51.16)	3,650,127	(63.67)
45 – 65 yrs	1,289,596	(22.52)	1,949	(31.00)	1,291,825	(22.53)
>65 yrs	790,234	(13.80)	1,122	(17.84)	791,495	(13.80)
≥1 ARI episode (n (%))	1,489,391	(26.01)	2231	(35.48)	1,491,622	(26.02)
≥ 1Pregnancy (n (%))	122,878	(2.15)	158	(2.51)	123,036	(2.15)
2009 Pandemic (n(%))	524059	(9.15)	901	(14.33)	524960	(9.16)
Seasonal 2009-2010 (n(%))	913,250	(15.95)	1660	(26.40)	914,910	(15.96)
Seasonal 2010-2011 (n(%))	789912	(13.79)	1529	(24.32)	791,411	(13.80)
Seasonal 2011-2012 (n(%))	856684	(14.93)	1747	(27.78)	856,684	(14.94)
Seasonal 2012-2013 (n(%))	817751	(14.28)	1702	(27.07)	819,453	(14.29)

The characteristics of the cases by vaccination status (seasonal and pandemic) are presented in table 2. Cases who were vaccinated with either pandemic or seasonal influenza vaccine were older and had more consultations for ARI during the study period. Moreover, cases who received seasonal influenza vaccines were more likely to be female. The distribution of Bell's palsy dates relative to vaccination dates is presented in figure 1.

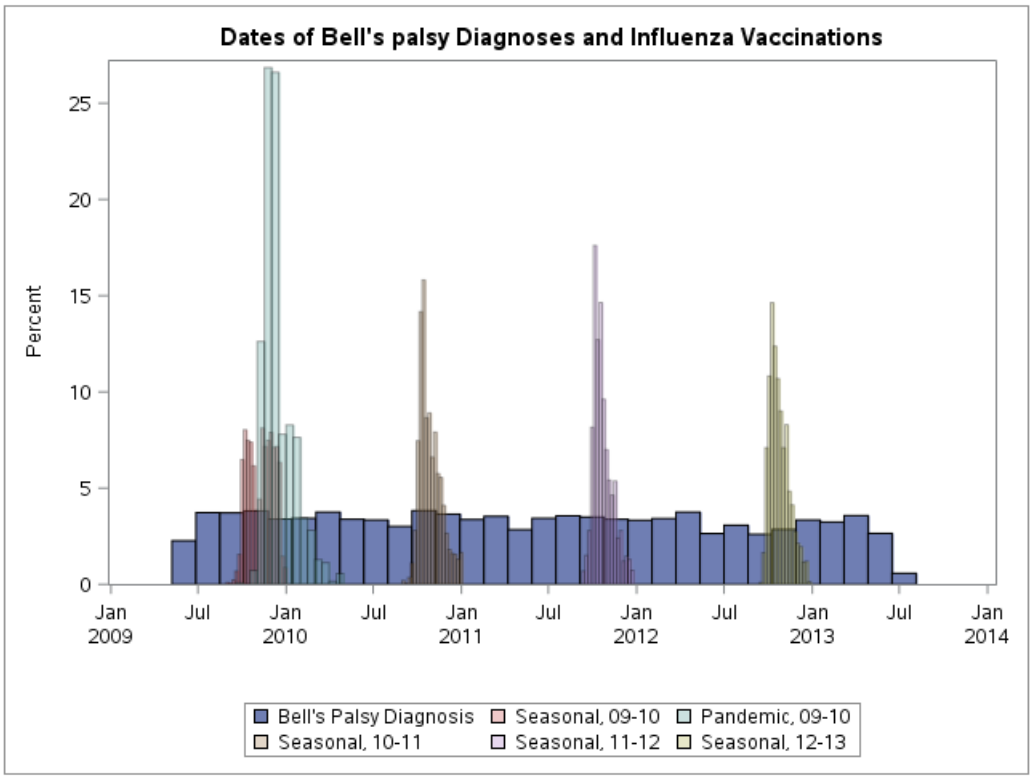


Figure 1. Distribution of Bell's palsy diagnosis dates and dates of vaccination during the observation period

Table 2. Main characteristics of cases occurring between 1 June 2009 and 30 June 2013 by vaccination status

	Received seasonal vaccine*					Received pandemic vaccine				
	Yes n=2408		No n= 3880		*p-value*	Yes n=901		No n= 5387		*p-value*
Demographics										
Female (n (%))	1313	(54.53)	1881	(48.48)	<.0001	454	(50.39)	2740	(50.86)	0.79
Mean age in years (SD)	58.59	(18.07)	36.56	(16.51)	<.0001	56.75	(19.64)	43.03	(19.61)	<.0001
Age (n (%))										
<45 yrs	532	(22.09)	2685	(69.20)		212	(23.53)	3005	(55.78)	
45 – 65 yrs	897	(37.25)	1052	(27.11)		357	(39.62)	1592	(29.55)	
>65 yrs	979	(40.66)	143	(3.69)	<.0001	332	(36.85)	790	(14.66)	<.0001

Of cases, 14% received the monovalent pandemic influenza vaccine whilst seasonal vaccines were received by 24 to 28% of cases dependent on the year. Thirty-five percent (2232 persons) experienced at least one episode of ARI during follow-up. In total, 3.5% (220 cases) received an influenza diagnosis. During follow-up 155 women had one pregnancy (4.85%) and three women had two pregnancies (0.09%).

We found that pregnancy was associated with a higher risk of Bell's palsy (RR 1.75, 95% CI 1.19 – 2.57) and that pregnant women had a higher likelihood of receiving an influenza vaccine (RR 5.05, 95% CI 3.26-7.82). An episode of ARI was strongly associated with Bell's palsy on the day of consultation (RR 6.99, 95% CI: 4.39- 11.13), but also in the 7 days following a consultation for ARI (RR 2.44, 95% CI 1.81 – 3.30). In addition, an episode of ARI was associated with an increased incidence of vaccination on the day of consultation for ARI (RR 2.93, 95% CI 1.58 – 5.46) and a reduced incidence of vaccination during the week following a consultation for ARI (RR 0.50, 95% CI 0.28 – 0.89). The distribution of ARI dates relative to vaccination dates over calendar time is given in figure 2.

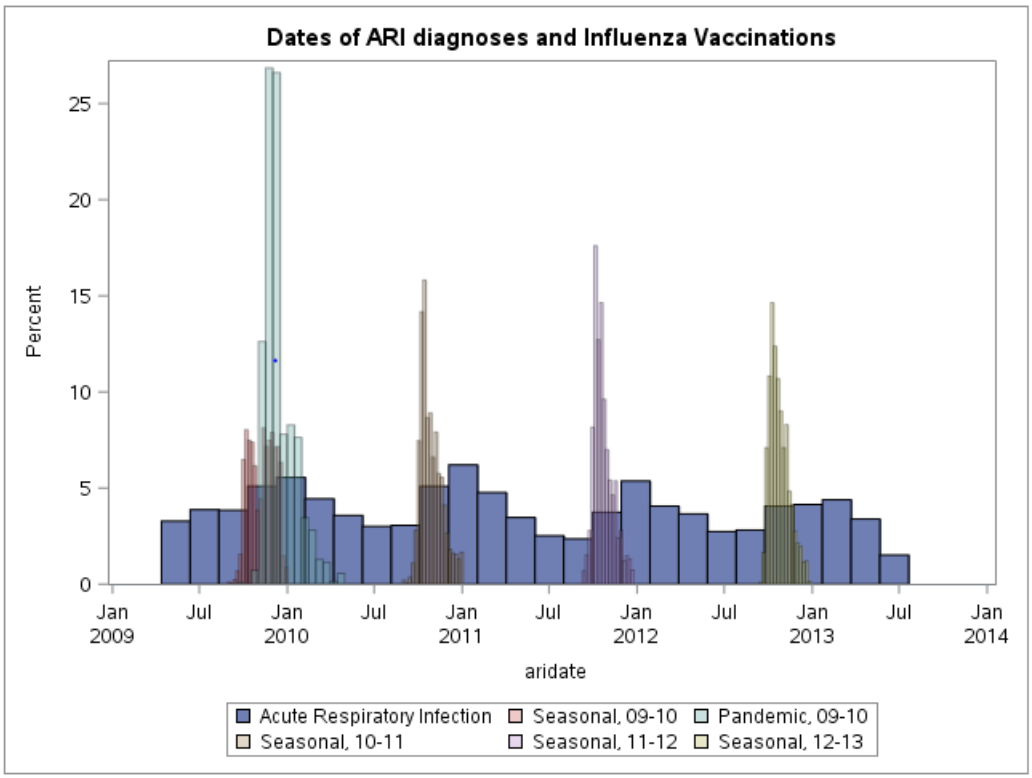


Figure 2. Distribution of acute respiratory infection dates and dates of vaccination during the observation period

There was no statistical evidence of an association between Bell's palsy and a consultation for influenza (RR 2.41, 95% CI 0.76 – 7.58).

The crude RI of Bell's palsy during the 42 days after vaccination with an influenza vaccine was 0.88 (95% CI: 0.74 – 1.04). On the day of vaccination the relative incidence was 2.15 (95% CI: 1.12 – 4.14). The RI was reduced in the fourteen days prior to vaccination, 0.70 (95% CI: 0.51 – 0.96). When adjusted for seasonality, episodes of ARI and pregnancies, the RI during the 42 days after vaccination with an influenza vaccine was 0.85 (95% CI: 0.72 – 1.01). At the date of vaccination the adjusted RI was 2.08 (95% CI: 1.08 - 4.01), during the 14 days preceding vaccination the adjusted RI was 0.68 (95% CI: 0.50 - 0.93).

When considering the type of vaccine (i.e. seasonal vs. pandemic) the adjusted RI was similar during the 42 days following seasonal vaccine (0.96, 95%CI: 0.82-1.13) or pandemic vaccine (0.73, 95%CI: 0.47-1.12).

The adjusted (for ARI and seasonality in men; for ARI, seasonality, and pregnancy in women) RI during the 42 days following influenza vaccination (any) was slightly lower in women (0.77, 95% CI: 0.61 – 0.99) compared to men (0.94, 95% CI: 0.74 – 1.19), with confidence intervals overlapping.

The adjusted relative incidence of Bell's palsy within 42 days of influenza vaccine stratified by vaccine and age can be found in Table 3.

Table 3. Age and season specific relative incidences (95% CI) of Bell's palsy within 42 days of influenza vaccination (adjusted for seasonality, ARI consultations and pregnancy)

Risk Period	All age groups				Age <45			Age 45 to 64			Age 65 +		
	Person Time (yrs)	N	RI	95% CI	N	RI	95% CI	N	RR	95% CI	N	RR	95% CI
*Any vaccine**													
Day -14 to -1	244	41	0.68	0.50-0.93	7	0.92	0.43-1.95	15	0.73	0.44-1.23	19	0.58	0.36-0.92
Day 0	17	9	2.08	1.08-4.01	2	3.58	0.89-14.40	2	1.36	0.34-5.44	5	2.12	0.88-5.15
Day 1 to 42	733	154	0.85	0.72-1.01	25	1.07	0.71-1.62	45	0.73	0.53-1.00	84	0.86	0.67-1.10
Non-Risk	23,031	5629											
2009 Pandemic													
Day -14 to -1	39	2	0.18	0.05-0.74	1	0.41	0.06-2.98	1	0.23	0.03-1.65	0	NA	NA
Day 0	3.2	0	NA	NA	0	NA	NA	0	NA	NA	0	NA	NA
Day 1 to 42	118	24	0.73	0.47-1.12	3	0.40	0.12-1.28	9	0.70	0.35-1.43	12	0.93	0.50-1.74
Non-Risk	5076	3681											
Season 2010-2011													
Day -14 to -1	52	7	0.64	0.29-1.39	0	NA	NA	1	0.36	0.05-2.68	6	0.90	0.37-2.17
Day 0	4.4	2	2.25	0.55-9.18	0	NA	NA	1	4.36	0.58-32.57	1	1.86	0.25-13.64
Day 1 to 42	198	53	1.28	0.90-1.80	11	1.58	0.73-3.42	16	1.44	0.75-2.76	26	1.03	0.62-1.72
Non-Risk	2705	635											
Season 2011-2012													
Day -14 to -1	54	17	1.23	0.73-2.08	3	1.28	0.36-4.53	7	2.17	0.91-5.16	7	0.76	0.34-1.72
Day 0	4.9	1	0.83	0.12-5.98	1	5.39	0.71-41.15	0	NA	NA	0	NA	NA
Day 1 to 42	224	45	0.81	0.57-1.16	5	0.59	0.21-1.60	14	1.12	0.58-2.19	26	0.67	0.41-1.10

Non-Risk	2522	537											
Season 2012-2013													
Day -14 to -1	54	8	0.62	0.30-1.30	3	1.82	0.51-6.50	2	0.34	0.08-1.44	3	0.58	0.17-1.93
Day 0	4.8	3	2.87	0.91-9.09	1	6.86	0.89-52.67	1	2.22	0.30-16.25	1	2.33	0.32-17.23
Day 1 to 42	217	44	0.90	0.63-1.28	9	1.33	0.58-3.07	9	0.43	0.21-0.88	26	1.28	0.74-2.20
Non-Risk	2627	547											

All Bell's palsy cases regardless of vaccination status included. All analyses adjusted for seasonality by quarter, ARI consultations, and pregnancy in strata that contained pregnant cases.

In the analysis in which only observation time after vaccination was included, exposure to 2009 pH1N1 vaccine with control for ARI, seasonality, and pregnancy produced a RI of 0.88 (95% CI: 0.47, 1.65) while exposure to 2010-11, 2011-12, and 2012-13 seasonal vaccines produced RIs of 1.56 (0.95, 2.57); 0.69 (0.45, 1.06); and 0.91 (0.57, 1.46), respectively.

The main analysis disregarded seasonal influenza vaccines during the 2009-2010 season. We considered the receipt of seasonal influenza vaccines during the 2009-2010 season as a separate risk factor in a sensitivity analysis. The results showed a RI of 1.14 (0.86, 1.51) in the 42 days following vaccination.

Discussion

Bell's palsy is a syndrome for which the exact cause is unclear. As a result it could have multiple triggers, of which – considering the hypothetical autoimmune aetiology – influenza and influenza vaccination could be one. Clusters of Bell's palsy cases have been reported following influenza vaccination in the past. An association was reported for Bell's palsy and Pandemrix, an AS03 adjuvanted pandemic influenza vaccine in Sweden (15), and a signal was reported from Taiwan (17). In this study, we evaluated the risk of Bell's palsy following vaccination with influenza vaccines containing A/California/7/2009 (H1N1)-like viral strains, including pandemic vaccines, in the UK.

The increased risk of Bell's palsy on the day of influenza vaccination was expected, based upon the findings of Stowe et al (13), and a likely opportunistic recording of cases.

We found no evidence of an increased incidence of Bell's palsy consultations following seasonal influenza vaccination overall, nor for monovalent pandemic influenza vaccine in 2009. Therefore our study does not confirm the results identified by Bardage et al (15) in Sweden. While Bardage et al. controlled for sex, age, and health utilization as measured by contacts within the year prior to vaccination, they were unable to control for unmeasured within-person confounders or for seasonality. Given the association we found here between ARI and Bell's palsy, failure to control for infection and/or seasonality could lead to the increased HR found by Bardage et al. When adjusted for seasonality, episodes of ARI and pregnancies, the RI during the 42 days after vaccination with an influenza vaccine was 0.85 (95% CI: 0.72 – 1.01).

Other than mild protective effects in women following exposure to any H1N1-containing vaccine and in 45-64 year olds following exposure to the 2012-13 seasonal vaccine, all estimated RIs associated with vaccine exposure were non-significant. Given the number of associations assessed together with the upper limits of the confidence

intervals being nearly equal to one, we did not further investigate these apparent protective effects, as these could reasonably be due to chance.

One of the more restrictive assumptions of the SCCS method is that the distribution of exposure after a certain time must be independent of the event history prior to that time (18). Bell's palsy is not a contra-indication for influenza vaccination. Nonetheless, it is possible that people will delay vaccination after Bell's palsy, which can represent a violation of the assumption of the SCCS. Generally, this delay in vaccination would bias RI estimates upward by producing a scarcity of cases in control intervals. In our main analysis we fixed the 14 days prior to vaccination as a separate risk period. The reduced incidence found during this risk period demonstrates that persons will delay vaccination when diagnosed with Bell's palsy. We assumed that a 14 day period would be sufficient to exclude any bias resulting from this delay. As evidenced by the sensitivity analysis which only considered observation time after vaccination and produced estimates very similar to those produced with a 14-day low risk period, this 14-day period was sufficient to control for a potential healthy vaccinee effect.

A second restrictive assumption of the SCCS method is that events are either recurrent and independent within individuals or not-recurrent and uncommon (18). Bell's palsy can recur, however this is rare (1) and is reported to do after a latency period of approximately 10 years (5, 13). In our study we considered any second consultation of Bell's palsy within 6 months to belong to a single episode. We found that 1.4% of persons had more than one episode within our relatively short observation period. As recurrent events are rare we assume the bias is negligible (23).

Our study has limitations that could impact the observed results.

Whilst the SCCS inherently deals with measured and unmeasured fixed confounding variables, time varying confounders will still need to be measured and adjusted for. We adjusted for seasonality by quarter, consultations for ARI and pregnancies, as these factors were identified as confounders in our study. Although both ARI and pregnancies were association with exposure and outcome, adjusting for them had minimal impact on estimates. As we identified pregnancies by date of delivery we did not capture all pregnancies. Similarly, consultations for ARI do not reflect all the ARIs actually occurring. We could not adjust for time varying factors that were not measured such as changes in medical coding practice, healthcare seeking behaviour, or vaccination policy.

Finally, persons who develop Bell's palsy may consult their GP only after prolonged persistence of symptoms or not at all, making incomplete reporting of cases possible in our study. If reporting was differential by vaccination status, meaning if persons who develop Bell's palsy shortly following vaccination were more likely or less likely to consult their GP, this would have introduced bias in this study.

In conclusion, our study did not provide evidence of an increased risk of Bell's palsy following vaccination with any influenza vaccine containing A/California/7/2009 (H1N1)-like viral strains, either pandemic or seasonal vaccines.

References

1. Gilden DH. Clinical practice. Bell's Palsy. The New England journal of medicine. 2004;351(13):1323-31. Epub 2004/09/24.

2. Rowlands S, Hooper R, Hughes R, Burney P. The epidemiology and treatment of Bell's palsy in the UK. Eur J Neurol. 2002;9(1):63-7. Epub 2002/01/11.

3. Rowhani-Rahbar A, Baxter R, Rasgon B, Ray P, Black S, Klein JO, et al. Epidemiologic and clinical features of Bell's palsy among children in Northern California. Neuroepidemiology. 2012;38(4):252-8.

4. Rath B, Linder T, Cornblath D, Hudson M, Fernandopulle R, Hartmann K, et al. All that palsies is not Bell's - the need to define Bell's palsy as an adverse event following immunization. Vaccine. 2007;26(1):1-14.

5. Greco A, Gallo A, Fusconi M, Marinelli C, Macri GF, de Vincentiis M. Bell's palsy and autoimmunity. Autoimmunity reviews. 2012;12(2):323-8.

6. Holland NJ, Weiner GM. Recent developments in Bell's palsy. BMJ. 2004;329(7465):553-7. Epub 2004/09/04.

7. Chou CH, Liou WP, Hu KI, Loh CH, Chou CC, Chen YH. Bell's palsy associated with influenza vaccination: two case reports. Vaccine. 2007;25(15):2839-41. Epub 2006/11/07.

8. Baxter R, Toback SL, Sifakis F, Hansen J, Bartlett J, Aukes L, et al. A postmarketing evaluation of the safety of Ann Arbor strain live attenuated influenza vaccine in adults 18-49 years of age. Vaccine. 2012;30(20):3053-60.

9. Baxter R, Toback SL, Sifakis F, Hansen J, Bartlett J, Aukes L, et al. A postmarketing evaluation of the safety of Ann Arbor strain live attenuated influenza vaccine in children 5 through 17 years of age. Vaccine. 2012;30(19):2989-98.

10. Izurieta HS, Haber P, Wise RP, Iskander J, Pratt D, Mink C, et al. Adverse events reported following live, cold-adapted, intranasal influenza vaccine. JAMA : the journal of the American Medical Association. 2005;294(21):2720-5.

11. Mutsch M, Zhou W, Rhodes P, Bopp M, Chen RT, Linder T, et al. Use of the inactivated intranasal influenza vaccine and the risk of Bell's palsy in Switzerland. The New England journal of medicine. 2004;350(9):896-903.

12. Zhou W, Pool V, DeStefano F, Iskander JK, Haber P, Chen RT, et al. A potential signal of Bell's palsy after parenteral inactivated influenza vaccines: reports to the Vaccine Adverse Event Reporting System (VAERS)--United States, 1991-2001. Pharmacoepidemiology and drug safety. 2004;13(8):505-10.

13. Stowe J, Andrews N, Wise L, Miller E. Bell's palsy and parenteral inactivated influenza vaccine. Human vaccines. 2006;2(3):110-2. Epub 2006/10/03.

14. Rowhani-Rahbar A, Klein NP, Lewis N, Fireman B, Ray P, Rasgon B, et al. Immunization and Bell's palsy in children: a case-centered analysis. American journal of epidemiology. 2012;175(9):878-85.

15. Bardage C, Persson I, Ortqvist A, Bergman U, Ludvigsson JF, Granath F. Neurological and autoimmune disorders after vaccination against pandemic influenza A (H1N1) with a monovalent adjuvanted vaccine: population based cohort study in Stockholm, Sweden. BMJ. 2011;343:d5956. Epub 2011/10/14.

16. Lee GM, Greene SK, Weintraub ES, Baggs J, Kulldorff M, Fireman BH, et al. H1N1 and seasonal influenza vaccine safety in the vaccine safety datalink project. American journal of preventive medicine. 2011;41(2):121-8.

17. Huang WT, Huang WI, Huang YW, Hsu CW, Chuang JH. The reporting completeness of a passive safety surveillance system for pandemic (H1N1) 2009 vaccines: a capture-recapture analysis. Vaccine. 2012;30(12):2168-72.

18. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Statistics in medicine. 2006;25(10):1768-97. Epub 2005/10/13.

19. Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. Biometrics. 1995;51:228-35.

20. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Informatics in primary care. 2011;19(4):251-5.

21. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiology and drug safety. 2007;16(4):393-401.

22. Department of Health. Pandemic H1N1 (swine) influenza vaccine uptake amongst patient groups in primary care in England 2009/10. In: DH/HPA28, editor. Oct 2010.

23. Weldeselassie YG, Whitaker HJ, Farrington CP. Use of the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. Epidemiology and infection. 2011;139(12):1805-17.

Supplementary Material

ANNEX 1: READ CODES FOR DATA EXTRACTION

Read code for Bell's palsy

F31..00	Facial nerve disorders
F310.00	**Bell's (facial) palsy**

AHD codes for influenza vaccination

1002090105	Influenza A H1N1v
1002090104	Influenza A H1N1v unknown brand (other health provider)
1002090103	Influenza A H1N1v (other health provider)
1002090102	Influenza A H1N1v (other health provider)
1002090101	Influenza A H1N1v
1002090100	Influenza A H1N1v unknown brand
1002090000	Influenza

READ codes for influenza vaccination

65E..00	Influenza vaccination
65E0.00	First pandemic influenza vaccination
65E1.00	Second pandemic influenza vaccination
65E2.00	Influenza vacc othr hlth prov
65E3.00	1st pan flu vac othr hlth prov
65E4.00	2nd pan flu vac othr hlth prov
65E5.00	CELVAPAN - first influenza A (H1N1v) 2009 vaccination given
65E6.00	CELVAPAN - second influenza A (H1N1v) 2009 vaccination given
65E7.00	CELVAPAN - 1st flu A (H1N1v) 2009 vacc by othr hlth provider
65E8.00	CELVAPAN - 2nd flu A (H1N1v) 2009 vacc by othr hlth provider

65E9.00	PANDEMRIX - first influenza A (H1N1v) 2009 vaccination given
65EA.00	PANDEMRIX - second influenza A (H1N1v) 2009 vaccination give
65EB.00	PANDEMRIX - 1st flu A (H1N1v) 2009 vac by othr hlth provider
65EC.00	PANDEMRIX - 2nd flu A (H1N1v) 2009 vac by othr hlth provider

READ codes for ARI

'H00..00'	'H03..00'	'H042000'	'H060700'	'H01yz00'	'H040000'	'H053.00'	'H061200'
'H00..11'	'H03..11'	'H042100'	'H060800'	'H01z.00'	'H040100'	'H055.00'	'H061300'
'H00..12'	'H03..12'	'H042z00'	'H060900'	'H02..00'	'H040200'	'H05y.00'	'H061400'
'H00..13'	'H030.00'	'H043.00'	'H060A00'	'H02..11'	'H040300'	'H05z.00'	'H061500'
'H00..15'	'H031.00'	'H043.11'	'H060B00'	'H02..12'	'H040400'	'H05z.11'	'H061600'
'H00..16'	'H032.00'	'H043000'	'H060C00'	'H02..13'	'H040500'	'H05z.12'	'H061z00'
'H01..00'	'H033.00'	'H043100'	'H060D00'	'H020.00'	'H040600'	'H06..00'	'H062.00'
'H01..11'	'H034.00'	'H043200'	'H060E00'	'H021.00'	'H040w00'	'H060.00'	'H06z.00'
'H010.00'	'H035.00'	'H043211'	'H060F00'	'H022.00'	'H040x00'	'H060.11'	'H06z000'
'H010.11'	'H035000'	'H043z00'	'H060v00'	'H023.00'	'H040z00'	'H060000'	'H06z011'
'H011.00'	'H035100'	'H044.00'	'H060w00'	'H023000'	'H041.00'	'H060100'	'H06z100'
'H012.00'	'H035z00'	'H04z.00'	'H060x00'	'H023100'	'H041000'	'H060200'	'H06z111'
'H013.00'	'H036.00'	'H05..00'	'H060z00'	'H023z00'	'H041100'	'H060300'	'H06z112'
'H014.00'	'H03z.00'	'H050.00'	'H061.00'	'H024.00'	'H041z00'	'H060400'	'H07..00'
'H01y.00'	'H04..00'	'H051.00'	'H061000'	'H025.00'	'H042.00'	'H060500'	'H0y..00'
'H01y000'	'H040.00'	'H052.00'	'H061100'	'H02z.00'	'H042.11'	'H060600'	'H0z..00'

READ codes for delivery date

'63...00'	'63...00'	'639..00'	'63A..00'	'632..00'	'63D..00'
'6331.00'	'633..00'	'633a.00'	'6341.00'	'6342.00'	'63E2.00'

'635..11'	'7F10.00'	'7F10000'	'7F10100'	'7F10y00'	'7F10z00'
'7F10z11'	'7F10z12'	'7F11.00'	'7F11000'	'7F11100'	'7F11200'
'7F11300'	'7F11y00'	'7F11z00'	'7F12.00'	'7F12000'	'7F12100'
'7F12111'	'7F12y00'	'7F12z00'	'7F13.00'	'7F13000'	7F13100 '
'7F13111'	'7F13200'	'7F13300'	'7F13y00'	'7F13z00'	7F14.00'
'7F14100'	'7F14y00'	'7F14z00'	'7F15.00'	'7F15000'	'7F14000'
'7F15100'	'7F15y00'	'7F15z00'	'7F16.00'	'7F16000'	'7F16200'
'7F16300'	'7F16400'	'7F16500'	'7F16600'	'7F16700'	'7F16800'
'7F16900'	'7F16A00'	'7F16B00'	'7F16y00'	'7F16z00'	'7F17.00'
'7F17000'	'7F17100'	'7F17200'	'7F17300'	'7F17y00'	'7F17z00'
'7F18.00'	'7F18000'	'7F18100'	'7F16100'	'7F18y00'	'7F18z00'
'7F19.00'	'7F19000'	'7F19100'	'7F19y00'	'7F19z00'	'7F1A.00'
'7F17.11'	'7F17.12'	'L34..00'	'L398.00'	'L398300'	'L398400'
'Ly0..00'					

READ codes for influenza

H27..	H27z.	H270z	H270.
Hyu06	H27yz	H27y.	H2710
H2711	H27yz	H27y.	H2710
H2711	Hyu07	H27y1	Hyu05
H271.	H27y0	H271z	

4.2 THE IMPACT AND LONGEVITY OF MEASLES-ASSOCIATED IMMUNE SUPPRESSION: A MATCHED COHORT STUDY USING DATA FROM THE THIN GENERAL PRACTICE DATABASE IN THE UNITED KINGDOM

Kartini Gadroen*

Caitlin Dodd*

Gwen M C Masclee

Maria A J de Ridder

Daniel Weibel

Michael J Mina

Bryan T Grenfell

Miriam C J M Sturkenboom

David A M C van de Vijver

Rik L de Swart

*Authors have contributed equally to the manuscript

BMJ open 8.11 (2018): e021465.

ABSTRACT

Objective

To test the hypothesis that measles infection increases susceptibility to and incidence of infectious diseases over a prolonged period of time.

Design

A population-based matched cohort study.

Data Sources

This study examined children aged 1 to 15 years in the Health Improvement Network (THIN) UK general practice medical records database. Participants included 2,228 patients with measles diagnosed between 1990 and 2014. They were matched on age, sex, GP-practice and calendar year with 19,930 children without a measles diagnosis. All controls had received at least one measles vaccination. Children with a history of immune-compromising conditions or with immune-suppressive treatment were excluded.

Primary outcome measures

Incidence rate ratio (IRR) of infections, anti-infective prescriptions and all-cause hospitalisations following measles in pre-determined periods using multivariate analysis to adjust for confounding variables.

Results

In children with measles, the incidence rate for non-measles infectious disease was significantly increased in each time period assessed up to 5 years post-measles: 43% in the first month (IRR: 1.43; 95%CI: 1.22 to 1.68), 20% from month one to the first year (IRR: 1.22; 95%CI: 1.14 to 1.31), 10% from year 1 to 2.5 years (IRR: 1.10; 95%CI: 1.02 to 1.19), and 15% (IRR: 1.15; 95%CI: 1.06 to 1.25) in years 2.5 to 5 years of follow-up. Children with measles were more than three times as likely to receive an anti-infective prescription in the first month and 15%-24% more likely between the first month and 5 years. The rate of hospitalization in children with measles was increased only in the month following diagnosis but not thereafter (IRR: 2.83; 95%CI: 1.72 to 4.67).

Conclusion

Following measles, children suffered from increased rates of diagnosed infections, requiring increased prescribing of antimicrobial therapies. This population-based matched cohort study supports the hypothesis that measles has a prolonged impact to increase susceptibility to non-measles infectious diseases.

Introduction

Measles is a highly contagious childhood disease.[1] During the pre-vaccine era, nearly every child acquired measles before the age of 15 years.[2] A key characteristic of the disease is a transient immune suppression, causing increased susceptibility to opportunistic infections. As a result, measles is often complicated by pneumonia, diarrhoea or otitis media, which may lead to severe and even fatal disease.[3 4] The introduction of measles-containing vaccines has reduced measles incidence,[1] as well as childhood mortality.[5] Interestingly, this reduction in childhood mortality is stronger than what would have been expected based on measles mortality in unvaccinated populations.[6] Recent studies into the mechanism of measles immune suppression, based on observations in experimentally infected non-human primates, showed that measles virus preferentially replicates in $CD150^{+}$ memory lymphocytes. It was hypothesised that viral cytotoxicity and immune-mediated clearance resulted in depletion of these cells, leading to a loss of acquired immunological memory.[4] Consistent with this hypothesis, a subsequent ecological study using population level data from England and Wales, the United States, and Denmark, found that rates of non-measles infectious disease mortality are tightly coupled to measles incidence – with a greater mortality rate at higher recent measles incidence. Mina et al measured a duration of measles-induced immunomodulation by assessing the association between measles incidence and childhood mortality. The results showed that measles was associated with increased mortality from other infectious diseases over a period of more than two years.[7] However, the study was based on population-level ecological association data, and the authors did not have access to case-based data.

Monovalent measles vaccination was introduced in England in 1968, and replaced in 1988 by the multivalent measles, mumps and rubella (MMR) vaccine. Initially MMR was offered only as a single dose at the age of 12 months. In 1996 a second dose was introduced and offered at age of 40 months. From 1996 to 2004 the number of reported measles cases in the UK was small. Following the publication of a subsequently discredited study linking autism and measles vaccination in 1998, coverage dropped for several years below herd protection level, and in 2007 measles was re-established in the UK. In response, an MMR catch-up campaign targeting individuals up to 18 years of age was implemented in 2008. In response to a mumps outbreak, Wales had already implemented a national MMR vaccination campaign targeting individuals aged between 11 and 25 years old in 2005.

In the present study, we have used individual-level data from a United Kingdom database to test whether measles results in prolonged increased susceptibility to other infections. The aim of our study was to assess whether measles is associated with increased frequency of non-measles infectious disease, anti-infective prescriptions, or hospitalisations over a prolonged period of time.

Methods

Data source

For this matched-cohort study we used data from The Health Improvement Network (THIN) database. THIN is a population-based general practice registry which contains prospectively collected, anonymized longitudinal electronic patient records from over 550 General Practitioner (GP) practices across the United Kingdom (UK), capturing health care data from more than 12 million patients (about 6% of the population).[8 9] Data recorded in THIN include demographic, socioeconomic, and clinical information, including chief complaint, symptoms, test results, diagnoses, prescriptions, and referrals to hospitals. The population covered has similar demographic characteristics to the national UK population, and the recording of consultations and prescriptions is comparable to national levels.[10 11] Diagnoses and symptoms are recorded in Read codes, a standard terminology, maintained by the UK National Health Service Centre for Coding and Classification (NHS CCC).[12] Information on drug prescription is recorded using British National Formulary (BNF) codes and the MULTILEX product dictionary. The specific codes

used for this study were selected by a medical doctor and reviewed by a virologist, medical doctor and epidemiologist for their relevance (see S1 List for selected read codes).

Study design and population

The source population consisted of all patients who had contributed longitudinal data to the database between January 1st, 1990 to September 30th, 2014, from the age of 6 months to 15 years. This study period captures the period of time when vaccination rates fell during the late 1990s, with increased measles cases in the following years. The measles group consisted of children with a measles diagnosis (whether or not laboratory confirmed) between the ages of 1 and 15 years. The date of measles diagnosis was taken as the index date. To each child with a measles diagnosis, up to 10 children free of measles were matched on age in years, sex, GP-practice, and calendar time in years. Children free of measles were required to have had at least one dose of measles-containing vaccine, prior to the matched case's index date. We considered that having received at least one dose of measles-containing vaccine would reduce the chance that children included in the "free of measles" group had ever had measles . Patients with a history of immune-compromising conditions (e.g. HIV infection, and organ, or bone marrow transplantation), or with immune suppressive treatment prior to the index date were excluded. See Table S2 for the STROBE statement of this study.

Patient involvement

No patient was involved in setting the research question, outcome measures, design or conduct of the study. The results were not disseminated to the patients, as the study was based on anonymised patient records.

Outcomes

Four clinical outcomes were considered: infections, anti-infective prescriptions, all cause hospitalisations, and all-cause mortality. The outcomes were defined by the relevant clinical codes for symptoms and diagnoses, or drug codes. Infections included all communicable diseases other than measles. Infections were required to be 14 days apart to be considered a new event. Anti-infective prescriptions included all systemic antibiotics, anti-mycotic, antivirals, and anti-parasitic medication. For anti-infective prescriptions and hospitalisations, any event occurring on a different day (at least 1 day apart) was considered a new event.

Follow-up

Follow-up started at the index date and continued for a period of five years, until date of transfer out the general practitioner's practice, the 15th birthday, or death, whichever came earliest. Each outcome was analysed in pre-determined periods following measles diagnosis: within the first month; ≥one month to <1 year; ≥1 year to <2.5 years; and ≥2.5 years to <5 years, to observe changes over time. Hazard ratios (HR) for hospitalisation were calculated with follow up starting at 30 days after the index date to avoid inclusion of hospitalisations due to initial complications related to measles.

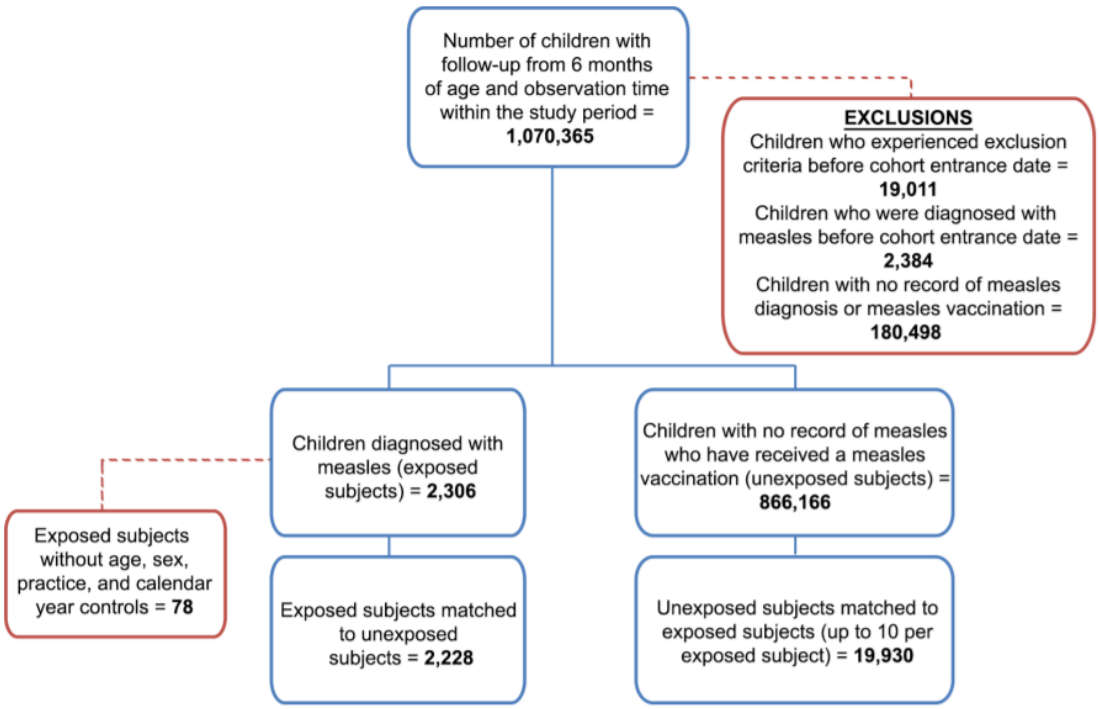


Figure 1 | Flowchart of study cohort selection. Starting from 1,070,365 eligible children in the THIN database, 2,228 measles patients and 19,930 matched controls were selected for this study.

Potential confounders and effect modifiers

We considered as potential confounders: chronic respiratory disease, cardiovascular disease, prior exposure to routine childhood vaccines other than measles containing vaccines, deprivation index, health care consumption, and occurrence of each outcome of interest in the year prior to index. Potential confounders were assessed at the index date. Vaccine adherence was defined as exposure to any dose of other routine childhood vaccines such as pertussis containing vaccines before the index date and coded as binary with vaccine adherence equal to one if any other childhood vaccine was received and zero otherwise. The Townsend deprivation score, a measure of social deprivation based on unemployment level, car ownership, home ownership, and household overcrowding levels by area, was used within a particular zip code.[13] Health care consumption, as a proxy for general health, was assessed by the rate of GP consultations in the year prior to the index date,[14] and categorized using quintile cut-off points. For a list of various types of consultations included to calculate GP consultation rate, please see supporting information S1. For each outcome, the number of events in the year prior to index was calculated.

Statistical analysis

Baseline characteristics were compared between children with measles and children free of measles using Student t-test, Mann-Whitney U-test, chi-square test, or Fisher exact test as appropriate. Observed incidence rates of measles diagnosis codes as well as measles notification codes were estimated by dividing the number of cases by the number of person-years (PYs) at risk within the database stratified by calendar year, and were compared with expected incidence rates, derived from publicly available official statistics from the UK National Archives.[15] The differences in incidence of the outcomes between children with measles and children free of measles were analysed for each period using Poisson regression. For this analysis, matching was relaxed due to uninformative matched strata for each outcome, with over 1000 uninformative strata for the hospitalization outcome. A stratified

analysis was therefore not conducted, and the analysis was adjusted for confounding using multivariable analysis. We submitted the following confounders: history of cardiovascular malformation, history of respiratory disease, exposure to childhood vaccinations other than measles containing vaccination, age, sex, and GP consultation rate. Exposure to childhood vaccinations other than measles containing vaccination was not retained in final models. In addition per outcome, we submitted rate of the outcome in the year prior to the index date. Absolute rates of each outcome per 1,000 person days were calculated with covariates fixed as follows: cardiovascular and respiratory history = No, Receipt of other childhood vaccines = Yes, Number of consults and events in the previous year = median, Age = 3 years, Sex = Female. Kaplan-Meier curves and log-rank tests were used to compare time to first hospitalisation between measles infected and control individuals, with follow-up beginning at 30 days after the index date (to avoid including codes related to the initial measles infection). A stratified Cox proportional hazards model, stratified by matched set and adjusted for confounding variables, was applied to estimate hazard ratios comparing children with measles and children free of measles. Assumptions of proportional hazards were assessed by inspecting the K-M curves and formally tested with inspection of a measles*time interaction term. Model selection was by backward covariate selection, with the criteria $P < 0.1$. Subsequently we verified automatically selected models using minimization of AIC. We also estimated the hazard ratios for the outcomes first infection and first prescription.

Sensitivity analysis

Children who have received vaccinations may be different in their underlying health status, social background, lifestyle, health care seeking behaviour and health care utilization from those who did not receive vaccinations. To examine the possible effect of these unmeasured confounders, we conducted a sensitivity-analysis, stratifying the data into matched sets in which all measles cases had received, or had not received a measles-containing vaccine (i.e. non-measles group vaccinated vs measles-group vaccinated). In post-hoc analyses, we assessed the IRR of each outcome over the entire study period in vaccine adherent vs non-adherent children for each outcome using Poisson regression. We also examined the correlation of the consultation rate the year before and after the index date in measles vs. control groups using linear regression. For data management and analysis we used SAS v9.3. The study was approved by the independent THIN Scientific Review Committee (SRC reference number: 15-006).

Results

From the database population of 1,070,365 children aged 1 to 15 years, we identified 2,228 eligible children with a measles diagnosis. These children were matched to 19,930 children free of measles. Figure 1 illustrates the composition of the study cohort. Table 1 describes baseline characteristics of children with measles and children free of measles. Median follow-up time was 5.0 years (IQR: 2.2 to 5.0). The incidence rate of measles and of measles notification as reported in the THIN database were similar to the expected overall confirmed measles incidence rate as reported by official UK Government statistics (see Figure S1). There was no significant difference in follow-up time between the children with measles and the children free of measles. Exposure to childhood vaccines other than measles containing vaccines prior to the index date was lower among children with measles (98.1% vs 99.8 %), but this difference was small compared to the difference in vaccination coverage of measles containing vaccines prior to the index date (54.4% in children with measles vs 100% in children free of measles, due to inclusion criteria). GP consultation rate in the year prior to index date was slightly higher in the measles group than in the non-measles group: mean 13.87 vs. 13.22 ($p < 0.001$) consults in the year prior, respectively. The Townsend deprivation index was similar in children with measles and children free of measles. The rate of infections and anti-infective prescriptions prior to index were similar between measles and non-measles subjects while hospitalisations prior to index were more frequent for subjects subsequently diagnosed with measles. Table 2 describes events of interest occurring during follow-up in measles and non-measles subjects.

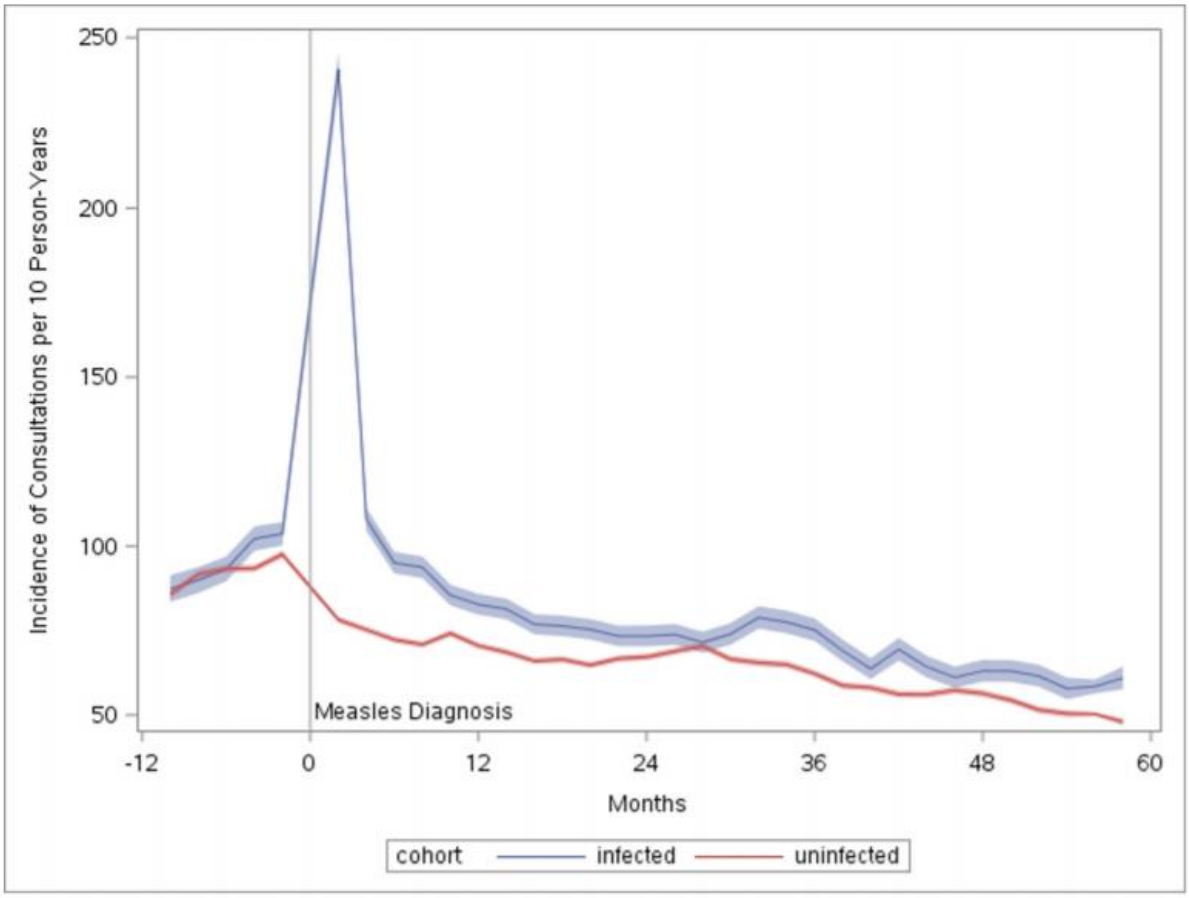


Figure 2 | Consultations in measles patients and matched controls. Incidence rates of consultations in children diagnosed with measles (blue lines) or matched controls (red lines) per 100,000 person-years, plotted by time (in months) before or after diagnosis of measles. The vertical dotted line indicates the time point of diagnosis in the measles patients. The shaded areas represent 95% confidence intervals.

Table 1. Baseline Characteristics of Enrolled Subject

		Measles group (n=2,228)		**Non-measles group (n = 19,930)**		
Variable	**Category**	**Mean ± SD or N(%)**	**Median (IQR)**	**Mean ± SD or N(%)**	**Median (IQR)**	**P-value**
Age at Case Diagnosis		3.06 ± 3.04	2 (1 to 4)	3.16 ± 3.01	2 (1 to 4)	0.1264
Person time (days)		1,379.9 ± 595.33	1,826 (849 to 1,826)	1,358.7 ± 611.54	1,826 (804 to 1,826)	0. 1186

		Measles group (n=2,228)		Non-measles group (n = 19,930)		
Variable	**Category**	**Mean ± SD or N(%)**	**Median (IQR)**	**Mean ± SD or N(%)**	**Median (IQR)**	**P-value**
Sex	Female	1,038 (46.59%)		9,275 (46.54%)		0.9643
Region	England	1,816 (81.51%)		16,291 (81.74%)		0.9871
	Northern Ireland	52 (2.33%)		448 (2.25%)		
	Scotland	167 (7.50%)		1,497 (7.51%)		
	Wales	193 (8.66%)		1,695 (8.50%)		
Experience of an excluding event during follow-up		125 (5.61%)		898 (4.51%)		**0.0219**
History of respiratory disease		84 (3.77%)		737 (3.70%)		0.8592
History of cardiovascular disease		18 (0.81%)		124 (0.62%)		0.3249
Townsend Deprivation Score	0	108 (4.85%)		1,001 (5.02%)		0.1696
	1	482 (21.63%)		4,553 (22.84%)		
	2	409 (18.36%)		3,483 (17.48%)		
	3	422 (18.94%)		3,947 (19.80%)		
	4	426 (19.12%)		3,727 (18.70%)		
	5	381 (17.10%)		3,198 (16.05%)		

		Measles group (n=2,228)		Non-measles group (n = 19,930)		
Variable	**Category**	**Mean ± SD or N(%)**	**Median (IQR)**	**Mean ± SD or N(%)**	**Median (IQR)**	**P-value**
	Missing	0 (0.00%)		21 (0.11%)		
Vaccine non-adherence	Yes	43 (1.93%)		29 (0.15%)		**<0.0001**
Measles vaccination before index date		1,212 (54.40%)		19,930 (100.00%)		**<0.0001**
Measles vaccination ever during observation		2,044 (91.74%)		19,930 (100.00%)		**<0.0001**
# Consults in the year before Index (continuous)		13.87 ± 11.54	11 (6 to 19)	13.22 ± 13.80	10 (5 to 17)	**<0.0001**
# Consults in the year before Index (categorical)	0-3	300 (13.46%)		3,731 (18.72%)		**<0.0001**
	4 to 7	427 (19.17%)		4,193 (21.04%)		
	8 to 11	443 (19.88%)		3,546 (17.79%)		
	12 to 19	542 (24.33%)		4,406 (22.11%)		
	>19	516 (23.16%)		4,054 (20.34%)		
Infections in the year prior to index		0.86 ± 1.27	1 (0 to 2)	0.87 ± 1.58	1 (0 to 2)	**0.7782**
Anti-infectives in the year prior to index		1.58 ± 1.97	0 (0 to 1)	1.53 ± 2.41	0 (0 to 1)	**0.2708**
Hospitalisations in the year prior to index		0.11 ± 0.51	0 (0 to 0)	0.07 ± 0.41	0 (0 to 0)	**0.0004**

Infectious disease

The most frequently occurring infectious diseases were upper respiratory infectious diseases (for details see Table S4). The incidence rate ratio (IRR; Table 3) of infections for children with measles compared to children free of measles was 43% higher in the first month (IRR: 1.43; 95%CI: 1.22 to 1.68), 20% higher from the first month to the first year (IRR: 1.22; 95%CI: 1.14 to 1.31), 10% higher from the first year to 2.5 years (IRR: 1.10; 95%CI: 1.02 to 1.19), and 15% higher (IRR: 1.15; 95%CI: 1.06 to 1.25) in the 2.5 to 5 years of follow-up (Figure 2).

The absolute rate of infections per 1,000 person days in the first month to first year was 1.7 (95%CI: 1.6 to 1.9) for children with measles and 1.33 (95%CI: 1.29 to 1.36) for children free of measles. The adjusted hazard ratio for non-measles infectious disease over the full follow-up period starting 30 days after measles diagnosis was 1.20 (95%CI: 1.13 to 1.28) (see Table S5).

Table 2. Descriptive statistics of events in enrolled measles exposed and non-exposed children

		Measles group (n=2,228)		**Non-measles group (n = 19,930)**		
Variable	**Category**	**Mean ± SD OR N(%)**	**Median (IQR)**	**Mean ± SD OR N(%)**	**Median (IQR)**	**P-Value**
# Infections (continuous)		1. 61 ± 2.17	1 (0 to 2)	1.28 ± 1.85	1 (0 to 2)	**<0.0001**
# Infections (categorical)	0	864 (38.78%)		9224 (46.28%)		**<0.0001**
	1 to 2	856 (38.42%)		7184 (36.05%)		
	3 to 5	377 (16.92%)		2852 (14.31%)		
	6 to 10	115 (5.16%)		591 (2.97%)		
	>10	16 (0.72%)		79 (0.40%)		
# Anti-infective Rx (continuous)		4.58 ± 5.45	3 (1 to 6)	3.35 ± 4.43	2 (0 to 5)	**<0.0001**
# Anti-infective Rx (categorical)	0	326 (14.63%)		5,104 (25.61%)		**<0.0001**
	1 to 2	631 (28.32%)		6,168 (30.95%)		
	3 to 5	651 (29.22%)		4,617 (23.17%)		
	6 to 10	393 (17.64%)		2,892 (14.51%)		
	11 to 20	187 (8.39%)		979 (4.91%)		
	>20	40 (1.80%)		170 (0.85%)		

# Hospitalisations (continuous)		0.16 ± 0.74	0 (0 to 0)	0.12 ± 0.63	0 (0 to 0)	**0.0001**
# Hospitalisations (categorical)	0	1,999 (89.72%)		18,369 (92.17%)		**0.0014**
	1 to 2	204 (9.16%)		1,396 (7.00%)		
	3 to 5	20 (0.90%)		134 (0.67%)		
	6 to 10	3 (0.13%)		24 (0.12%)		
	>10	2 (0.09%)		7 (0.04%)		
Death	No	2,226 (99.91%)		19,921 (99.95%)		0.305
	Yes	2 (0.09%)		9 (0.05%)		

Prescriptions

Children with measles received more anti-infective prescriptions than children without measles in all periods (Table 3, Figure 4, Table S6). The absolute rate of anti-infective prescriptions per 1,000 person days in the first month to first year was 0.55 (95%CI: 0.51 to 0.59) for children with measles and 0.45 (95%CI: 0.43 to 0.47) for children free of measles. The adjusted hazard ratio for anti-infective prescription over the full follow-up period starting 30 days after measles diagnosis was 1.24 (95%CI: 1.18 to 1.31). Within the first month of follow-up, children with measles had more than a threefold increase in use of anti-infective drugs as compared to controls (IRR: 3.60; 95%CI: 3.31 to 3.91). Following the first month, children who had measles continued to use more anti-infective drugs over the entire duration of the follow-up: 1 month to 1 year (IRR 1.24; 95%CI: 1.18 to 1.32); 1 year to 2.5 years (IRR 1.21; 95%CI: 1.13 to 1.29), 2.5 years to 5 years (IRR 1.15; 95%CI: 1.07 to 1.24).

Table 3. Incidence rate ratios (IRRs) of events of interest in pre-defined time periods following measles infection

Time Period	**Analysis**	**Incidence Rate Ratio (95% Confidence Interval)**		
		Infections	**Anti-infective prescriptions**	**Hospitalisation**
Days 0 to 31	Primary	1.43(1.22 to 1.68)	3.60 (3.31 to 3.91)	2.83 (1.72 to 4.67)
	Unadjusted	1.57 (1.34 to 1.84)	3.77 (3.48 to 4.08)	3.24 (2.03 to 5.19)
	Sensitivity (vaccinated measles subjects only)	1.47 (1.17 to 1.86)	4.65 (4.20 to 5.14)	1.92 (0.89 to 4.14)
	Sensitivity (unvaccinated measles subjects only)	1.33 (1.07 to 1.65)	2.45 (2.12 to 2.82)	3.30 (1.60 to 6.82)
Days 32 to 365	Primary	1.22(1.14 to 1.31)	1.25 (1.18 to 1.32)	1.14 (0.88 to 1.48)
	Unadjusted	1.31 (1.21 to 1.41)	1.31 (1.24 to 1.39)	1.29 (0.94 to 1.77)
	Sensitivity (vaccinated measles subjects only)	1.15 (1.04 to 1.27)	1.25 (1.16 to 1.35)	0.95 (0.61 to 1.46)

	Sensitivity (unvaccinated measles subjects only)	1.26 (1.15 to 1.39)	1.24 (1.15 to 1.35)	1.29 (0.92 to 1.81)
Days 366 to 913	Primary	1.10 (1.02 to 1.19)	1.21 (1.13 to 1.29)	1.08 (0.80 to 1.47)
	Unadjusted	1.15 (1.06 to 1.24)	1.25 (1.17 to 1.34)	1.19 (0.85 to 1.66)
	Sensitivity (vaccinated measles subjects only)	1.10 (0.99 to 1.22)	1.21 (1.11 to 1.32)	1.26 (0.79 to 2.04)
	Sensitivity (unvaccinated measles subjects only)	1.09 (0.99 to 1.21)	1.22 (1.12 to 1.34)	0.93 (0.64 to 1.35)
Days 914 to 1826	Primary	1.15 (1.06 to 1.25)	1.15 (1.07 to 1.24)	1.24 (0.92 to 1.67)
	Unadjusted	1.23 (1.13 to 1.35)	1.22 (1.13 to 1.31)	1.38 (1.07 to 1.78)
	Sensitivity (vaccinated measles subjects only)	1.06 (0.94 to 1.20)	1.25 (1.13 to 1.37)	1.08 (0.76 to 1.54)
	Sensitivity (unvaccinated measles subjects only)	1.21 (1.07 to 1.35)	1.07 (0.96 to 1.19)	1.37 (0.87 to 2.17)

*Primary and sensitivity analyses were adjusted for: Frequency of consultations in the year prior to index, frequency of the outcome of interest in the year prior to index, history of cardiovascular malformation, history of respiratory disease, age, and sex.

Hospitalisation

Despite smaller sample sizes, the analysis on hospitalisations also showed increased IRRs, although these were significant during the first period only (Figure 5). Confounder selection using either backward selection, or minimization of the AIC resulted in the same model, namely control for the hospitalisation rate prior to the index date, the GP consultation rate in the year prior to index date and history of cardiac malformation. The absolute rate of hospitalisations per 1,000 person days in the first month to first year was equal at 0.2 (95%CI: 0.1 to 0.2) for children with measles and children free of measles. The adjusted HR of hospitalisation for measles vs. non-measles subjects was 1.12 (95%CI: 0.96 to 1.31).

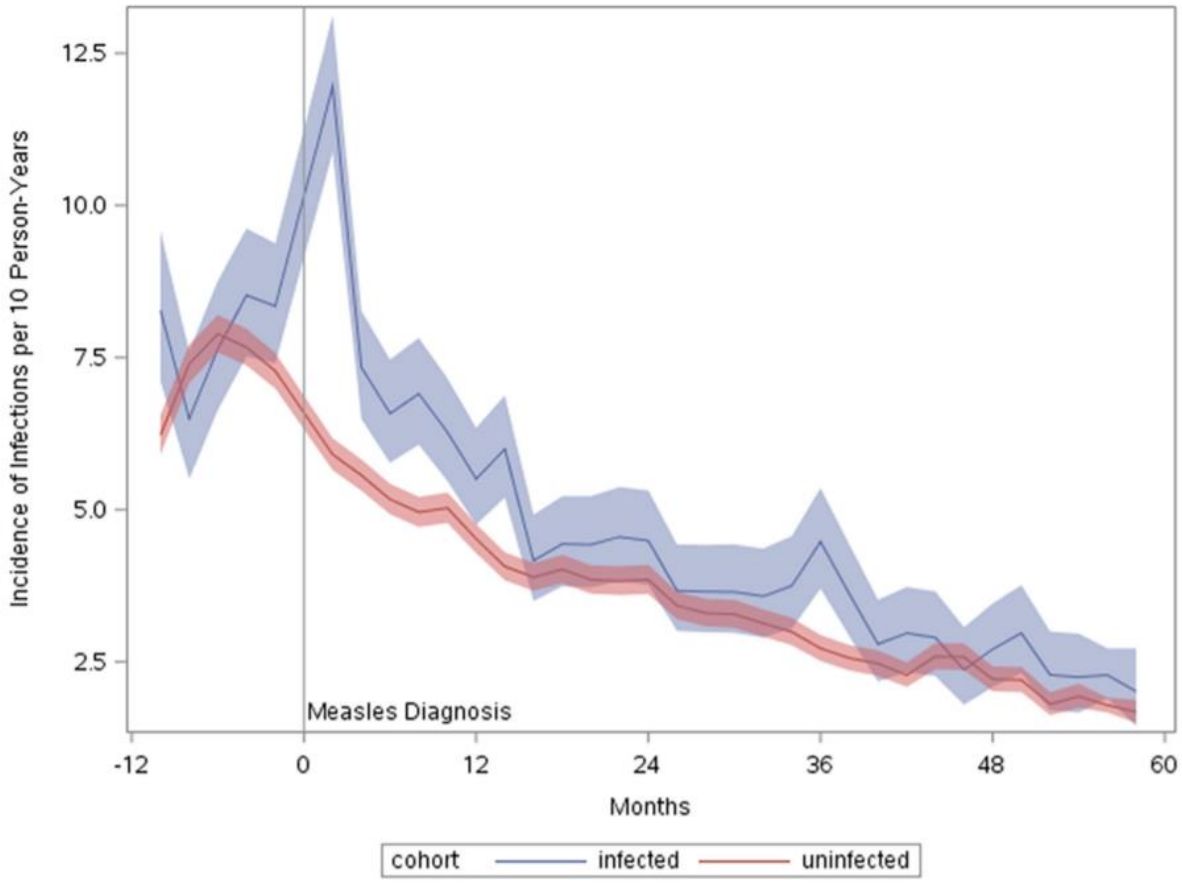


Figure 3 | Infections in measles patients and matched controls. Incidence rates of infections in children diagnosed with measles (blue lines) or matched controls (red lines) per 100,000 person-years, plotted by time (in months) before or after diagnosis of measles. The vertical dotted line indicates the time point of diagnosis in the measles patients. The shaded areas represent 95% confidence intervals.

Sensitivity analysis

Results of the sensitivity analysis were partially in agreement with findings from the main analysis. When we restricted the analysis to only those children who had received measles vaccination prior to receiving a diagnosis of measles (54.4% of all eligible children with a measles diagnosis), differences to the main analyses were not observed for anti-infective prescriptions. However, an increased rate of hospitalisations was no longer detected in any time period and an increased rate of infections no longer extended beyond one year post-diagnosis. In the analysis restricted to those children who had not had a measles vaccination prior to receiving a diagnosis of measles (45.6 % of all eligible children with a measles diagnosis) the results were in line with the main findings for hospitalizations, infections and anti- infective prescriptions with the exception that increased risk for anti-infective prescriptions did not extend into the period 2.5 to 5 years following measles.

Post-hoc analysis of the impact of vaccine adherence regardless of measles status revealed that vaccine non-adherent children were 42% more likely to receive an anti- infective prescription than vaccine-adherent children. There was no difference in risk of infections, or hospitalisations. Regressing post-index consults on pre-index

consults and measles, or non-measles status revealed that both groups showed similar trends with the rate of consultation before index date higher than that after the index date.

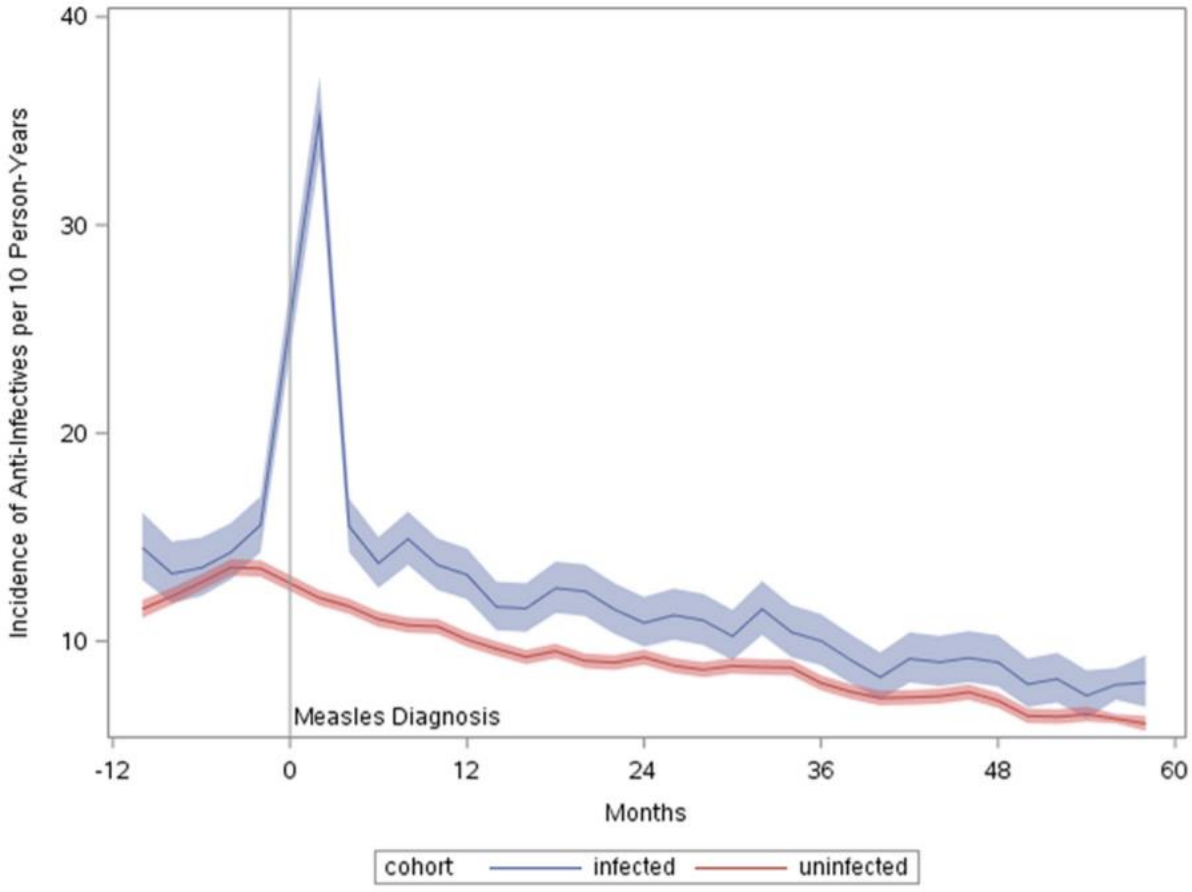


Figure 4 | Anti-infective prescriptions in measles patients and matched controls. Incidence rates of anti-infective prescriptions in children diagnosed with measles (blue lines) or matched controls (red lines) per 100,000 person-years, plotted by time (in months) before or after diagnosis of measles. The vertical dotted line indicates the time point of diagnosis in the measles patients. The shaded areas represent 95% confidence intervals.

Discussion

To our knowledge, this is the first matched-cohort study to investigate the longevity of measles-associated immune suppression in a high-income country. The results of this study are in strong agreement with previous non-clinical and ecological studies, also in high-income countries.[7] We found that rates of diagnosed infections and anti-infective prescriptions are elevated following measles infection for up to five years. While increased risk of infections and anti-infective prescriptions remained statistically significant over the full five-year study period, the effect size diminished particularly after the first year and statistical significance is partly explained

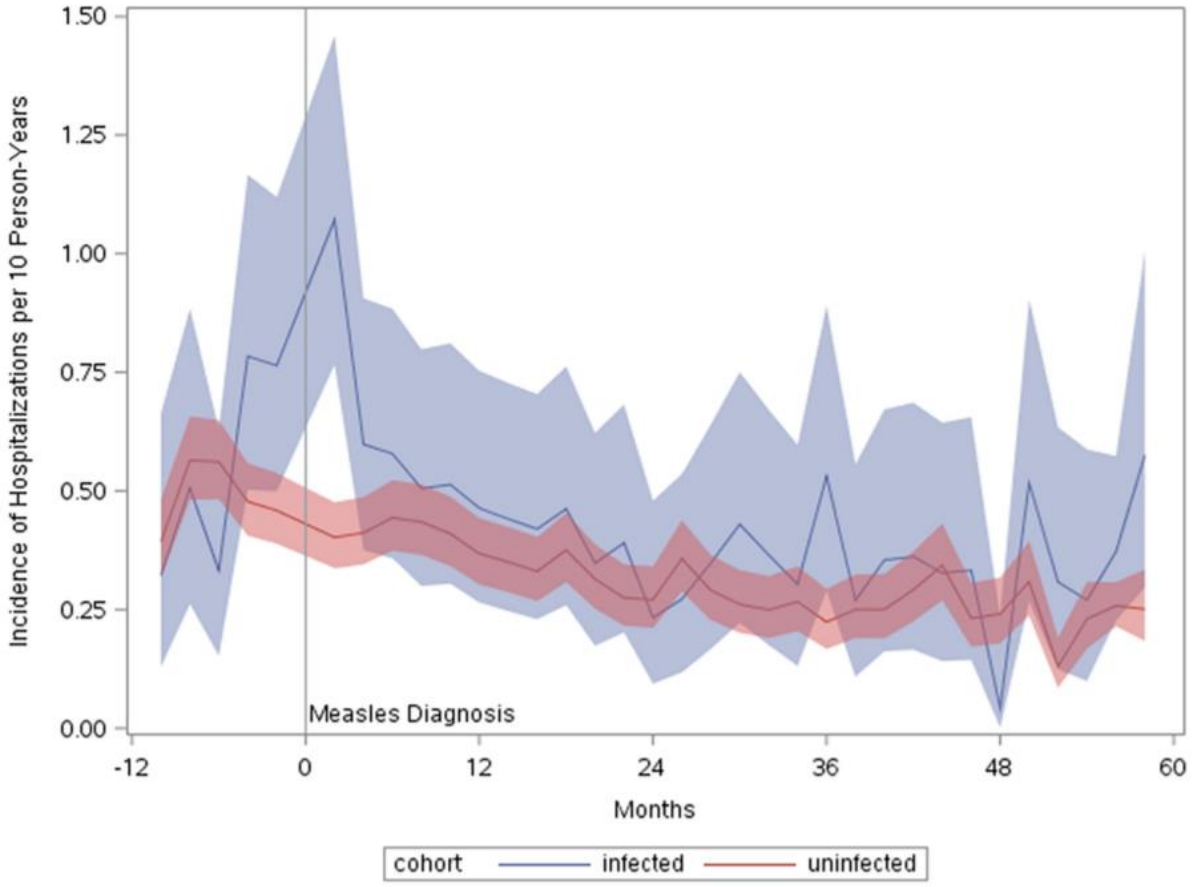


Figure 5 | Hospitalisations in measles patients and matched controls. Incidence rates of hospitalisations in children diagnosed with measles (blue lines) or matched controls (red lines) per 100,000 person-years, plotted by time (in months) before or after diagnosis of measles. The vertical dotted line indicates the time point of diagnosis in the measles patients. The shaded areas represent 95% confidence intervals.

by our large sample size. Children diagnosed with measles were hospitalized more frequently than children free of measles although this was only significant in the first month following infection. When we excluded the first month post measles, the time to first hospitalisation did not differ between the measles group and the non-measles group. This could be explained, at least in part, by a survival bias, whereby a disproportionately large number of measles cases entered the hospital during the first month, and these may have represented the most severe cases. Additionally, a lack of effect on hospitalization after the first month was likely a result of the low overall number of hospitalisations in our cases and controls. We acknowledge that the first interval spanning one month to one year post-measles is wide and have conducted analysis using smaller intervals, the results of which can be found in supplementary material.

The incidence rates of infections, anti-infective prescriptions and hospitalisations in the measles group appear to increase prior to the index date, that is, before they got measles (Figures 3-5). This could partially be explained by a lag time between a suspected diagnosis and a definite diagnosis. In some instances, a GP may have coded a definite diagnosis on the date a confirmation had been received either from the lab or from the hospital. For some outcomes however the rise in incidence begins months before diagnosis. Validation studies to assess the accuracy of the date of diagnosis using this type of database are lacking.

To be considered a new event, prescriptions only had to be given on a different day. Acknowledging that a prescription can be changed if there is poor response or allergy to the first drug, we also examined the effect of anti-infective prescription, considering a 14-day interval between anti-infective prescriptions. This did not change the significance or direction of any result (results not shown). Both groups revealed similar trends with the rate of consultation before index date higher than after index date.[16 17] This is most likely related to age. Although measles is a statutory notifiable infectious disease under EU legislation,[18] an underreporting of (severe) cases, who might have by-passed the GP and gone directly to the hospital cannot be ruled out. Also it is possible that a mild measles infection would not have prompted a visit to the GP and may have gone undetected as well.[19 20] This means that we may have missed some children with measles. To provide additional assurance that controls were children truly free of measles, controls had to have at least one measles-containing vaccination prior to the index date. An advantage of this type of observational study is that it is not necessary to identify all outcomes in all children in order to obtain an unbiased estimate. A key assumption however, is comparability of children with measles and children free of measles. In order to ensure that the children with measles and the children free of measles were comparable we matched them on confounding factors such as age, sex, GP practice, and calendar time. We also considered including experiencing an excluding event (i.e. an immune-compromising condition, or immune suppressive treatment) as a censoring variable but determined this was not consistent with our matching strategy – the groups were matched to be comparable at index. Nevertheless, we acknowledge that it is possible that confounding due to differences in underlying health status, social background, lifestyle, health seeking behavior, and health care utilization between children with measles and children free of measles may have occurred. The complexity of these factors makes them difficult to control. We attempted to overcome the confounding effect of underlying health status by excluding children with a history of immune-compromising conditions, and controlling for co-morbidities such as cardiovascular disease, and respiratory disease. We assessed social background and lifestyle by testing for differences in social deprivation within a particular zip code and matching on practice. Certain children may have had a lower threshold for visiting the GP and therefore may have had a higher likelihood of receiving a diagnosis of measles (particular during an outbreak) and may also have been diagnosed more frequently with other infectious diseases and/or may have received a prescription for anti-infectives more frequently. To investigate this, we included GP consultation rate in the year prior to cohort entry as a covariate in each of our models. In the unmatched Poisson analyses, we did not control for all potential confounders. Because 472 unique practices were represented in the cohort, it was impossible to control for practice. Similarly, the 25 years included in the study period make control for calendar year infeasible unless calendar year is treated as a continuous variable, which would require the assumption of a linear relationship between year and log(events). To address the potential effect of calendar time, we have conducted analyses stratified by calendar period (before 2005 and after 2004) and included these results in supplementary material.

Because vaccinated and unvaccinated children may differ in their health seeking behaviour, or likelihood of acquiring infectious disease, we conducted a sensitivity analysis in two strata: 1) restricting to only those children who had received a measles vaccination prior to the index date, and 2) restricting to only those children who were unexposed to measles vaccination prior to the index date. Results from both sub analyses were in line with the findings from the main analysis with the exceptions that the period of increased risk for infections did not extend past one year and no increased risk for hospitalizations was detected when analysis was limited to measles vaccinated children.

We did not adjust for measles vaccination after index date in the context of post exposure prophylaxis because many exposed persons are not identified until more than 72 hours after initial exposure, which is too late for prophylaxis with measles vaccine.[2] Post-hoc analysis of vaccine adherent vs non-adherent children revealed an increased rate of anti-infective prescriptions in non-adherent children but no difference for other outcomes.

We conclude that our results support the hypothesis that infection with measles is associated with long-term increased risk of other infectious diseases, and that by preventing measles, vaccination is associated with non-specific heterologous improvements in health. However, because all of the non-measles controls received vaccination, we cannot rule out a direct benefit of vaccination to boost heterologous immune function, as has been suggested[21 22]. Nonetheless the results fit with what would be expected from animal models and what has been shown in ecological studies, and warrants further investigation into the long-term consequences of viral infections, particularly those with heightened tropism for immune memory cells, on host resistance.

Contributors

Conceived and designed the study: MM, BG, DV, RS. Code selection: KG, GM, DV, RS. Data extraction and statistical analysis: CD, GM, MR. Interpretation of data: KG, CD, GM, MR, DW, MM, BG, MS, DV, RS. Authored draft paper: KG, CD. Critical revisions of manuscript: KG, CD, GM, MR, DV, RS, MM, BG. Study supervision: DV, RS, MS. Obtained funding: None.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval

Scientific Review Committee, Cegedim Strategic Data Medical Research UK, now IMS Health (THIN SRC reference number: 15-006).

Data sharing statement

No additional unpublished data from the study are available. All data are contained in the manuscript and the supplementary data files.

References

1. Rota PA, Moss WJ, Takeda M, de Swart RL, Thompson KM, Goodson JL. Measles. Nat Rev Dis Prim 2016; 2:16049. PubMed PMID: 27411684.

2. Strebel PM, Papania MJ, Fiebelkorn AP, Halsey NA. 20 - Measles vaccine A2 - Plotkin, Stanley A. In: Orenstein WA, Offit PA, editors. Vaccines (Sixth Edition). London: W.B. Saunders; 2013. p. 352-87.

3. Griffin DE. Measles virus-induced suppression of immune responses. Immunol Rev 2010; 236:176-89. PubMed PMID: 20636817.

4. De Vries RD, de Swart RL. Measles immune suppression: functional impairment or numbers game? PLoS Pathog 2014; 10(12):e1004482. PubMed PMID: 25522010.

5. Simons E, Ferrari M, Fricks J, Wannemuehler K, Anand A, Burton A, et al. Assessment of the 2010 global measles mortality reduction goal: results from a model of surveillance data. Lancet 2012; 379(9832):2173-8. PubMed PMID: 22534001.

6. Aaby P, Bukh J, Lisse IM, Smits AJ. Measles vaccination and reduction in child mortality: a community study from Guinea-Bissau. J Infect 1984; 8(1):13-21. PubMed PMID: 6699411.

7. Mina MJ, Metcalf CJ, de Swart RL, Osterhaus AD, Grenfell BT. Long-term measles-induced immunomodulation increases overall childhood infectious disease mortality. Science 2015; 348(6235):694-9. PubMed PMID: 25954009.

8. IMS Health UK: IMS Health Incorporated; 2015 [cited 2016 09 Jun 2016]. Available from: http://csdmruk.cegedim.com/.

9. Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. Pharmacoepidemiology and drug safety. 2007; 16(4):393-401. PubMed PMID: 17066486.

10. Bourke A, Dattani H, Robinson M. Feasibility study and methodology to create a quality-evaluated database of primary care data. Informatics in primary care. 2004; 12(3):171-7. PubMed PMID: 15606990.

11. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Informatics Prim Care 2011;19(4):251-5. PubMed PMID: 22828580.

12. Booth N. What are the Read Codes? Health Lib Rev 1994;11(3):177-82. PubMed PMID: 10139676.

13. Townsend P, Phillimore P, Beattie A. Health and deprivation. Inequality and the North. London: Chroom Helm Ltd; 1987.

14. Goldstein BA, Bhavsar NA, Phelan M, Pencina MJ. Controlling for Informed Presence Bias Due to the Number of Health Encounters in an Electronic Health Record. Am J Epidemiol 2016; 184(11):847-55. PubMed PMID: 27852603.

15. UK Government Web Archive. Measles, Epidemiological data: 1996-2013:; 2014 [updated 04 March 2014; cited 2016 09 Jun 2016]. Available from: http://webarchive.nationalarchives.gov.uk/20140505192926/http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Measles/EpidemiologicalData/

16. Janicke DM, Finney JW, Riley AW. Children's health care use: a prospective investigation of factors related to care-seeking. Medical Care 2001 Sep;39(9):990-1001. PubMed PMID: 11502956.

17. Freid VM, Makuc DM, Rooks RN. Ambulatory health care visits by children: principal diagnosis and place of visit. Vital and health statistics Series 13, Data from the National Health Survey. 1998 May(137):1-23. PubMed PMID: 9631643.

18. EU. Decision No. 2119/98/EC of the European Parliament and of the Council of 24 September 1998 setting up a network for the epidemiological surveillance and control of communicable diseases in the Community. 1998; Available from: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A31998D2119

19. Sonoda S, Nakayama T. Detection of measles virus genome in lymphocytes from asymptomatic healthy children. J Med Virol 2001; 65(2):381-7. PubMed PMID: 11536248.

20. Sanders R, Dabbagh A, Featherstone D. Risk analysis for measles reintroduction after global certification of eradication. J Infect Dis 2011; 204 Suppl 1:S71-7. PubMed PMID: 21666216.

21. Netea MG, Quintin J, van der Meer JW. Trained immunity: a memory for innate host defense. Cell Host Microbe 2011;9(5):355-361.

22. Blok BA, Arts RJ, van Crevel R, Benn CS, Netea MG. Trained innate immunity as underlying mechanism for the long-term, nonspecific effects of vaccines. J. Leukoc Biol 2015; 98(3):347-356.

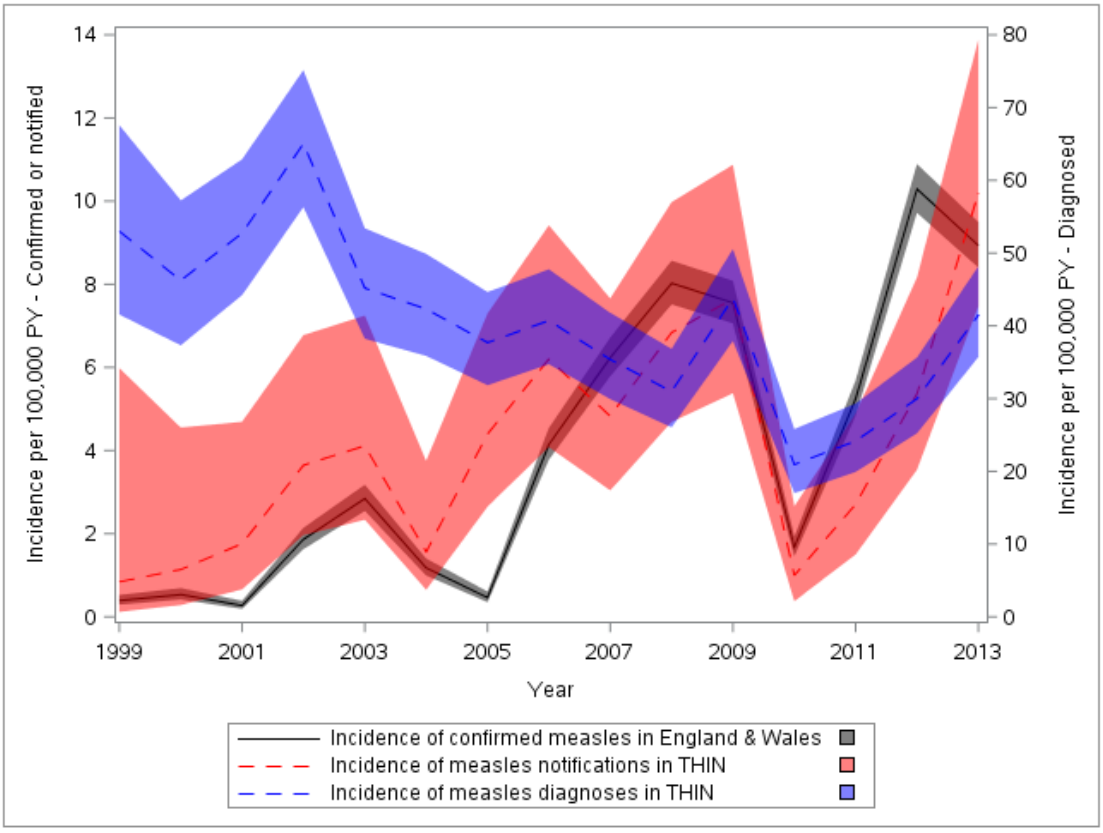


Fig S1 Observed IR in the THIN database and in the population of England and Wales [15]

S3 Table. Most commonly diagnosed infectious diseases by cohort

Measles Infected Subjects	N	Percent	Non-Measles Subjects	N	Percent
Upper respiratory tract infection NOS	1,344	35.6688	Upper respiratory tract infection NOS	9,337	34.2429
Upper respiratory infection NOS	877	23.2749	Upper respiratory infection NOS	6,316	23.1635
Chickenpox - varicella	319	8.4660	Chickenpox - varicella	2,568	9.4180
Viral infection NOS	268	7.1125	Viral infection NOS	1,916	7.0268
Molluscum contagiosum	152	4.0340	Molluscum contagiosum	1,408	5.1638
Non-specific viral rash	118	3.1316	Non-specific viral rash	803	2.9450
Viral illness	103	2.7335	Viral upper respiratory tract infection NOS	625	2.2921
Viral upper respiratory tract infection NOS	81	2.1497	Viral illness	608	2.2298
Viral gastroenteritis	72	1.9108	Viral gastroenteritis	507	1.8594
Chickenpox	56	1.4862	Coryza - acute	430	1.5770
Common cold	55	1.4597	Chickenpox	416	1.5257
Coryza - acute	48	1.2739	Flu like illness	329	1.2066
Flu like illness	47	1.2473	Common cold	305	1.1186
Slapped cheek syndrome	26	0.6900	Slapped cheek syndrome	177	0.6491
Acute bronchiolitis	20	0.5308	Acute bronchiolitis	133	0.4878
Streptococcal tonsillitis	15	0.3981	Rhinitis – acute	110	0.4034

Scarlet fever - scarlatina	13	0.3450	Herpes simplex	107	0.3924
Rhinitis - acute	13	0.3450	Scarlet fever - scarlatina	83	0.3044
Diarrhoea & vomiting -infect	7	0.1858	Influenza	68	0.2494
Fifth disease	7	0.1858	Diarrhoea & vomiting -infect	62	0.2274
Nasal catarrh - acute	7	0.1858	Streptococcal tonsillitis	57	0.2090
Rubella	6	0.1592	Nasal catarrh - acute	56	0.2054
Roseola infantum	6	0.1592	Pneumonia due to unspecified organism	48	0.1760

S4 Table. Most commonly prescribed anti-infectives by cohort

Measles Infected Subjects	N	Percent	Non-Measles Subjects	N	Percent
Amoxicillin 125mg/5ml oral suspension sugar free	2,246	0.2194	Amoxicillin 125mg/5ml oral suspension sugar free	14,871	0.2184
Amoxicillin 125mg/5ml oral suspension	1,238	0.1209	Amoxicillin 125mg/5ml oral suspension	7,500	0.1101
Chloramphenicol 0.5% eye drops	606	0.0592	Chloramphenicol 0.5% eye drops	4,111	0.0604
Trimethoprim 50mg/5ml oral suspension sugar free	567	0.0554	Phenoxymethylpenicillin 125mg/5ml oral solution	3,527	0.0518
Phenoxymethylpenicillin 125mg/5ml oral solution	458	0.0447	Fusidic acid 2% cream	2,921	0.0429
Erythromycin ethyl succinate 125mg/5ml oral suspension	440	0.0430	Trimethoprim 50mg/5ml oral suspension sugar free	2,791	0.0410
Amoxicillin 250mg/5ml oral suspension sugar free	322	0.0314	Erythromycin ethyl succinate 125mg/5ml oral suspension	2,476	0.0364
Flucloxacillin 125mg/5ml oral solution	235	0.0230	Amoxicillin 250mg/5ml oral suspension sugar free	2,041	0.0300
Co-amoxiclav 125mg/31mg/5ml oral suspension sugar free	234	0.0229	Flucloxacillin 125mg/5ml oral solution	1,874	0.0275
Fusidic acid 2% cream	196	0.0191	Chloramphenicol 1% eye ointment	1,292	0.0190

S5 Table. Hazard Ratio of measles versus non-measles, for outcomes first infection and first prescription

Outcome	Hazard Ratio (95% Confidence Interval)	*P*-value
Infections	1.20* (1.13 to 1.28)	<0.0001
Anti-infective prescriptions	1.24** (1.18 to 1.31)	<0.0001

*Controlled for number of diagnosed infections in the year prior to index and number of consultations in the previous year.
**Controlled for history of respiratory illness, number of anti-infective prescriptions in the year prior to index, and number of consultations in the year prior to index categorized by quintiles

S6 Table. Incidence rate ratios (IRRs) of events of interest in shorter time periods within the first year following measles infection

Time Period	Analysis	Incidence Rate Ratio (95% Confidence Interval)		
		Infections	Anti-infective prescriptions	Hospitalisation
Days 0 to 31	Primary	1.43(1.22 to 1.68)	3.60 (3.32 to 3.92)	2.85 (1.73 to 4.70)
Days 32 to 64	Primary	1.19 (0.99-1.43)	1.34 (1.18 to 1.52)	1.45 (0.76 to 2.76)
Days 65 to 182	Primary	1.28 (1.16 to 1.41)	1.26 (1.16 to 1.36)	1.13 (0.78 to 1.66)
Days 183 to 365	Primary	1.18 (1.07 to 130)	1.20 (1.12 to 1.29)	1.08 (0.77 to 1.51)
Days 366 to 913	Primary	1.10 (1.02 to 1.18)	1.21 (1.13 to 1.28)	1.08 (0.80 to 1.46)
Days 914 to 1826	Primary	1.15 (1.06 to 1.25)	1.15 (1.07 to 1.24)	1.24 (0.93 to 1.67)

* Primary and analyses were adjusted for: Frequency of consultations in the year prior to index, frequency of the outcome of interest in the year prior to index, history of cardiovascular malformation, history of respiratory disease, age, and sex.

S7 Table. Incidence rate ratios (IRRs) of events of interest in pre-specified time periods following measles infection, stratified by calendar year (pre 2005 vs. 2005 and later)

Calendar Period	Time Period	Analysis	Incidence Rate Ratio (95% Confidence Interval)		
			Infections	Anti-infective prescriptions	Hospitalisation
1999-2004	Days 0 to 31	Primary	1.73 (1.39 to 2.16)	3.55 (3.12 to 4.04)	1.55 (0.63 to 3.82)
	Days 32 to 365	Primary	1.29 (1.17 to 1.42)	1.29 (1.18 to 1.40)	1.13 (0.74 to 1.74)
	Days 366 to 913	Primary	1.13 (1.02 to 1.24)	1.26 (1.15 to 1.38)	1.16 (0.70 to 1.93)
	Days 914 to 1826	Primary	1.22 (1.09 to 1.36)	1.20 (1.09 to 1.32)	1.27 (0.77 to 2.10)
2005-2013	Days 0 to 31	Primary	1.17 (0.93 to 1.48)	3.63 (3.26 to 4.06)	3.63 (1.99 to 6.59)
	Days 32 to 365	Primary	1.15 (1.04 to 1.26)	1.21 (1.12 to 1.31)	1.06 (0.77 to 1.46)
	Days 366 to 913	Primary	1.07 (0.95 to 1.19)	1.16 (1.07 to 1.27)	0.91 (0.63 to 1.31)
	Days 914 to 1826	Primary	1.05 (0.93 to 1.20)	1.09 (0.98 to 1.22)	1.08 (0.77 to 1.51)

Figure S2. Kaplan-Meier plot for Infections

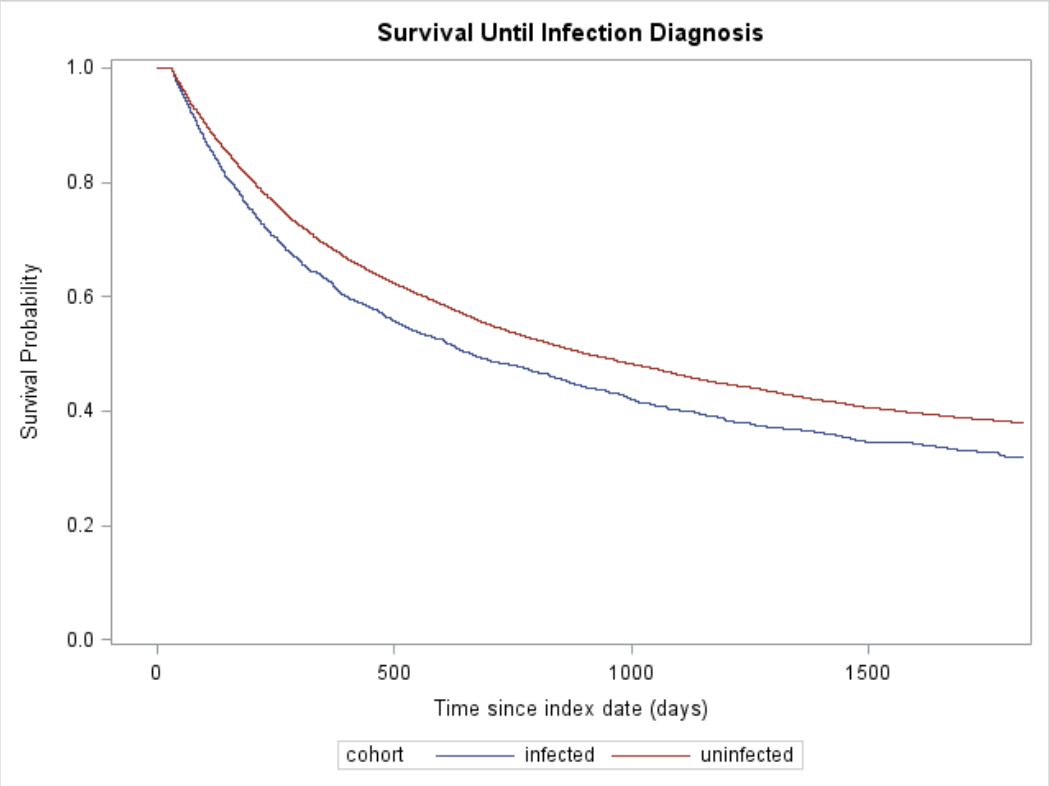


Figure S2. Kaplan-Meier plot for Anti-infective Prescriptions

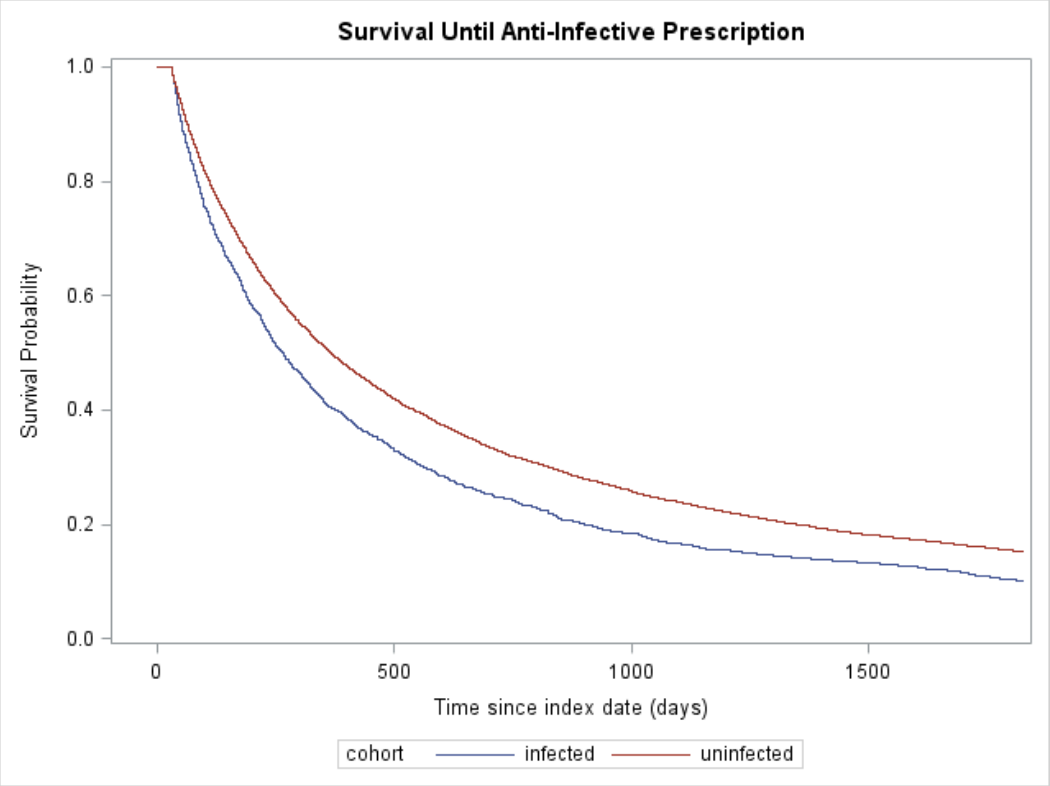
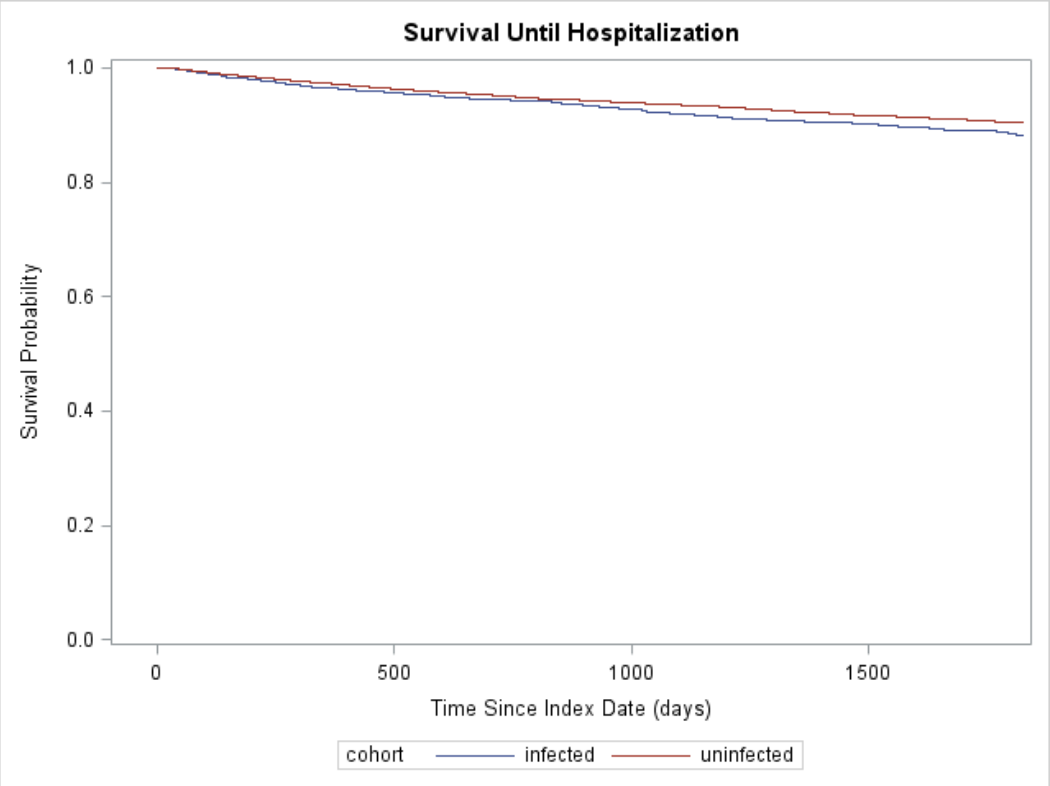


Figure S3. Kaplan-Meier plot for Hospitalizations



CHAPTER 5 COLLABORATIVE STUDIES TO ASSESS VACCINE EFFECTS

5.1 International collaboration to assess the risk of Guillain-Barré Syndrome following influenza A (H1N1) 2009 monovalent vaccines

Caitlin N. Dodd, Silvana A. Romio

Steven Black, Claudia Vellozzi

Nick Andrews, Miriam Sturkenboom

Patrick Zuber, Wei Hua

Jan Bonhoeffer, Jim Buttery

Nigel Crawford, Genevieve Deceuninck

Corinne de Vries, Philippe De Wals

Patrick Garman, M. Victoria Gutierrez Gimeno

Harald Heijbel, Kwan Hur

Anders Hviid, Jeffrey Kelman

Tehri Kilpi, SK Chuang

Kristine Macartney, Melisa Rett

Vesta Richardson Lopez-Callada, Daniel Salmon

Francisco Gimenez Sanchez, Nuria Sanz

Barbara Silverman, Jann Storsaeter

Umapathi Thirugnanam, Nicoline van der Maas

Katherine Yih, Tao Zhang

Hector Izurieta, and the Global H1N1 GBS Consortium

Vaccine 31.40 (2013): 4448-4458.

Abstract:

Background:

The global spread of the 2009 novel pandemic influenza A (H1N1) virus led to the accelerated production and distribution of monovalent 2009 Influenza A (H1N1) vaccines (pH1N1). This pandemic provided the opportunity to evaluate the risk of Guillain-Barré syndrome (GBS), which has been an influenza vaccine safety concern since the swine flu pandemic of 1976, using a common protocol among high and middle-income countries. The primary objective of this project was to demonstrate the feasibility and utility of global collaboration in the assessment of vaccine safety, including countries both with and without an established infrastructure for vaccine active safety surveillance. A second objective, included *a priori,* was to assess the risk of GBS following pH1N1 vaccination.

Methods:

The primary analysis used the self-controlled case series (SCCS) design to estimate the relative incidence (RI) of GBS in the 42 days following vaccination with pH1N1 vaccine in a pooled analysis across databases and in analysis using a meta-analytic approach.

Results:

We found a relative incidence of GBS of 2.42 (95% CI 1.58-3.72) in the 42 days following exposure to pH1N1 vaccine in analysis of pooled data and 2.09 (95% CI 1.28-3.42) using the meta-analytic approach.

Conclusions:

This study demonstrates that international collaboration to evaluate serious outcomes using a common protocol is feasible. The significance and consistency of our findings support a conclusion of an association between 2009-10 H1N1 vaccination and GBS. Given the rarity of the event the relative incidence found does not provide evidence in contradiction to international recommendations for the continued use of influenza vaccines.

Introduction

Assessment of vaccine safety post-licensure requires well-designed epidemiological studies, which can be challenging for many countries due to scarcity of available data. Therefore, spontaneous reporting systems are more commonly used for post-marketing safety monitoring[1]. Traditionally, vaccines have been manufactured and introduced in the United States (US) and Europe before introduction in other countries, hence US and European vaccine safety monitoring capacity has served the global need to evaluate the safety of new vaccines[1]. However, vaccines are now being manufactured and introduced in several countries outside the US and Europe[2], requiring the development of vaccine safety monitoring systems globally to assure the safety of the world's vaccine supply and maintain trust in immunization programs. International vaccine safety collaborations can help build vaccine safety monitoring infrastructure and capacity and provide a means to assess rare adverse events following immunization (AEFI) in countries that now have limited capacity[3].

To demonstrate that international collaboration is feasible for vaccine safety studies to investigate rare, serious and clinically complex AEFI, a group of vaccine safety researchers conducted a proof of concept collaborative vaccine safety study using a standard protocol[4-6]. A steering group[1] from the World Health Organization (WHO), United States Food and Drug Administration (FDA) and Centers for Disease Control and Prevention (CDC), European CDC, Erasmus Medical Center, Cincinnati Children's Hospital, and the Brighton Collaboration[7], provided standardized methods and definitions for a study that included investigators from Australia, Canada, China, Denmark, Finland, France, Israel, Mexico, The Netherlands, Norway, Singapore, Spain, Sweden, the United Kingdom, and the United States.

The global spread of the 2009 novel pandemic influenza A (H1N1) virus[8] led to the accelerated production of monovalent 2009 Influenza A (H1N1) vaccines (pH1N1) by manufacturers in the Americas, Europe, and Asia[9]. Rapid and extensive vaccine administration was implemented worldwide. This pandemic provided the opportunity to evaluate the risk of Guillain-Barré syndrome (GBS), an acute polyradiculoneuropathy, following receipt of these vaccines using a common protocol among high and middle-income countries and to assess the feasibility of this collaborative effort.[10] Several factors contributed to choosing this vaccine and this adverse event (GBS) to test the new consortium: First, GBS has been an influenza vaccine safety concern since 1976, when an elevated risk of GBS was identified following the "swine-flu" influenza vaccine[11]; second, case definitions and classifications for GBS are available, providing a tool for standardized assessment across sites[12]; third, since almost all GBS cases are hospitalized, unbiased case ascertainment could be achieved using hospital databases; and finally, since GBS is rare, assessment of risk would benefit from the increased sample size and statistical power that could result from an international collaboration.

The primary objective of this project was to demonstrate the feasibility and utility of global collaboration in the assessment of vaccine safety, including countries both with and without an established infrastructure for vaccine safety active surveillance. A second objective, included *a priori,* was to assess the relative risk of GBS following pH1N1 vaccination.

Methods

We chose the self-controlled case series (SCCS) design[13] to estimate the relative incidence (RI) of GBS in the 42 days following vaccination with pH1N1 vaccine. We chose this case-only analytic approach because it can be implemented in populations with varying levels of infrastructure for conducting epidemiologic studies; specifically,

[1] Steven Black, Caitlin Dodd (Cincinnati Children's Hospital), Hector Izurieta (FDA), Patrick Zuber (WHO)

it does not require the availability of accurate population denominators which are difficult to obtain in many countries[9,10]. The case series approach includes only individuals who experienced the event of interest (GBS) in the analysis. Each individual's person-time during follow-up is divided into predefined vaccine exposed and non-exposed periods. Each GBS case then falls into a risk or non-risk window and contributes exposed and non-exposed time. Unvaccinated GBS cases contribute to the estimation of other time-varying covariates such as seasonality. Data are analyzed by conditional Poisson regression. The SCCS design requires that cases be ascertained completely and in an unbiased manner and that the probability of exposure is not affected by occurrence of the event of interest. Apart from its intrinsic resource efficiency, this design also controls for measured or unmeasured within-person non-time dependent confounding characteristics, including demographics and chronic co-morbid conditions, genetic susceptibility, and others[10].

Study population

As shown in Table 1, 15 countries with available data and willingness to participate contributed data for this study: Cases that met inclusion criteria for this study from Australia, Canada, China, Israel, Mexico, Singapore, Spain, and the United States and from the European Vaccine Adverse Event Surveillance and Communication (VAESCO) consortium[14] (http://vaesco.net) (Denmark, Finland, France, The Netherlands, Norway, Sweden, and the United Kingdom) were included. Australian data were provided by hospitals in the state of Victoria (including Melbourne), Sydney, Perth, and Adelaide. Canadian data were provided from the entire province of Quebec; Chinese cases were contributed by sentinel hospitals in Hong Kong and Shanghai. Israeli data were provided by Maccabi, a national health maintenance organization (HMO) and Mexican data were contributed from Mexico City and surrounding rural areas. Singapore data were provided from one rural and one central hospital. Spanish data were provided by hospitals in Almeria, Barcelona, and Valencia. US data were contributed from the Department of Defense (DoD), the Department Veterans Affairs (VA), the Vaccine Safety Datalink (combined hospitalization and vaccination data from a collaboration of 8 health care organizations), Medicare, and the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) Program[15, 16].

Case ascertainment and classification

The specific method of case ascertainment varied from country to country with some countries identifying potential cases through administrative databases whereas other countries reviewed hospital discharge logs manually (Table 2). Databases from all US sites other than the DoD contained only vaccinated cases and were limited to post-vaccination follow up time. Each site independently defined an observation period, ranging from 4 to 18 months, during which cases were obtained. The number of medical records requested and reviewed was not reported by the participating sites. Countries that did not identify and classify cases using a procedure compatible with the common protocol were excluded from this study (Table 1). For countries in which cases were ascertained through active surveillance, only those cases with verified hospital admission were included.

Diagnostic codes for GBS (ICD-9 code of 357.0, ICD-10 code of G61.0, or Read codes F370*) were used to identify potential cases for review. Cases identified exclusively through specialty network reporting or through other passive reporting method were excluded. The specific method of case ascertainment used and number of cases identified are shown in Table 2.

All cases were classified locally according to Brighton Collaboration criteria [12] (Table 3). All cases meeting Brighton level 1, 2, 3 were considered confirmed and included in the primary analysis. A secondary analysis also included Brighton categories 4 and 4A. Category 4A was specifically defined for this international study and included cases diagnosed by a neurologist but for whom the medical chart did not provide sufficient information for the study reviewers to classify the case according to Brighton Collaboration criteria (Table 3).

Table 3: Brighton Collaboration Case Definition for Guillain-Barré Syndrome

Level	Requirements
1	Clinical, electrophysiological, AND cerebrospinal fluid (CSF) data
2	Clinical, data and electrophysiological OR cerebrospinal fluid (CSF) data
3	Clinical data
4	Information available is insufficient for levels 1-3, but no other diagnosis is apparent or warranted.
4A	Diagnosis was made by a neurologist, insufficient diagnostic data available in the medical chart (adopted specifically for this study)

Vaccination status

Vaccination status was obtained through automated immunization registries or databases when available or through a review of the patient's vaccination record where it was not. In some cases, receipt of vaccine was obtained through self-report and subsequently verified in the vaccination record. The method used for each participating site is shown in Table 2. All subjects had presence or absence of pH1N1 vaccine recorded with the date of exposure. Patients for whom no record of seasonal influenza vaccination were available were considered unexposed to the seasonal vaccine.

Covariates

Presence or absence of risk factors for GBS, including preceding gastrointestinal infections and respiratory infections, were collected for the 30 days prior to diagnosis through chart review or recall by the subject when chart review was not possible. While a standard abstraction form was not used, a standard case report form was used to record all data used in analysis. Since presence or absence of the potential infective episodes in the 30 days preceding GBS rather than exact dates of the episodes was collected, these infections could not be controlled for in the analysis but could be studied as potential effect modifiers. To control for circulation of the pandemic influenza virus we used seasonality as a proxy. This was accomplished by defining the peak of influenza season for each site as the period during which > 15% of all surveillance influenza laboratory tests were positive, and estimating the relative incidence of GBS in this peak season. This produced time periods for each site defined either as "high influenza circulation" or "low influenza circulation". Although the seasonality of GBS is not strictly related to influenza infection, influenza surveillance data provided an efficient means by which to uniformly define seasonal periods across continents and hemispheres. This formulation allowed for a peak influenza season specific to each site while also allowing for a common estimate of the effect of seasonality across sites. Data to determine these periods of influenza circulation were obtained from publicly available governmental influenza surveillance, where available. For some sites, it was necessary to obtain these data from influenza surveillance conducted at the site.

Data collection and sharing

All data, with the exception of data from the VAESCO consortium, were uploaded to a secure WHO workspace where it was checked for quality and completeness by study group statisticians; VAESCO data were maintained by

the VAESCO data management center at Erasmus Medical Center. All data was de-identified prior to submission. Institutional Review Board approval was obtained for those sites at which the study was not considered exempt. Inclusion and exclusion criteria as well as protocol and statistical analysis considerations were discussed on bi-weekly telephone conferences with all sites beginning in January, 2010.

Table 1: Database Inclusion and Exclusion in Primary and Sensitivity Analyses by Country

Excluded Databases

Country	Database	Criteria For Exclusion
France		Patient Consent Required (potential bias)
Israel	Maccabi	Brighton Collaboration criteria not provided.
Mexico	Mexico City	Relative incidence found to be an outlier compared to all other study site relative incidence findings
	Mexican States	Data were obtained solely from a specialist network (potential bias)
Norway		Data were obtained solely from a specialist network (potential bias)
Sweden		Patient consent required with potential bias

Included Databases

Country	Database	Analyses in which database is included	Criteria For Exclusion where applicable
Australia	Adelaide	- Primary Analysis - All sensitivity analyses	
	MCRI[a]	- Primary Analysis - All sensitivity analyses	
	Sydney	- Primary Analysis - All sensitivity analyses	
Canada	Quebec	- Primary Analysis - All sensitivity analyses	
China	Hong Kong	- Primary Analysis - All sensitivity analyses	
	Shanghai	- Analyses through Brighton 4A	All reported cases were Brighton level 4 or 4A
Denmark		- Primary Analysis - All sensitivity analyses	
Finland		- Primary Analysis - All sensitivity analyses	
The Netherlands	IPCI[a]	- Primary Analysis - All sensitivity analyses	
Singapore	NNI/CGH[a]	- Primary Analysis - All sensitivity analyses	
	NNI/TTSH[a]	- Analyses through Brighton 4A	All reported cases were Brighton level 4 or 4A
Spain	Almeria	- Analyses through Brighton 4A	All reported cases were Brighton level 4 or 4A
	Barcelona	- Primary Analysis - All sensitivity analyses	
	Valencia	- Primary Analysis	

		- All sensitivity analyses	
The United Kingdom	CPRD[a]	- Analyses through Brighton 4A	All reported cases were Brighton level 4 or 4A
The United States	DoD[a]	- Primary Analysis - All sensitivity analyses	
	Medicare	- Vaccinated-cases only analyses, excluded from all others	Database contained only cases post-vaccination
	PRISM[a]	- Vaccinated-cases only analysis, excluded from all others	Database contained only cases post-vaccination
	VA[a]	- Vaccinated-cases only analysis, excluded from all others	Database contained only cases post-vaccination
	VSD[a]	- Vaccinated-cases only analysis, excluded from all others	Database contained only cases post-vaccination

[a] Murdoch Children's Research Institute (MCRI), Maccabi Health Maintenance Organization (Maccabi), Integrated Primary Care Information Database (ICPI), National Neuroscience Institute Singapore General Hospital (NNI/CGH), National Neuroscience Institute Tan Tock Seng Hospital (NNI/TTSH), CPRD (Clinical Practice Research Datalink), Department of Defense (DoD), Post-Licensure Rapid Immunization Safety Monitoring (PRISM), Department of Veterans Affairs (VA), Vaccine Safety Datalink (VSD)

Table 2: Characteristics of Databases Included in Primary or Sensitivity Analyses by Country

Country	Database	Dates of Observation	Number of Cases	Case Ascertainment	Vaccination Status Ascertainment
Australia	Adelaide	9/30/2009 – 9/30/2010	1	Administrative Database, active prospective surveillance	Vaccine Registry, Self-Report, Outpatient Chart Review
	MCRI[a]	9/30/2009 – 9/30/2010	54	Administrative Database, active prospective surveillance	Vaccine Registry, Self-Report, Outpatient Chart Review
	Sydney	9/30/2009 – 9/30/2010	5	Administrative Database, active prospective surveillance	Vaccine Registry, Self-Report, Outpatient Chart Review
Canada	Quebec	10/13/2009 – 3/31/2010	80	Administrative Database, active prospective surveillance	Vaccine registry
China	Hong Kong	12/21/2009 – 6/30/2010	20	Hospital log review	Outpatient Chart Review, Self-Report
	Shanghai	1/1/2009 – 7/1/2010	22	Administrative Database	Outpatient Chart Review

Denmark		11/1/2009 – 11/1/2010	31	National Patient Register using primary discharge diagnoses	Vaccine Registry
Finland		11/1/2009 – 11/1/2010	29	Hospital discharge and hospital outpatient records, primary diagnoses	Vaccine registry
The Netherlands	IPCI[a]	11/1/2009 – 11/1/2010	80	Identified prospectively through neurologists. Completeness verified retrospectively against claims codes in each hospital.	GP[b] medical record
Singapore	NNI/CGH[a]	11/5/2009 – 8/31/2010	6	Administrative Database	Outpatient Chart Review, Self-Report
	NNI/TTSH[a]	11/5/2009 – 8/31/2010	13	Administrative Database	Hospital Medical Records
Spain	Almeria	11/1/2009 – 4/30/2010	8	Administrative Database	Outpatient Chart Review
	Barcelona	11/1/2009 – 4/30/2010	14	Administrative Database	Outpatient Chart Review
	Valencia	11/1/2009 – 4/30/2010	10	Administrative Database	Vaccine Registry
The United Kingdom	CPRD[a]	11/1/2009 – 11/1/2010	40	Automated GP records	GP[b] records
The United States	DoD[a]	11/1/2009 – 4/30/2010	6	Administrative Database Electronic Medical Records	Vaccine Registry
	Medicare	11/1/2009 – 4/30/2010	39	Administrative Database	Vaccine Registry
	PRISM[a]	10/22/2009 – 8/7/2010	8	Vaccine Registries and Claims Databases	Electronic Medical Claims
	VA[a]	11/1/2009 –	2	Administrative Database	Vaccine Registry and Administrative

		4/30/2010			Electronic Medical Records	database
	VSD[a]	8/1/2009 4/30/2010	–	11	Administrative Database Electronic Medical Records	Vaccine Registry and Administrative database.

[a]Murdoch Children's Research Institute (MCRI), Integrated Primary Care Information Database (ICPI), National Neuroscience Institute Singapore General Hospital (NNI/CGH), National Neuroscience Institute Tan Tock Seng Hospital (NNI/TTSH), CPRD (Clinical Practice Research Datalink), Department of Defense (DoD), Post-Licensure Rapid Immunization Safety Monitoring (PRISM), Department of Veterans Affairs (VA), Vaccine Safety Datalink (VSD)

[b] General Practitioner

Analysis and Statistical Methods

Data were analyzed using the SCCS method to investigate whether pH1N1 vaccination was associated with an increased risk of GBS during the pre-specified high-risk time window of days 1-42 post vaccination. This period of increased risk was chosen because of evidence from previous published studies[11]. We conducted two co-primary analyses: an analysis pooling all individuals across sites and an analysis using a meta-analytic approach in which estimates of rate ratios from each database were weighted based upon within and between-study errors and subsequently merged. While the pooled analysis provides more power, the meta-analytic approach is more conservative in its estimation as it weighs results from sites with less variability more heavily, thus providing greater assurance that outlying observations or sites with highly variable data will not bias the overall RI estimate. All analyses using the standard SCCS approach excluded the two weeks preceding vaccination from the background period to account for a possible healthy vaccinee effect[17], and controlled for seasonality as defined by periods of circulating influenza.

Only cases meeting Brighton Collaboration criteria level 1-3[12] from databases which included pre and post-vaccination time were included in the primary analysis, which also included a time-varying covariate to assess the effect of seasonality. The date of diagnosis or hospitalization was used as the index date for GBS.

The standard SCCS method is only valid if the occurrence of an event (GBS in this case) does not alter the probability of subsequent exposure. This assumption may be violated since knowledge on the part of practitioners and patients regarding the *a priori* association between GBS and swine flu vaccine[11] may influence vaccination practices, and patients may forego or delay vaccination after GBS diagnosis[18]. For this reason, we also evaluated GBS risk using modifications of the standard SCCS approach to analyze data in which the event-dependent exposures assumption may have been violated. The first of these is the vaccinees-only approach, in which only vaccinated subjects are included and the observation window begins at the date of vaccination. The removal of non-vaccinated cases reduces power to estimate the effect of time-varying covariates such as seasonality[19]. The second modification is the pseudo-likelihood approach, a novel method which considers all cases (vaccinated or not) included in the analysis[20]. In this extension to the standard SCCS, we estimated risk under a counterfactual in which every vaccine exposure is treated as the last possible exposure for that subject. The inclusion of all cases during the entire observation period retains optimal power for the estimation of time-varying covariates[20]. The pseudo-likelihood approach was used to estimate the effect of pH1N1 vaccine with adjustment for seasonality. The

vaccinated-cases only and pseudo-likelihood approaches are equivalent when only one exposure is considered and no time-dependent covariates are included in the analysis. Because there was less than 20% difference between the pH1N1 vaccine-associated relative incidence estimates from the standard and pseudo-likelihood approaches (the difference decided a priori as evidence of bias from contraindication), all subsequent analyses were conducted using the standard approach.

A series of sensitivity analyses were conducted. To assess the effect of seasonal influenza vaccination, the exposure dates were included when known and subjects with missing data were assumed to be non-recipients of seasonal vaccine. We also assessed the risk window in more detail using days 1-7, 8-21, and 22-42 and estimated risks for each different window simultaneously within the same model. Subsequently, cases meeting Brighton criteria through level 4 and 4A were included and an analysis using the date of onset rather than the date of diagnosis was conducted. An analysis limited to Brighton criteria levels 1 and 2 was also conducted. To understand possible effect modifiers and confounders, analyses stratified by sex, age category (< 5, 5-9, 10-18, 19-49, 50-64, and 65+ years), history of GBS, and presence of recent infections were also performed. To capitalize on the diversity of vaccine types and manufacturers in the data set, we also stratified by adjuvanted and non-adjuvanted vaccines.

In the meta-analytic approach, we adjusted for seasonality using month-long periods rather than seasonal peaks since, as data were not being pooled across databases, a common measure of seasonality was not necessary. Estimates of the exposure to pH1N1 vaccine in each database, considering only first dose as exposure of interest, were subsequently combined using a meta-analytic approach with a random effects model in which the estimate from each site is weighted by the inverse of its variance plus the variance of estimates between databases[21]. All analyses were conducted using SAS 9.2 (SAS Institute, Cary NC).

Table 4: Number and Characteristics of Guillain-Barré syndrome cases and Relative Incidence following pH1N1 vaccination by Database from Ten Countries[1]

Country	Database	Number of Cases	Number of pH1N1[2] exposed cases	Number of Cases meeting Brighton Criteria 1-4A	Number of Exposed Cases meeting Brighton Criteria 1-4A	Age (Mean, SD[3])	Sex = M (Frequency, %)	Database-specific GBS[4] RI[5] (CI[6]), Brighton Levels 1-4A
Australia	Adelaide	1	0	1	0	4 (NA[7])	1 (100%)	NA
	MCRI[8]	54	10	54	10	49.2 (23.9)	30 (56%)	2.10 (0.40, 11.05)
	Sydney	5	0	5	0	5.8 (4.7)	3 (60%)	NA
Canada	Quebec	80	43	80	43	49.5 (21.9)	55 (69%)	1.45 (0.70, 3.00)
China	Hong	20	5	20	5	57.8	12 (60%)	0.88 (0.07,

	Kong					(13.3)		11.32)
	Shanghai	22	0	22	0	42.0 (18.1)	15 (68%)	NA
Denmark		31	4	31	4	49.2 (20.2)	14 (45.2)	4.08 (0.48, 34.83)
Finland		29	13	29	13	54.4 (20.8)	12 (41.4)	2.59 (0.77, 8.68)
The Netherlands	IPCI[8]	80	29	79	28	45.0 (20.8)	32 (40.0)	2.81 (1.07, 7.34)
Singapore	NNI/CGH[8]	6	1	6	1	36.3 (16.9)	5 (83%)	3.60 * 10^9 (0, infinity)
	NNI/TTSH[8]	13	2	13	2	54.9 (16.7)	9 (69%)	3.60 * 10^9 (0, infinity)
Spain	Almeria	8	1	8	1	45.9 (20.3)	5 (63%)	1.27 x 10^9 (0, infinity)
	Barcelona	14	0	14	0	38.9 (23.7)	8 (57%)	NA
	Valencia	10	0	10	0	47.8 (22.8)	6 (60%)	NA
The United Kingdom	CPRD[8]	40	3	40	3	45.4 (20.4)	17 (42.5)	10.92 (0.92, 130.13)
The United States	DoD[8]	6	6	6	6	28.8 (6.8)	6 (100%)	8.39 (0.73, 97.00)
	Databases with Vaccinated Cases Only							
	Medicare	39	39	35	35	72.8 (8.5)	25 (64%)	2.04 (0.99, 4.20)
	PRISM[8]	8	8	7	7	48 (33.5)	4 (50%)	2.27 (0.44, 11.77)
	VA[8]	2	2	1	1	60 (12.7)	2 (100%)	NA
	VSD[8]	11	11	11	11	51.5 (24.2)	4 (36%)	3.78 (0.92, 15.61)

[1] Descriptive statistics are for all cases regardless of Brighton Collaboration Criteria unless otherwise specified.

[2] monovalent 2009 (H1N1) A vaccines

[3] Standard Deviation

[4] Guillain-Barré syndrome

[5] Relative Incidence

[6] Confidence Interval

[7] Not Applicable

[8] Murdoch Children's Research Institute (MCRI), Integrated Primary Care Information Database (ICPI), National Neuroscience Institute Singapore General Hospital (NNI/CGH), National Neuroscience Institute Tan Tock Seng Hospital (NNI/TTSH), CPRD (Clinical Practice Research Datalink), Department of Defense (DoD), Post-Licensure Rapid Immunization Safety Monitoring (PRISM), Department of Veterans Affairs (VA), Vaccine Safety Datalink (VSD)

Results

Pooled Data Analysis

In the primary analysis of pooled data limited to Brighton Collaboration criteria levels 1-3 (Table 5), we found a RI of 2.86 (95% CI 1.88- 4.34). In country-specific analyses for the meta-analytic approach, analysis of the data contributed by the Mexico City database was found to have a very high RI of 39.19 (3.74, 410.41). When we excluded Mexican cases from the primary analysis, the estimate of the RI for the primary analysis decreased to 2.42 (1.58, 3.72) (Table 5). Based upon these results, cases from Mexico were excluded from all analyses along with cases from those databases with potential ascertainment bias, resulting in 10 countries contributing cases to the analysis data set (Table 1).

The vaccinated cases only approach produced an estimated RI of 2.37 (1.47, 3.85). The pseudo-likelihood approach produced a similar point estimate of 2.23 (1.42, 3.52) (Table 5).

In sensitivity analyses, inclusion of cases through Brighton criteria levels 4 and 4A increased the RI estimate from the primary analysis (Brighton criteria 1-3) using the standard SCCS and pseudo-likelihood approaches to 2.83 (1.91, 4.19) and 2.59 (1.72, 3.90), respectively. Limiting included cases to Brighton criteria levels 1 and 2 in analysis using the standard SCCS only slightly reduced the estimate to 2.34 (1.48, 3.70). Adjusting for seasonal influenza vaccine exposure led to no change in the pH1N1-associated estimate, 2.57 (1.68, 3.93) (p value vs. primary analysis RI = 0.85) and found no increase in relative incidence associated with seasonal influenza vaccine exposure [0.77 (0.28, 2.14)]. Using recorded date of onset as opposed to the date of diagnosis as the index date produced almost no change, likely due to the fact that the risk interval is long (data not shown). The analyses for multiple risk periods following vaccination yielded estimates of 2.61 (1.27, 6.35), 3.11 (2.18, 6.46), and 1.91 (1.31, 3.98) for risk windows of 1-7, 8-21 and 22-42 days following vaccination, respectively. Excluding subjects with a reported history of GBS led to a slightly reduced estimate of 2.27 (1.47, 3.51). Excluding patients with reported influenza like illness or upper respiratory illness in the 30 days before onset of GBS slightly increased the pH1N1-associated estimate to 2.88 (1.79, 4.65). The exclusion of those with reported gastrointestinal illness also increased the vaccine-associated estimate to 2.73 (1.75, 4.26).

In age-stratified analyses, using the standard SCCS showed that the RI in days 1-42 following exposure increased with age: children age < 19 years, the RI was 0.73 (0.16, 3.46), adults age 19-49 years, RI = 1.56 (0.51, 4.71), age 50-64 years, RI = 2.78 (1.36, 5.68), and 4.30 (2.18, 8.50) in those 65 and older. These confidence intervals overlap, suggesting a trend rather than a statistically significant difference in relative incidence by age. In standard SCCS analysis stratified by sex, the estimated RI was slightly higher in males, 2.75 (1.65, 4.57) than in females, 2.34 (1.09, 5.04) although the difference was not statistically significant (p = 0.73).

The estimate of adjuvanted vaccines, performed using the vaccinated cases only approach, yielded a RI estimate of 1.88 (1.03, 3.41) while the non-adjuvanted estimate was higher at 2.97 (1.13, 7.84). This difference was not statistically significant (p=0.43).

Because data on cases exposed to vaccines containing the MF-59 adjuvant were limited to one database, we were unable to reliably compare the RI associated with each of the two adjuvants.

Meta-analytic approach

Results from the meta-analytic approach were similar to those from the pooled analysis but the magnitude of the estimates was decreased (Table 6). The standard SCCS approach produced an estimate of 2.09 (1.28, 3.42) while the vaccinated cases only approach produced an estimate of 2.33 (1.5, 3.62). Analysis of adjuvanted and non-adjuvanted vaccines using the meta-analytic approach yielded RI estimates of 1.65 (0.86, 3.19) and 3.10 (1.70, 5.65), respectively. This difference between adjuvanted and non-adjuvanted exposures was not statistically significant (p=0.16).

Table 5: Relative Incidence of GBS following pH1N1 Vaccination in Data Pooled across Twenty Databases from Ten Countries

Analysis	Risk Window(s)	Exclusions	Brighton Criteria Levels	Relative Incidence	Confidence Interval
Primary Analysis					
Standard Self-controlled Case Series (SCCS)	Days 1-42	Databases (DBs) with Vaccinated Cases Only	1-3	2.42	(1.58, 3.72)
Sensitivity Analyses					
Standard SCCS	Days 1-42	DBs with Vaccinated Cases Only	1-4A	2.83	(1.91, 4.19)
Standard SCCS	Days 1-42	DBs with Vaccinated Cases Only	1-2	2.34	(1.48, 3.70)
Standard SCCS	Days 1-42	DBs with Vaccinated Cases Only Cases with reported URI or ILI in the 30 days before diagnosis	1-3	2.88	(1.79, 4.65)

Standard SCCS	Days 1-42	DBs with Vaccinated Cases Only Cases with reported GI in the 30 days before diagnosis	1-3	2.73	(1.75, 4.26)
Pseudo-Likelihood	Days 1-42	DBs with Vaccinated Cases Only	1-3	2.23	(1.42, 3.52)
Pseudo-Likelihood	Days 1-42	DBs with Vaccinated Cases Only	1-4A	2.59	(1.72, 3.90)
Vaccinated Cases Only	Days 1-42	Unvaccinated cases Cases vaccinated after diagnosis	1-3	2.37	(1.47, 3.85)
Standard SCCS	Days 1-7* 8-21* 22-42* *Modeled Simultaneously	DBs with Vaccinated Cases Only	1-3	2.61 3.11 1.91	(1.17, 5.84) (1.77, 5.47) (1.07, 3.42)
Vaccinated Cases Only **Adjuvanted**	Days 1-42	Unvaccinated cases Cases vaccinated after diagnosis Non-adjuvanted vaccine recipients	1-3	1.88	(1.04, 3.41)
Vaccinated Cases Only **Non-Adjuvanted**	Days 1-42	Unvaccinated cases Cases vaccinated after diagnosis Adjuvanted vaccine recipients	1-3	2.97	(1.13, 7.84)

Table 6: Relative Incidence of GBS following pH1N1 Vaccination in Results from Twenty Databases in Ten Countries, pooled using a Meta-analytic approach

Analysis	Risk Window(s)	Exclusions	Brighton Criteria Levels	Relative Incidence	Confidence Interval
Self-controlled Case Series (SCCS)	Days 1-42	Databases (DBs) with Vaccinated Cases Only	1-3	2.09	(1.28, 3.42)
Vaccinated Cases Only	Days 1-42	Unvaccinated cases excluded	1-3	2.33	(1.50, 3.62)

		Cases vaccinated after diagnosis			
Vaccinated Cases Only **Adjuvanted**	Days 1-42	Unvaccinated cases Cases vaccinated after diagnosis Non-adjuvanted vaccine recipients	1-3	1.65	(0.86, 3.19)
Vaccinated Cases Only **Non-Adjuvanted**	Days 1-42	Unvaccinated cases Cases vaccinated after diagnosis Adjuvanted vaccine recipients	1-3	3.10	(1.70, 5.65)

Discussion

We have shown that international collaboration to evaluate serious rare outcomes using a common protocol is feasible and offers some advantages compared to single country or site analyses. Because GBS following vaccination is very rare with reported rates between 0.04 and 0.17 cases per 100,000 vaccinations[22], this combined analysis included a much larger number of cases than any published single country analysis and allowed inclusion of data from sites that did not have enough cases for a site-specific analysis. This provided both increased power to evaluate the outcome but also sufficient power to conduct sub-analyses by vaccine type. Secondly, the availability of data from several countries allowed us to identify a site (Mexico), which had a RI of GBS following pH1N1 vaccine much higher than that at any other site; had the analysis been conducted only in Mexico, conclusions regarding the risk of GBS following vaccination could have been inappropriately generalized to other populations.

We have found an increased risk of GBS following receipt of pH1N1 influenza vaccine. This risk is consistent with the level of risk reported by others. Estimates from single-country studies ranged from 1.05 to 4.70, the majority of which reported statistically significant increased risk [23-32]. Estimates were lower in studies of adjuvanted vaccines (1.05-3.04) [25, 31] than in non-adjuvanted vaccines (1.57-4.70) [26-30, 32].

Because we knew *a priori* that both adjuvanted and non-adjuvanted vaccines would be used within our study population, we included a comparative analysis in our analysis plan. In all our primary and sensitivity analyses, the risk of GBS following administration of non-adjuvanted vaccines was significantly elevated. The increased risk found for adjuvanted vaccines was not as consistent. It was significantly elevated in our pooled analyses but became non-significant in the meta-analysis. Moreover, the point estimates of the risk for adjuvanted vaccines were consistently lower than those for non-adjuvanted vaccines in all our analyses. This preliminary finding is reassuring, given general concerns regarding the use of adjuvanted vaccines for influenza, and the fact that these vaccines use less influenza antigen per dose, a useful advantage for pandemic situations during which the amount of available

antigen for vaccine production may initially be limited. We hypothesize that one possible explanation for the apparent (non-significant) risk difference between adjuvanted and non-adjuvanted vaccines is the higher amount of influenza antigen in non-adjuvanted vaccines, although other factors could have contributed. Another possible explanation is increased protection from influenza in those who have received adjuvanted vaccines [33] and a subsequent reduction in GBS due to influenza infection, which may have confounded our results. The trend described was seen for both MF-59 and AS03 adjuvanted vaccine. However, only one country in the study (The Netherlands) used MF-59 adjuvanted vaccine, so our ability to compare adjuvants was limited. While we believe our results to be reassuring, they are by no means definitive. Although the difference between the RI estimates for adjuvanted and non-adjuvanted vaccines are not different in our pooled analysis (p=0.43) or in our meta-analysis (p=0.16), the finding warrants further investigation.

Results obtained through the primary analysis (the standard self-controlled case series), the pseudo-likelihood approach, and vaccinated-cases only up to Brighton level 3 were very similar, indicating there was likely little bias introduced if a history of GBS impacted vaccination practices during the 2009-2010 season . Exclusion of Brighton level 3 cases produced very little change in the RI estimate, suggesting lack of a diagnostic bias in the absence of electrophysiological or cerebrospinal fluid data. Inclusion of cases meeting Brighton Criteria levels 4 and 4A increased the relative incidence estimate. Given that in these cases the diagnosis could not be reliably verified, it is unclear if this increase reflects a more complete capture of true GBS cases or the inclusion of non-true cases occurring near the date of vaccination because of diagnosis bias. Future studies in which the time and resources for centralized adjudication are available may be able to answer this question.

Inclusion of the reported date of seasonal influenza vaccination led to little change in the estimate. This could be attributed to minimal risk associated with seasonal influenza vaccine[34], or to under-reporting or confusion regarding whether seasonal or pandemic vaccine that was received leading to misclassification bias. In the analysis of multiple risk periods following vaccination, the increased incidence in the pre-specified high risk period including days 8-21 supports the 1976 finding of highest risk in weeks 2-3 following vaccination[11]. It has been hypothesized that the risk peaks during this interval because this is when the humoral immune response to the vaccine is highest[35]. Previous findings of increased risk in males as well as increasing risk with increasing age were also supported in our analyses[36]. The background incidence of GBS has also been shown to be about 40% higher in males than in females and to increase with increasing age[37].

It is interesting to note that the exclusion of those subjects who experienced influenza-like illness or gastrointestinal illness resulted in small, non-significant increases in vaccination-associated relative incidence, with p values vs. primary analysis RI = 0.59 and 0.70, respectively. While these non-significant findings appear to be inconsistent with those published by the VAESCO consortium [14] and other studies [27], it is important to note that the SCCS is a methodology based upon an underlying timeline. As we did not have data on dates of influenza-like illness or gastrointestinal illness and could not include these as time-varying covariates, exclusion of those with an infection in the 30 days before diagnosis likely excluded infection-induced cases which should have occurred at similar rates within and outside of the vaccine-associated risk window. This exclusion would therefore serve to increase the estimated relative incidence associated with pH1N1 vaccine exposure.

An additional finding in our study is the increased GBS risk following pH1N1 vaccination with the increased age of vaccine recipients. This may be the result of a lower immune response in older individuals[38] with consequent increased susceptibility to H1N1 infection and subsequent infection-induced GBS in this age group. It may also be related to higher pre-vaccination antibody titers in those previously exposed to H1N1-like viruses[39] potentially leading to greater immune response following vaccination. This possible age effect requires further investigation.

As we have indicated, database-specific analysis of the data contributed by Mexico City produced a very high RI associated with pH1N1 exposure. We hypothesize that this could be due to a longer period of H1N1 circulation in Mexico prior to vaccine introduction with a high likelihood of vaccinated individuals having already been exposed to the H1N1 wild type virus[40, 41] which may have induced a greater immune response upon receipt of pH1N1 vaccine due to already elevated antibody titers. Additionally, cases of confirmed H1N1 infection in Mexico tended to have more severe clinical presentation and to result in death more frequently as compared to other countries, perhaps indicating greater virulence of the virus in the early stages of the pandemic[40]. However, control for seasonality in the Mexican database did not reduce the pH1N1 vaccine associated RI estimate. Further studies to elucidate the reason for this much higher risk level in Mexico are warranted.

The meta-analytic approach for pooling of database-specific relative incidences weights those databases with a large degree of variation less heavily than those with less variation. Therefore, databases with only one or two exposed cases and consequently with large standard errors, have much less weight in the combined estimate of 2.09 (1.28, 3.42). While this estimate is attenuated through the weights applied in the meta-analysis, the relative incidence of GBS in the 42 days following vaccination remains significantly elevated with a confidence interval very similar to that of the pooled estimate of 2.42 (1.58, 3.72). Interestingly, results from stratification of adjuvanted and non-adjuvanted vaccines in the meta-analytic analyses produced a non-significant relative incidence for adjuvanted vaccines and an increased relative incidence for non-adjuvanted vaccines. This can be interpreted as evidence that the trend of increased risk in non-adjuvanted vaccines as opposed to adjuvanted vaccines is not being driven by a set of influential databases.

In this proof of concept, we have learned that international collaborative database studies to evaluate vaccine safety are feasible, even across continents. However, the requirement that participating sites have access to databases from which cases could be ascertained in an unbiased manner limited participation to high and middle income sites with existing infrastructure to conduct active surveillance. However, new vaccines are now being introduced either exclusively in the developing world or concurrently with their release in developed countries. In addition, some newer vaccines, such as the malaria vaccine currently in phase three trials in Africa[42], will mainly target the developing world. These changes indicate the need for improved vaccine safety assessment in low and middle income countries to ensure the safety of new vaccines.

Developing capacity outside of developed countries to evaluate vaccine safety signals that arise out of passive surveillance systems or from other sources is necessary both to assure the safety of the world's vaccine supply and also to prevent vaccine safety scares from undermining successful programs. The only low to middle income country (LMIC) remaining in the final analysis was China, evidencing a need to increase infrastructure in LMIC for ascertainment of vaccination status and adjudication of adverse events following vaccination. In addition, evaluation of very rare event can be facilitated by the increased statistical power that could be achieved through international collaboration.

Although the protocol was common among sites, the degree to which sites were able to review charts and ascertain important covariates such as infections varied from site to site. Future collaborative studies would benefit from centralized case adjudication, improved data quality control and closer supervision of data abstraction and case ascertainment.

This study had several limitations. While the results of the analyses stratified by presence or absence of an adjuvant are intriguing, the use of adjuvanted or non-adjuvanted vaccines was generally homogeneous within each country. Because of this homogeneity, it was not possible to estimate the difference between adjuvanted and non-adjuvanted vaccine associated relative incidence within the same population and it was not possible to

separate the effect of the vaccine formulation from the unknown effect of the country and its associated characteristics. Also, the observed association may have been modified by infections, some of which are known to increase the risk of GBS. While we attempted to control for concomitant infections such as influenza like illness and gastrointestinal illness, it was not possible within the limitation of the current study to include these infections as time varying covariates, which would be the ideal approach in the self-controlled case series methodology. Controlling for such time varying covariates may have attenuated the observed associations, as shown in previous studies [43-45]. Additionally, the use of reported seasonal peaks of circulating influenza was a means of estimating seasonal effects across sites but may not have been as accurate as the standard approach of estimating fixed-length periods when analyzing data from one geographic location. Given budget and data-sharing constraints, case verification was performed by each site and quality control was performed on site; a pooled review by a single expert group was not conducted. In spite of the use of common criteria provided by the Brighton Collaboration, the adjudication process may have varied from site to site.

Conclusion

We have demonstrated that multinational studies are feasible and can provide a useful platform to evaluate future vaccine safety concerns especially for rare, serious events. We look forward to the development of a sustainable global infrastructure in both developed and developing countries to meet global needs. The finding of much higher risk in Mexico and our ability to contrast this risk with that found in other countries using data submitted under a common protocol is a strength of this multinational study and supports the need for international collaboration in vaccine safety monitoring.

The significance and consistency of our findings support a conclusion of an association between 2009 (H1N1) vaccination and GBS. Nonetheless, given the rarity of the event the relative incidence found suggests that the vaccine would be responsible for very few excess GBS cases. Although we are not able to estimate attributable risk using the SCCS methodology, we know from other studies that the background risk of GBS is approximately 0.9[46] cases per one million individuals and that the relative risk associated with the 1976-77 swine influenza vaccination campaign was 7.60[11]. A relative incidence of 2-3 following vaccination would mean approximately 1-2 excess cases per one million vaccinees. Due to this minimal increase in incidence, our findings do not provide evidence in contradiction of international recommendations for the continued use of influenza vaccines.

This large collaborative multinational study has made possible the generation of a number of new hypotheses related to possible differences in risk of vaccine-associated GBS by age and by the use of adjuvants, which will require further investigation.

Acknowledgements:

The authors thank Bruce Fireman and Heather Whitaker for methodological input, Bart Spiessens for assistance with SAS code, and Judith Badoo and David Cho for facilitating collaboration. We also thank collaborators from the Vaccine Safety Datalink (VSD) (US) for their data contribution and methodological input.

The PRISM authors acknowledge the contributions of Joaquim Fernandes, Yihai Liu, Cheryl N. McMahill-Walraven, and Claire Spettell of Aetna, and Amy Ball, David Nau, Jane Stacy, and Yihua Xu of Humana.

References

1. World Health Organization, *Global Vaccine Safety Blueprint*, in *WHO/IVB*. 2012.

2. Mahoney, R.T., et al., *The introduction of new vaccines into developing countries. IV: Global Access Strategies.* Vaccine, 2007. **25**(20): p. 4003-11.

3. Jacobson, R.M., et al., *Adverse events and vaccination-the lack of power and predictability of infrequent events in pre-licensure study.* Vaccine, 2001. **19**(17–19): p. 2428-2433.

4. Hua, W., Guoying Sun, Caitlin N. Dodd, Silvana A. Romio, Heather J. Whitaker, Hector S. Izurieta, Steven Black, Miriam CJM Sturkenboom, Robert L. Davis, Genevieve Deceuninck, and N. J. Andrews, *A Simulation Study to Compare Three Self-Controlled Case Series Approaches: Correction for Violation of Assumption and Explanation of Bias*. 2012, Office of Biostatistics and Epidemiology, Center for Biologics Evaluation and Research, Food and Drug Evaluation: Rockville, MD. p. 22.

5. Dodd, C., Izurieta HS, Zuber P, Romio SA, et al. *International Collaborative Case Series Safety Monitoring for Pandemic 2009 H1N1 Vaccines: Estimation of the Risk of Guillain-Barre Syndrome*. in *International Conference on Pharmacoepidemiology and Therapeutic Risk Management*. 2012. Barcelona, Spain.

6. Izurieta HS, B.S., Zuber P, Dodd C, Bonhoeffer J, Sturkenboom M, Romio S, et al., *A proof-of-concept collaborative international study for the post-licensure epidemiological assessment of the safety of newly introduced vaccines*, in *Vaccine Safety Update, Brighton collaboration symposium*. 2012: Shanghai Marriott Hotel City Centre, Shanghai, China.

7. Bonhoeffer, J., et al., *The Brighton Collaboration: addressing the need for standardized case definitions of adverse events following immunization (AEFI).* Vaccine, 2002. **21**(3-4): p. 298-302.

8. Garske, T., et al., *Assessing the severity of the novel influenza A/H1N1 pandemic.* BMJ, 2009. **339**.

9. Partridge, J. and M.P. Kieny, *Global production of seasonal and pandemic (H1N1) influenza vaccines in 2009–2010 and comparison with previous estimates and global action plan targets.* Vaccine, 2010. **28**(30): p. 4709-4712.

10. Whitaker, H.J., et al., *Tutorial in biostatistics: the self-controlled case series method.* Statistics in Medicine, 2006. **25**(10): p. 1768-1797.

11. Schonberger, L.B., et al., *Guillain-Barré syndrome following vaccination in the national influenza immunization program, United States, 1976–1977.* American Journal of Epidemiology, 1979. **110**(2): p. 105-123.

12. Sejvar, J.J., et al., *Guillain-Barre syndrome and Fisher syndrome: case definitions and guidelines for collection, analysis, and presentation of immunization safety data.* Vaccine, 2011. **29**(3): p. 599-612.

13. Farrington, C.P., *Relative incidence estimation from case series for vaccine safety evaluation.* Biometrics, 1995. **51**(1): p. 228-35.

14. Dieleman, J., et al., *Guillain-Barre syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe.* BMJ, 2011. **343**: p. d3908.

15. Salmon, D., et al., *Success Of Program Linking Data Sources To Monitor H1N1 Vaccine Safety Points To Potential For Even Broader Safety Surveillance.* Health Aff (Millwood), 2012. **31**(11): p. 2518-2527.

16. Nguyen, M., et al., *The Food and Drug Administration's Post-Licensure Rapid Immunization Safety Monitoring program: strengthening the federal vaccine safety enterprise.* Pharmacoepidemiol Drug Saf, 2012. **21 Suppl 1**: p. 291-7.

17. Fine, P.E. and R.T. Chen, *Confounding in studies of adverse reactions to vaccines.* Am J Epidemiol, 1992. **136**(2): p. 121-35.

18. Zhou, W., et al., *Surveillance for safety after immunization: Vaccine Adverse Event Reporting System (VAERS)--United States, 1991-2001.* MMWR Surveill Summ, 2003. **52**(1): p. 1-24.

19. Maclure, M., et al., *When should case-only designs be used for safety monitoring of medical products?* Pharmacoepidemiology and Drug Safety, 2012. **21**: p. 50-61.

20. Farrington, C.P., H.J. Whitaker, and M.N. Hocine, *Case series analysis for censored, perturbed, or curtailed post-event exposures.* Biostatistics, 2009. **10**(1): p. 3-16.

21. DerSimonian, R. and N. Laird, *Meta-analysis in clinical trials.* Control Clin Trials, 1986. **7**(3): p. 177-88.

22. Haber, P., et al., *Guillain-Barre syndrome following influenza vaccination.* JAMA, 2004. **292**(20): p. 2478-81.

23. Crawford, N.W., et al., *Guillain-Barre syndrome following pandemic (H1N1) 2009 influenza A immunisation in Victoria: a self-controlled case series.* Med J Aust, 2012. **197**(10): p. 574-8.

24. Burwen, D.R., et al., *Surveillance for Guillain-Barre syndrome after influenza vaccination among the Medicare population, 2009-2010.* Am J Public Health, 2012. **102**(10): p. 1921-7.

25. De Wals, P., et al., *Risk of Guillain-Barre syndrome following H1N1 influenza vaccination in Quebec.* JAMA, 2012. **308**(2): p. 175-81.

26. Greene, S.K., et al., *Risk of Confirmed Guillain-Barré 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**(11): p. 1100-1109.

27. Wise, M.E., et al., *Guillain-Barre syndrome during the 2009-2010 H1N1 influenza vaccination campaign: population-based surveillance among 45 million Americans.* Am J Epidemiol, 2012. **175**(11): p. 1110-9.

28. Yih, W.K., et al., *Surveillance for adverse events following receipt of pandemic 2009 H1N1 vaccine in the Post-Licensure Rapid Immunization Safety Monitoring (PRISM) System, 2009-2010.* Am J Epidemiol, 2012. **175**(11): p. 1120-8.

29. Tokars, J.I., et al., *The risk of Guillain-Barre syndrome associated with influenza A (H1N1) 2009 monovalent vaccine and 2009-2010 seasonal influenza vaccines: results from self-controlled analyses.* Pharmacoepidemiol Drug Saf, 2012. **21**(5): p. 546-52.

30. Bardage, C., et al., *Neurological and autoimmune disorders after vaccination against pandemic influenza A (H1N1) with a monovalent adjuvanted vaccine: population based cohort study in Stockholm, Sweden.* BMJ, 2011. **343**: p. d5956.

31. Andrews, N., et al., *Guillain-Barre syndrome and H1N1 (2009) pandemic influenza vaccination using an AS03 adjuvanted vaccine in the United Kingdom: self-controlled case series.* Vaccine, 2011. **29**(45): p. 7878-82.

32. Salmon, D., Michael Proschan, Richard Forshee, Paul Gargiullo, William Bleser, Dale Burwen, Freanscesca Cunningham, Patrick Garman, Sharon Greene, Grace Lee, Claudia Vellozzi, Katherine Yih, Bruce Gellin, Nichole Lurie, *A Meta-Analysis of teh Association between Guillain-Barre Syndrom and Influenza A (H1N1) 2009 Monovalent Inactivated Vaccines in the United States.* The Lancet, 2013. **In Press**.

33. Gilca, V., et al., *Antibody persistence and response to 2010-2011 trivalent influenza vaccine one year after a single dose of 2009 AS03-adjuvanted pandemic H1N1 vaccine in children.* Vaccine, 2011. **30**(1): p. 35-41.

34. Stowe, J., et al., *Investigation of the Temporal Association of Guillain-Barré Syndrome With Influenza Vaccine and Influenzalike Illness Using the United Kingdom General Practice Research Database.* American Journal of Epidemiology, 2009. **169**(3): p. 382-388.

35. Schonberger, L.B., et al., *Guillain-Barré syndrome: Its epidemiology and associations with influenza vaccination.* Annals of Neurology, 1981. **9**(S1): p. 31-38.

36. Hughes, R.A.C. and J.H. Rees, *Clinical and Epidemiologic Features of Guillain-Barré Syndrome.* Journal of Infectious Diseases, 1997. **176**(Supplement 2): p. S92-S98.

37. Shui, I.M., et al., *Guillain-Barre syndrome incidence in a large United States cohort (2000-2009).* Neuroepidemiology, 2012. **39**(2): p. 109-15.

38. Weinberger, B., et al., *Biology of immune responses to vaccines in elderly persons.* Clin Infect Dis, 2008. **46**(7): p. 1078-84.

39. Hancock, K., et al., *Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus.* N Engl J Med, 2009. **361**(20): p. 1945-52.

40. Perez-Padilla, R., et al., *Pneumonia and Respiratory Failure from Swine-Origin Influenza A (H1N1) in Mexico.* New England Journal of Medicine, 2009. **361**(7): p. 680-689.

41. Echevarría-Zuno, S., et al., *Infection and death from influenza A H1N1 virus in Mexico: a retrospective analysis.* The Lancet, 2009. **374**(9707): p. 2072-2079.

42. Abdulla, S., et al., *Randomized, controlled trial of the long term safety, immunogenicity and efficacy of RTS,S/AS02D malaria vaccine in infants living in a malaria-endemic region.* Malar J, 2013. **12**(1): p. 11.

43. Andrews, N., et al., *Post-licensure comparison of the safety profile of diphtheria/tetanus/whole cell pertussis/haemophilus influenza type b vaccine and a 5-in-1 diphtheria/tetanus/acellular pertussis/haemophilus influenza type b/polio vaccine in the United Kingdom.* Vaccine, 2010. **28**(44): p. 7215-20.

44. Huang, W.T., et al., *Lack of association between acellular pertussis vaccine and seizures in early childhood.* Pediatrics, 2010. **126**(2): p. 263-9.

45. Andrews, N., et al., *Does oral polio vaccine cause intussusception in infants? Evidence from a sequence of three self-controlled cases series studies in the United Kingdom.* Eur J Epidemiol, 2001. **17**(8): p. 701-6.

46. Lasky, T., et al., *The Guillain-Barre syndrome and the 1992-1993 and 1993-1994 influenza vaccines.* N Engl J Med, 1998. **339**(25): p. 1797-802.

5.2 Narcolepsy and adjuvanted pandemic influenza A (H1N1) 2009 vaccines: a Multi-country assessment

Daniel Weibel, Miriam Sturkenboom
Steven Black, Maria de Ridder
Caitlin Dodd, Jan Bonhoeffer
Ann Vanrolleghem, Nicoline van der Maas
Gert Jan Lammers, Sebastiaan Overeem
Angela Gentile, Norberto Giglio
Vanesa Castellano, Jeffrey C. Kwong
Brian J. Murray, Karen Cauch-Dudek
Diana Juhasz, Michael Campitelli
Alexandre N. Datta, Ulf Kallweit
Wan-Ting Huang, Yu-Shu Huang
Chung-Yao Hsu, Hsi-Chung Chen
Maria Giner-Soriano, Rosa Morros
Carles Gaig, Ester Tió
Silvia Perez-Vilar, Javier Diez-Domingo
Francisco Javier Puertas, Lawrence W. Svenson
Salaheddin M. Mahmud, Bruce Carleton
Monika Naus, Lisen Arnheim-Dahlström
Lars Pedersen, Frank DeStefano
Tom T. Shimabukuro

Vaccine 36.41 (2018): 6202-6211.

Background:

In 2010, a safety signal was detected for narcolepsy following vaccination with Pandemrix, an AS03-adjuvanted monovalent pandemic H1N1 influenza (pH1N1) vaccine. To further assess a possible association and inform policy on future use of adjuvants, we conducted a multi-country study of narcolepsy and adjuvanted pH1N1 vaccines.

Methods:

We used electronic health databases to conduct a dynamic retrospective cohort study to assess narcolepsy incidence rates (IR) before and during pH1N1 virus circulation, and after pH1N1 vaccination campaigns in Canada, Denmark, Spain, Sweden, Taiwan, the Netherlands, and the United Kingdom. Using a case-control study design, we evaluated the risk of narcolepsy following AS03- and MF59-adjuvanted pH1N1 vaccines in Argentina, Canada, Spain, Switzerland, Taiwan, and the Netherlands. In the Netherlands, we also conducted a case-coverage study in children born between 2004 and 2009.

Results:

No changes in narcolepsy IRs were observed in any periods in single study sites except Sweden and Taiwan; in Taiwan incidence increased after wild-type pH1N1 virus circulation and in Sweden (a previously identified signaling country), incidence increased after the start of pH1N1 vaccination. No association was observed for Arepanrix-AS03 or Focetria-MF59 adjuvanted pH1N1 vaccines and narcolepsy in children or adults in the case-control study nor for children born between 2004 and 2009 in the Netherlands case-coverage study for Pandemrix-AS03.

Conclusions:

Other than elevated narcolepsy IRs in the period after vaccination campaigns in Sweden, we did not find an association between AS03- or MF59-adjuvanted pH1N1 vaccines and narcolepsy in children or adults in the sites studied, although power to evaluate the AS03-adjuvanted Pandemrix brand vaccine was limited in our study.

Introduction

In fall 2009, large-scale vaccination campaigns were implemented globally in response to the influenza A (H1N1) pandemic (pH1N1). Different inactivated monovalent vaccines were used; the United States used unadjuvanted vaccines, whereas in Europe, MF59- (Focetria, Novartis Vaccines and Diagnostics) and AS03-adjuvanted (Pandemrix, GSK biologicals) vaccines were mostly used. In Canada, AS03-adjuvanted Arepanrix (ID Biomedical Corp., a subsidiary of GSK Biologicals) was used. Arepanrix and Pandemrix had similar pH1N1 antigens and used the same AS03 adjuvant; however, there were slight differences in the manufacturing processes for the two vaccines. In August 2010, case reports from Sweden and Finland emerged describing narcolepsy in children following vaccination with Pandemrix.[1-6]

Narcolepsy is a chronic debilitating sleep disorder with a suspected autoimmune etiology. Genetic predisposition also appears to play a role; narcolepsy with cataplexy (with hypocretin deficiency) is highly associated with HLA-DQB1*06:02, while narcolepsy without cataplexy (and normal hypocretin levels) is less associated with HLA-DQB1*06:02, but still more so than in the general population. Narcolepsy/cataplexy is characterized by excessive daytime sleepiness (EDS), cataplexy, sleep paralysis, hypnagogic hallucinations and fragmented nighttime sleep. Symptoms often emerge gradually and may initially be non-specific.[7-10] The insidious onset can result in diagnostic delays of months to years, making it challenging to study exposures that might cause or contribute to disease occurrence.[11]

Several European studies were initiated to rapidly evaluate the possible association between Pandemrix-AS03 and narcolepsy.[1,2,4,5,12-18] These studies are summarized in reviews by Verstraeten *et al*.[19], Sturkenboom[20] and Sarkanen *et al*.[21] Several studies showed an increased risk, but results were variable within and across studies and subject to methodological challenges due to narcolepsy epidemiology and increased awareness about the association. Most studies focused on Pandemrix-AS03 and data were limited on other adjuvanted pH1N1 vaccines, including Arepanrix-AS03 and Focetria-MF59.[22,23]

To better understand the relationship between narcolepsy and different adjuvanted pH1N1 vaccines, we organized an international research network, including study sites within and outside of Europe, where adjuvanted vaccines were used and little or no substantial concerns were raised about an association with narcolepsy in local media. Drawing on the expertise of clinicians, research scientists and public health officials within the network, we implemented the Systematic Observational Method for Narcolepsy and Influenza Immunization Assessment (SOMNIA) study.

Methods

Study site selection

We used a stepwise process to identify, recruit, and select participating study sites for incidence rates and case-control analyses. First, we identified countries that used adjuvanted 2009 pH1N1 vaccine using information obtained from the World Health Organization and from the two vaccine manufacturers. We excluded countries where compensation programs for narcolepsy associated with pH1N1 vaccination existed because we believed this could potentially bias case ascertainment with vaccinated cases potentially being evaluated sooner or differently than non-vaccinated cases. Finland, Norway and Sweden were excluded from the case-control study on this basis. In addition, Finland and Sweden had been signaling countries. Of note Sweden, as a signaling country was included in the incidence rate study to provide a reference comparison, but data from Sweden were not pooled in the incidence rate analysis. Working with national, regional and local health officials, academics, and sleep centers, we assessed whether potential study sites had acceptable availability and accessibility of vaccination and outcome

data. Israel, South Korea and Cuba, were excluded because of lack of exposure information. Brazil was eliminated because obtaining timely administrative approvals was not feasible. Finally, we engaged in discussions with prospective investigators to gauge willingness and ability to participate in the incidence rate or case-control studies and confirm existence of acceptable data and data systems to articipate using the common study protocol. After completing these steps and obtaining the necessary clearances we included the final set of countries in our analysis.

All sites used the same protocol, the same data collection materials and common analytics, but sites could implement the protocol based upon their local processes and health care structure. For quality control, the study coordination team verified implementation of the protocol with all sites during monthly calls. The data management team further verified all the data and discussed potential biases with the sites. Sites were responsible for verification of their results and decisions on inclusion of the data and could access remotely the secure data sharing environment. These distributed data management procedures have been previously described. [24]

Narcolepsy incidence rates analysis

We estimated narcolepsy incidence rates (IR) at ten study sites (in seven countries) using population-based electronic health record databases from general practitioners (GP) (Spain [Valencia and Catalonia], the Netherlands, and the United Kingdom) or claims/record linkage databases (Canada [Manitoba, Alberta and British Columbia], Denmark, Sweden, and Taiwan). Study populations included individuals registered in the database for at least one year prior to start of follow-up. Follow-up started at the begining of the study period (January 1, 2003) or the date of registration and ended at the earliest of the following: death, the patient moving, the end of the study period (December 31st, 2013) or outcome occurrence. Cases were captured in these databases by identifying individuals with newly diagnosed narcolepsy with or without cataplexy. (Database codes and algorithms are shown in Table 1). In GP databases in the Netherlands and from records of sleep medicine specialists in Valencia, case finding algorithms were validated using the Brighton Collaboration case definition criteria for narcolepsy.[25] In other sites, we required both diagnostic codes for narcolepsy along with reimbursement claims for a multiple sleep latency test (MSLT) to reduce the risk of false positives. We calculated IRs by year and month and categorized them into three periods: 1) pre-pH1N1, 2) during wild-type pH1N1 virus circulation until the start of pH1N1 vaccination (country specific), and 3) from the start of pH1N1 vaccination through the end of the study in December 2013. Periods of wild-type pH1N1 virus circulation and vaccination varied by sites. We stratified IRs by age and sex and pooled aggregated person-time and case counts for further analysis. IR data from Sweden were analyzed separately since Sweden had been a priori identified as a signaling country, and hence, served as a comparator for other sites in our IR analysis.[26] We estimated incidence rate ratios (IRR), comparing the two latter periods to the pre-pH1N1 period, using Poisson regression.

Table 1: Overview of study sites for the IR, case-control and case-coverage analyses**.**

	Study site (vaccines)	IR		Case-control or case-coverage		
		Data source	**codes and algorithms**	**Case identification**	**Controls source**	**pH1N1 vaccine exposure**
Europe	**Switzerland** (Pandemrix-AS03, Focetria-MF59)			14 hospitals & sleep centers, no consent required	Matched in same hospital of case. No consent required	general practitioner (GP) medical record
	Spain, Catalonia (Pandemrix-AS03, Focetria-MF59)	SIDIAP general practitioners' database	ICD-10 code G47.4	Sleep units at 13 public hospitals until end of 2013 & also in SIDIAP. No consent required	SIDIAP database No consent required	SIDIAP database
	Spain, Valencia (Pandemrix-AS03, Focetria-MF59)	SIA regional general practitioners' databases	ICD-9CM codes 347.* with Manual validation	Identified from 24 sleep centers and electronic registries (inpatient and outpatient databases)	SIA general practitioners' database. No consent required	Obtained from electronic registry
	The Netherlands (Pandemrix-AS03, Focetria-MF59)	IPCI general practitioners' databases	Free text narcolepsy & MSLT. cases manually validated	Four sleep centers (academic and non-university hospitals). Consent required	IPCI general practitioners' database. No consent required	1. Electronic medical GP records for all patients 2. For cases also from vaccination card and public health agency to estimate completeness in GP records
	The United Kingdom (Pandemrix AS03)	THIN general practitioners' databases	Read codes F27.00;F270.00; F271.00; F27z.00			
	Denmark (Pandemrix AS03)	Danish Civil Registration System covering the Northern and Central Region of Jutland in	ICD-10 code G47.4, Inpatient, ambulatory care and emergency room diagnosis			

		Denmark linked to Danish National Patient Register				
	Sweden	Patient register at the National board of health Population from population register at Statistics Sweden	ICD-10 code G47.4 diagnosis in and outpatient			
South America	**Argentina** Focetria-MF59			13 MSLT sites, pediatric, adult neurology and respiratory centers in Buenos Aires. Verbal consent	General practitioners. Verbal informed consent conducted	Vaccination cards from cases and controls
North America (Canada)	**Ontario** Arepanrix-AS03			patient charts at sleep units after using physician billing claims to generate initial list of MSLTs performed. Limited to maximum age ≤24	Matched from ICES provincial database of all residents with health insurance. No consent required	1. primary care physician (PCP), family physicians and pediatricians charts; 2. public health unit records (electronic databases, and physician billing claims data)
North America (Canada)	**Alberta** Arepanrix-AS03	Cases indentified and populaiton denominator established with claims/hospital record linkage databases	ICD9-CM codes 347.* with MSLT procedure			

	Manitoba Arepanrix-AS03	Population-based Hospital and Physician Claims databases Linked to Manitoba Health Population Registry	ICD9-CM codes 347.* with MSLT procedure			
	British Columbia Arepanrix-AS03	Cases originating from British Columbia Medical Services Plan database ; denominator information established through national statistics.	ICD9-CM codes 347.* following MSLT procedure code			
Asia	**Taiwan** (Focetria-MF59, AdimFlu-S unadjuvanted)	National population NHI ENROLL data linked to NHI claims data	Referral MSLT & ≥3 ICD-9-CM codes (347*) in outpatient, ED, or inpatient after an MSLT referral	Recruited from three largest sleep centers in Taiwan. Identified initially by MLST referral from electronic data	Matched from National Health Insurance Database. Consent not required	national registry and National Health Insurance Database. Data on H1N1 vaccination missing for 39% in schoolchildren 7-17 years and 44% for persons ≥18 years (all non-adjuvanted)

GP = general practitioner, BC = Brighton Collaboration, PCP = primary care physician

* Linked Medical Records = Population based medical records (GP and specialist diagnoses), directly linked; Population-based registry = Population based registries (emergency room, in and out patient diagnoses); Medical Record diagnoses + Census Population = In and outpatient diagnoses, case counts and population counts (census);

Case-control analysis

Seven sites in six countries met our criteria for inclusion in the case-control study: Argentina, Canada (Ontario), Spain (Valencia and Catalonia), Switzerland, Taiwan, and the Netherlands (Table 1). All sites collected information in the same electronic case report forms (Chameleon, Erasmus Medical Center, Rotterdam, the Netherlands). Depending on data sharing restrictions, local investigators either transferred aggregated data to a secure remote research environment for further analysis and one stage pooling or ran the same analyses locally and transferred coefficients and counts for inclusion in a two-stage meta-analysis (Figure 1).

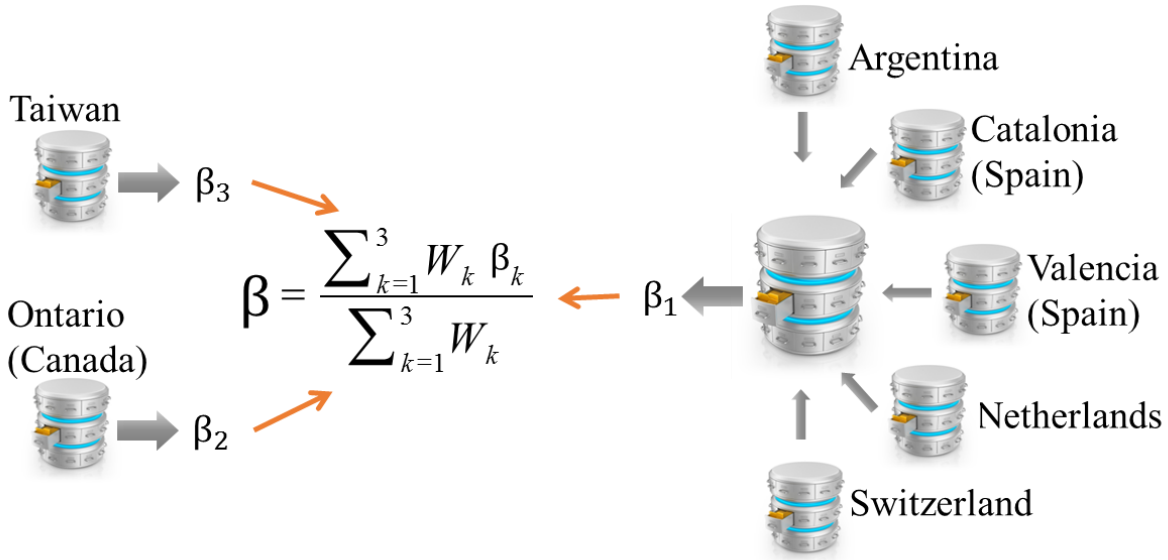


Figure 1. Two-stage hybrid approach for pooling case-control data from study sites. The two-stage hybrid approach pooled case-control data to estimate an odds ratio from European Union country sites and Argentina (β_1). Odds ratios from Taiwan and Ontario (β_2 and β_3) were analyzed in a subsequent meta-analytic approach including the pooled odds ratio from European Union country sites and Argentina (W_k = weight related to estimate β_k, being the reciprocal of its variance, Σ = summation operator, β = meta-analysis result for the odds ratio).

Cases and controls

Local investigators identified narcolepsy cases at sleep centers using diagnosis lists or diagnostic test outcomes. Investigators abstracted medical records, blinded for pH1N1 vaccination status, and classified cases into certainty levels using the Brighton Collaboration narcolepsy case definition.[25] Cases were included if they were classified as Brighton Collaboration level 1-4 for persons ≥16 years or level 1-2 for persons <16 years, and had an MSLT referral and a diagnosis both made after March 31, 2009 (start of the H1N1 pandemic). Brighton Collaboration classification designates level 1 as the highest level of diagnostic certainty; levels have different cut-off values for MSLT results starting at age 16 years, which were maintained in the study. Primary index date was date of referral for diagnostic MSLT. Sites identified all cases diagnosed from April 1, 2009 until the end of 2015, but this varied by site based on feasibility (Table 1). The Netherlands required consent from cases because of the need to collect data from both GPs as well as the National Public Health Agency.

Up to 20 controls were matched to each case by site on age (year of birth), sex and index date. As per protocol, controls were selected from the population giving rise to the cases, identification could be implemented in different ways based on feasibility and health care structure. Ontario, Valencia, Catalonia, the Netherlands, and Taiwan sampled controls from population-based health record databases. Argentina identified controls from primary care facilities in the same geographic area as the cases. Switzerland recruited controls from the same hospitals as cases, using auxiliary diseases not related to vaccination.

Exposure

The main exposure of interest was adjuvanted pH1N1 vaccination (i.e., Pandemrix-AS03, Arepanrix-AS03 or Focetria-MF59). We obtained information on vaccines similarly for cases and controls from medical records, vaccination registries, insurance databases, or vaccination cards. Only written or electronic records of immunization were accepted. Risk windows for pH1N1 vaccine exposure were any time prior to index date and further split in: 1 to 180 days, 181 days to 2 years, and >2 years before index date. The Ontario site provided the only data for Arepanrix-AS03; hence, there was no pooling of the Arepanrix-AS03 data as Ontario was unique. Taiwan, Argentina, the Netherlands, Switzerland, Valencia, and Catalonia contributed data for Focetria-MF59. The Netherlands, Switzerland, and Valencia contributed data for Pandemrix-AS03.

Case-control analysis

Odds ratios (OR) and 95% confidence intervals (CI) were calculated using conditional logistic regression or exact logistic regression if zero cells occurred. The reference category for estimation of the pH1N1 vaccination effect was no pH1N1 vaccination. Due to low exposure levels, we only present the analysis for the risk window of any time prior to the index date because this provided the maximum power.

To address potential awareness bias in the European Union (EU) sites, we analyzed data from two time periods, a "restricted" period and a "total" period. The restricted period analysis included cases from participating EU sites in the Netherlands, Valencia, Catalonia, and Switzerland only when they were diagnosed prior to onset of awareness about the narcolepsy signal in Europe (August 2010), and also cases from sites outside the EU diagnosed anytime during the entire study period. The total period analysis included cases from all sites for the entire study period, including cases from EU sites diagnosed after media attention.

We pooled data from sites using a hybrid approach (one stage pooling of matched case and control pairs for EU sites and Argentina, which could share data, and two-stage pooling with Taiwan and Ontario, which shared case counts and coefficients which were subsequently meta-analyzed with the one stage pooled data from EU sites and Argentina) (Figure 1). Children were defined as ≤18 years of age and adults as ≥19 years. Due to incomplete pH1N1 vaccination information from GPs in children born between 2004 and 2009 in the Netherlands (because these children were vaccinated at local health agencies instead of GPs where we obtained exposure information for all study subjects), these cases and controls were excluded from the case-control analysis. However, cases born between 2004 and 2009 in the Netherlands were included in a post hoc case-coverage analysis (see below). In Switzerland, only child cases were included because of potential selection and information biases that was detected in adult cases. (Tables 1 and 2). In addition to using diagnostic MSLT referral as index date, we conducted sensitivity analyses using EDS onset date (requiring EDS starting after March 31, 2009). We used SAS v9.2 and statistical significance was set at p-value <0.05.

Table 2. Narcolepsy case characteristics for the case-control and case-coverage analysis by study site

	Netherlands	Switzerland	Spain, Catalonia	Spain, Valencia	Argentina	Canada, Ontario	Taiwan	Total
Children								
Total cases/controls [N]	**22*/205**	**22/132**	**5/100**	**11/220**	**11/86**	**28/55**	**51/510**	**150/1308**
Brighton level [%]								
1	31.8	36.4	0	0	9.1	0	7.8	13.3
2	68.2	54.5	100	63.6	81.8	57.1	66.7	65.3
3	0	9.1	0	36.4	9.1	42.9	17.7	18.7
4a	0	0	0	0	0	0	7.8	2.7
Cataplexy present [%] **	95.5	86.4	100	63.6	90.9	57.1	72.5	76.7
Age at Diagnosis [%]								
< 6 yrs	4.5	0	0	9.1	0	0	0	1.3
6-12 yrs	50.0	40.9	60.0	90.9	63.6	21.4	37.3	43.3
13-18 yrs	45.5	59.1	40.0	0	36.4	78.6	62.7	55.3
pH1N1 vaccination coverage cases [%]								
Focetria-MF59	0	0	0	0	27.3	0	≤45.1	≤18.4
Pandemrix- and Arepanrix-AS03	31.8*	0	0	0	0	≤17.9	0	≤33.1
Unadjuvanted	0	0	0	0	0	0	≤45.1	≤16.3
pH1N1 vaccination coverage controls [%]								
Focetria-MF59	3.4	3.8	0	2.3	10.5	0	0.6	2.1
Pandemrix- and Arepanrix-AS03	0	0	0	0	0	17.3	0	0.8
Unadjuvanted	0	0	0	0	0	0	37.3	15.9
Adults								
Total cases/controls [N]	**32/280**		**13/260**	**36/720**	**4/12**	**39/75**	**86/860**	**210/2207**
Brighton level [%]								
1	53.1		7.7	0	0	0	1.2	9.1
2	31.3		84.6	38.9	50.0	≤25.6	39.5	≤38.6
3	12.5		7.7	55.6	50.0	69.2	31.4	38.6
4a	3.1		0	5.6	0	≤12.8	27.9	≤15.2
Cataplexy present [%] **	78.1		84.6	41.7	50.0	25.6	41.9	47.2
Age at Diagnosis [%]								
19-59 yrs	96.9		100	97.2	100	100	100	99.0
60 +	3.1		0	2.8	0	0	0	1.0
pH1N1 vaccination coverage cases [%]								
Focetria-MF59	3.1		0	0	0	0	≤14.0	≤6.2
Pandemrix- and Arepanrix-AS03	0		0	0	0	≤12.8	0	1.4
Unadjuvanted	0		0	2.8	0	0	≤14.0	≤6.2
pH1N1 vaccination coverage controls [%]								
Focetria-MF59	2.5		0	0.1	33.3	0	1.1	1.0
Pandemrix- and Arepanrix-AS03	0		0.4	1.4	0	8.0	0	0.8
Unadjuvanted	0		0.4	0	0	0	8.8	3.5

*Nine child cases were included in a separate case-coverage study for reasons described in the methods. The child case total for the Netherlands includes nine from the case-coverage study and 13 from the case-control study.

**By definition, Brighton Collaboration (BC) narcolepsy case levels 1 and 2 have unambiguous cataplexy; BC narcolepsy case levels 3 and 4 may not have cataplexy. For inclusion in this study, children ages <16 years are BC levels 1 and 2, children 17-18 years are BC levels 1-4, and adults ≥19 years are BC levels 1-4.

**Cell counts that represent case counts of five or fewer (for Ontario, Canada) or two or fewer (for Taiwan) cannot be displayed due to national patient privacy regulations. Therefore these are represented as range (i.e. ≤ n).

Case-coverage analysis

In the Netherlands, cases born from 2004 through 2009 were analyzed using a case-coverage design. During the 2009 H1N1 pandemic these children were vaccinated with Pandemrix-AS03 at Municipal Health Services, and not through their GPs. Vaccinations for these children were registered in a nationwide database, Influsys, managed by the National Public Health Institute. Exposure in cases was therefore obtained from Influsys and exposure prevalence in the population for children born in the same year was obtained by calendar week and year of birth and used in the analysis, similar to the method used in the United Kingdom by Stowe et al.[27] This post-hoc analytic approach allowed us to include information from the Netherlands, where individual exposure data was not available, to complement other data on Pandemrix-AS03.

Ethics committee approval

This study was conducted under the principles of the Helsinki declaration.[28] Each site was responsible for obtaining appropriate ethical approvals, the overall study was approved by the Institutional Review Board at Cincinnati Children's Hospital Medical Center.

Role of the funding source

Two investigators (TS and FD) from the sponsoring organization, the U.S. Centers for Disease Control and Prevention, participated in development of the study design, analysis and interpretation of data, writing the report, and in the decision to submit the paper for publication.

Results

Narcolepsy incidence rates analysis

540 million person-years from ten sites in seven countries contributed to the narcolepsy IR analysis. In Sweden, a previously identified signaling country, IRs increased significantly after the start of its pH1N1 vaccination campaign for children 5-19 years (IRR=9.01; 95% CI 6.89-11.80) and adults 20-59 years (IRR=1.69; 95% CI 1.46–1.95). From 2011 onwards, narcolepsy IRs decreased in Sweden (Figure 2). In other EU sites, narcolepsy IRs in the post-vaccination period did not change significantly compared to the pre-pH1N1 vaccination period. In Canada, where Arepanrix-AS03 was used, no changes in IRs were observed in any province sites in any age category. In Taiwan, pre-pH1N1 period narcolepsy IR was lower (0.29 per 100,000 person-years, 95% CI 0.27-0.32) than in EU and Canadian sites (varying between 0.5-1.5 per 100,000 person-years) and we observed significant IR increases in children 5-19 years (IRR=2.50; 95% CI 1.46-4.28) and adults 20-59 years (IRR=2.23; 95% CI 1.26-3.94) during wild-type pH1N1 virus circulation prior to the vaccination campaign (also previously presented in Dodd et al.[29]).

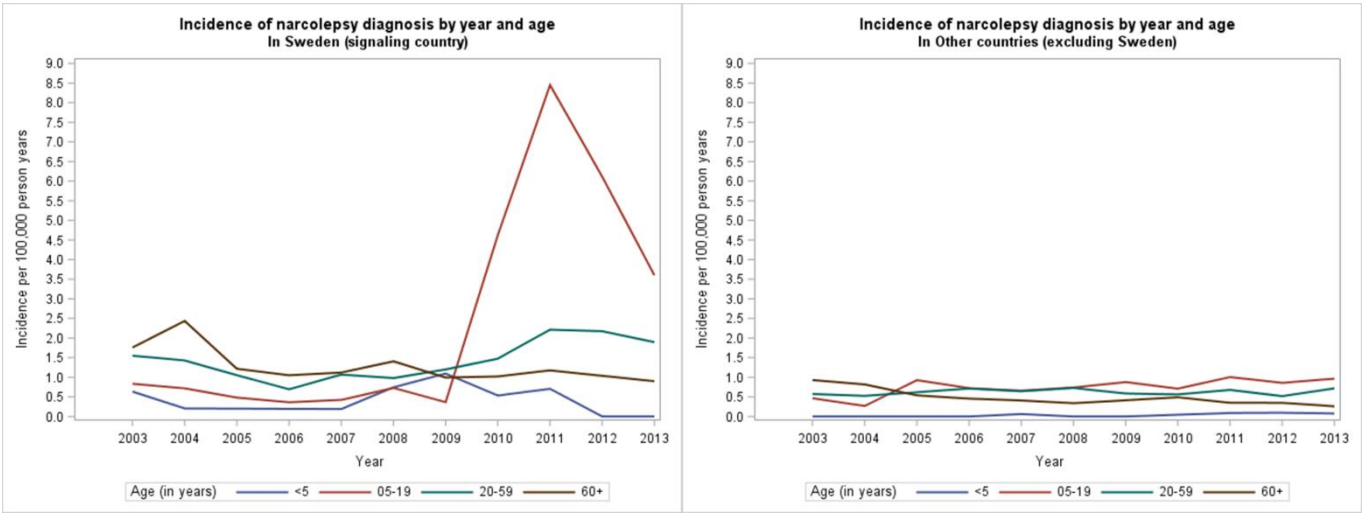


Figure 2. Incidence rates of narcolepsy by age group and year from Sweden (left, a signaling country) and other study sites excluding Sweden (right, pooled data).

Alberta, Canada (2003-2013); British Columbia, Canada (2003-2013), Manitoba, Canada (2003-2010); Catalonia, Spain (2007-2013), Valencia, Spain (2009-2013), Denmark (2003-2013), the Netherlands (2003-2013); Sweden (2003-2013), Taiwan (2003-2012), United Kingdom (2003-2013).

Case-control and case-coverage analysis

We included 360 narcolepsy cases with MSLT referral during the study period: 150 were children ≤18 years and 210 were adults ≥19 years, which were matched to a total of 3515 controls (online supplement Table 2). For the restricted period analysis (excluding cases diagnosed after awareness in the EU), 96 child and 121 adult cases were included. For the total period analysis, 141 child and 210 adult cases were included. Nine child cases born between 2004 and 2009 from the Netherlands (all diagnosed after awareness) were only included in the case-coverage analysis.

Brighton Collaboration narcolepsy case definition diagnosis levels varied by site, with EU sites having more level 1 cases (23.40% in EU vs. 2.74% outside EU) and more cataplexy (73.06% in EU vs. 50.68% outside EU) (Table 2). Median delay between EDS onset and narcolepsy diagnosis was longer in adults compared to children and varied between sites, with a very short delay for Taiwan (online supplement Table 1). Shortening of delay was seen in children in some sites in the EU but not outside the EU following media and public awareness. In all sites, exposure to pH1N1 vaccine was low in cases and controls, except in Dutch children born between 2004 and 2009, with seven out of nine cases exposed to Pandemrix-AS03.

In the meta-analysis for the restricted period, exposure to any type of adjuvanted pH1N1 vaccine was not associated with narcolepsy in children or adults. The OR of the restricted period analysis in children ≤18 years of age was 0.80 (95% CI 0.21-3.01) for Arepanrix-AS03 and 4.12 (95% CI 0.99-17.16) for Focetria-MF59. The OR of the restricted period analysis in adults was 1.00 (95% CI 0.21-4.81) for Arepanrix-AS03 and 0.71 (95% CI 0.16-3.14) for Focetria-MF59. The risk with Pandemrix-AS03 could not be estimated in the restricted period analysis due to the paucity of cases.

The total period analyses including all cases, as well as the separate case-coverage analysis in the Netherlands did not reveal an increased risk of narcolepsy in children or adults. The OR of the total period analysis in children ≤18

years was 0.80 (95% CI 0.21-3.01) for Arepanrix-AS03 and 1.40 (95% CI 0.43-4.64) for Focetria-MF59. In the case-coverage analysis, of the nine child cases born from 2004 through 2009, seven were exposed to Pandemrix-AS03, yielding an OR of 1.44 (95% CI 0.30-6.98). The OR of the total period analysis in adults was 1.00 (95% CI 0.21-4.81) for Arepanrix-AS03, 0.65 (95% CI 0.14-2.95) for Focetria-MF59, and 0.66 (95% CI 0-3.3) for Pandemrix-AS03 (Figure 3).

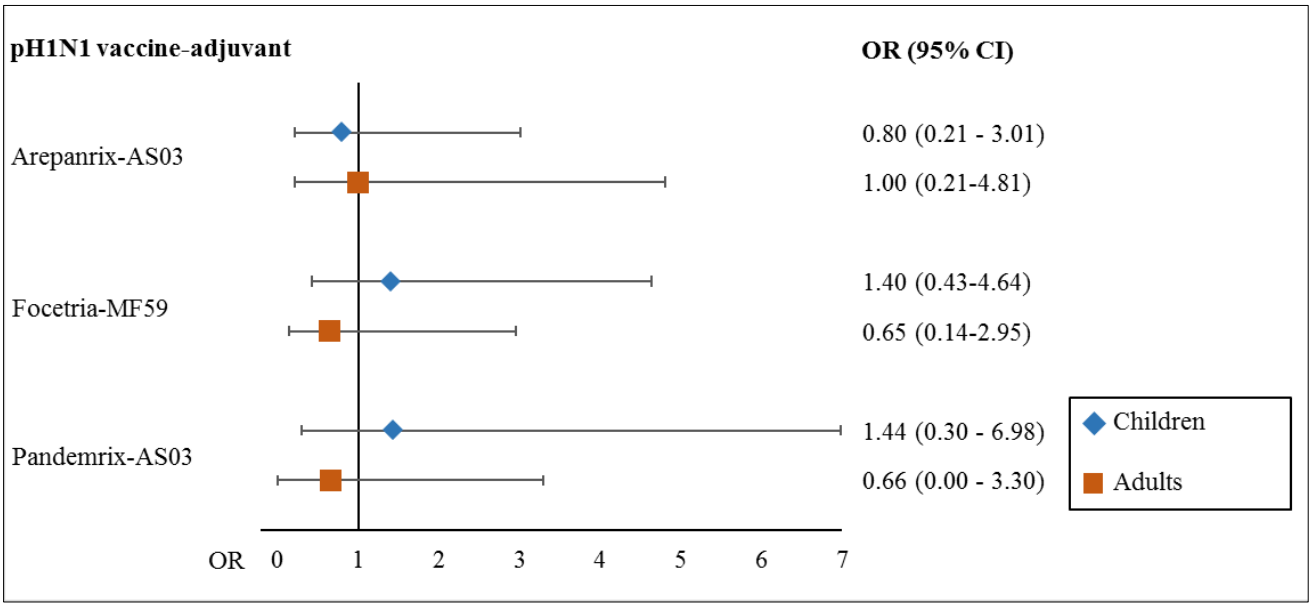


Figure 3: Odds ratios (OR) for narcolepsy and 95% confidence intervals (CI) by vaccine brands in the total period analysis for children (≤18 years) and adults (≥19 years). Arepanrix-AS03: case-control study in Ontario, Canada. Focetria-MF59: two-stage random effects meta-analysis of data from Taiwan, Argentina, the Netherlands, and Valencia and Catalonia, Spain. Pandemrix-AS03: case-coverage study in the Netherlands for children and case-control study in Valencia, Spain, for adults.

Sensitivity analyses using EDS onset as index date for the total period reduced the number of cases considerably, either due to an EDS date before April 1, 2009 or missing EDS onset date. However, this analysis did not substantially alter the main findings of the total period analysis, which used MSLT referral as index date. For Arepanrix-AS03, OR estimates lowered to 0.29 (95% CI 0.03-2.65) in children and remained 1.00 (95% CI 0.05-18.9) in adults. For Focetria-MF59, the pooled estimate ORs were 2.06 (95% CI: 0.63-6.72) in children and 0.65 (95% CI: 0.14-2.95) in adults. For Pandemrix-AS03, ORs were 0.48 (95% CI 0.15-1.58) in children and 1.12 (95% CI 0-6.1) in adults.

Discussion

The SOMNIA study, a multi-country effort that included data from sites on four continents, did not find an increase in narcolepsy IRs associated with pH1N1 vaccination campaigns (except in Sweden, a previously identified signaling country) nor did it detect significant associations between narcolepsy following any adjuvanted pH1N1 vaccines studied. To our knowledge, this is the largest and most geographically diverse study with the longest study period examining the association between adjuvanted pH1N1 vaccines and narcolepsy. Oil-in-water adjuvants like AS03 and MF59 increase immunogenicity of influenza vaccines making them attractive (or possibly necessary) for use in future pandemic influenza vaccines, and they may have the potential to improve the performance of seasonal influenza vaccines.[30] Assessing the safety of these adjuvants has substantial public health and clinical importance.

In our study, the IR of narcolepsy increased in Sweden beginning in summer 2010 and declined in 2011, especially in children 5-19 years. No similar increase in IR of narcolepsy was observed at any of the other participating sites. However, in Taiwan, a site with a very short lag time between symptom onset and diagnosis of narcolepsy, a significant increase in IR of narcolepsy was observed during circulation of wild-type pH1N1 virus but prior to pH1N1 vaccination, a phenomenon also reported in China, which had low (unadjuvanted) vaccination coverage.[31]

Ontario, where vaccination coverage was 32.2%[32], provided case-control data on Arepanrix-AS03; no association with narcolepsy was found in children or adults, which contrasts with a prior finding from a study in Quebec, where a small increase in risk was found, although with large confidence intervals.[23] Since Arepanrix-AS03 has the same adjuvant and a similar pH1N1 antigen as the Pandemrix-AS03 vaccine used in Finland and Sweden, it appears that the association for Pandemrix-AS03 and narcolepsy observed in some European countries is not likely due to the AS03 adjuvant or the pH1N1 antigen in the vaccine alone. Focetria-MF59 was not associated with narcolepsy in children or adults in this study, although the upper limit of the confidence interval cannot exclude a small increase in risk in children. This is consistent with prior observations of a lack of association (or just a few case reports) in countries using this vaccine.[33] Because of low vaccination coverage in sites for Pandemrix-AS03, we were constrained in our ability to evaluate the possibility of an association in our case-control study; our data from the Netherlands are based on a small number of children who were between six months and five years of age at vaccination.

The SOMNIA study contributes to our understanding of the epidemiology of narcolepsy before, during, and after the 2009 H1N1 influenza pandemic and pH1N1 vaccination campaigns. Importantly, it contributes to the body of evidence regarding the possible association between adjuvanted pH1N1 vaccines and narcolepsy as an adverse event following immunization. Prior data, mostly from Europe, predominantly involved Pandemrix-AS03 exposure and included many cases diagnosed shortly after increased media attention and public awareness. Inclusion of the cases diagnosed in the time period immediately after heightened media attention, may overestimate an association, which is supported by the simulations of Wijnans et al.[35] Awareness about a possible association with Pandemrix-AS03 may have resulted in vaccinated cases being diagnosed sooner than they normally would have been, resulting in apparent clusters of narcolepsy.[36] We attempted to address this potential bias by including countries outside the EU and countries in the EU that experienced less media and public awareness; we also conducted an analysis restricting cases in Europe to those diagnosed before attention arose. Additionally, we minimized potential bias from accelerated diagnosis of vaccinated cases by recruiting cases up to five years after knowledge of a possible association became widespread. The findings of our SOMNIA study indicate that the impact of media and public awareness, and the resulting detection bias, may have diminished over time. Our risk estimates of narcolepsy following adjuvanted pH1N1 vaccination for the restricted period analysis compared to the total period analysis did not differ substantially. This is in contrast to the prior Vaccine Adverse Event Surveillance & Communication (VAESCO) study of EU countries, which had a much shorter case accrual period.[36] Overall, our study did not find an increased risk of narcolepsy following vaccination with AS03- or MF59-adjuvanted pH1N1 vaccines based on the total period analysis.

Although we did not detect any significant associations between adjuvanted pH1N1 vaccination and narcolepsy beyond Sweden, we acknowledge the overall trend of evidence of an increased risk associated with Pandemrix-AS03.[5,14,15,23] A biologic mechanism to explain this observation has not been established, but it has been postulated that an interaction involving the immune responses to administration of Pandemrix-AS03 and infection with wild-type pH1N1 virus could be a contributing factor.[20] This would explain the apparent presence of an association in Finland, Sweden and Norway where wild-type virus circulated coincident with the vaccination program, whereas no association was seen in Ontario where wild-type virus was no longer circulating at the time of the vaccination

program. We were not able to address this hypothesis in the SOMNIA study, but it remains a focus area that will likely require a global cooperative research effort.

This retrospective and observational study has several limitations. Despite multiple study sites, statistical power in the restricted period analysis was limited due to fewer cases and low pH1N1 vaccination coverage (online supplement Table 2). For Focetria-MF59, the only country with high coverage in the at-risk age group was Argentina, but the number of cases was low. Although we used standardized case definitions, case misclassification could have occurred as diagnostic procedures and information capture differed across sites. Nonetheless, our application of the Brighton Collaboration case definition criteria for all cases decreased the likelihood of misclassification. Misclassification of exposure could have happened from incomplete recording of pH1N1 vaccinations due to lack of access to school vaccination records in Taiwan. Non-exhaustive inclusion of cases in Ontario (due to distance) and the Netherlands (timeliness of consent) led to incomplete inclusion of cases at the data lock point. For the Netherlands, where consent for study participation was needed, vaccinated cases might be more likely to give consent (given public awareness), which could overestimate the risk; however no significant association was observed. For Ontario, data collection had not been totally completed by the end of the study, but this was not related to vaccination status and therefore impacted power alone. Selection bias was detected in Switzerland for the adult cases (not for children), when we compared our cases to a published series of exposed cases,[37] demonstrating lack of inclusion of exposed cases in adults – this led to exclusion of adult cases and controls in Switzerland. Finally, although we used a common protocol across all sites, it was necessary to provide flexibility for implementation at the local study site level for identifying cases and controls. This might have introduced variability; however, the main criterion in selecting controls was that they be representative of individuals receiving pH1N1 vaccination in the general population and that exposure information was obtained similarly for cases and controls.

Conclusion

We did not observe increases in population-based narcolepsy IRs associated with pH1N1 vaccination campaigns at participating sites, except for Sweden. In the case-control analysis, we did not detect evidence of a significant increased risk of narcolepsy following any of the adjuvanted pH1N1 vaccines in children or adults in our study population, although upper limits of estimates do not exclude small to moderate risks. The SOMNIA study highlights the usefulness of international collaboration in the evaluation of vaccine safety signals for rare adverse events. Using a common protocol and methods reduces heterogeneity, permits contribution of data from countries across the world, and allows for combining data for increased statistical power necessary to address questions about rare events.

References

1. Medical Products Agency. A registry based comparative cohort study in four Swedish counties of the risk for narcolepsy after vaccination with Pandemrix - A first and preliminary report, by the Medical Products Agency.2011. https://lakemedelsverket.se/upload/nyheter/2011/PandemrixRegReport110328.pdf

2. Partinen M, Saarenpaa-Heikkila O, Ilveskoski I, et al. Increased incidence and clinical picture of childhood narcolepsy following the 2009 H1N1 pandemic vaccination campaign in Finland. PLoS One 2012;7:e33723.

3. European Medicines Agency. Pandemrix: EPAR - Product Information2016 10/06/2016. http://www.ema.europa.eu/ema/index.jsp?curl=pages/includes/document/document_detail.jsp?webContentId=WC500038121&mid=WC0b01ac058009a3dc

4. Bardage C, Persson I, Ortqvist A, Bergman U, Ludvigsson JF, Granath F. Neurological and autoimmune disorders after vaccination against pandemic influenza A (H1N1) with a monovalent adjuvanted vaccine: population based cohort study in Stockholm, Sweden. BMJ 2011;343:d5956.

5. Nohynek H, Jokinen J, Partinen M, et al. AS03 adjuvanted AH1N1 vaccine associated with an abrupt increase in the incidence of childhood narcolepsy in Finland. PLoS One 2012;7:e33536.

6. Gadroen K, Straus SM, Pacurariu A, Weibel D, Kurz X, Sturkenboom MC. Patterns of spontaneous reports on narcolepsy following administration of pandemic influenza vaccine; a case series of individual case safety reports in Eudravigilance. Vaccine 2016.

7. Kornum BR, Faraco J, Mignot E. Narcolepsy with hypocretin/orexin deficiency, infections and autoimmunity of the brain. Current opinion in neurobiology 2011;21:897-903.

8. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E. Hypocretin (orexin) deficiency in human narcolepsy. Lancet 2000;355:39-40.

9. Scammell TE. Narcolepsy. N Engl J Med 2015;373:2654-62.

10. Overeem S, Black JL, 3rd, Lammers GJ. Narcolepsy: immunological aspects. Sleep Med Rev 2008;12:95-107.

11. Luca G, Haba-Rubio J, Dauvilliers Y, et al. Clinical, polysomnographic and genome-wide association analyses of narcolepsy with cataplexy: a European Narcolepsy Network study. J Sleep Res 2013;22:482-95.

12. Dauvilliers Y, Arnulf I, Lecendreux M, et al. Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. Brain : a journal of neurology 2013;136:2486-96.

13. Heier MS, Gautvik KM, Wannag E, et al. Incidence of narcolepsy in Norwegian children and adolescents after vaccination against H1N1 influenza A. Sleep medicine 2013;14:867-71.

14. Miller E, Andrews N, Stellitano L, et al. Risk of narcolepsy in children and young people receiving AS03 adjuvanted pandemic A/H1N1 2009 influenza vaccine: retrospective analysis. BMJ 2013;346:f794.

15. O'Flanagan D, Barret AS, Foley M, et al. Investigation of an association between onset of narcolepsy and vaccination with pandemic influenza vaccine, Ireland April 2009-December 2010. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 2014;19:15-25.

16. Persson I, Granath F, Askling J, Ludvigsson JF, Olsson T, Feltelius N. Risks of neurological and immune-related diseases, including narcolepsy, after vaccination with Pandemrix: a population- and registry-based cohort study with over 2 years of follow-up. Journal of internal medicine 2014;275:172-90.

17. Szakacs A, Darin N, Hallbook T. Increased childhood incidence of narcolepsy in western Sweden after H1N1 influenza vaccination. Neurology 2013;80:1315-21.

18. Oberle D, Pavel J, Mayer G, Geisler P, Keller-Stanislawski B, German Narcolepsy Study G. Retrospective multicenter matched case-control study on the risk factors for narcolepsy with special focus on vaccinations (including pandemic influenza vaccination) and infections in Germany. Sleep medicine 2017;34:71-83.

19. Verstraeten T, Cohet C, Dos Santos G, et al. Pandemrix and narcolepsy: A critical appraisal of the observational studies. Hum Vaccin Immunother 2015:1-7.

20. Sturkenboom MC. The narcolepsy-pandemic influenza story: can the truth ever be unraveled? Vaccine 2015;33 Suppl 2:B6-B13.

21. Sarkanen TO, Alakuijala APE, Dauvilliers YA, Partinen MM. Incidence of narcolepsy after H1N1 influenza and vaccinations: Systematic review and meta-analysis. Sleep Med Rev 2018;38:177-86.

22. Choe YJ, Bae GR, Lee DH. No association between influenza A(H1N1)pdm09 vaccination and narcolepsy in South Korea: an ecological study. Vaccine 2012;30:7439-42.

23. Montplaisir J, Petit D, Quinn MJ, et al. Risk of narcolepsy associated with inactivated adjuvanted (AS03) A/H1N1 (2009) pandemic influenza vaccine in Quebec. PLoS One 2014;9:e108489.

24. Trifiro G, Coloma PM, Rijnbeek PR, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? Journal of internal medicine 2014;275:551-61.

25. Poli F, Overeem S, Lammers GJ, et al. Narcolepsy as an adverse event following immunization: case definition and guidelines for data collection, analysis and presentation. Vaccine 2013;31:994-1007.

26. Wijnans L, Lecomte C, de Vries C, et al. The incidence of narcolepsy in Europe: before, during, and after the influenza A(H1N1)pdm09 pandemic and vaccination campaigns. Vaccine 2013;31:1246-54.

27. Stowe J, Andrews N, Kosky C, et al. Risk of Narcolepsy after AS03 Adjuvanted Pandemic A/H1N1 2009 Influenza Vaccine in Adults: A Case-Coverage Study in England. Sleep 2016;39:1051-7.

28. Association WM. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects2013. http://www.wma.net/en/30publications/10policies/b3/index.html.pdf?print-media-type&footer-right=[page]/[toPage]

29. Dodd CN, de Ridder M, Huang W, et al. Incidence rates of narcolepsy diagnoses in Taiwan, Canada, and Europe: methods for rapid assessment of potential safety issues on a population level in the SOMNIA study. PLoS One 2018;In Press.

30. Even-Or O, Samira S, Ellis R, Kedar E, Barenholz Y. Adjuvanted influenza vaccines. Expert Rev Vaccines 2013;12:1095-108.

31. Han F, Lin L, Warby SC, et al. Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. Ann Neurol 2011;70:410-7.

32. Gilmour H, Hofmann N. H1N1 vaccination. Health reports 2010;21:63-9.

33. Ahmed SS, Schur PH, MacDonald NE, Steinman L. Narcolepsy, 2009 A(H1N1) pandemic influenza, and pandemic influenza vaccinations: what is known and unknown about the neurological disorder, the role for autoimmunity, and vaccine adjuvants. Journal of autoimmunity 2014;50:1-11.

34. Mereckiene J, Cotter S, Weber JT, et al. Influenza A(H1N1)pdm09 vaccination policies and coverage in Europe. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 2012;17.

35. Wijnans L, Dodd C, de Ridder M, et al. Pandemic influenza vaccine & narcolepsy: simulations on the potential impact of bias. Expert Rev Vaccines 2016;15:573-84.

36. ECDC. European Centre for Disease Prevention and Control. Narcolepsy in association with pandemic influenza vaccination (a multi-country European epidemiological investigation). Stockholm: ECDC; 2012 September 2012. http://ecdc.europa.eu/en/publications/Publications/Vaesco%20report%20FINAL%20with%20cover.pdf

37. Kallweit U, Mathis J, Jenni OG, et al. Post-H1N1 Flu Vaccination Narcolepsy in Switzerland: A Retrospective Survey in the 30 Sleep-Certified Swiss Centers. Eur Neurol 2016;75:105-8.

Supplementary Material

Supplementary Table 1: Number of cases with a diagnosis of narcolepsy in the period before and after awareness in the EU and MSLT referral after study start and among them availability of EDS/cataplexy date (for those only EDS dates after April 1, 2009 are included)

	Nether-lands	Switzerland	Spain, Catalonia	Spain, Valencia	Argentina	Canada, Ontario	Taiwan	Total
Children								
Total cases included [N]	**22***	**22**	**5**	**11**	**11**	**28**	**51**	**150**
Diagnosis before awareness in EU (August 2010) & MSLT referral after study start (April 1, 2009)	3	1	0	2	1	11	10	28
EDS/cataplexy date available and after study start (April 1, 2009)	1	0	0	0	1	<6	10	<19
Delay between EDS and diagnosis (Median days)#	1735	453		878		481	140	
Diagnosis after awareness in EU (Jul 31, 2010) & MSLT referral after study start (April 1, 2009)	**19**	21	5	9	10	17	41	122
EDS/cataplexy date available and after study start (April 1, 2009)	15	13	2	6	9	10	39	100
Delay between EDS and diagnosis (Median days)#	808	571	761	619	238	596	127	
Adults								
Total cases included [N]	**32**		**13**	**36**	**4**	**39**	**86**	**210**
Diagnosis before awareness in EU (August 2010) & MSLT referral after study start (April 1, 2009)	7		5	6	0	11	20	49
EDS/cataplexy date available for (Dx cases before Aug 2010)	0		0	0	0	<6	20	<26
Delay between EDS and diagnosis (Median days)#	2066		-	870	-	738	118	
Diagnosis after awareness in EU (Jul 31, 2010) & MSLT referral after study start (April 1, 2009)	25		8	30	4	28	66	161

	Nether-lands	Switzerland	Spain, Catalonia	Spain, Valencia	Argentina	Canada, Ontario	Taiwan	Total
EDS/cataplexy date available and after study start (April 1, 2009)	6		3	23	3	8	61	106
Delay between EDS and diagnosis (Median days)#	2654		1181	595	952	713	135	

*Nine child cases born between 2004 and 2009 were included in a case-coverage study for reasons described in the methods nd not the case control The child case total for the Netherlands includes nine from the case-coverage study and 13 from the case-control study.

**Cell values that represent case counts of five or fewer (for Ontario, Canada) or two or fewer (for Taiwan) cases may not be reported as absolute numers due to patient privacy regulations and are represented as range (i.e. ≤ n).

Median delays calculated between date of available EDS date and Diagnosis date in subjects diagnosed after start of study period (April 1, 2009), not considering date of MSLT or EDS to be after study entry

Supplementary Table 2: Overview of national immunization programs for H1N1pdm09

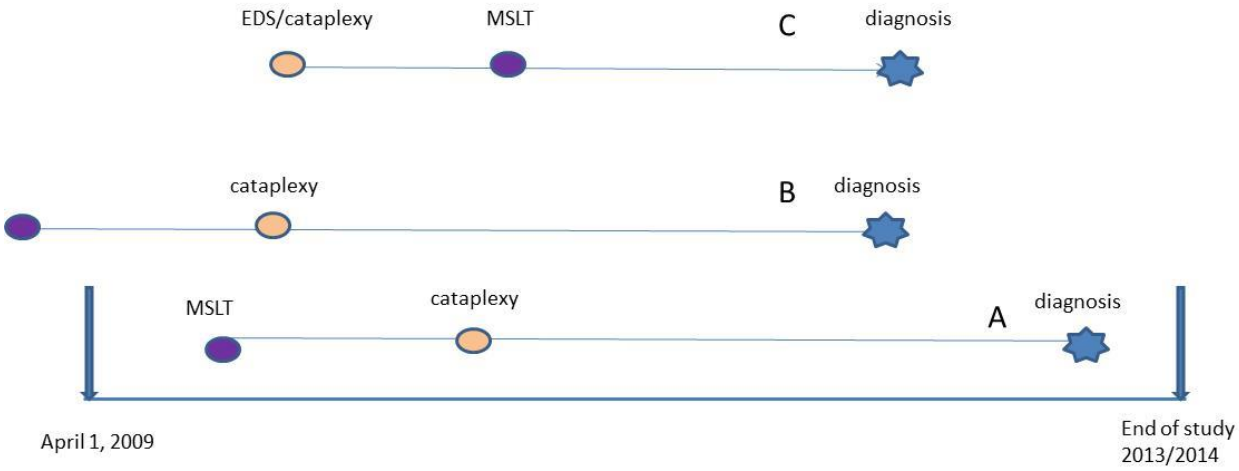
Country	Vaccine(s) Used	Target Population	Population Based overage Rates
Argentina	Focetria	Risk groups	< 4 yo in risk groups=86% > 4 yo in risk groups= 99%
Canada	Arepanrix, Unadjuvanted for pregnant women only, Panvax H1N1 (CSL) for pregnant women late in program	Entire population	In those aged ≥12 years: Ontario: 32.2% (30.3%-34.0%) Manitoba: 37.2% (33.2%-41.2%) Alberta: 37.1% (33.9%-40.2%) British Columbia: 35.6% (32.8%-38.4%)
Denmark	Pandemrix	Risk groups	6% among adults
Netherlands	Pandemrix (<5 yo) Focetria (> 6mo) >5 yo Focetria, Pandemrix (family of children < 5 yo)	Risk groups	< 5 yo= 75% Risk groups 70%
Spain Valencia	Focetria 6 -17 years Pandemrix 18 to 59 y-o, Focetria > 60, Panenza for pregnant women	Risk groups	**SIDIAP:** < 18: 0.82%, all population 3.5 % **Valencia:** 6 months-14 years (with risk factors):11% 15-59 years (with risk factors): 13% 60 years or older (with risk factors): 28% Pregnant women= 9% Healthcare workers & professionals providing special services: 30%
Sweden	Pandemrix	Entire population	< 18 yo = 12% > 18 yo = 12-14%

Switzerland	Focetria (<6mo-18 yoa, pregnant women) Pandemrix (>18yoa), Celtura (>3yoa), Fluzone (>3months)	Risk groups	< 18 yo= 10% > 18 yo = 20%
Taiwan	AdimFlu-S, unadjuvanted (≥1 yo) Focetria (≥6 mo)	Entire population	6 mo–18 yo = 67% ≥19 yo = 12%

Supplementary table 3: Table Population based frequency of the HLA DQB1*0602 polymorphism at selected sites.

Country	HLA DQB1*0602 Polymorphism Population Based Frequency	Source
Taiwan	3.4%	Chen P.L., et al., Comprehensive Genotyping in Two Homogeneous Graves' Disease Samples Reveals Major and Novel HLA Association Alleles. PLoS ONE 2011
Argentina	15.2%	Caputo M., et al., GENOTIPIFICACION DEL GEN HLA DQB1 EN DIABETES AUTOINMUNE DEL ADULTO (LADA). MEDICINA (Buenos Aires) 2005
Canada (Ontario)	18 - 24.8% (depending on source)	Personal communication, Kathryn Tinckam, University of Toronto, ON, Canada Kotb M., et al., An immunogenetic and molecular basis for differences in outcomes of invasive group A streptococcal infections, *Nature Medicine* 2002
Canada (BC)	25.33	Poirier G., et al., HLA Antigens in Narcolepsy and Idiopathic Central Nervous System Hypersomnolence. Sleep 1986
The Netherlands	24%	Tafti M., et al., DQB1 Locus Alone Explains Most of the Risk and Protection in Narcolepsy with Cataplexy in Europe. Sleep 2014
Spain (IDIAP)	15%	Balas A., et al., Allelic and haplotypic HLA frequency distribution in Spanish hematopoietic patients. Implications for unrelated donor searching. Tissue Antigens 2011
Spain (FISABIO)	14.5%	Crespi C., et al., HLA polymorphism in a Majorcan population of Jewish descent. Tissue Antigens 2002
Switzerland	12-21% (depending on source)	Personal communication, Jan Bonhoeffer, University Children's' Hospital Basel, Switzerland. Buhler S., et al., The Heterogeneous HLA Genetic Makeup of the Swiss Population. PLoS ONE 2012

Supplementary Figure 1: Graphic representation of patient inclusion for primary index date analysis: Patient A has MSLT after April 2009 and is included, patient B will be excluded since MSLT onset is prior to start of study period. Patient C is included



5.3 Enhancing global vaccine pharmacovigilance: Proof-of-concept study on aseptic meningitis and immune thrombocytopenic purpura following measles-mumps containing vaccination.

Silvia Perez-Vilar

Daniel Weibel

Miriam Sturkenboom

Steven Black

Christine Maure

Jose Luis Castro

Pamela Bravo-Alcántara

Caitlin N. Dodd

Silvana A. Romio

Maria de Ridder

Swabra Nakato

Helvert Felipe

Molina-León

Varalakshmi Elango

Patrick L.F. Zuber

WHO Global Vaccine Safety-Multi Country Collaboration

Vaccine 36.3 (2018): 347-354.

Abstract

New vaccines designed to prevent diseases endemic in low and middle-income countries (LMICs) are now being introduced without prior record of utilization in countries with robust pharmacovigilance systems. To address this deficit, our objective was to demonstrate feasibility of an international hospital-based network for the assessment of potential epidemiological associations between serious and rare adverse events and vaccines in any setting. This was done through a proof-of-concept evaluation of the risk of immune thrombocytopenic purpura (ITP) and aseptic meningitis (AM) following administration of the first dose of measles-mumps-containing vaccines using the self-controlled risk interval method in the primary analysis. The World Health Organization (WHO) selected 26 sentinel sites (49 hospitals) distributed in 16 countries of the six WHO regions. Incidence rate ratios (IRR) of 5.0 (95% CI: 2.5-9.7) for ITP following first dose of measles-containing vaccinations, and of 10.9 (95% CI: 4.2-27.8) for AM following mumps-containing vaccinations were found. The strain-specific analyses showed significantly elevated ITP risk for measles vaccines containing Schwarz (IRR: 20.7; 95% CI: 2.7-157.6), Edmonston-Zagreb (IRR: 11.1; 95% CI: 1.4-90.3), and Enders´Edmonston (IRR: 8.5; 95% CI: 1.9-38.1) strains. A significantly elevated AM risk for vaccines containing the Leningrad-Zagreb mumps strain (IRR: 10.8; 95% CI: 1.3-87.4) was also found. This proof-of-concept study has shown, for the first time, that an international hospital-based network for the investigation of rare vaccine adverse events, using common standardized procedures and with high participation of LMICs, is feasible, can produce reliable results, and has the potential to characterize differences in risk between vaccine strains. The completion of this network by adding large reference hospitals, particularly from tropical countries, and the systematic WHO-led implementation of this approach, should permit the rapid post-marketing evaluation of safety signals for serious and rare adverse events for new and existing vaccines in all settings, including LMICs.

Introduction

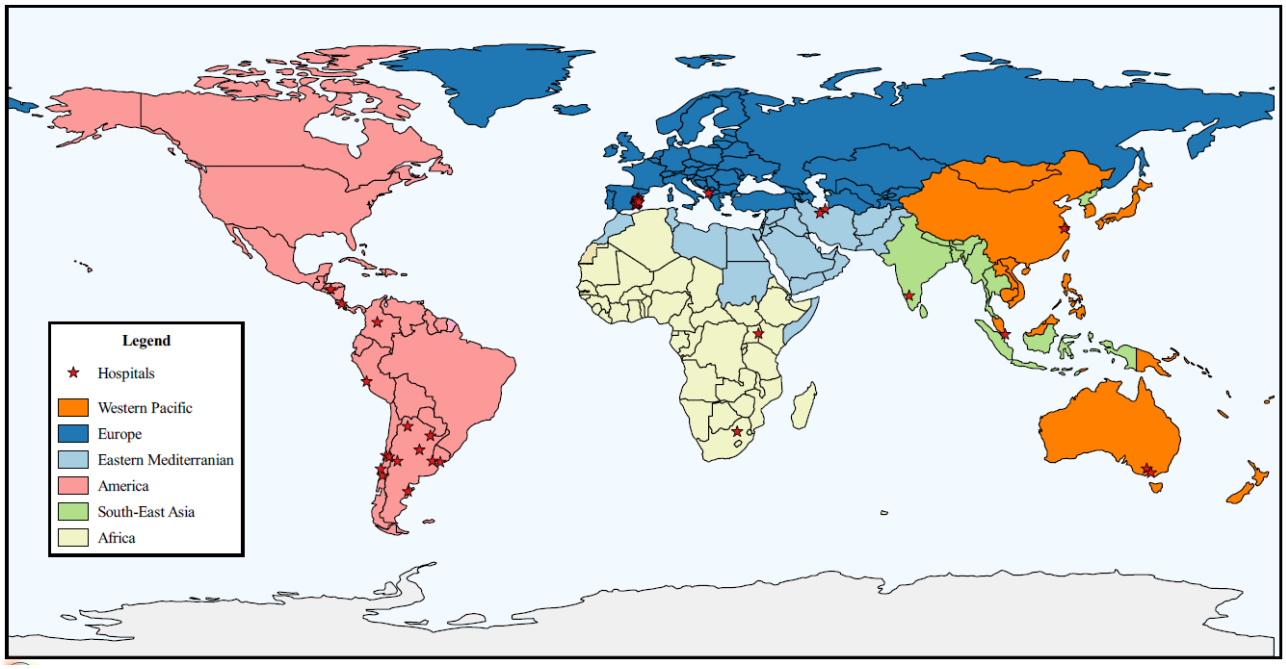
With increasing number of vaccine products available, expansion of vaccine manufacturing capabilities, and availability of new vaccines targeted against diseases highly prevalent in low and middle-income countries (LMICs) (1), there is a need to enhance vaccine pharmacovigilance infrastructures globally (2). Many countries do not have technical capacity and/or large enough populations to permit the evaluation of rare adverse events following immunization (AEFI) (2, 3). Enhancement of vaccine pharmacovigilance capabilities is a key activity for the World Health Organization (WHO) Global Vaccine Safety Initiative (GVSI) (4-6). A previous international pilot study sponsored by WHO and the Food and Drug Administration (FDA), to evaluate the safety of the 2009-10 pandemic influenza vaccine, demonstrated that multinational hospital-based vaccine safety studies were feasible and could provide a useful framework for the evaluation of safety concerns (7). Optimization of operational models, centralization of case adjudication, improvements in data quality control, closer supervision of data abstraction, and demonstration of the feasibility of such international collaborations, with high participation from LMICs, were identified by WHO as issues to be resolved (7). Thus, for a subsequent demonstration project, it was important to reach higher participation from LMICs, select a vaccine widely used, and an AEFI that, at least in severe cases, would require hospitalization (2). It was also essential to select an AEFI known to be associated with some of the vaccine strains being used.

Measles-containing vaccines are live-attenuated, often given in combination with mumps and rubella vaccines. The first dose is usually given at one year of age, although it is administered at nine months of age in countries with ongoing measles transmission (8). The second dose is either given at 15-18 months of age, at 4-6 years of age, or in campaigns. Our objective was to demonstrate feasibility of an international hospital-based network for assessing epidemiological associations between rare adverse events and vaccines in any setting, including LMICs. Two well-established associations were chosen: risk of aseptic meningitis (AM) following first dose of mumps-containing vaccines (9-11), and risk of immune thrombocytopenic purpura (ITP) following first dose of measles-containing vaccines (8, 12-14).

Methods

International hospital-based retrospective observational study conducted as proof-of-concept for the investigation of rare AEFI using two analytical case-only methods: self-controlled risk interval (SCRI) and case-crossover (15, 16). For this purpose, WHO selected 26 sentinel sites (49 hospitals) distributed in 16 countries of the six WHO regions (Figure 1). Selection criteria and capability assessments are described elsewhere (Bravo-Alcántara P, Perez-Vilar S, Molina-León HF et al. (accepted for publication in Vaccine)).

Figure 1. Geographical distribution of participating hospitals in the WHO regions



Disclaimer: Lines on the map represent approximate border lines for which there may not yet be full agreement.

Study population

The study population included children ages 9-23 months admitted to a network-participating hospital during January 2010-March 2014, with a discharge diagnosis of either AM or ITP. Only individuals living in the pre-defined catchment area of the hospital, or, for those hospitals without a pre-specified catchment area, in the same city in which the hospital was located, were eligible.

Case ascertainment and classification

Participating hospitals identified potential cases through hospital discharge databases using pre-specified ICD-9/ICD-10 codes (Supplementary material; Table S-1) whereas hospitals not using a discharge codification system or not having electronic databases used free text. A trained physician or nurse blinded to vaccination status reviewed medical records of potential cases according to established case definitions (Supplementary material; Tables S-2 and S-3). Potential cases for which medical records were not available were excluded. Only first episodes of AM or ITP were considered.

Potential AM cases were excluded if they met criteria for encephalitis (17) (Supplementary material; Table S-4), the medical records showed that a physician ruled out a diagnosis of AM, a meningitis pathogen other than mumps virus was identified in cerebrospinal fluid (CSF), CSF protein concentration (in absence of traumatic lumbar puncture or intracerebral event) was ≥50mg/dL with ≥10 leukocytes/mm^3 and glucose ≤40mg/dL in CSF, or if polymorphonuclear leukocytes (PMNs) in the CSF were >1,000/mm3 with glucose ≤40mg/dL (modified from Lussiana et al.) (18).

Potential ITP cases were excluded if classified as chronic (defined as lasting >6 months) (12, 14), with onset of symptoms occurring >42 days prior to hospital admission, or if a physician diagnosis in the medical records ruled

out the diagnosis of ITP or thrombocytopenia. ITP cases with medical conditions associated with higher ITP risk (congenital/hereditary thrombocytopenia, aplastic anemia, defibrination syndrome, acquired hemolytic anemia, chronic liver disease, malignancy, or drug-induced thrombocytopenia) were also excluded. For the analyses presented here, patients treated with platelet-depleting medications (amiodarone, heparin, carbamazepine, phenytoin, valproic acid, quinidine, quinine, rifampicin, ethambutol, sulfisoxazole, vancomycin, ampicillin, trimethoprim-sulfamethoxazole, naproxen, or ranitidine) during hospitalization or in the 42 days prior, unless there was evidence that the drug was administered after disease onset date, were also excluded.

All cases were classified as either confirmed (Level 1-3 of diagnosis certainty) or non-confirmed (Supplementary material; Tables S-2 and S-3). Only confirmed cases entered the analyses.

The event date for AM cases was onset date of signs and symptoms suggestive of meningitis, admission date, or date of first physician diagnosis, whichever occurred earlier. The event date for ITP cases was onset date of spontaneous bleeding (19), date of first laboratory result with a platelet count <50,000/µL performed within 42 days prior to hospital admission or during hospitalization, admission date, or date of first physician diagnosis, whichever occurred earlier.

Vaccination status

Vaccination status was retrieved, for confirmed cases only, from vaccine registries, vaccination cards, and medical records. The exposure of interest was first dose of measles/mumps-containing vaccine. Patients were considered as non-vaccinated when any other vaccinations, but not measles-containing vaccines, were registered in the consulted sources. Individuals without any vaccination record were excluded from the study.

Data collection and sharing

Sites collected data using a common protocol, and transferred them into electronic case report forms using the purpose-built Chameleon® system (Erasmus Medical Center (EMC)). Chameleon® classified the cases automatically according to their level of diagnostic certainty. Outcome and exposure-coded datasets containing non-identifiable time interval-only data created by Chameleon® were uploaded to a central remote research environment, located at EMC, through a secure connection.

Quality assurance

In parallel with the study protocol and manual of procedures, a quality assurance plan was developed. It included roles and responsibilities for feasibility assessment, protocol development, data collection/transformation, analysis and reporting. The coordination team trained investigators through on-site and/or virtual meetings and through a simulation exercise using dummy cases, reviewed data submitted using standardized procedures, and sent reports to the sites detailing inconsistencies and missing data found. Following these communications, sites were asked to submit final data for analyses. Detailed information on quality assurance activities implemented and operating procedures followed for data collection, entry, and submission can be found elsewhere (Bravo-Alcántara P, Perez-Vilar S, Molina-León HF et al. (accepted for publication in Vaccine)).

Statistical analyses

The risks of AM following mumps-containing vaccination and ITP following measles-containing vaccination were estimated using self-controlled risk interval (SCRI) analyses (15, 20, 21). The observation period started on the day following first-dose vaccination and ended on day 84 post-vaccination. Days 8-35 were considered the risk period, days 1-7 and 36-42 washout periods, and days 43-84 the non-risk period. Thus, only vaccinated cases for which the

event occurred within 84 days following vaccination were included. Poisson regression conditioned on the fact that the event occurred was used to estimate the incidence rate ratio. Differential risk of AM and ITP in the risk and non-risk windows due to circulation of wild viruses linked to the diseases of interest and age were adjusted for in the models as follows: (1) cut-off points for seasonality were March 31, June 30, September 30, and December 31; (2) age was controlled for with periods ending at 365, 457, 549, 641 days, and 732 days of age.

Per protocol, a case-crossover design was chosen as secondary analysis (22). The observation period was 84 days prior to event occurrence (case window: days -1 to -42; control window: days -43 to -84). Thus, cases without at least 84 days of follow-up prior to the event were excluded, regardless of vaccination status. The risk periods were days -8 to -35 for the case window and days -50 to -77 for the control window. The remaining periods were considered washout periods. Crude odds ratios were estimated using conditional logistic regression.

One site did not collect complete vaccination dates for any of the confirmed cases; thus, the day of vaccination was randomly imputed by Chameleon® within the month and year provided. Because of the importance of having exact vaccination dates for case-only methods, analyses with and without cases from this site (Iran-01) were performed.

Because the risks for AM and ITP may vary by virus strain, (8-11, 23-25), exploratory analyses were performed by mumps and measles strain received, respectively. The two participating Iranian sites reported that three measles-mumps-rubella (MMR) vaccines, manufactured by Razi Vaccine, Serum Institute of India and Sanofi Pasteur, were used in the country during the study period, but they could not identify which specific product was administered to an individual patient. Thus, a separate analysis for the two Iranian sites was also conducted. Measles/mumps strains included in the vaccine products used by participating countries are shown in Table 1.

Table 1. Measles and mumps strains included in the vaccine products used by the participating countries during the study period

Vaccine product	Measles strain	Mumps strain
Priorix®, GlaxoSmithKline Biologicals	Schwarz	RIT 4385*
Priorix Tetra®, GlaxoSmithKline Biologicals	Schwarz	RIT 4385*
MMR, Shanghai Institute of Biological Products, Co., Ltd.	Shanghai-191	S79
Measles, Lanzhou Institute of Biological Products Co., Ltd.	Shanghai-191	-
Measles-Rubella, Beijing Tiantan Biological Products, Co.,Ltd.	Shanghai-191	-
M-M-R-II®, Merck Sharp & Dohme Corp.	Enders´ Edmonston	Jeryl Lynn (Level B)
MMR, Razi Vaccine and Serum Research Institute	AIK-C	Hoshino
M-M-RVAXPRO®, Sanofi Pasteur-MSD	Enders´Edmonston	Jeryl Lynn (Level B)
Trimovax®, Sanofi Pasteur	Schwarz	Urabe Am9
Measles, Serum Institute of India Pvt. Ltd	Edmonston-Zagreb	-
Measles-Rubella, Serum Institute of India Pvt. Ltd	Edmonston-Zagreb	-
MMR, Serum Institute of India Pvt. Ltd	Edmonston-Zagreb	Leningrad-Zagreb
Tresivac®, Serum Institute of India Pvt. Ltd	Edmonston-Zagreb	Leningrad-Zagreb
Rouvax®, Sanofi Pasteur	Schwarz	-

Abbreviations: MMR (measles-mumps-rubella); * Derived from Jeryl Lynn strain

All analyses were conducted using SAS 9.4 (SAS Institute, Inc., Cary, NC). The WHO Ethics Review Committee and all local Ethics Committees approved the study and provided a waiver of informed consent according to article 32 of the Declaration of Helsinki (26). Given the need for accurate information on vaccination status, a waiver to contact parents or legal representatives in case of lack of vaccination information was also obtained.

Results

A total of 84 confirmed AM cases and 183 confirmed ITP cases were eligible for inclusion in the case-only analyses. Number of confirmed cases successfully linked to vaccination records by site/country, level of diagnosis certainty, and site characteristics, including case ascertainment methods, vaccination data sources, and identifiers used to link exposures and outcomes, are shown in Table 2.

Table 2. Characteristics of participating sentinel sites

Site[1]	Beds	Case ascertainment		Vaccination status ascertainment				Common outcome-exposure identifier			Confirmed aseptic meningitis cases[2]			Confirmed ITP cases[2]		
	(n)	ICD codes	Free text	Electronic vaccine registry	Vaccination cards	Medical records	Parents contacted[3]	Unique identification number	Clinical history number	National identity card	Level 1 (n)	Level 2 (n)	Level 3 (n)	Level 1 (n)	Level 2 (n)	Level 3 (n)
Albania	240	ICD-9	-	✔	✔	-	-	-	✔	-	1	-	-	5	-	-
Argentina-01	330	ICD-10	-	✔	-	-	✔	-	✔	✔	1	-	-	6	-	1
Argentina-02	78	ICD-10	-	✔	-	-	✔	-	✔	✔	-	-	-	1	-	-
Argentina-03	380	ICD-10	-	✔	-	✔	✔	-	✔	✔	-	-	-	4	-	-
Argentina-04	246	ICD-10	-	✔	-	-	-	-	✔	✔	-	-	-	-	-	-
Argentina-05	224	ICD-10	-	✔	-	-	-	-	✔	✔	-	-	-	4	-	-
Argentina-06	61	ICD-10	-	✔	-	-	✔	-	✔	✔	-	-	-	2	-	-
Australia-01	334	ICD-10	-	✔	-	-	-	✔	✔	-	2	5	-	5	-	2
Australia-02	184	ICD-10	-	✔	✔	-	-	✔	✔	-	-	-	1	4	-	-
Chile-01	440	ICD-10	-	✔	-	-	-	-	-	✔	3	-	-	2	-	-
Chile-02	300	ICD-10	-	✔	-	-	✔	-	-	✔	5	-	1	4	-	-
Chile-03	704	ICD-10	-	✔	-	-	-	-	-	✔	-	-	-	6	-	-
Chile-04	876	ICD-10	-	✔	-	-	✔	-	-	✔	-	1	-	5	-	-
China	500+	ICD-10	-	✔	-	-	-	-	-	-	-	-	1	7	-	-
Colombia	340	ICD-10	-	✔	-	-	-	-	-	✔	-	-	-	2	-	-
Costa Rica	313	ICD-10	-	✔	-	-	✔	-	-	✔	1	2	1	13	-	-
Honduras	1,109	ICD-10	-	-	-	✔	✔	-	✔	✔	-	-	-	1	-	-
India	1,200	ICD-9, ICD-10	-	-	-	✔	✔	✔	-	-	3	5	-	1	1	-
Iran-01	246	ICD-10	-	-	-	✔	-	-	-	-	8	16	2	14	3	-
Iran-02	340	ICD-10	-	-	-	✔	✔	-	-	-	9	6	1	20	-	-
Peru	465	ICD-10	-	-	-	✔	✔	-	-	✔	-	-	-	7	-	-
Singapore	830	ICD-9, ICD-10	-	✔	-	✔	-	✔	✔	✔	-	-	-	17	1	2
South Africa	3,200	ICD-10	-	-	-	✔	✔	-	-	-	-	1	2	-	-	-
Spain[4]	10,987	ICD-9	-	✔	-	-	-	✔	-	-	2	-	3	32	2	6
Uganda	254	-	✔	-	✔	✔	✔	-	-	-	-	-	-	-	-	-
Uruguay	245	ICD-10	-	✔	-	-	-	-	-	✔	1	-	-	3	-	-

[1]The study period was January 2010 to March 2014, except for Australia, which retrospectively included the first 25 most recent cases that fulfill inclusion criteria for each condition (for both sites combined)

[2] Only the highest level of diagnosis certainty achieved applies. The cases correspond to confirmed cases for which a link to vaccination data was available. Confirmed cases for which vaccination status was unknown were excluded from the study

[3] Parents contacted were asked to provide a copy of the vaccination cards

[4] Spain was designated as one site, but included all public hospitals of the Valencia Region, its hospital beds correspond to the total number of beds from the combined hospitals

Among 84 AM cases, 80 (95%) received a first dose of mumps-containing vaccines (Table 3). A total of 51 (61%) and a total of 73 (87%) were eligible for inclusion in the SCRI and case-crossover analyses, respectively. The risk of AM following mumps containing vaccines was 10.9 (95% CI 4.2-27.8) with the SCRI analysis. Sensitivity analyses excluding Iran-01 resulted in an IRR estimate of 11.7 (95% CI 3.5-39.3). Intervals between first dose of mumps-containing vaccine and aseptic meningitis onset for cases included in the strain-specific SCRI analyses are shown in Figure 2a. A significantly increased AM risk was found for the Leningrad-Zagreb mumps strain (IRR: 10.8; 95% CI: 1.3-87.4). Risk estimates for S79, UrabeAm9 and RIT 4385/Jeryl-Lynn strains could not be assessed given small numbers. For the vaccine products used in Iran (Hoshino/Leningrad-Zagreb/UrabeAm9), an IRR of 20.3 (95% CI: 4.8-85.2) was identified (Table 4). Case-crossover analysis produced an overall unadjusted OR of 35.0 (95% CI: 4.8-255.5). When cases from Iran-01 were excluded, the OR estimate was 22.0 (95% CI: 3.0-163.2).

Table 3. Characteristics of children with confirmed aseptic meningitis or immune thrombocytopenic purpura (ITP)

Characteristic	Confirmed aseptic meningitis cases n=84	Confirmed ITP cases n=183
Male sex (n, %)	54 (64%)	98 (54%)
Age at onset in months (median; IQR)	13 (12-15)	15 (12-19)
Mumps-containing first dose vaccination (n, %)	80 (95%)	-
Exact date known (n, %)	60 (75%)	-
Vaccine brand known (n, %)	41 (51%)	-
Age at vaccination in months (median; IQR)	12 (11-12.5)	-
Measles-containing first dose vaccination (n, %)	-	172 (94%)
Exact date known (n, %)	-	159 (92%)
Vaccine brand known (n, %)	-	125 (73%)
Age at vaccination in months (median; IQR)	-	12 (12-15)

Two aseptic meningitis cases died during the observation period, one case in Spain 78 days after disease onset date and another case in Australia 608 days following disease onset. None ITP case was known to die during the observation period.

Table 4. Risk of aseptic meningitis following mumps-containing vaccination and risk of immune thrombocytopenic purpura (ITP) following measles-containing vaccination; overall and by vaccine strain

	SCRI analyses				
Mumps vaccine strain[1]	**Eligible confirmed aseptic meningitis cases[2]**		**Follow-up (days)**	**Relative incidence (IRR)**	
	Event in risk period (8-35 days)	**Event in non-risk period (43-84 days)**	**Median (P25-P75)**	**Unadjusted (95% CI)**	**Adjusted 95% CI**
Overall	35	5	85 (85, 85)	10.9 (4.2-27.8)	10.8 (4.0-29.2)
Overall[3]	22	3	85 (85, 85)	11.7 (3.5-39.3)	12.4 (3.1-49.1)
Hoshino/Leningrad-Zagreb/UrabeAm9	27	2	85 (85, 85)	20.3 (4.8-85.2)	Non-estimable
Hoshino/Leningrad-Zagreb/UrabeAm9*	14	0	85 (85, 85)	Non-estimable	Non-estimable
Leningrad-Zagreb	7	1	85 (85, 85)	10.8 (1.3-	6.4 (0.3-

				87.4)	124.4)
RIT 4385/Jeryl Lynn (Level B)	0	1	85 (85, 85)	Non-estimable	Non-estimable

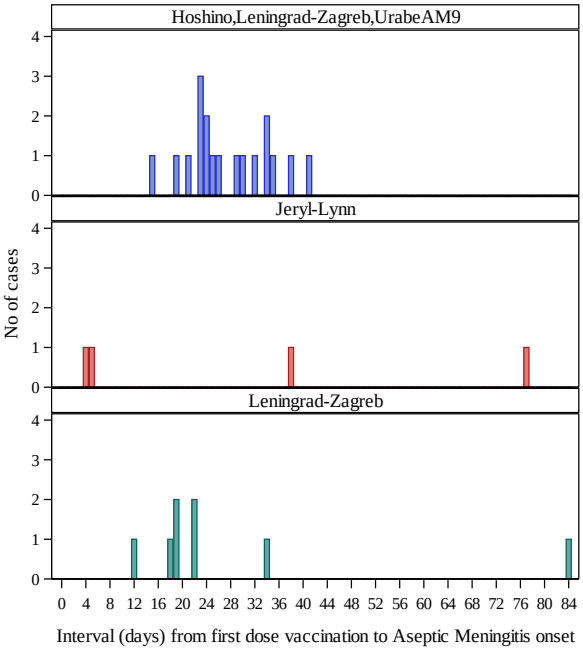
Measles vaccine strain	**Eligible confirmed ITP cases[1]**		**Follow-up (days)**	**Relative incidence (IRR)**	
	Event in risk period (8-35 days)	**Event in non-risk period (43-84 days)**	**Median (P25-P75)**	**Unadjusted 95% CI**	**Adjusted 95% CI**
Overall	36	12	85 (70, 85)	5.0 (2.5-9.7)	5.6 (2.7-11.9)
Overall[3]	36	8	85 (70, 85)	7.7 (3.5-17.3)	9.1 (3.7-22.3)
AIK-C/ Edmonston-Zagreb /Schwarz	2	5	85 (85, 85)	0.51 (0.10-2.54)	0.54 (0.08-3.55)
Edmonston-Zagreb	7	1	85 (67, 85)	11.1 (1.4-90.3)	8.4 (0.7-100.3)
Enders´ Edmonston	11	3	85 (43, 85)	8.5 (1.9-38.1)	28.7 (1.9-443.5)
Schwarz	14	1	85 (76, 85)	20.7 (2.7-157.6)	Non-estimable
Shanghai-191	0	1	85 (85, 85)	Non-estimable	Non-estimable

[1] There were no cases within days 8-35 or days 43-84 following first dose vaccination with mumps strains S79 or Urabe Am9.

[2] The remaining cases occurred during the washout periods (days 1-7, days 36-42 following vaccination)

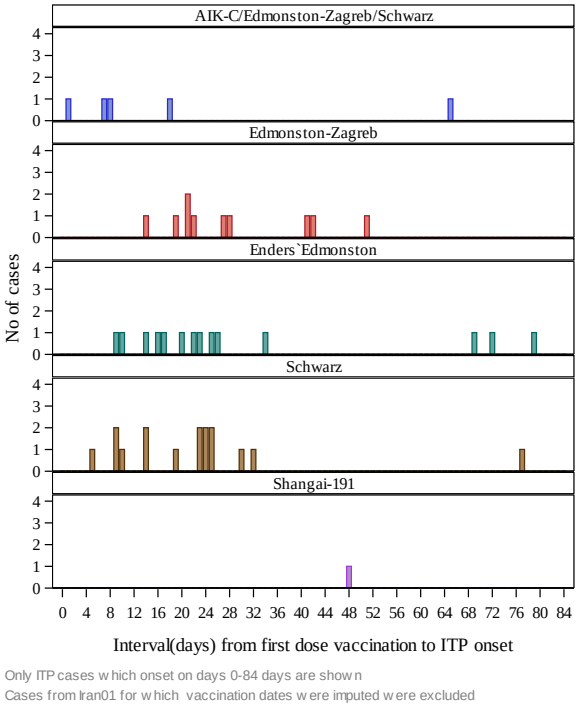
[3] Excluding cases from Iran-01 since this site did not provide exact vaccination dates

Figure 2a. Interval between first dose of mumps-containing vaccines and aseptic meningitis onset by mumps vaccine strain



Among 183 ITP cases, 172 (94%) were vaccinated with first dose of measles-containing vaccines. Of them, 55 (30%) and 152 (83%) were eligible for inclusion in the SCRI and case-crossover analyses, respectively. The risk of ITP following measles vaccination was 5.0 (95% CI: 2.5-9.7); exclusion of cases from Iran-01 resulted in an IRR estimate of 7.7 (95% CI: 3.5-17.3). Intervals between first dose of measles-containing vaccine and ITP onset for cases included in the strain-specific SCRI analyses are shown in Figure 2b. This analysis showed a significantly elevated ITP risk for measles vaccines containing Schwarz (IRR: 20.7; 95% CI: 2.7-157.6), Edmonston-Zagreb (IRR: 11.1; 95% CI: 1.4-90.3), and Enders´Edmonston (IRR: 8.5; 95% CI: 1.9-38.1) strains. Risk estimates for Shanghai-191 could not be assessed because of small numbers. Our estimates for the vaccine product/s used in Iran (AIK-C/ Edmoston-Zagreb/Schwarz) did not show an increased risk of ITP (IRR: 0.51; 95% CI: 0.10-2.54) (Table 4). The case-crossover analysis produced an overall unadjusted OR of 4.7 (95% CI: 2.1-10.7). When cases from Iran-01 were excluded, the OR estimate was 6.6 (95% CI: 2.6-16.9).

Figure 2b. Interval between first dose of measles-containing vaccines and immune thrombocytopenic purpura) ITP onset by measles vaccine strain



Discussion

The success of this proof-of-concept study in obtaining participation and data useful for analysis from sites located in all regions of the world using a common protocol has demonstrated the feasibility of international collaborative hospital-based studies, with high participation of LMICs, for the investigation of serious and rare AEFI. Moreover, the study has confirmed increased risks of AM following first dose of mumps-containing vaccines, and of ITP following first dose of measles-containing vaccines. It has also shown, potential risk differences between vaccine strains for both associations. The elevated risk estimates found for the Leningrad-Zagreb mumps strain are consistent with previous studies (27, 28). Regarding Jeryl-Lynn-derived strain vaccines, although the study did not have enough power to confirm the absence of risk for these strains, our finding of zero cases in the risk window was consistent with the hypothesis of no association (25, 29). The two Iranian sites reported that three vaccine products, containing the mumps strains Hoshino, Leningrad-Zagreb and UrabeAm9 were used during the study period, but they did not differentiate between them. Therefore, we could not assign the high risk of AM identified in Iran to one or other of these three strains (23, 24, 27, 28, 30-32). This would require further investigation in subsequent studies, particularly to determine the risk associated with the Hoshino strain, given the limited literature available on its safety profile (33-36). AM usually occurs within 2-5 weeks following mumps vaccination (9, 11, 31, 32, 37, 38); therefore, our study used a risk window of 8-35 days post-vaccination. Our study found a statistically significant risk when the washout period (days 1-7 and days 36-42 post-vaccination) was compared to the non-risk periods (days 43-84 post-vaccination) for the vaccine/s products used in Iran (IRR: 12.9; 95% CI: 2.8-59.7), which suggests the possibility of an increased risk also for the washout period, that deserves investigation in future studies.

The elevated risk of ITP following measles-containing vaccination is consistent with the literature (12-14). Our strain-specific unadjusted analysis showed a significantly elevated ITP risk for measles vaccines containing the Schwarz, Edmonston-Zagreb, and Enders´Edmonston strains. No risk of ITP was identified in Iran, which reported the concurrent distribution of three vaccine products including the AIK-C, Edmonston-Zagreb and Schwarz strains, without distinguishing between them. Among 172 vaccinees included in this study, at least 155 (90%) received MMR or measles-rubella vaccines. Given the known association between wild rubella infection and ITP (39), and the existence of a few studies showing mostly mild thrombocytopenia following rubella vaccination in some adults (19, 40-42), a potential contribution of the rubella component of the vaccine to our findings may not be excluded.

Case-only methods can be efficient epidemiological designs for use in vaccine safety, particularly for LMICs, given that population denominators or separate controls are not required; moreover, time-fixed confounders are inherently adjusted for (16). Self-controlled case series (SCCS) methods have been successfully implemented in similar international collaborations, such as the hospital-based international collaborative investigation of Guillain-Barre syndrome following the H1N1 2009-2010 pandemic influenza vaccination (7), and the investigation of the association between intussusception and rotavirus in Mexico and Brazil (43). In our study, some of the participating sites could not identify end of the follow-up period independently of the event being investigated, thus, modifying the duration of the observation period in ways that could potentially bias results (44). The SCRI approach simplifies the SCCS design by reducing the length of the control interval (21). The selection of shorter non-risk periods, as done in our study under the assumption that participants were not lost to follow-up during this 84-day period, not only may solve this limitation for LMICs, but may also decrease the effect of time-varying confounders on the risk estimates, because risk variations in such a short period may be negligible (21). Nonetheless, adjustments for age group and seasonality were performed, when possible. For comparison purposes, we used case-crossover as a secondary analysis, given that it does not require follow-up after case occurrence; to decrease the possibility of bias associated with variations in the distributions of exposures over time, only one control window of the same duration as the case window was selected (16). The method requires the same underlying probability of vaccination in all time intervals, which is unlikely to hold true for pediatric vaccines, which are usually administered according to pre-specified schedules (16). However, our case-crossover unadjusted risk estimates for ITP following measles-containing vaccines and for AM following mumps-containing vaccines were comparable to those obtained using the SCRI method, although the latter estimate was less stable due to limited study power.

Case-only methods demand careful determination of event onset and vaccination dates. Therefore, we were particularly thorough in training site investigators. Given that one site could not provide exact vaccination dates (only month/year of vaccination were recorded), we performed analyses both excluding and including this site (using imputed dates for the site). Although these analyses showed differences in point estimates, all results were significant and the confidence intervals overlapped. Since SCRI uses data only from vaccinees, the approach minimizes potential misclassification due to incomplete/absent data on vaccination status, another frequent shortcoming in LMICs. Nonetheless, a possible limitation in the approach used here is that site variability may be a potential source of selection bias as the sites may have differences in access to vaccination records and in patient's health-seeking behavior. Bias could also be associated with site differences in diagnosis capabilities and quality of medical records. Also, our use of self-controlled analytical methods did not permit estimations of absolute risk (20).

Our results show that collaborative studies for the investigation of different vaccine products by strain and potentially by manufacturer are feasible. The power to do so, and to investigate risk by country/region (Supplementary material; Tables S-5 and S-6) will increase when additional large hospitals with medical specialties for rare and difficult to diagnose events, high quality medical records and easy access to vaccination records are included (2). The inclusion of large referral hospitals with electronic discharge databases should decrease per case

investigation costs by reducing efforts associated with data extraction, study coordination, training, data quality assessment, and provide quality medical records and higher reliability in disease codification. The use of large hospitals would also reduce the likelihood of having participating hospitals that do not contribute cases to the analysis, as has occurred in some of our sites. Because easy and unequivocal linkages between hospital and vaccination records and proven access to vaccination information would increase data quality and efficiency, it is important to carefully select the participating sites, particularly in LMICs. Given the current interest on the development of vaccines for diseases such as dengue, malaria, and Zika, prioritization should be given to the addition of sites from tropical/sub-tropical areas in LMICs for future studies.

Conclusions

This collaboration has demonstrated, for the first time, that a multi-country hospital-based network with high participation of LMICs, using a common protocol and standardized procedures, permits the investigation of rare vaccine adverse events, can produce reliable results, and has the potential to characterize risk differences between vaccine strains. The completion of this network with the addition of large referral hospitals, including from tropical/subtropical countries, and the systematic implementation of this hospital-based approach, should permit the rapid and sustainable evaluation of safety signals for serious and rare AEFI for new and existing vaccines in all settings, and the comparison of safety profiles for vaccine products.

Funding

Center for Biologics Evaluation and Research (CBER)-U.S. Food and Drug Administration (FDA) funded this project. GRiP, Global Research in Pediatrics, European Union Seventh framework Programme (FP7/2007-2013) provided additional funding under grant agreement n° 261060.

Conflicts of interest

DW has received honoraria from GlaxoSmithKline Biologicals (GSK) for consultancies on malaria vaccine safety studies and implementation unrelated to the content of this manuscript. SB is a consultant for GSK. MS is heading a research group that conducts PASS studies for pharmaceutical companies including GSK.

All other authors confirm that there are no known conflicts of interest associated with this publication.

Acknowledgements

The authors thank Hector S. Izurieta for his valuable contributions to the conception of the project, Mees Mosseveld and Peter Rijnbeek for development of data collection and data sharing tools, Sergio Castillo-Pérez for mapping, and personnel from participating sentinel sites for facilitating the study implementation.

References

1. Gavi. The Vaccine Alliance 2017 [cited 2017 February 27]. Available from: http://www.gavi.org.

2. Izurieta HS, Zuber P, Bonhoeffer J, Chen RT, Sankohg O, Laserson KF, et al. Roadmap for the international collaborative epidemiologic monitoring of safety and effectiveness of new high priority vaccines. Vaccine. 2013;31(35):3623-7.

3. World Health Organization. Global vaccine safety blueprint. The landscape analysis. WHO/IVB/12.04: Quality, Safety, and Standards unit of the Department of Immunizations, Vaccines and Biologicals; 2012 [cited 2016 February, 7]. Available from: http://apps.who.int/iris/bitstream/10665/70854/1/WHO_IVB_12.04_eng.pdf.

4. Amarasinghe A, Black S, Bonhoeffer J, Carvalho SM, Dodoo A, Eskola J, et al. Effective vaccine safety systems in all countries: a challenge for more equitable access to immunization. Vaccine. 2013;31 Suppl 2:B108-14.

5. World Health Organization. Global vaccine safety blueprint. WHO/IVB/12.07: Quality, Safety, and Standards unit of the Department of Immunization, Vaccines and Biologicals; 2012 [cited 2016 February 7]. Available from: http://extranet.who.int/iris/restricted/bitstream/10665/70919/1/WHO_IVB_12.07_eng.pdf?ua=1.

6. Maure CG, Dodoo AN, Bonhoeffer J, Zuber PL. The Global Vaccine Safety Initiative: enhancing vaccine pharmacovigilance capacity at country level. Bull World Health Organ. 2014;92(9):695-6.

7. Dodd CN, Romio SA, Black S, Vellozzi C, Andrews N, Sturkenboom M, et al. International collaboration to assess the risk of Guillain Barre Syndrome following Influenza A (H1N1) 2009 monovalent vaccines. Vaccine. 2013;31(40):4448-58.

8. Measles vaccines: WHO position paper. Wkly Epidemiol Rec. 2009;84(35):349-60.

9. Tapiainen T, Prevots R, Izurieta HS, Abramson J, Bilynsky R, Bonhoeffer J, et al. Aseptic meningitis: case definition and guidelines for collection, analysis and presentation of immunization safety data. Vaccine. 2007;25(31):5793-802.

10. Mumps virus vaccines. Wkly Epidemiol Rec. 2007;82(7):51-60.

11. Bonnet MC, Dutta A, Weinberger C, Plotkin SA. Mumps vaccine virus strains and aseptic meningitis. Vaccine. 2006;24(49-50):7037-45.

12. O'Leary ST, Glanz JM, McClure DL, Akhtar A, Daley MF, Nakasato C, et al. The risk of immune thrombocytopenic purpura after vaccination in children and adolescents. Pediatrics. 2012;129(2):248-55.

13. Mantadakis E, Farmaki E, Buchanan GR. Thrombocytopenic purpura after measles-mumps-rubella vaccination: a systematic review of the literature and guidance for management. J Pediatr. 2010;156(4):623-8.

14. France EK, Glanz J, Xu S, Hambidge S, Yamasaki K, Black SB, et al. Risk of immune thrombocytopenic purpura after measles-mumps-rubella immunization in children. Pediatrics. 2008;121(3):e687-92.

15. Maclure M, Fireman B, Nelson JC, Hua W, Shoaibi A, Paredes A, et al. When should case-only designs be used for safety monitoring of medical products? Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:50-61.

16. Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods. Vaccine. 2004;22(15-16):2064-70.

17. Sejvar JJ, Kohl KS, Bilynsky R, Blumberg D, Cvetkovich T, Galama J, et al. Encephalitis, myelitis, and acute disseminated encephalomyelitis (ADEM): case definitions and guidelines for collection, analysis, and presentation of immunization safety data. Vaccine. 2007;25(31):5771-92.

18. Lussiana C, Loa Clemente SV, Pulido Tarquino IA, Paulo I. Predictors of bacterial meningitis in resource-limited contexts: an Angolan case. PLoS One. 2011;6(10):e25706.

19. Wise RP, Bonhoeffer J, Beeler J, Donato H, Downie P, Matthews D, et al. Thrombocytopenia: case definition and guidelines for collection, analysis, and presentation of immunization safety data. Vaccine. 2007;25(31):5717-24.

20. Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Stat Med. 2006;25(10):1768-97.

21. Li R, Stewart B, Weintraub E. Evaluating efficiency and statistical power of self-controlled case series and self-controlled risk interval designs in vaccine safety. J Biopharm Stat. 2016;26(4):686-93.

22. Maclure M, Mittleman MA. Should we use a case-crossover design? Annu Rev Public Health. 2000;21:193-221.

23. Ki M, Park T, Yi SG, Oh JK, Choi B. Risk analysis of aseptic meningitis after measles-mumps-rubella vaccination in Korean children by using a case-crossover design. Am J Epidemiol. 2003;157(2):158-65.

24. Fujinaga T, Motegi Y, Tamura H, Kuroume T. A prefecture-wide survey of mumps meningitis associated with measles, mumps and rubella vaccine. Pediatr Infect Dis J. 1991;10(3):204-9.

25. Black S, Shinefield H, Ray P, Lewis E, Chen R, Glasser J, et al. Risk of hospitalization because of aseptic meningitis after measles-mumps-rubella vaccination in one- to two-year-old children: an analysis of the Vaccine Safety Datalink (VSD) Project. Pediatr Infect Dis J. 1997;16(5):500-3.

26. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191-4.

27. da Cunha SS, Rodrigues LC, Barreto ML, Dourado I. Outbreak of aseptic meningitis and mumps after mass vaccination with MMR vaccine using the Leningrad-Zagreb mumps strain. Vaccine. 2002;20(7-8):1106-12.

28. da Silveira CM, Kmetzsch CI, Mohrdieck R, Sperb AF, Prevots DR. The risk of aseptic meningitis associated with the Leningrad-Zagreb mumps vaccine strain following mass vaccination with measles-mumps-rubella vaccine, Rio Grande do Sul, Brazil, 1997. Int J Epidemiol. 2002;31(5):978-82.

29. Makela A, Nuorti JP, Peltola H. Neurologic disorders after measles-mumps-rubella vaccination. Pediatrics. 2002;110(5):957-63.

30. Dourado I, Cunha S, Teixeira MG, Farrington CP, Melo A, Lucena R, et al. Outbreak of aseptic meningitis associated with mass vaccination with a urabe-containing measles-mumps-rubella vaccine: implications for immunization programs. Am J Epidemiol. 2000;151(5):524-30.

31. Sugiura A, Yamada A. Aseptic meningitis as a complication of mumps vaccination. Pediatr Infect Dis J. 1991;10(3):209-13.

32. Mumps meningitis and MMR vaccination. Lancet. 1989;2(8670):1015-6.

33. Dorreh FHM. Adverse events associated with MMR vaccines in Arak. Journal of Iranian Clinical Research. 2015;1:6-10.

34. Esteghamati A, Keshtkar A, Heshmat R, Gouya MM, Salar Amoli M, Armin S, et al. Adverse reactions following immunization with MMR vaccine in children at selected provinces of Iran. Arch Iran Med. 2011;14(2):91-5.

35. Ueda K, Miyazaki C, Hidaka Y, Okada K, Kusuhara K, Kadoya R. Aseptic meningitis caused by measles-mumps-rubella vaccine in Japan. Lancet. 1995;346(8976):701-2.

36. Sood A, Mitra M, Joshi HA, Nayak US, Siddaiah P, Babu TR, et al. Immunogenicity and safety of a novel MMR vaccine (live, freeze-dried) containing the Edmonston-Zagreb measles strain, the Hoshino mumps strain, and the RA 27/3 rubella strain: Results of a randomized, comparative, active controlled phase III clinical trial. Hum Vaccin Immunother. 2017:1-8.

37. Kimura M, Kuno-Sakai H, Yamazaki S, Yamada A, Hishiyama M, Kamiya H, et al. Adverse events associated with MMR vaccines in Japan. Acta Paediatr Jpn. 1996;38(3):205-11.

38. Miller E, Andrews N, Stowe J, Grant A, Waight P, Taylor B. Risks of convulsion and aseptic meningitis following measles-mumps-rubella vaccination in the United Kingdom. Am J Epidemiol. 2007;165(6):704-9.

39. Toltzis P. 50 Years Ago in The Journal of Pediatrics: Some Recently Recognized Manifestations of the Rubella Syndrome. J Pediatr. 2015;167(2):441.

40. Forrest JM, Honeyman MC, Lovric VA. Rubella vaccination and thrombocytopenia. Aust N Z J Med. 1974;4(4):352-5.

41. Freestone DS, Prydie J, Smith SG, Laurence G. Vaccination of adults with Wistar RA 27/3 rubella vaccine. J Hyg (Lond). 1971;69(3):471-7.

42. Bartos HR. Thrombocytopenia associated with rubella vaccination. N Y State J Med. 1972;72(4):499.

43. Patel MM, Lopez-Collada VR, Bulhoes MM, De Oliveira LH, Bautista Marquez A, Flannery B, et al. Intussusception risk and health benefits of rotavirus vaccination in Mexico and Brazil. N Engl J Med. 2011;364(24):2283-92.

44. 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):1805-17.

Supplementary material

Table S-1. Identification of probable aseptic meningitis and immune thrombocytopenic purpura (ITP) cases through electronic databases

	ICD-9 codes in first discharge diagnosis position	ICD-10 codes in first discharge diagnosis position
Aseptic meningitis probable cases	047 (047.0-047.9) Meningitis due to enterovirus 049.0-049.1 Other non-arthropod-borne viral meningitis 072.1 Mumps meningitis 321.2 Meningitis due to viruses not elsewhere classified 322.0, 322.1, 322.9 Meningitis of unspecified cause	A87.0 Meningitis due to enterovirus A87.1 Adenoviral meningitis A87.2 Lymphocytic choriomeningitis A87.8 Other viral meningitis A87.9 Viral meningitis, unspecified B26.1 Mumps meningitis G02.0 Meningitis due to viruses not elsewhere classified G03.0, G03.8, G03.9 Meningitis of unspecified cause
ITP probable cases	287.30-287.39 Primary thrombocytopenia 287.41-287.49 Secondary thrombocytopenia 287.5 Thrombocytopenia, unspecified	D69.3, D69.4 (D69.41-D69.43) Primary thrombocytopenia D69.5 (D69.51, D69.59) Secondary thrombocytopenia D69.6 Thrombocytopenia, unspecified

Abbreviations: ITP (Immune Thrombocytopenic Purpura)

Table S-2. Aseptic meningitis case definition

LEVEL 1 OF DIAGNOSTIC CERTAINTY (9) *An aseptic meningitis case without exclusion criteria for which the medical record review found:* *Clinical evidence of acute meningitis such as fever, headache, vomiting, bulging fontanelle, nuchal rigidity or other signs of meningeal irritation* *AND* *Pleocytosis in cerebrospinal fluid (CSF) determinated as >5 leukocytes/mm^3 (µL)* *AND* *Absence of any microorganism on Gram stain of CSF* *AND* *Negative bacterial culture of CSF in the absence of antibiotic treatment before obtaining the first CSF sample*
LEVEL 2 OF DIAGNOSTIC CERTAINTY (9) *An aseptic meningitis case without exclusion criteria for which the medical record review found:* *Clinical evidence of acute meningitis such as fever, headache, vomiting, bulging fontanelle, nuchal rigidity or other signs of meningeal irritation* *AND* *Pleocytosis in cerebrospinal fluid (CSF) determinated as >5 leukocytes/mm3 (µL)* *AND* *Absence of any microorganism on Gram stain of CSF* *AND* *No bacterial culture of CSF obtained OR negative culture in the presence of antibiotic treatment before obtaining the first CSF sample*
LEVEL 3 OF DIAGNOSTIC CERTAINTY (this criteria is in addition to the existing Brighton Collaboration criteria (9)) *An aseptic meningitis case without exclusion criteria for which the medical record review found:* *A physician diagnosis of aseptic meningitis with no evidence to the contrary identified in the medical records*
INSUFFICIENT EVIDENCE (9) *If the evidence available for an event is insufficient because information is missing*
NO CASE (9) *An event does not meet the case definition if the investigation reveals a negative finding of a necessary criterion for classification in Levels 1-2-3 or if a different event is the final diagnosis*

Table S-3. Immune thrombocytopenic purpura (ITP) case definition

LEVEL 1 OF DIAGNOSTIC CERTAINTY (Adapted from O´Leary et al. (12)) *A physician diagnosis of ITP with no indication that this diagnosis is differential or a rule out diagnosis and with no indication that it was drug-induced* *AND* *A laboratory result with a platelet count less than 50,000/µL,* *AND* *Evidence of spontaneous bleeding, including purpura, hemorrhagic oozing of skin lesions including rashes, hematoma, bruising, hematemesis, hemotochezia, occult bleeding per rectum, epistaxis, hemoptysis, hematuria, vaginal bleeding, conjunctival bleeding, intracranial bleeding*
LEVEL 2 OF DIAGNOSTIC CERTAINTY *A physician diagnosis of ITP with no indication that this diagnosis is differential or a rule out diagnosis and with no indication that it was drug-induced* *AND* *A laboratory result with a platelet count less than 50,000/µL*
LEVEL 3 OF DIAGNOSTIC CERTAINTY *A physician diagnosis of thrombocytopenia with no indication that this diagnosis is differential or a rule out diagnosis and with no indication that it was drug-induced* *AND* *A laboratory result with a platelet count less than 50,000/µL*
INSUFFICIENT EVIDENCE *If the evidence available for an event is insufficient because information is missing*
NO CASE *An event does not meet the case definition if the investigation reveals a negative finding of a necessary criterion for classification in Levels 1-2-3 or if a different event is the final diagnosis*

Table S-4. Encephalitis case definition: exclusion criteria

LEVEL 1 OF DIAGNOSTIC CERTAINTY (17) *Demonstration of acute inflammation of central nervous system parenchyma (±meninges) by histopathology*
LEVEL 2 OF DIAGNOSTIC CERTAINTY (Adapted from Brighton Collaboration (17)) *Encephalopathy (e.g. depressed or altered level of consciousness, lethargy, or personality change lasting>24h)*

AND one or more of the following:

- *Decreased or absent response to environment, as defined by response to loud noise or painful stimuli*
- *Decreased or absent eye contact*
- *Inconsistent or absent response to external stimuli*
- *Decreased arousability*
- *Seizure associated with loss of consciousness (as described in the medical records)*

OR focal or multifocal findings referable to the central nervous system, including one or more of the following:

- *Focal cortical signs (including but not limited to: aphasia, cortical blindness)*
- *Cranial nerve abnormality/abnormalities*
- *Visual field defect/defects*
- *Presence of primitive reflexes (Babinski´s sign, glabellar reflex, snout/sucking reflex)*
- *Motor weakness (either diffuse or focal; more often focal)*
- *Sensor abnormalities (either positive or negative; sensory level)*
- *Altered deep tendon reflexes (hypo-o hyperreflexia, reflex asymmetry)*
- *Cerebellar dysfunction, including ataxia, dysmetria, cerebellar nystagmus*

AND two or more of the following indicators of inflammation of the central nervous system (CNS):

- *Fever (≥38ºC)*
- *CSF pleocytosis (>5 leukocytes/mm3 (µL))*
- *Electroencephalography (EEG) findings consistent with encephalitis*
- *Neuroimaging consistent with encephalitis*

LEVEL 3 OF DIAGNOSTIC CERTAINTY (Adapted from Brighton Collaboration (17))

Encephalopathy (e.g. depressed or altered level of consciousness, lethargy, or personality change lasting>24h)

AND one or more of the following:

- *Decreased or absent response to environment, as defined by response to loud noise or painful stimuli*
- *Decreased or absent eye contact*
- *Inconsistent or absent response to external stimuli*
- *Decreased arousability*

- *Seizure associated with loss of consciousness (as described in the medical records)*

OR focal or multifocal findings referable to the CNS, including one or more of the following:

- *Focal cortical signs (including but not limited to: aphasia, cortical blindness)*
- *Cranial nerve abnormality/abnormalities*
- *Visual field defect/defects*
- *Presence of primitive reflexes (Babinski´s sign, glabellar reflex, snout/sucking reflex)*
- *Motor weakness (either diffuse or focal; more often focal)*
- *Sensor abnormalities (either positive or negative; sensory level)*
- *Altered deep tendon reflexes (hypo-o hyperreflexia, reflex asymmetry)*
- *Cerebellar dysfunction, including ataxia, dysmetria, cerebellar nystagmus*

AND one of the following indicators of inflammation of CNS:

- *Fever (≥38ºC)*
- *CSF pleocytosis (>5 leukocytes/mm3 (µL))*
- *Electroencephalography (EEG) findings consistent with encephalitis*
- *Neuroimaging consistent with encephalitis*

LEVEL 3A OF DIAGNOSTIC CERTAINTY (Adapted from Brighton Collaboration (17))

Insufficient information is available to distinguish case between acute encephalitis or acute disseminated encephalomyelitis

LEVEL 4 OF DIAGNOSTIC CERTAINTY (this criteria is in addition to the existing Brighton Collaboration criteria (17))

The medical record review found a physician diagnosis of encephalitis or encephalopathy with no indication in the medical records that this diagnosis is a differential or a rule out diagnosis will be excluded.

Table S-5. Aseptic meningitis cases following mumps-containing vaccination by country eligible for the SCRI analyses

Country	Eligible confirmed aseptic meningitis cases (n)			Follow-up (days)
	Event in risk period (days 8-35 days)	Event in washout period (days 1-7 and 36-42)	Event in non-risk period (days 43-84)	Median (P25-P75)
Albania	0	1	0	85 (85, 85)
Argentina	0	0	1	57 (57, 57)
Australia	0	0	1	85 (85, 85)
Chile	5	0	0	85 (85, 85)
China	0	0	0	--
Colombia	0	0	0	--
Costa Rica	2	0	0	85 (85, 85)
Honduras	0	0	0	--
India	0	0	0	--
Iran	27	9	2	85 (85, 85)
Peru	0	0	0	--
Singapore	0	0	0	--
South Africa	0	0	0	--
Spain	0	2	0	85 (85, 85)
Uganda	0	0	0	--
Uruguay	1	0	0	85 (85, 85)

Table S-6. Immune thrombocytopenic purpura (ITP) cases following measles-containing vaccination by country eligible for the SCRI analyses

Country	Eligible confirmed ITP cases (n)			Follow-up (days)
	Event in risk period (days 8-35 days)	Event in washout period (days 1-7 and 36-42)	Event in non-risk period (days 43-84)	Median (P25-P75)

Albania	4	0	0	85 (85, 85)
Argentina	1	1	1	85 (85, 85)
Australia	5	1	0	85 (85, 85)
Chile	2	1	0	85 (85, 85)
China	0	0	1	85 (85, 85)
Colombia	0	0	0	--
Costa Rica	4	0	0	85 (76, 85)
Honduras	0	1	0	85 (85, 85)
India	0	0	0	--
Iran	2	2	5	85 (85, 85)
Peru	1	0	1	85 (85, 85)
Singapore	5	1	1	85 (85, 85)
South Africa	0	0	0	--
Spain	12	0	3	85 (85, 85)
Uganda	0	0	0	--
Uruguay	0	0	0	--

CHAPTER 6 METHODS FOR DEALING WITH HETEROGENEITY AND BIAS

6.1 PANDEMIC INFLUENZA VACCINE & NARCOLEPSY: SIMULATIONS ON THE POTENTIAL IMPACT OF BIAS

Leonoor Wijnans*

Caitlin Dodd*

Maria de Ridder

Silvana Romio

Daniel Weibel

Sebastiaan Overeem

Gert Jan Lammers

Jan Bonhoeffer

Steve Black

Miriam C J M Sturkenboom

*Authors have contributed equally to the manuscript

Expert review of vaccines 15.5 (2016): 573-584.

Abstract

Several studies have identified an association between Pandemrix^{TM}, an AS03 adjuvanted pandemic influenza A(H1N1) vaccine, and narcolepsy, a rare and under-diagnosed sleep disorder with a median onset-to-diagnosis interval of ten years. This paper reviews potential sources of bias in published studies and aims to provide, through simulation, methodological recommendations for assessment of vaccine safety signals.

Our simulation study showed that in the absence of an association between the vaccine and the outcome, presence of detection bias and differential exposure misclassification could account for elevated risk estimates. These may play a major role, particularly in alert situations when observation times are limited and the disease has a long latency period. Estimates from the case-control design were less inflated than those from the cohort design when these biases were present. Overall, these simulations provide useful insights for the design and interpretation of future studies.

Background

In August 2010, case reports linking the occurrence of narcolepsy in children aged 5 to 19 years to an AS03 adjuvanted H1N1pdm09 (pH1N1) vaccine, Pandemrix™ (GlaxoSmithKline, Middlesex, United Kingdom) were published in Finland and Sweden [1,2]. In the European Union, Pandemrix was widely used, with over 30 million doses administered. Coverage was particularly high in the Nordic countries [3]. Following reports from Sweden and Finland, the European Medicines Agency initiated a review procedure [4] which eventually led to the restriction of indication for Pandemrix [5].

Narcolepsy is a chronic sleep disorder that is severely debilitating. The dysregulation of the sleep-wake cycle is caused by the destruction of hypocretin forming neurons in the hypothalamus, which is thought to result from an auto-immune process [6].Symptoms include excessive daytime sleepiness (EDS) and cataplexy [7]. Symptoms usually emerge gradually and can initially be non-specific. Consequently, symptoms can be attributed to other diagnoses resulting in a delay of narcolepsy diagnosis and treatment [8-11]. Despite significant improvements in the speed and accuracy of narcolepsy diagnosis [10-12], a recent study found that the median delay between onset and diagnosis remains approximately 10 years [9].

As of May 2015, eight epidemiological studies testing the association between Pandemrix and clusters of narcolepsy cases [13-21] have been published reporting risk estimates ranging from 1.6 to 14.4. An overview of the main characteristics of these studies is presented in table 1. Generally, published studies were meticulous in their methods and applied sensitivity analyses to evaluate the presence of biases. Nonetheless, studies were inevitably observational and, as studies were mostly initiated rapidly after the signal emerged, they had limited time for case capture. Combined with the often nonspecific symptoms and onset of narcolepsy resulting in delayed diagnosis these studies are particularly prone to bias. Five years after the original signal emerged it remains unclear if and how potential sources of bias affected the estimates from the association studies. Consequently it is still unknown what the exact association between Pandemrix and narcolepsy is [22,23].

Table 1. Main characteristics of studies testing the association between narcolepsy and Pandemrix

Ref	Year	Country, region	Age Range	Case definition	Index Date	Exclusion criteria	Primary study period	Dealt with confounding	Dealt with detection bias	Relative Rate reported
Retrospective cohort studies										
[18]	2012	Finland	4-19 yrs	BC definition level 1-3	First contact for EDS	-	01/01/2009 - 15/08/2010	No	Limit observation period to start media attention. Sensitivity analysis on start professional attention.	RR: 12·7 (95% CI: 6·1 - 30·8)
[21]	2012 /2014	Ireland		BC definition level 1-3	First contact for EDS	Those with symptom onset before April 2009	01/04/ 2009 - 31/12/2010	No	Sensitivity analysis considering period before the increased media attention in Sweden and Finland	RR: 13·0 (95% CI: 4·8 - 34·7)
[14]	2011	Sweden; Stockholm, Skane, Vastra Gotaland, Ostergotland		Presence diagnosis code (non-validated)	Diagnosis	Those with onset before October 2009	01/10/2009 - 31/12/2010	No	No	HR: 4·19 (95% CI: 1·76 - 12·1)
[17]	2013	Sweden; Stockholm, Skane, Vastra Gotaland, Ostergotland)		Presence diagnosis code (non-validated)	Diagnosis	Prevalent cases (looking back 5 years)	01/10/2009 - 31/12/2010	Adjusted for age (5-year bands), gender, county, education, income, secondary health care use, pregnancy status & presence of diagnoses (ICD-10), ethnicity.	No	HR: 2·92 (95% CI: 1·78 - 4·79)
Case Control Studies										

[20]	2013	France		BC definition level 1-3	Diagnosis	Non-consenting cases, cases with onset before January 1, 2005, Vaccinated cases with onset before vaccination	01/10/2009 - 30/04/2011	Controls were matched on age, gender and geographic location	Sensitivity analyses were performed that considered as index date (i) the date of referral for polysomno-graphy - MSLT; and (ii) the date of first symptom onset.	OR: 4·55 (95% CI: 2·34 - 8·88)
15	2012	Finland, Sweden, Norway, Italy, Denmark, UK, Netherlands, France	<19 yrs	BC definition level 1-4	MSLT referral	-	01/04/ 2009 - 30/06/2010	Controls were matched on year of birth, sex and index date (i.e. the date of onset of narcolepsy) and in Norway, Italy and UK also by region/practice.	Sensitivity analyses with different time periods (before media & professional attention); analysis of changes in lag times	Signal-generating countries OR (children and adolescents): 14·2 (95%CI: 2·5–infinity) Non-signal-generating countries OR (children and adolescents): 1·6 (95%CI: 0·5–6·1)
19	2013	England	4 - 18 yrs	International classification of sleep disorders criteria (definite & probable)	EDS Onset	Cases with onset before 2008 and/or diagnosis after July 2011	01/10/2009 - 31/12/2010	Adjusted for age, time period & clinical conditions that were indications for pandemic vaccination.	Choice of index date, sensitivity analyses included diagnosed case only and varying of index date (first healthcare contact or index date).	OR: 14·4 (95% CI :4·3 - 48·5)
13	2011	Sweden	0- <19 yrs	American Academy of Sleep Medicine criteria for narcolepsy with	EDS Onset	Cases without cataplexy	01/01/2009 - 31/12/2010	No	Choice of index date, analysis by different time windows after vaccination (3	RR: 6·6 (95% CI: 23·1 - 14·5)

				cataplexy					months).	
16	2014	Quebec province, Canada	>6 mnths (on 01/07/2009)	BC definition level 1-3	Date of onset, determined through questionnaire and interview	Those with symptom onset before January 2009	01/01/2009 - 31/12/2010	Adjusted for age, gender, seasonality and circulation of H1N1 virus (Cohort). Matched by age and gender (Case Control).	Sensitivity analyses including different observation periods to account for circulation virus and vaccination	RR (Cohort): 4·58 (95% CI: 1·59-11·77) OR (CC): 1·48 (95% CI: 0·37 - 7·03) RR (SCCS): 2·07 (95% CI: 0·70 - 6·17)

It is not unthinkable that a similar scenario could unfold in the future, i.e. a safety signal involving a difficult to diagnose condition with a delayed onset is linked to exposure with a new vaccine. Indeed, a similar situation has occurred in the past, when clusters of cases of Guillain-Barré syndrome were detected after the introduction of a new swine flu vaccine [24]. Using the example of narcolepsy and Pandemrix, we explore, in the absence of a formal hypothesis, the potential impact of two sources of bias that are likely to occur in similar scenarios.

Detection bias. The first source of bias is a type of selection bias. Awareness of a potential association between narcolepsy and vaccination amongst physicians and the general public could result in earlier diagnosis for vaccinated cases compared to unvaccinated cases, making vaccinated cases more likely to be included in observational studies with limited observation time. [15]. We refer to this as 'detection bias'.

Differential exposure misclassification. A second source of bias we consider is a form of recall bias, in which the onset of symptoms is misattributed, resulting in misclassification of onset dates to the period following vaccination. As narcolepsy symptoms often develop gradually and onset of symptoms is not always clearly identifiable, studies into narcolepsy are particularly prone to recall bias. We hypothesize that recalling onset of EDS with knowledge of a putative association between vaccination and narcolepsy could lead a patient to recall that symptoms started after vaccination [25]. We refer to this as 'differential exposure misclassification'.

Methods

We considered the impact of detection bias and differential exposure misclassification as defined above on the association measure between Pandemrix and narcolepsy.

Simulation

We simulated a population of 100,000 subjects < 19 years of age on April 1st 2009 to mimic the signal-generating population. We subsequently simulated dates of birth and death (based upon average lifespans in western Europe) to create a simulated lifetime for each subject. EDS onset dates were assigned over the lifespan of subjects based upon the reported age and gender specific incidence rates of narcolepsy with cataplexy onset [26]. Given these EDS onset dates, initial narcolepsy diagnosis dates were assigned using a random value drawn from a distribution of narcolepsy onset-to-diagnosis intervals which was assumed to have a gamma distribution chosen to mimic the distribution of onset-to-diagnosis intervals reported in the literature: a median of 10 years with a range of 0 to 40 years [11]. Additionally, since the underlying onset-to-diagnosis interval in children is potentially shorter [10], alternate gamma distributions with medians of 3 (range 0-13) and 7 (range 0-27) years were also used. All onset-to-diagnosis intervals were simulated to be at least 40 days long.

Overall vaccination coverage in this population was simulated at 25, 50 and 75%.. Vaccination dates were assigned independent of the age of a subject using a beta distribution of administration times mimicking real-life Pandemrix administration dates between October 12, 2009 and February 12, 2010 [27].

A null association (RR=1) was assumed for the actual relation between vaccine exposure and outcome.

Detection bias. Reduction in the onset-to-diagnosis interval was applied only to vaccinated cases for whom initial diagnosis would occur after the date of media attention (simulated to be August 15, 2010). If EDS onset occurred before August 15, 2010, the reduction was applied only to the interval from August 15, 2010 to the initial date of diagnosis. The date of narcolepsy diagnosis was reset in this way using values drawn from logit-normal distributions with medians of 30, 60, and 90% (See Figure in appendix) to produce reductions of the interval, with the restriction that the interval still was at least 40 days. Data with no reduction (i.e. 0% reduction) in the interval were also simulated (Figure 1).

Differential exposure misclassification. Misattribution of EDS onset dates to the period following vaccination was applied with probability equal to values drawn from logit-normal distributions with medians of 30 and 60% (Figure 1) to subjects who were diagnosed with narcolepsy after vaccination and after the start of media attention. In this case the onset date was reset to a random date between the vaccination date and the minimum of diagnosis date and vaccination date plus 180 days, based upon the six month risk period used by Miller *et al.* in their self-controlled case series analysis [19]. Data with no misattribution of onset dates were also simulated (Figure 2).

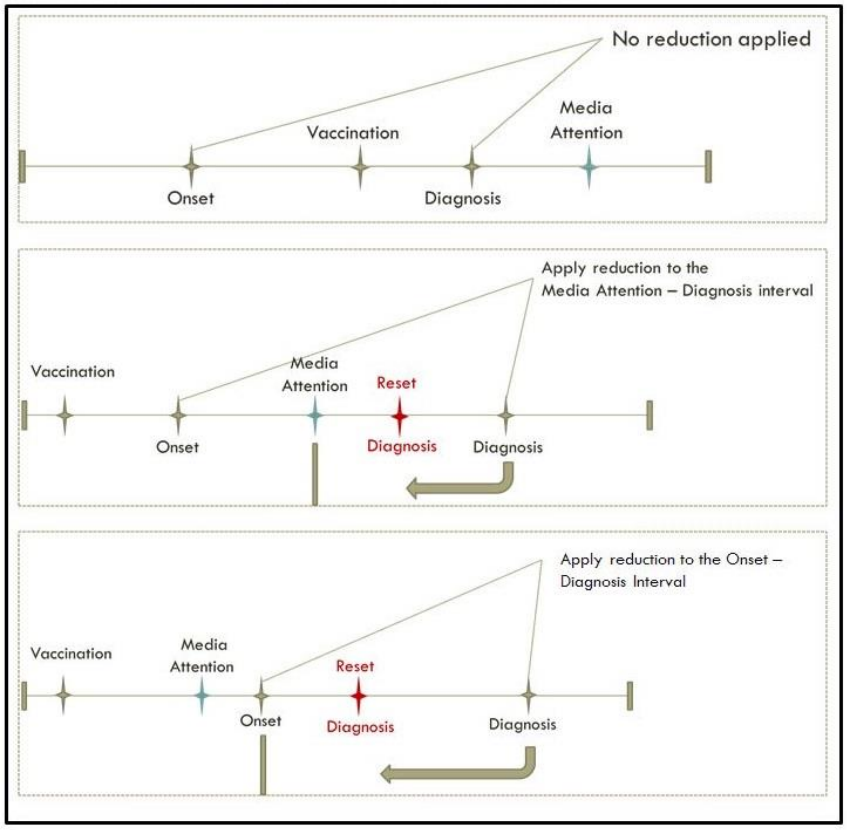


Figure 1. Application of detection bias

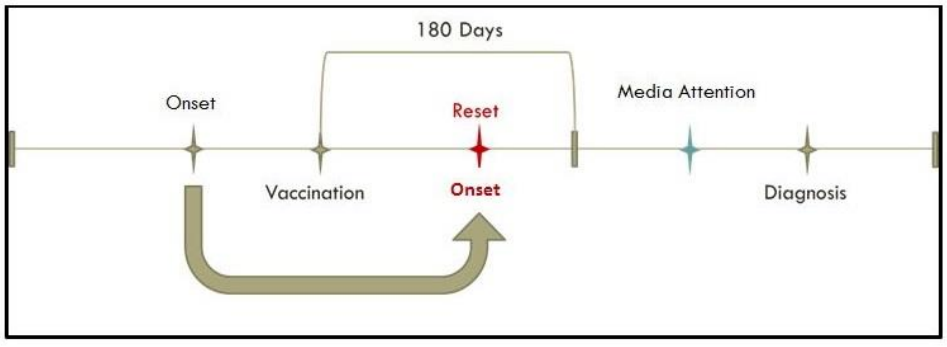


Figure 2. Application of differential exposure misclassification

We simulated 9 combinations of the underlying population settings: gamma scale (baseline onset to diagnosis interval, 3 different values) and vaccination coverage (3 values), to which we applied 12 combinations of the simulated sources of bias: detection bias (4 values), and differential exposure misclassification parameters (3

values) for a total of 108 combinations of simulation parameters. Variation in the underlying population settings (baseline onset to diagnosis interval and vaccination coverage) was conducted in the absence of a hypothesis regarding the impact of these changes on effect estimates.

Analysis

The association between vaccination and narcolepsy in children aged 4 - <19 years during the study period was analyzed using dynamic cohort and case-control designs. In the primary analyses a case capture (study period) of April 1, 2009 to December 1, 2010 was used in line with several published studies. We calculated absolute incidence rates in 6-month periods and calculated case counts during exposed and unexposed person time to investigate how incidence would change over time in the presence of detection bias. We additionally calculated the number of onset dates in exposed and unexposed person time at each level of each of the two bias parameters. In the comparative cohort analysis, the incidence rate of narcolepsy was compared between dynamic cohorts of vaccinated and non-vaccinated persons. All person time after the date of vaccination was considered exposed, whereas the entire case-capture period of non-vaccinated persons as well as the pre-vaccination time in vaccinated subjects contributed to non-exposed person time. Rate ratios were calculated based on Poisson regression. In the case-control analysis, cases were matched to 10 controls on sex, age in years and onset date. Odds ratios were calculated using conditional logistic regression.

We conducted several analyses to investigate the effects of different design choices and ways to mitigate bias. All sensitivity analyses were conducted using vaccination coverage of 50% and the baseline onset-to-diagnosis interval distribution described in literature with median 10 years, range 0-40 years. To study the effect of the length of case capture period, analyses with observation periods as long as 50 years were conducted. To study the effect of exclusion of cases possibly affected by awareness of a putative association, in one of the settings we excluded the cases with onset dates and diagnosis dates after August 15, 2010. Each of these sensitivity analyses was conducted in the absence of a hypothesis.

For each set of simulation parameters, 500 replications were analyzed, each producing an estimate and 95% confidence interval. Reported results are the exponentiated median of these 500 estimates calculated on the log scale and medians of the lower and upper confidence limits. All analyses were conducted using SAS 9·2.

Results

Application of onset-to-diagnosis interval reduction (detection bias) and differential exposure misclassification over three coverage rates and three baseline onset-to-diagnosis intervals increased the number of narcolepsy onset dates observed in the study period. Figure 3 shows, for exposed and unexposed children, the number of onset dates associated with narcolepsy diagnosed cases in scenarios with different percentages of differential exposure misclassification (columns), vaccination coverage (rows) and levels of detection bias (X-axis in each plot), using a baseline onset-to-diagnosis interval with a median of 10 (range 0-40) years. The number of observed narcolepsy onset dates increases at approximately the same rate in exposed and unexposed person time with an increasing detection bias in the absence of differential exposure misclassification (within column 1, Figure 3) except when vaccine coverage is 25% in which case no onset dates are observed in exposed person time. With the introduction of differential exposure misclassification in exposed subjects, new narcolepsy diagnoses occur more often in post-vaccination person time. The number of onset dates within unexposed person time also increases with increased reduction in the onset to diagnosis interval because, in these cases, the bias is being applied to vaccinated cases who experienced onset prior to vaccination. Figure 4 shows the effects of reduction of EDS onset-to-diagnosis date on the shape of incidence rates over calendar time in this cohort of 0-19 year olds in 2009. With a reduction of 60 or 90% in lag time, a clear peak in incidence of narcolepsy diagnoses occurs after media attention. These rates then return to the baseline rate or fall below the baseline rate due to depletion of cases through early diagnosis. The primary study period of April 1, 2009 to

December 1, 2010 is indeed a period of marked increase in newly diagnosed cases with reduction in time from onset to diagnosis.

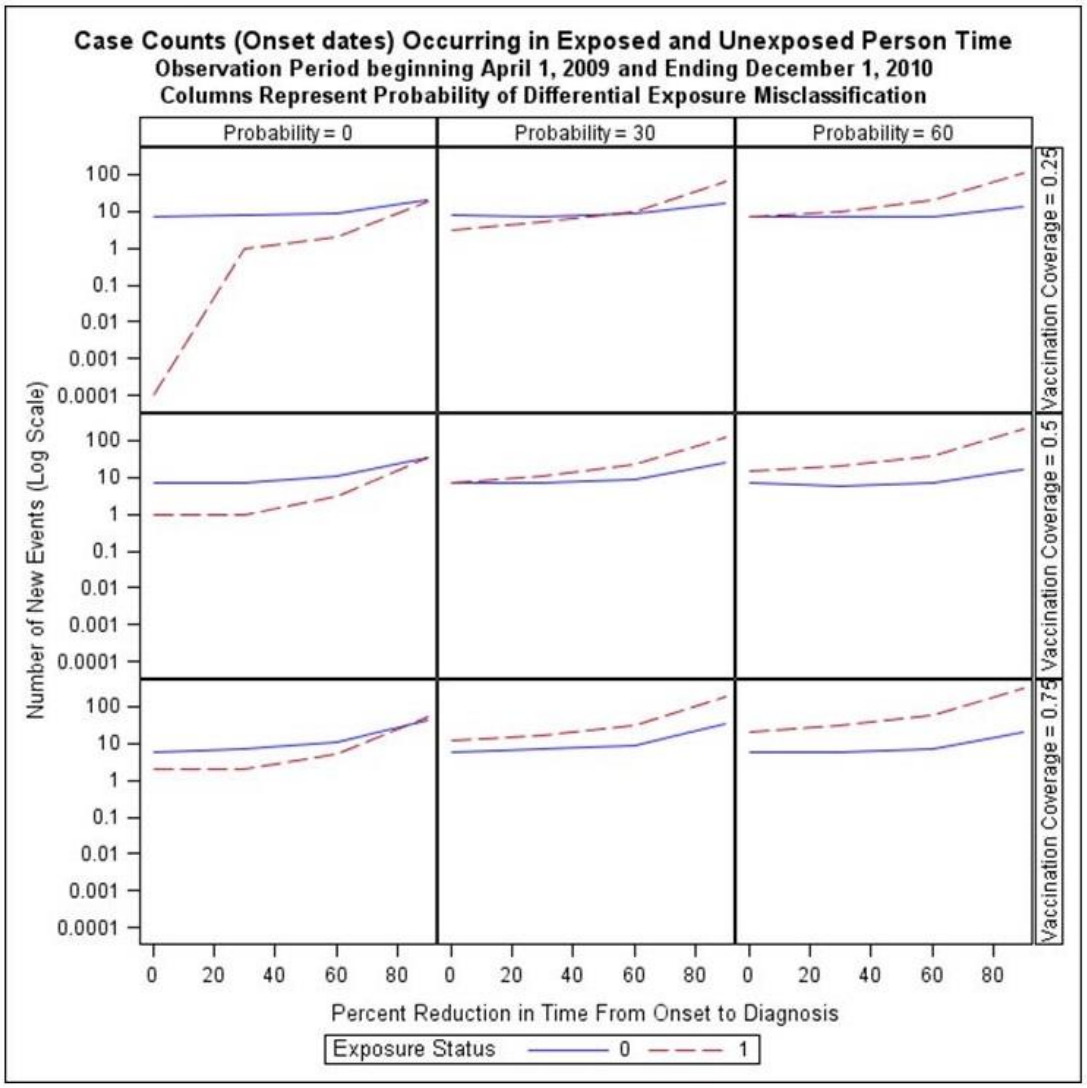


Figure 3. Case Counts (Onset dates) occurring in exposed and unexposed person time

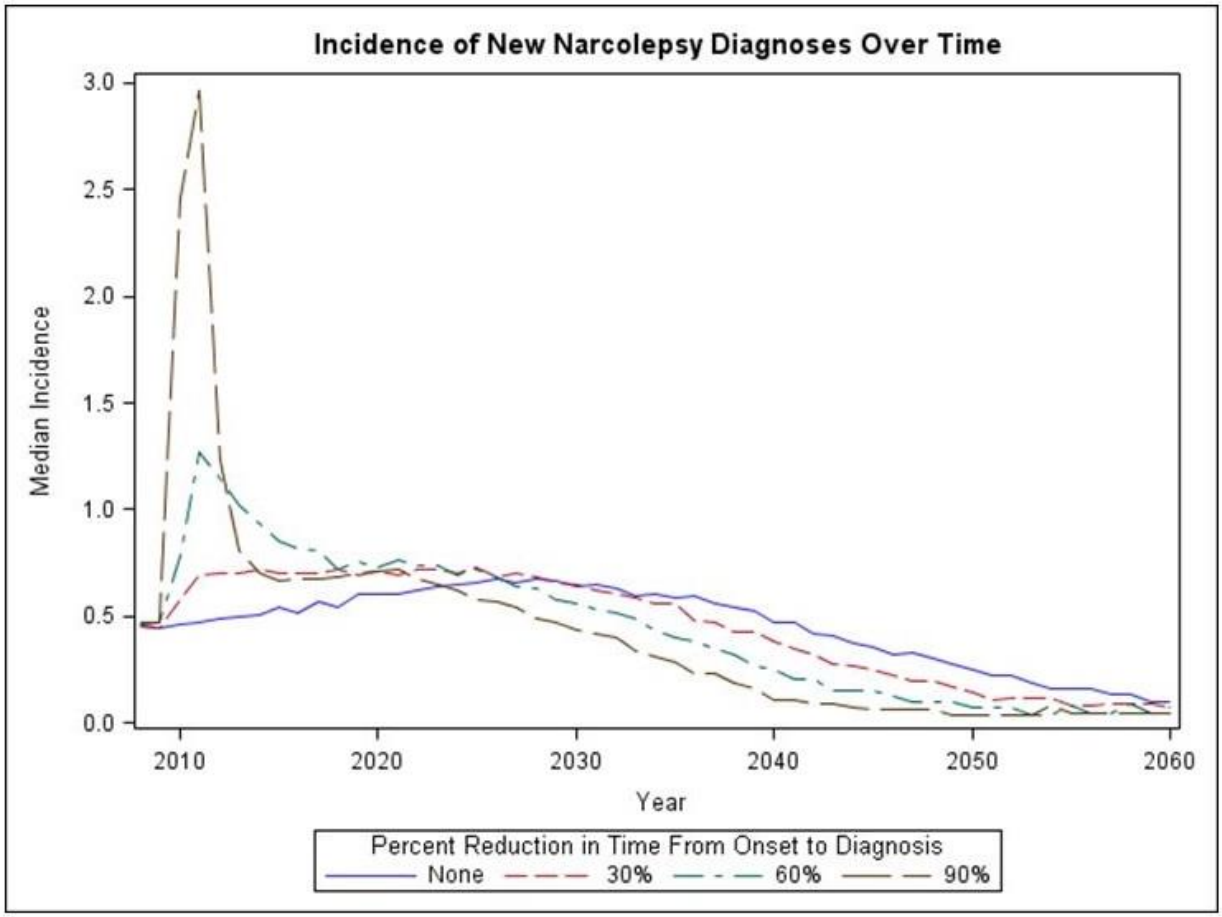


Figure 4. Incidence of new narcolepsy diagnoses over time.

Table 2 shows the results of cohort and case-control analyses of all 108 different parameter settings.

Table 2. Relative Risks and Odds Ratios in primary cohort and case-control analyses

Baseline Onset-to-diagnosis interval	Differential Exposure Misclassification Bias	Detection Bias	**Coverage = 25%**		**Coverage = 50%**		**Coverage = 75%**	
			Cohort RR (95% CI)	Case-Control OR (95%CI)	Cohort RR (95%CI)	Case-Control OR (95%CI)	Cohort RR (95%CI)	Case-Control OR (95%CI)
Median 3 years, Range 0-13 years (Gamma Scale Parameter = 2)	0	0	0.42 (0.16, 1.12)	0.97 (0.34,2.77)	0.38 (0.19, 0.77)	0.99 (0.47 ,2.11)	0.31 (0.17, 0.57)	1.02 (0.53 ,1.94)
		30	0.57 (0.25, 1.32)	1.31 (0.54 ,3.13)	0.5 (0.27, 0.93)	1.19 (0.61 ,2.31)	0.42 (0.25, 0.71)	1.14 (0.65 ,1.98)
		60	1.05 (0.57, 1.92)	2.06 (1.06 ,4.04)	0.82 (0.52, 1.31)	1.69 (1.03 ,2.76)	0.61 (0.41, 0.91)	1.33 (0.87 ,2.04)
		90	3.61 (2.62, 5.02)	5.03 (3.5 ,7.17)	2.15 (1.68, 2.76)	2.83 (2.18 ,3.69)	1.32 (1.07, 1.63)	1.43 (1.19, 1.68)
	30	0	1.09 (0.57, 2.08)	2.1 (1.02 ,4.21)	0.99 (0.6, 1.62)	1.74 (1.02 ,2.93)	0.88 (0.56, 1.4)	1.43 (0.88 ,2.31)
		30	1.51	2.6 (1.37	1.4	2.08	1.11	1.58

			(0.88, 2.6)	,4.87)	(0.89, 2.18)	(1.29 ,3.31)	(0.74, 1.66)	(1.04 ,2.42)
		60	2.75 (1.76, 4.27)	3.92 (2.42 ,6.34)	2.27 (1.59, 3.27)	2.73 (1.85 ,4.01)	1.8 (1.29, 2.52)	1.89 (1.34 ,2.67)
		90	8.09 (6.11, 10.73)	8.11 (5.97 ,11.05)	5.27 (4.18, 6.63)	4.57 (3.58, 5.83)	3.39 (2.76, 4.17)	2.55 (2.06, 3.17)
	60	0	1.81 (1.07, 3.09)	2.93 (1.63 ,5.15)	1.64 (1.06, 2.55)	2.2 (1.37 ,3.51)	1.54 (1.02, 2.31)	1.71 (1.12 ,2.64)
		30	2.61 (1.63, 4.23)	3.62 (2.15 ,6.08)	2.34 (1.58, 3.48)	2.69 (1.79 ,4.07)	2.07 (1.42, 3.02)	1.95 (1.32 ,2.91)
		60	4.71 (3.2, 6.9)	5.41 (3.55 ,8.28)	4.1 (2.94, 5.75)	3.76 (2.64 ,5.35)	3.43 (2.48, 4.79)	2.54 (1.8 ,3.57)
		90	14.36 (10.88, 18.95)	12.49 (9.24 ,16.84)	10.39 (8.13, 13.29)	7.48 (5.78, 9.69)	7.18 (5.71, 9.04)	4.19 (3.30, 5.35)
Median 7 years, Range 0-27 years (Gamma Scale Parameter = 4)	0	0	0.35 (0.05, 2.92)	0.9 (0.11 ,9.73)	0.31 (0.07, 1.38)	0.97 (0.22 ,4.5)	0.26 (0.08, 0.83)	1.03 (0.31 ,3.47)
		30	0.53 (0.12, 3.11)	1.16 (0.19 ,7.65)	0.43 (0.13, 1.41)	1.29 (0.36 ,4.33)	0.37 (0.14, 0.98)	1.16 (0.41 ,3.32)
		60	0.93 (0.31, 2.95)	2.18 (0.63 ,7.85)	0.75 (0.31, 1.75)	1.74 (0.72 ,4.22)	0.58 (0.28, 1.19)	1.36 (0.63 ,2.93)
		90	4.13 (2.58, 6.67)	5.71 (3.37 ,9.75)	2.24 (1.57, 3.22)	2.99 (2.03 ,4.38)	1.27 (0.93, 1.72)	1.74 (1.26 ,2.4)
	30	0	1.88 (0.75, 4.78)	3.1 (1.1 ,9.02)	1.73 (0.8, 3.74)	2.35 (1.02 ,5.38)	1.5 (0.72, 3.11)	1.76 (0.83 ,3.75)
		30	2.71 (1.19, 6.25)	3.89 (1.56 ,9.89)	2.16 (1.11, 4.27)	2.75 (1.31 ,5.71)	1.89 (0.98, 3.59)	2.01 (1.02 ,3.94)
		60	4.73 (2.47, 9.05)	5.57 (2.75 ,11.61)	3.79 (2.16, 6.73)	3.74 (2.06 ,6.81)	3.06 (1.78, 5.26)	2.39 (1.38 ,4.14)
		90	14.96 (9.84, 22.58)	12.7 (8.19 ,19.61)	8.43 (6.05, 11.7)	6.44 (4.57 ,9.12)	5.12 (3.83, 6.85)	3.28 (2.44 ,4.44)
	60	0	3.45 (1.57, 7.59)	4.42 (1.82 ,10.78)	3.29 (1.66, 6.5)	3.24 (1.59 ,6.66)	3.01 (1.54, 5.83)	2.37 (1.19 ,4.75)
		30	5.12 (2.56, 10.3)	5.77 (2.62 ,12.43)	4.42 (2.41, 8.24)	3.83 (2.06 ,7.45)	3.94 (2.14, 7.31)	2.89 (1.51 ,5.52)
		60	8.97 (4.97, 16.34)	8.55 (4.48 ,16.2)	8.44 (4.84, 14.56)	5.79 (3.27 ,10.28)	6.63 (3.85, 11.22)	3.74 (2.14 ,6.64)
		90	32.43 (20.97, 49.31)	22.68 (14.57 ,35.31)	20.09 (13.87, 29.04)	12.95 (8.84 ,19.04)	12.69 (9.1, 17.65)	6.52 (4.44, 8.81)

Median 10 years, Range 0-40 years (Gamma Scale Parameter = 6)	0	0	0 (0, NA)	0 (0, NA)	0.31 (0.04, 4.43)	1.03 (0.16 ,12.32)	0.22 (0.03, 1.83)	1.15 (0.18 ,6.7)
		30	0.48 (0.06, 9.09)	1.12 (0.11 ,61.51)	0.36 (0.05, 2.98)	1.15 (0.2 ,8.23)	0.34 (0.08, 1.55)	1.27 (0.26 ,5.98)
		60	1.03 (0.2, 6.25)	2.1 (0.34 ,15.68)	0.72 (0.21, 2.65)	1.83 (0.5 ,6.71)	0.55 (0.19, 1.57)	1.43 (0.45 ,4.46)
		90	4.37 (2.35, 8.25)	6.08 (2.93 ,12.35)	2.24 (1.39, 3.62)	2.99 (1.79 ,5)	1.26 (0.84, 1.88)	1.73 (1.14 ,2.62)
	30	0	2.34 (0.67, 8.67)	3.76 (0.9 ,16.51)	2.29 (0.8, 6.82)	2.91 (0.93 ,8.74)	2.02 (0.76, 5.43)	1.95 (0.7 ,5.61)
		30	3.82 (1.22, 11.13)	4.96 (1.49 ,16.85)	2.87 (1.13, 7.32)	3.41 (1.25 ,9.25)	2.77 (1.11, 6.53)	2.19 (0.87 ,5.58)
		60	6.19 (2.48, 15.72)	7.47 (2.67 ,20.57)	5.65 (2.56, 12.36)	4.52 (2 ,10.46)	4.08 (1.92, 8.38)	2.76 (1.26 ,5.85)
		90	19.32 (11.4, 33.15)	15.55 (8.85 ,27.5)	10.36 (6.83, 15.73)	7.55 (4.86 ,11.72)	6.11 (4.24, 8.82)	3.76 (2.56 ,5.47)
	60	0	4.76 (1.66, 13.52)	5.78 (1.81 ,19.46)	4.31 (1.68, 10.74)	4.16 (1.54 ,11.17)	4.54 (1.69, 11.83)	2.85 (1.05 ,7.6)
		30	6.82 (2.71, 17.06)	7.24 (2.54 ,21.47)	6.44 (2.67, 15.22)	5.22 (2.04 ,12.88)	5.71 (2.34, 13.4)	3.17 (1.32 ,7.97)
		60	12.92 (5.7, 29.72)	11.3 (4.53 ,27.27)	10.63 (5.08, 22.61)	7.73 (3.44 ,16.83)	8.65 (4.14, 18.89)	4.83 (2.18 ,10.74)
		90	46.02 (25.69, 82.4)	32.09 (17.78 ,58.46)	28.4 (17.13, 47.12)	16.98 (10.15 ,28.1)	17.29 (11.15, 26.99)	7.74 (4.98, 12.12)

Because a 10 year onset-to-diagnosis interval has been reported in the literature, we have chosen to illustrate our results using underlying populations with this onset-to-diagnosis interval and the intermediate vaccine coverage of 50%.

Using a cohort analysis on this underlying population to which the maximum reduction of time from EDS onset-to-diagnosis (90%) has been applied in the absence of differential exposure misclassification produced a median RR of 2.24 (95% CI: 1.39, 3.62). In case-control analysis, the same simulation parameter settings produced an OR of 2.99 (95% CI: 1.79, 5.00) (Table 2).

In the absence of a reduction in the EDS onset-to-diagnosis interval, differential exposure misclassification resulted in a RR of 4.31 (95% CI: 1.68, 10.74) when vaccination coverage was 50% and the EDS date was attributed to the post-vaccination period with a median probability of 60% for vaccinated cases. In the case-control analysis, the same simulation parameter settings produced an OR of 4.16 (95% CI: 1.54, 11.17) (Table 1).

When combining the effect of detection bias and differential exposure misclassification , the estimates were higher in cohort analyses than in case-control analyses as the biases became more pronounced. In the most

extreme scenario, with a median 90% reduction in the onset-to-diagnosis interval in vaccinated cases and a median probability of differential misclassification equal to 60% in vaccinated cases, we found a RR of 28.4 in the cohort analysis (95% CI: 17.13, 47.12). The same parameter settings produced an OR of 16.98 (95% CI: 10.15, 43.85) in the case-control analysis (Table 2). In the absence of either source of bias, median RR estimates from the cohort analysis for all scenarios were less than one when observation time was limited. However, with extension of observation time up to 25 years, the RR was estimated to equal to the simulated RR of one.

Results from case-control analyses were less inflated when detection bias and differential exposure misclassification were present. For both case-control and cohort designs, increased vaccination coverage and a shorter baseline onset-to-diagnosis interval lead to RR estimates closer to the true rate of one when biases are present (Table 1).

Extension of the case capture period reduces the bias (Figure 5). With each extension, the rate of narcolepsy in vaccinated subjects converges toward the background rate. As illustrated in figure 4, reduction in time from onset to diagnoses leads to incidences greater than the background rate in the period following awareness of the association in vaccinated cases, followed by reduction in the incidence rate to levels below the background rate.

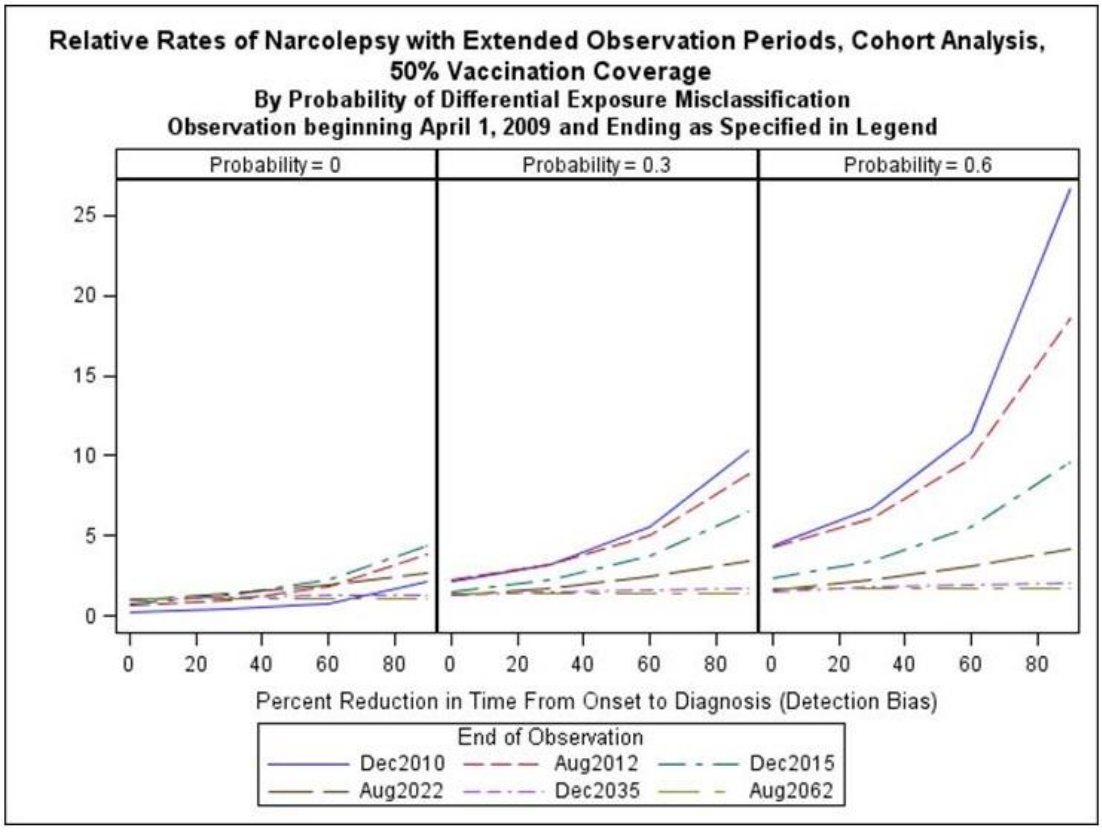


Figure 5. Relative rates of narcolepsy with extended observation periods.

When cases with an onset date after August 15, 2010 were excluded, the RR was 1.87 (95% CI: 1.15, 3.05) in cohort analyses for the extreme setting of detection bias in the absence of differential exposure misclassification. Similarly, the RR was 4.45 (1.76, 11.67) with exclusion of cases with onset after August 15, 2010 at the most extreme setting of differential exposure misclassification in the absence of detection bias. Exclusion of cases with EDS onset dates after media attention, with a 90% reduction in the onset-to-diagnosis

interval and a 60% probability of differential exposure misclassification, produced an RR of 27.10 (95% CI: 16.52, 44.11) while the estimate was 28.4 (95% CI: 17.13, 47.12) when these cases were not excluded. Exclusion of all cases with a diagnosis of narcolepsy after media attention resulted in estimates less than one and confidence intervals including one for all parameter settings. Excluding these cases nullified the effect of differential misclassification bias because only those cases diagnosed after media attention were simulated to misattribute their date of EDS onset to the period following vaccination.

Discussion

Our results indicate that, in the absence of a real association between Pandemrix and narcolepsy, the presence of detection bias or differential exposure misclassification elevates risk estimates.

In the absence of either source of bias, median RR estimates from the cohort analysis for all scenarios were less than the expected value of one. Our explanation for this observation is as follows. The study observation period is limited and the interval between onset and diagnosis can be longer than the study observation time, therefore, as diagnosis is the criteria for case inclusion, a number of cases with onset within the observation period will not be included as cases. However, exposed and unexposed person time within the cohort is fixed. When we analyzed all cases with onset within the observation period regardless of their diagnosis date, the RR was equal to one. In the absence of either bias, using diagnosis dates for case capture, an observation period as long as 25 years would be necessary to obtain the true RR of one.

We found that biased attribution of EDS onset (differential exposure misclassification) has a greater impact on the estimates than a reduction in the EDS onset-to-diagnosis interval (detection bias) both in the cohort and case-control designs. While detection bias increases the relative risk estimates, the effect is not discernible until the onset-to-diagnosis interval is so reduced that many additional cases can be detected in a short observation period. The simultaneous presence of detection bias and differential exposure misclassification increases RRs more rapidly than could be expected by the effect size of each bias in isolation.

In an attempt to exclude detection bias, several published studies limited their primary observation period for EDS onset to the period before media attention [15,18] or included sensitivity analyses using such a reduced study period [21]. Additionally, studies used primary index dates that were thought to be less susceptible to such a bias including onset of symptoms [14,19], first contact with health care [18,21] or referral to specialist care [15,20]. In line with observations from our simulations, limiting analysis to subjects with an onset date prior to media attention will not eliminate the effect of detection bias, since all patients need to be diagnosed to be included, which is where the bias arises. To illustrate this, when limiting cases to those with an EDS date before media attention, Nohynek *et al*. found that the RR increased from 11·4 to 12·7 [18] and O'Flanagan *et al*. found that the RR increased from 13·0 to 14·5 [21]. Since only diagnosed subjects can be included as cases, detection bias will be unavoidable if the onset-to-diagnosis interval is shorter in vaccinated individuals. The only way to circumvent the combined effects of detection bias and differential exposure misclassification would be to select only patients diagnosed before media attention. This will result in limited observation time and limited case inclusion as illustrated by our simulations and as was shown in the VAESCO study (13). We are not aware of any existing statistical methods to control for detection bias although quantitative bias analysis could adjust for hypothesized biases [28].

With limited observation time, we found that, in the presence of detection bias and differential exposure misclassification, estimates from the case-control design are less inflated than those from the cohort design. The resilience of the case-control in this scenario has several causes: the outcome is rare and the pool of controls, matched only by sex and age at onset, is large; also, the invariability of exposed person time, which is limited by observation time and vaccine coverage in the cohort approach, is avoided. Additionally, in this simulated scenario, we were able to sample controls from the same population as the cases and to assess their exposure without error, thereby avoiding the most problematic sources of bias in case-control studies. The only

study to date in which data were analyzed using both a case-control and a cohort design found lower estimates in the case-control than in the cohort design [16]. Applying these findings to the interpretation of all published studies, however, presents a challenge as each study differed not only in design choice but in many other ways including underlying population, diagnostic practices, inclusion and exclusion criteria, and many others. In general, however, estimates from case-control studies were similar to those from those cohort studies in which diagnosis was used as an index date. Those cohort studies which used onset of symptoms as an index date produced much higher estimates, suggesting presence of bias, particularly differential exposure misclassification, a true association in those populations, or both. However, given the complexity of the interplay of design choices and underlying populations, a meaningful comparison between designs implemented in published studies is not feasible.

Increased vaccination coverage reduced the bias in cohort and case-control analyses. In cohort analyses, this is explained by an increase in the person time denominator for vaccinated cases with a smaller increase in events and, simultaneously, a decrease in the person time denominator for unvaccinated cases with a smaller decrease in the number of events. In case-control analyses, this could be attributed to a greater probability of matching to vaccinated controls as vaccination coverage increases, leading in turn to more informative strata in a conditional analysis.

When a shorter interval from onset to diagnosis was assumed, the impact of simulated biases was less pronounced. This is due to the fact that, with a shorter onset-to-diagnosis interval, more cases, whether vaccinated or not, are being captured during the study period.

We chose to simulate only those sources of bias for which data in the absence of a vaccine safety signal exists and for which simulated variables could be modified to mimic the bias. Our simulations therefore do not reflect all of the biases that could potentially affect estimates of an association between Pandemrix and narcolepsy. For example, it is possible that non-vaccinated cases also experienced a reduction in the onset to diagnosis interval due to increased awareness of narcolepsy. However, inclusion of additional simulation parameters such as this would have required the making of additional assumptions for which we had no basis in published data. By focusing on biases that could be evaluated without making untenable assumptions, these simulations provide insights that can improve rapid evaluation of vaccine safety signals by decision makers. There were several uncertainties, including the true background rate of narcolepsy and the true interval between onset of symptoms and diagnosis, for which we made assumptions in order to conduct our simulations. The validity of these assumptions will ultimately determine the robustness of our simulations.

The introduction of a new vaccine, or an existing vaccine in new populations, requires the assessment of vaccine safety. Large numbers of people can be exposed in a relatively short period providing a challenge to real-time safety surveillance. In such situations, as illustrated by the experience with Pandemrix and narcolepsy during the 2009/2010 H1N1 pandemic, it can be difficult to determine if a safety finding is a true association or not. Despite these challenges, the timely and accurate assessment of potential associations between adverse events and vaccination are crucial to ensure vaccine safety and maintain the public's confidence. We believe that our simulations provide useful insights for the design and interpretation of future studies. Importantly, our results illustrate that in future analyses of safety signals for diseases with long latency periods for which observation times are limited the effect of limited case capture together with fixed person time denominators should be recognized. Similarly, the changes in exposed and unexposed person time denominators with changing vaccination coverage should be also taken into account. As we have shown, the case-control design provides less biased estimates in these circumstances as it does not require the calculation of person time. Moreover, our simulations illustrate the importance of not only understanding background rates of adverse events of special interest prior to vaccination campaigns, but also having insight in the background onset-to-diagnosis interval.

Recommendations:

Because rapid assessment of a vaccine safety signal, by definition, means limited case capture time, not only the background incidence of events of interest but the background onset-to-diagnosis interval should be understood for proper interpretation of risk estimates.

The impact of differential exposure misclassification in these simulations underlines the need for accurate and linkable vaccine registries as well as blinded assessment of cases.

When person-time is fixed and the outcome is rare, a case-control design is more resilient to bias and should be considered.

Population cohorts should continue to be followed over time to monitor how rates of narcolepsy change following the H1N1 pandemic. If these biases were indeed present, we would expect to see incidence eventually fall back to or even below the background rate.

To conclude, our results indicate that, in the absence of a real association between the vaccine and narcolepsy, presence of detection bias and differential exposure misclassification could account for elevated RRs in vaccinees in association studies. While this does not exclude a real increased risk of narcolepsy following Pandemrix, it is possible that the levels of increased risk observed were at least partially due to bias.

Expert Commentary

When the narcolepsy signal emerged in Sweden and Finland in 2010, studies had to be started rapidly across Europe in order to address this signal. Possibilities to evaluate this safety signals were limited to observational studies which, as a biological mechanism for the observed adverse event was not known, involved a great deal of guess work on risk windows, potential confounding factors and alternative explanations for observed association. As an answer on the potential association between Pandemrix and narcolepsy was needed rapidly, studies performed had limited follow-up time. They had to do with the resources available at the time which meant that they could not always rely on blinded, prospectively collected data on vaccination for the study population or on detailed prospectively collected data on potential confounding variables such as underlying comorbidities.

To add to this, suspicion of a potential association between narcolepsy and Pandemrix was already spreading amongst healthcare professionals in Finland from as early as February 2010 and was general knowledge after August 2010 when regulatory agencies published on the association which was picked up by the media. Knowledge on the association with vaccination may have resulted in a reduction of the onset to diagnosis interval in vaccinated individuals, whereas this would not happen to the same extent in non-vaccinated subjects. Knowledge of a putative association between vaccination and a specific event could also have resulted in patients placing symptom onset after vaccination.

The simulations described in this article illustrate that in the absence of a real association between the vaccine and narcolepsy presence of detection bias and differential exposure misclassification could account for the elevated risks detected. Moreover, the simulations also suggest that it would be too early to dismiss an elevated risk of narcolepsy following other influenza vaccines or influenza infection based upon absence of associations in observational studies alone. The veracity of the association between narcolepsy and Pandemrix will become clearer as studies are conducted with longer follow-up times, especially when studies into potential mechanisms are taken into account.

The uncertainties surrounding the role Pandemrix may have played in the surge of narcolepsy diagnoses seen in several European countries which still exist to date do underline the need to improve the infrastructures available in Europe to monitor vaccine safety and evaluate vaccine safety signals if these were to emerge. Moreover, they point towards the need to further develop methods for rapid safety assessment, such as sequential monitoring, and the need to develop methods which can adjust for stimulated diagnosis in the

presence of awareness. Finally, a systematic assessment of potential sources of bias and their impact should be an integral part of any assessment of a safety signal that relies on observational studies. Without such an assessment great caution should be exerted before drawing any conclusions from these studies.

Five Year View

It is not possible to predict when and how a future pandemic will evolve. Although we might now have considerable experience to inform the safety profile of the existing influenza vaccines, with a new pandemic virus and a new mass-vaccination programme we need to be prepared for the occurrence of new safety signals. The experience of narcolepsy has taught us that it is very helpful to have a good understanding not only of the epidemiology of potential adverse events in Europe but also of the diagnostic process for these events. It is necessary to know if there is potential for under diagnosis and what delays in diagnosis can be expected in different age groups and in different countries. Although impossible to pinpoint what adverse events will be of interest, considering the experience of influenza and (adjuvanted) influenza vaccines focus should be on neurological events and disorders with a potential auto-immune aetiology. The inter-pandemic period should be used to collect data on diagnosis rates in different age groups, populations, countries and improve the understanding of differences between European countries in recording and diagnosing these conditions.

References

1. National Institute for Health and Welfare (THL). National Institute for Health and Welfare recommends discontinuation of Pandemrix vaccinations. (Ed.^(Eds) (25 Aug 2010)

2. Medical Products Agency. The MPA investigates reports of narcolepsy in patients vaccinated with Pandemrix. (Ed.^(Eds) (2010)

3. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) Assessment report - Pandemrix Influenza vaccine (H1N1) (split virion, inactivated, adjuvanted) A/California/7/2009 (H1N1)v like strain (x-179a) - Procedure No.: EMEA/H/C/000832/A20/0045. (Ed.^(Eds) (21 July 2011)

4. European Medicines Agency. European Medicines Agency starts review of Pandemrix. . (Ed.^(Eds) (27 August 2010)

5. European Medicines Agency. Press release: European Medicines Agency recommends restricting use of Pandemrix (Ed.^(Eds) (21 July 2011)

6. Kornum B, Faraco J, Mignot E. Narcolepsy with hypocretin/orexin deficiency, infections and autoimmunity of the brain. *Curr Opin Neurobiol.*, 21, 897-903 (2011).

7. Nishino S, Ripley B, Overeem S, Lammers GJ, Mignot E. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet*, 355(9197), 39-40 (2000).

8. Overeem S, Black JL, 3rd, Lammers GJ. Narcolepsy: immunological aspects. *Sleep medicine reviews*, 12(2), 95-107 (2008).

9. Luca G, Haba-Rubio J, Dauvilliers Y *et al.* Clinical, polysomnographic and genome-wide association analyses of narcolepsy with cataplexy: a European Narcolepsy Network study. *Journal of sleep research*, 22(5), 482-495 (2013).*

10. Thorpy MJ, Krieger AC. Delayed diagnosis of narcolepsy: characterization and impact. *Sleep medicine*, 15(5), 502-507 (2014).

11. Morrish E, King MA, Smith IE, Shneerson JM. Factors associated with a delay in the diagnosis of narcolepsy. *Sleep medicine*, 5(1), 37-41 (2004).*

12. Singh AK, Mahlios J, Mignot E. Genetic association, seasonal infections and autoimmune basis of narcolepsy. *Journal of autoimmunity*, 43, 26-31 (2013).

13. Medical Products Agency. Narcolepsy – Case inventory study in Sweden 2009-2010. (Ed.^(Eds) (2011)

14. Medical Products Agency. A registry based comparative cohort study in four Swedish counties of the risk for narcolepsy after vaccination with Pandemrix - A first and preliminary report, by the. (Ed.^(Eds) (2011)

15. European Centre for Disease Prevention and Control. Narcolepsy in association with pandemic influenza vaccination (a multi-country European epidemiological investigation). ECDC (Ed.^(Eds) (Stockholm, September 2012) **

16. Montplaisir J, Petit D, Quinn MJ *et al.* Risk of Narcolepsy Associated with Inactivated Adjuvanted (AS03) A/H1N1 (2009) Pandemic Influenza Vaccine in Quebec. *PloS one*, 9(9), e108489 (2014).

17. Persson I, Granath F, Askling J, Ludvigsson JF, Olsson T, Feltelius N. Risks of neurological and immune-related diseases, including narcolepsy, after vaccination with Pandemrix: a population- and registry-based cohort study with over 2 years of follow-up. *Journal of internal medicine*, (2013).

18. Nohynek H, Jokinen J, Partinen M *et al.* AS03 adjuvanted AH1N1 vaccine associated with an abrupt increase in the incidence of childhood narcolepsy in Finland. *PloS one*, 7(3), e33536 (2012).*

19. Miller E, Andrews N, Stellitano L *et al.* Risk of narcolepsy in children and young people receiving AS03 adjuvanted pandemic A/H1N1 2009 influenza vaccine: retrospective analysis. *BMJ*, 346, f794 (2013).

20. Dauvilliers Y, Arnulf I, Lecendreux M *et al.* Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. *Brain : a journal of neurology*, 136(Pt 8), 2486-2496 (2013).

21. O'Flanagan D, Barret A, Foley M *et al.* Investigation of an association between onset of narcolepsy and vaccination with pandemic influenza vaccine, Ireland April 2009-December 2010. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 19(17) (2014).

22. Barker CI, Snape MD. Pandemic influenza A H1N1 vaccines and narcolepsy: vaccine safety surveillance in action. *The Lancet. Infectious diseases*, 14(3), 227-238 (2014).

23. Sturkenboom MC. The narcolepsy-pandemic influenza story: Can the truth ever be unraveled? *Vaccine*, 33 Suppl 2, B6-B13 (2015).

24. Schonberger LB, Bregman DJ, Sullivan-Bolyai JZ *et al.* Guillain-Barre syndrome following vaccination in the National Influenza Immunization Program, United States, 1976--1977. *American journal of epidemiology*, 110(2), 105-123 (1979).

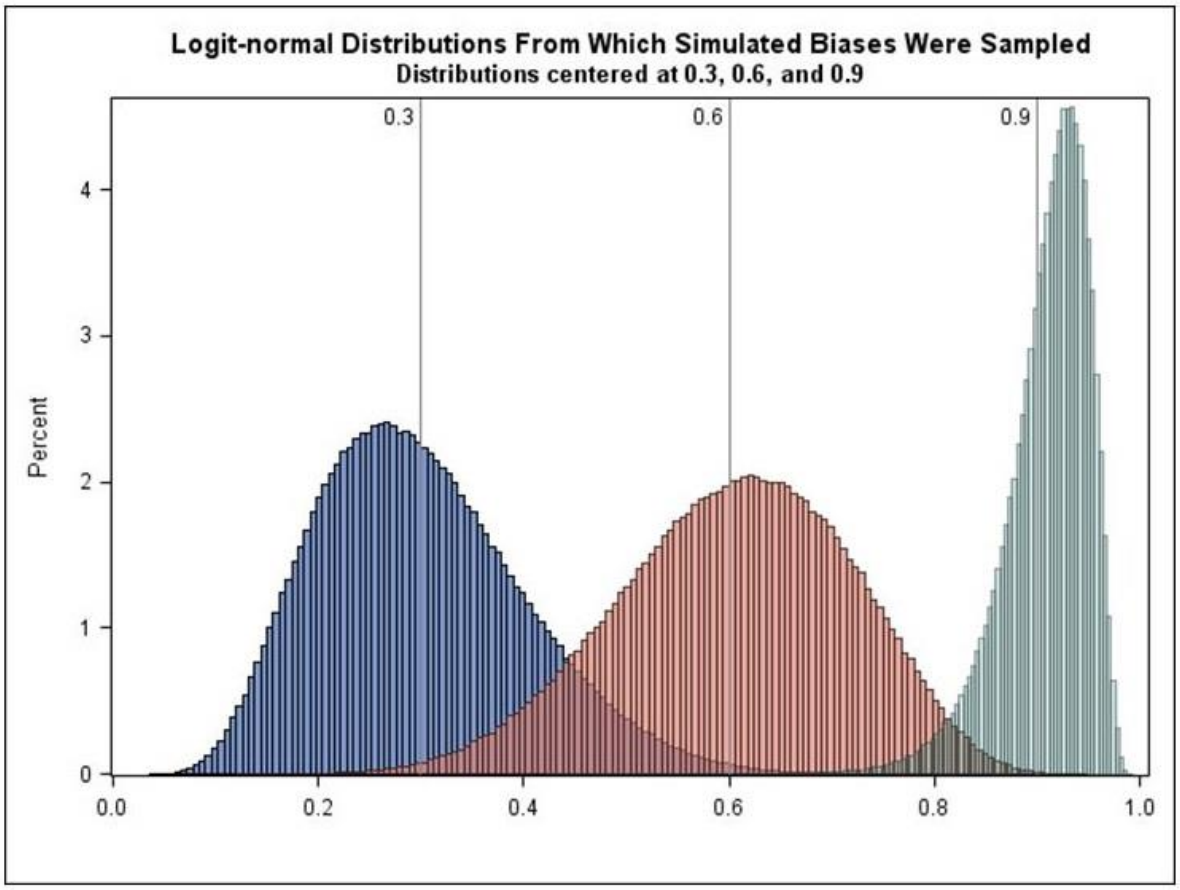
25. Andrews N, Miller E, Taylor B *et al.* Recall bias, MMR, and autism. *Archives of disease in childhood*, 87(6), 493-494 (2002).

26. Silber MH, Krahn LE, Olson EJ, Pankratz VS. The epidemiology of narcolepsy in Olmsted County, Minnesota: a population-based study. *Sleep*, 25(2), 197-202 (2002).*

27. Mereckiene J, Cotter S, Weber JT *et al.* Influenza A(H1N1)pdm09 vaccination policies and coverage in Europe. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 17(4) (2012).

28. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *International journal of epidemiology*, 43(6), 1969-1985 (2014).**

Supplementary Material



Supplementary Figure 1. Logit-normal distributions from which simulated biases were sampled

6.2 QUANTIFYING OUTCOME MISCLASSIFICATION IN MULTI-DATABASE STUDIES: THE CASE STUDY OF PERTUSSIS IN THE ADVANCE PROJECT

Rosa Gini, Caitlin N Dodd

Kaatje Bollaerts, Claudia Bartolini

Giuseppe Roberto, Consuelo Huerta-Alvarez

Elisa Martín-Merino, Talita Duarte-Salles

Gino Picelli, Lara Tramontan

Giorgia Danieli, Ana Correa

Chris McGee, Benedikt F H Becker

Charlotte Switzer, Sonja Gandhi-Banga

Jorgen Bauwens, Nicoline A T van der Maas

Gianfranco Spiteri, Emmanouela Sdona

Daniel Weibel, Miriam Sturkenboom

Submitted

Abstract

Background

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines using European healthcare databases. Event misclassification can result in biased estimates and contribute to heterogeneity in results. Here we report the impact of different event-finding algorithms for *Bordetella pertussis* (BorPer) on the estimated incidence rates (IRs) and algorithm validity.

Methods

Four participating databases retrieved data from primary care (PC) setting: (BIFAP: Spain), THIN and RCGP RSC: UK) and PEDIANET: Italy); the fifth SIDIAP (Spain) from both PC and hospital settings. The algorithms were defined by setting, data domain (diagnoses, drugs, or tests) and concept sets (specific or unspecified pertussis). BorPer IRs were estimated in children aged 0-14 years enrolled in 2012 and 2014 and followed up until the end of each calendar year and compared with IRs of confirmed pertussis from the ECDC surveillance system (TESSy).

Results

The number of cases and the estimated BorPer IRs per 100,000 person-years in PC, using data representing 3,173,268 person-years, were 0 (IR=0.0), 21 (IR=4.3), 21 (IR=5.1), 79 (IR=5.7), and 2 (IR=2.3) in BIFAP, SIDIAP, THIN, RCGP RSC and PEDIANET respectively. The IRs for combined specific/unspecified pertussis were higher and were comparable with data from TESSy, except PEDIANET. In SIDIAP the estimated IR was 45.0 when discharge diagnoses were included. The sensitivity and positive predictive value of combined PC specific and unspecific diagnoses for BorPer cases in SIDIAP were 85% and 72%, respectively, based on overlap between hospital and PC diagnoses (adjusted IR=35.5).

Conclusion

This study demonstrated the value of quantifying the impact of different event-finding algorithms across databases and the possibility of benchmarking with disease surveillance data as well as assessing validity estimates when data from different settings can be linked.

1. Introduction

ADVANCE is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines using existing healthcare databases in Europe [1] (see Appendix for list of consortium members). These databases have proven very useful for studying drug effects and are commonly used in pharmacoepidemiology [2].

Identifying events, such as vaccine-preventable diseases, adverse events of interest, co-morbidities and exposure to vaccination, is a pivotal first step in vaccine B/R studies. Since there is limited or no control over the primary data collection when using existing healthcare databases, event retrieval is usually not perfect. Individuals who experienced the event might not be retrieved, for example if an individual is admitted to hospital for the event but no primary care (PC) diagnosis is recorded, the event will not be retrieved from PC databases and some individuals might be identified as having the event when in fact they did not. In a PC database, this typically happens when the physician had only a suspicion, or if it was a ruled-out diagnosis.

Researchers who access these databases usually develop their own methods to identify events of interest, which are not always fully transparent [3, 4]. Events may be retrieved by combining information from different settings (e.g., PC and hospital) and data domains, for example diagnostic codes, drugs as proxies (e.g. in the case of diabetes), or laboratory measurements. Use of information from more than one data domain, compared with using diagnoses information only, can alter the sensitivity and positive predictive value (PPV) of the event-finding algorithm. This alteration may happen differently in different databases, due to the local characteristics of the database, the population, or the healthcare system.

It is well established that misclassification of events (false positives or false negatives) can introduce bias in epidemiological studies, which can be corrected, to some extent, using statistical methods [5-7]. However, to correct this bias, some validity parameters such as sensitivity and PPV are required [8]. For this a gold standard and chart reviews are required, which generally make it costly and time-consuming.

In an attempt to develop a systematic approach to quantifying the impact of using different event identification algorithms in multi-national, multi-database studies, the *component algorithm* strategy was developed (Roberto 2016): a set of standardized algorithms, called *components*, are defined and applied in each database. The impact of different algorithms on the resulting estimates of disease occurrence is subsequently measured [9]. In this study we aimed to refine this strategy by further standardizing the process, by developing and applying novel formulas, by using benchmark data from another source and by using a data source which had data from two settings. Since the proof-of-concept studies of ADVANCE focused on pertussis, we used this event as case study.

2. Methods

2.1. Bordetella pertussis disease information

Bordetella pertussis causes pertussis, a vaccine-preventable infectious disease of the respiratory tract. Symptoms include paroxysms of cough typically lasting from 1 to 6 weeks or more and these may be milder in adolescents or immunised children [10, 11]. Several tests are available to confirm *Bordetella pertussis* infection, including culture (which takes up to 14 days), serology and nucleic acid amplification tests. Pertussis is a notifiable infectious disease and cases should be reported to the national surveillance system in all the countries involved in ADVANCE. European Union member states are required to report available data on pertussis cases to the European Centre for Disease Prevention and Control (ECDC). A standardised case definition is used which classifies cases based on clinical, epidemiological and laboratory criteria [12]. All national reports are submitted to the European Surveillance System database (TESSy) managed by the ECDC [13].

2.2. Data sources

We assessed the impact of different event-finding algorithms using five databases that participated in the ADVANCE proof-of-concept studies: BIFAP and SIDIAP (Spain), PEDIANET (Italy) and RCGP RSC and THIN (United Kingdom). All databases were population-based with data from electronic medical records in the PC setting. In SIDIAP, the analyses were restricted to the population in this PC database that could be linked to hospital discharge records. Surveillance data on pertussis were obtained from the TESSy surveillance system through ECDC, a partner of ADVANCE.

2.3. Study population and study design

We used a dynamic cohort study design to study the impact of different event-finding algorithms on the estimated pertussis IRs. Due to the methodological nature of this study, to enable us to explore in more detail a number of strategies, we included a larger cohort in the study population than that in the other ADVANCE studies. Therefore, children aged 0 to 14 years who were registered in the participating databases entered the study cohorts on 1st January 2012 and 1st January 2014, and were followed up during 2012 and 2014, respectively. Children who were born during 2012 or 2014 were followed up from birth until the end of the calendar year. Children who were older than 14 years at any point in the follow-up were excluded. To exclude any previous cases that had been notified before the start of the study period, children who had a record of one of the components of pertussis during the two years prior to one of the cohort entry dates were excluded, unless the component referred to the data domain of drugs (see below for more details on the component definition).

2.4. Selection of component algorithms

A component algorithm is a standardised event-finding algorithm specified by three characteristics: the *setting* of primary data collection (PC or hospital), the *data domain* involved in the algorithm, and the *set of concepts* used to find the codes used to query the database [9]. The sets of concepts were created by aggregating the codes that were obtained from an initial proposed list, completed with a literature review and pertussis case definitions [2, 13]. The CodeMapper tool was used to support the process [14]. Labelling and classification of identified concepts, as well as the construction of the components, were conducted by one of the authors who is a pertussis expert (NvdM). As a result, seven concept sets were created (**Table 1**) [15, 16]. In particular, two sets of concepts belonged to the diagnoses data domain: the set labelled '(*Bordetella pertussis*)' included three concepts which specifically indicated *Bordetella pertussis* as the causative agent of the infection, while the set labelled '(pertussis unspecified)' included five concepts indicating unspecified pertussis. The corresponding codes and free text keywords are given in **Supplementary Table 1**.

Table 1. Sets of concepts selected for the component algorithms.

Each set of concept is indicated with a label and described with a text. Each set of concepts contains one or more concepts, each described with a text and, if available, with a Concept Unique Identifier of the Unified Medical Language System.

Concept set label	Concept set description	Concept	Concept Unique Identifier
(Bordetella pertussis)	Concepts referring to diagnoses specifically mentioning pertussis induced by an infection of *Bordetella pertussis*	Bordetella pertussis	C0043167
		Whooping cough due to Bordetella pertussis without pneumonia	C2887068
		Whooping cough due	C2887069

		to Bordetella pertussis with pneumonia	
(Pertussis unspecified)	Concepts referring to diagnoses which refer to pertussis, but without a specific indication that *Bordetella pertussis* is responsible for the infection	Whooping cough due to unspecified organism	C0043168
		Bordetella Infections	C0006015
		Whooping cough-like syndrome	C0343485
		Notification of whooping cough	
		Pneumonia in pertussis	C0155865
(Symptoms compatible with pertussis)	This set of concepts was introduced because the Spanish translation of 'whooping cough' was found to be considered by physicians as a symptom, not as a diagnosis	Concept of 'tos pertusoide' in Spanish general practice	
(Symtoms in infants)	Concepts referring to symptoms that were found to be predictive of pertussis in infants (Hurtado-Mingo 2013, Bellettini 2014)	Apnea	C0003578
		Cyanosis	C0010520
		Post-tussive vomiting	C1740793
		Paroxysms of coughing	C0231911
(Macrolides)	Use of macrolides	Macrolides	
(Bordetella pertussis test)	The concepts listed in this set indicate the prescription of tests that are considered to be confirmatory of a Bordetella pertussis infection	polymerase chain reaction test	
		culture or serology	
		isolation of Bordetella pertussis from a clinical specimen	
(Positive result from a Bordetella pertussis test)	The concepts listed in this set indicate a positive result from a tests confirmatory of a Bordetella pertussis infection	positive polymerase chain reaction test	
		positive culture or serology	
		positive isolation of Bordetella pertussis from a clinical specimen	

The primary components associating concepts with settings (PC and hospital) are described in **Table 2**. Some secondary components, combining primary components in pre-defined temporal relations (e.g., symptoms in the presence of a drug prescription in the previous 30 days) were also created.

Table 2. Components for pertussis.

The concept sets referred to by the words in round parentheses can be found in Table 1

Name	**Setting**	**Data domain**	**Concept set**
PC diagnosis, specific	Primary care practice	Diagnosis	(Bordetella pertussis)

Inpatient diagnosis, specific	Hospital	Diagnosis	(Bordetella pertussis)
PC diagnosis, unspecified	Primary care practice	Diagnosis	(Pertussis unspecified)
Inpatient diagnosis, unspecified	Hospital	Diagnosis	(Pertussis unspecified)
Symptoms	Primary care practice	Diagnosis or sign/symptoms	(Symptoms compatible with pertussis)
Symptoms in infants	Primary care practice	Diagnosis or sign/symptoms	(Symptoms in infants)
Test	Any setting where a test can be prescribed, or facility where the test is administered	Laboratory test	(Bordetella pertussis test)
Positive laboratory results	Any setting where a health professional records the results of a test, or facility where the results of the test are generated	Results from laboratory test	(Positive result from a Bordetella pertussis test)
Drug use	Facility dispensing medications or primary care practice issuing prescriptions	Drug	(Macrolides)
Secondary components			
Symptoms and drugs within 30 days	A patient is positive if he/she has both a record of Symptoms and of Drug use, and the interval between the dates is less than 30 days		
Symptoms in infants and drugs within 30 days	A patient is positive if he/she is 0 or 1 and has both a record of Symptoms in infants and of Drug use, and the interval between the dates is less than 30 days		

2.5. Analysis

Each database manager received an R-coded programme (quality checked by double-coding against Stata) which was programmed using the pre-specified common data model [1]. These programmes produced aggregated outputs, which were then transferred to the remote research environment. Event-finding algorithms were created as logical combinations of individual components using Boolean operators. For example, the two components 'PC diagnosis, specific' and 'PC diagnosis, unspecified' were combined in one component: 'PC specific OR unspecified diagnoses', which detected all individuals that were positive for either of the original components. Based on the different event-finding algorithms, incidence rates (IRs) were

estimated using the number of persons retrieved with the respective events as numerator and the follow-up person-time as denominator (see **Supplementary File 1**).

Age and country-specific incidences per 100,000 person-years of confirmed BorPer for both 2012 and 2014 were calculated for children aged 0-14 years. The calculations used the reported confirmed cases in the TESSy surveillance system in 2012 and 2014 as the numerator, and person-time from population distributions in EUROSTAT for 2012 and 2014 as the denominator [17]. Exact Poisson confidence intervals (95% CI) were calculated [18].

Some formulae link the true proportion of BorP and/or validity indices with each other and with the observed proportion of the component algorithms (**Table 3**). These formulas are explained in **Supplementary File 2**.

Table 3. Analytic formulas linking true proportion of pertussis and validity indices of one or two algorithms.

In the formulas, Π is the true proportion of cases of pertussis, P is the proportion of cases detected by the algorithm, SE is the sensitivity and PPV is the positive predictive value of the algorithm.

One algorithm		
# of formula	**Parameters known**	**Formula to derive another parameter**
Formula 1	PPV and SE	$\pi = \frac{P \times PPV}{SE}$
Formula 2	PPV and Π	$SE = \frac{P \times PPV}{\pi}$
Formula 3	SE and Π	$PPV = \frac{SE \times \pi}{P}$
Two algorithms A and B		
# of Formula	**Parameters known**	**Formula to derive another parameter**
Formula 4	SE of A, of B, and of A **AND** B	$SE_{A\ \mathbf{OR} B} = SE_A + SE_B - SE_{A\ \mathbf{AND} B}$
Formula 5	Π and PPV of A, of B, and of A **AND** B	$SE_{A\ \mathbf{OR}\ B} = \frac{P_A \times PPV_A}{\pi} + \frac{P_B \times PPV_B}{\pi} - \frac{P_{A\ \mathbf{AND}\ B} \times PPV_{A\ \mathbf{AND}\ B}}{\pi}$
Formula 6	SE of A **OR** B, and PPV of A, of B, and of A **AND** B	$\pi = \frac{P_A \times PPV_A + P_B \times PPV_B - P_{A\ \mathbf{AND}\ B} \times PPV_{A\ \mathbf{AND}\ B}}{SE_{A\ \mathbf{OR}\ B}}$
Formula 7	PPV of A, of B, and of A **AND** B	$PPV_{A\ \mathbf{OR}\ B} = \frac{P_A \times PPV_A + P_B \times PPV_B - P_{A\ \mathbf{AND}\ B} \times PPV_{A\ \mathbf{AND}\ B}}{P_{A\ \mathbf{OR}\ B}}$

In this study we considered Π = IR (see **Supplementary File 1**) and we assumed that for all algorithms A and B, the proportion of true positives among those detected by both algorithms (PPV of A **AND** B), was the same as the PPV of A OR B, whichever was the highest, which may be considered the most conservative assumption.

Since the concept set labelled '*Bordetella pertussis*' was composed of codes explicitly mentioning the bacterium, we considered that components based on this had a high likelihood of extracting true cases. Therefore we considered it was conservative to assume that the PPV for 'PC diagnosis, specific' and for 'inpatient diagnosis, specific' was 90%. We explored two scenarios for the cases extracted by the components associated with the concept set labelled 'pertussis unspecified', assuming that the PPV was 70% or 50%. The PPV for the component 'positive laboratory results' was assumed to be 100%. Finally, we assumed that all true cases in SIDIAP were recorded in at least one of the diagnosis or laboratory-based components: this assumption may overestimate sensitivity. Based on this and on the formulae in Table 3, we derived sensitivity and PPV estimates for the algorithm 'PC specific OR unspecified diagnosis' in SIDIAP, and the adjusted IR of BorPer in the study population.

3. Results

3.1. Study population

We followed 3,173,268 person-years of children during the study period: 488,847 from the SIDIAP database, 796,324 from BIFAP, 88,754 from PEDIANET, 1,387,939 from THIN and 411,404 from RCGP RSC (**Table 4**). The percentages of children aged 0 or 1 years in the population aged 0-14 years in Spain were 12.1% and 16.1% in SIDIAP in BIFAP, respectively, compared with 13.5% in the EUROSTAT Spanish population. In the UK the percentages were 15.1% and 14.8% in RCGP RSC and THIN 13.0%, respectively, compared with 14.3% in the EUROSTAT UK population and in PEDIANET (vs 12.9%); in (vs 14.3%).

3.2. Incidence rates estimated by the algorithms

The IRs for the component and composite algorithms, as well as the benchmark IRs from the TESSy surveillance system are illustrated in Figure 1 and documented in **Table 4**. The IRs estimated from the TESSy surveillance system in 2012 and 2014 for children aged 0-14 years were 21.2 (95% CI: 20.5; 22.0) for Spain, 13.4 (95% CI: 13.0; 13.9) for the United Kingdom, and 5.4 (95% CI: 5.1; .8) for Italy. The number of cases of 'PC diagnosis, specific' (and IRs per 100,000 PY) were 0 (0.0), 21 (4.3), 21 (5.1), 79 (5.7), and 2 (2.3) in the BIFAP, SIDIAP, RCGP RSC, THIN and PEDIANET databases, respectively. The component 'PC diagnosis, unspecified' had a higher IR in all databases, and combining the two components (one OR the other) increased the number of cases detected and the IRs to 135 (IR=17.0), 194 (IR=39.6), 39 (IR=43.9), 246 (IR=17.7), and 91 (IR=22.1), respectively. In BIFAP, SIDIAP, RCGP RSC and THIN, when taking into account that the unspecified component may have captured some false positives, the IRs were comparable with the corresponding IR from the TESSy surveillance system (17.0 vs 21.2; 39.6 vs. 21.2; 22.1 vs. 13.4; 17.7 vs. 13.4, respectively). In PEDIANET the composite IR was much higher than the IR from the TESSy database (43.9 vs 5.4).

SIDIAP was the only database in which data from both the PC and hospital settings could be linked. The total number of cases in 'PC OR inpatient diagnosis' in SIDIAP was 220 (IR=45.0), including 26 (12%) that had not been identified in the PC setting. Unlike in the PC setting, where most of the diagnoses were unspecified, in the inpatient setting there were around half specific and half unspecified diagnoses.

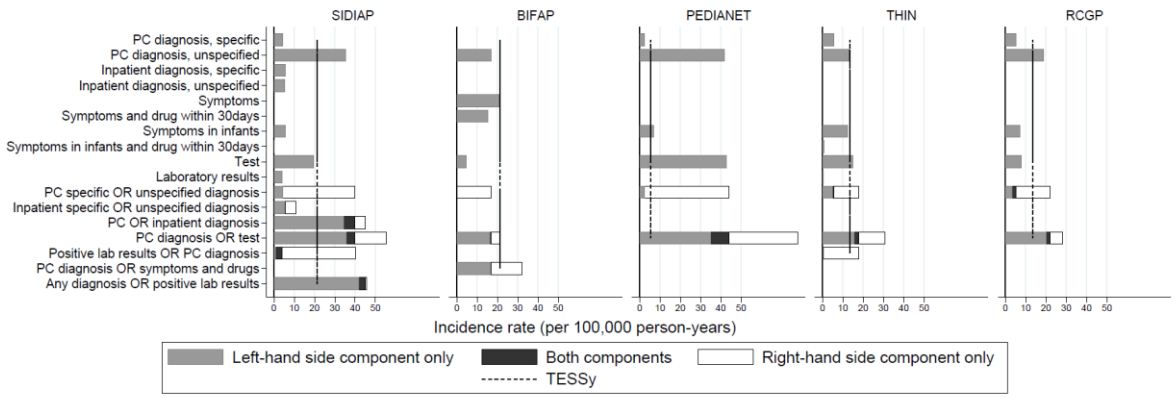


Figure 1. Study results for the incidence of tested component and composite algorithms.

For each component algorithm the incidence rate per 100,000 is represented. In composite algorithms, incidence rates are stratified per type of case: cases detected only by the left-hand component (indicated in the label before the keyword 'OR'), cases detected by both components, and cases detected by the right-hand component (indicated in the label after the keyword 'OR'). The line represents the national incidence rate per 100,000 based on TESSy data. Years 2012 and 2014.

DB	SIDIAP (Spain)				BIFAP (Spain)				PEDIANET (Italy)				THIN (United Kingdom)				RCGP (United Kingdom)			
Person-years	488,847				796,324				88, 754				1,387,939				411,404			
TESSy (IR and 95% CI)	21.2 (20.5-22.0)				21.2 (20.5-22.0)				5.4 (5.1-5.8)				13.4 (13.0-13.9)				13.4 (13.0-13.9)			
Component algorithms (N and IR per 100,000 PYs)																				
PC diagnosis, specific	21 (4.3)				0 (0.0)				2 (2.3)				79 (5.7)				21 (5.1)			
PC diagnosis, unspecified	173 (35.4)				135 (17.0)				37 (41.7)				178 (12.8)				77 (18.7)			
Inpatient diagnosis, specific	27 (5.5)																			
Inpatient diagnosis, unspecified	26 (5.3)																			
Symptoms					166 (20.8)															
Symptoms and drug within 30days					122 (15.3)															
Symptoms in infants	27 (5.5)								6 (6.8)				172 (12.4)				30 (7.3)			
Symptoms in infants and drug within 30days	1 (0.2)												8 (0.6)							
Test	96 (19.6)				38 (4.8)				38 (42.8)				209 (15.1)				32 (7.8)			
Positive laboratory results	19 (3.9)				0 (0.0)								3 (0.2)							
Composite algorithms																				
	N (IR)	N (IR) in left-hand component only	N (IR) In both components	N (IR) In right-hand component only	N (IR)	N (IR) in left-hand component only	N (IR) In both components	N (IR) In right-hand component only	N (IR)	N (IR) in left-hand component only	N (IR) In both components	N (IR) In right-hand component only	N (IR)	N (IR) in left-hand component only	N (IR) In both components	N (IR) In right-hand component only	N (IR)	N (IR) in left-hand component only	N (IR) In both components	N (IR) In right-hand component only

PC specific OR unspecified diagnosis	194 (39.6)	21 (4.3)	0 (0.0)	173 (35.4)	135 (17.0)	0 (0.0)	0 (0.0)	135 (17.0)	39 (43.9)	2 (2.2)	0 (0.0)	37 (41.7)	246 (17.7)	68 (4.9)	11 (0.8)	167 (12.0)	91 (22.1)	14 (3.4)	7 (1.7)	70 (17.0)
Inpatient specific OR unspecified diagnosis	52 (10.6)	25 (5.1)	1 (0.2)	26 (5.3)																
PC OR inpatient diagnosis	220 (45.0)	168 (34.3)	26 (5.3)	26 (5.3)																
PC diagnosis OR test	271 (55.4)	77 (15.8)	19 (3.9)	175 (35.8)	168 (21.1)	33 (4.1)	5 (0.6)	130 (16.3)	69 (77.7)	30 (33.8)	8 (9.0)	31 (34.9)	426 (30.7)	181 (13.0)	29 (2.1)	217 (15.6)	115 (28.0)	24 (5.8)	8 (1.9)	83 (20.2)
Positive lab results OR PC diagnosis	197 (40.3)	3 (0.6)	16 (3.3)	178 (36.4)	135 (17.0)	0 (0.0)	0 (0.0)	135 (17.0)					247 (17.8)	1 (0.1)	2 (0.1)	244 (17.6)				
PC diagnosis OR symptoms and drugs					255 (32.0)	133 (16.7)	2 (0.3)	120 (15.1)												
Any diagnosis OR positive lab results	223 (45.6)	204 (41.7)	16 (3.3)	3 (0.6)																

Table 4. Study results.. *Number of person-years entering the study in each database. Incidence rates of pertussis per 100,000 children aged 0-14, with 95% confidence interval (CI), from the TESSy surveillance system in the corresponding country. For each component algorithm the incidence rate per 100,000 is represented. In composite algorithms, incidence rates are stratified per type of case: cases detected only by the left-hand component (indicated in the label before the keyword 'OR'), cases detected by both components, and cases detected by the right-hand component (indicated in the label after the keyword 'OR'). Years 2012 and 2014*

In BIFAP, the 'symptoms and drugs within 30 days' component identified 122 cases with an IR of 15.3 per 100,000 PYs. When this component was combined with PC diagnoses, the IR increased to 32.0, which was higher than the reference IR which was 21.2. Almost none of the children aged 0 or 1 year old in 'symptoms in infants' in any database had a corresponding prescription or dispensing of macrolides in the 'symptoms in infants and drugs within 30 days' component.

The 'test' component was available in all databases and had a relatively high IR (from 4.8 in BIFAP to 42.8 in PEDIANET). 'Positive laboratory results' were only available in SIDIAP and THIN, with only 19 and 3 cases, respectively. In SIDIAP, 3 of the 19 cases were not captured by a diagnosis in either primary care or hospital settings.

In **Supplementary Figure 1** and **Supplementary Table 2**, the analysis was repeated for infants (children aged 0 or 1). The IRs in this subpopulation were around three times higher than the IRs in the overall study population. The findings confirmed the relationship between components observed in the general study population, with the exection of 'PC OR inpatient diagnosis' in SIDIAP (n=98), where 25.5% (n=25) were not retrieved from the PC setting, vs 11.8% in the overall study population.

In SIDIAP we explored two scenarios, corresponding to different assumptions for PPV of 'PC diagnosis, unspecified' and of 'inpatient diagnosis, unspecified': in the first, this was 70%, in the second 50%. As a consequence, in the first scenario 'PC specific OR unspecified diagnosis' had a PPV of 72% (or, in the second: 54%) and a sensitivity of 85% (or, in the second: 83%). Based on this estimate, the adjusted IR of BorPer in the SIDIAP study population was 35.5 per 100,000 PY (or, in the second scenario: 25.9) vs the TESSy surveillance system IR 21.2.

4. Discussion

We assessed several algorithms as potential strategies to detect cases of pertussis and thus estimate the IR in five European healthcare databases. The IRs estimated by these algorithms were heterogeneous within and between databases. However, there was at least one IR estimated by the algorithms in each database that was comparable with the reference value from the TESSy surveillance system, although some false positives were probably included. Based on a few assumptions, that may have overestimated sensitivity, it was estimated that the PPV and sensitivity of the algorithm detecting PC diagnoses in SIDIAP ranged from 54% to 72% and from 83% to 85%, respectively, and that the IRs of *Bordetella pertussis* in the corresponding population ranged from 25.9 to 35.5 per 100,000 person-years, against the TESSy surveillance system estimate of 21.2.

4.1. General comments

Three components were expected to have a high PPV: PC and inpatient specific diagnoses, and positive laboratory results. Two were expected to have lower PPV (PC and inpatient unspecified diagnoses). One was expected to be sensitive (prescription of a laboratory test), two were very unspecific (symptoms and symptoms in infants) and were planned to be used only in combination with the last component (prescription or use of macrolides) in a 30-days window of time.

In all the databases, at least one composite algorithm estimated a number of cases that was compatible with the number expected from the TESSy surveillance system, but this was not with the combination of the components which was expected to have a high PPV (specific diagnoses and laboratory tests) in any of the databases. One possible explanation could be that it takes several days to confirm the diagnosis of pertussis after the disease is suspected, and there may be no opportunity for the specific diagnosis to be recorded if the patient does not return to the healthcare facility. Another possible explanation may be that the medical personnel may not see the need to update the record for the purposes of clinical care. This attitude may be influenced by the level of awareness of possible reuse of electronic records for research purposes. These potential explanations may have varying levels of impact in the different databases. For example, in some

databases we observed that among the cases detected by a diagnostic component (unspecified or specific), the specific diagnosis was more frequent, indicating that some clinicians might have been more aware about potential research uses of the databases and therefore entered specific diagnoses rather than free text, which was common for unspecified diagnoses.

Based on the results of this study, in all the databases it is now possible to design sensitivity analysis using a more specific (but less sensitive) definition of pertussis. In case of heterogeneity in the results of a study on pertussis, designing such sensitivity analyses should be considered as a valid option. On the other hand, in all the databases there is now a possible choice among different sensitive algorithms: we explored several of them, among which 'unspecified diagnoses' (the most conservative) and 'test' (the least conservative). Even though these algorithms are likely to have lower PPVs, they may still be useful for sensitivity analyses, especially if there are reasons to think that a specific algorithm could be affected by differential misclassification. For example, pertussis may be more readily suspected and tested for in unvaccinated children, and therefore would be recorded in a more accurate manner.

We developed a component for infants that we though would be sensitive and, although it was likely to have a low PPV, it was less prone to differential misclassification, because it captured symptoms that physicians may not think of as being related with pertussis. However this component proved to be unusable; in reality, when we added a secondary component for concurrent macrolide use there were very few cases that would have been expected to be found in infants with an infection. In contrast, we developed a component specifically for the symptom 'pertussis-like cough' *(tos pertusoide* in Spanish language) that was apparently specific for pertussis cases that were only found in the BIFAP database. Not only did the majority of cases have a concurrent record of prescription of macrolides, but a manual review of a sample of 100 records including physician free text comments, found 2 cases of unspecified pertussis and 2 cases of suspected pertussis. Therefore, this component may be considered for sensitivity analysis.

4.2. Compatibility with TESSy and seroprevalence surveys

In this study we were able to compare the IRs estimated for paediatric cohorts in five databases using the various algorithms with the national IR estimates from ECDC's TESSy surveillance database. The cases captured by the two types of systems were expected to be slightly different, for various reasons. First, TESSy provides estimates at the national level using census denominators, while three of the databases participating in this study had a regional/multiregional scope (SIDIAP, BIFAP and PEDIANET) and two were based on a representative sample of the national population (THIN, RCGP RSC). Therefore it is possible that some clusters of the infectious disease might be under or over-represented in these database. Second, we collected only confirmed cases from TESSy, while some true cases captured by a PC database with a sensitive algorithm may never be confirmed (under ascertainment), or may never be notified (underreporting) [19, 20]. Thus the databases may be a complementary source of true cases which are not notified, while adding potentially false positive cases. Finally, the TESSy data for pertussis may also be affected by under ascertainment and underreporting.

The IR found for PEDIANET, which was much higher than the IR estimate from TESSy for Italy (43.9 vs 5.4), may be explained by a combination of both phenomena discussed above. PEDIANET collects data from PC physicians working in the Italian region Veneto, in the North East of the country. The Regional Office for Infectious Diseases of the Veneto Region provided an estimated IR of 10.0 to the data custodians of PEDIANET. This shows that the region had a higher pertussis notification rate than at the national level for 2012 and 2014, although almost all the diagnoses in PEDIANET were unspecified. However, the regional estimate could be underestimated because of under ascertainment. Finally, as in the other databases, many cases in PEDIANET could be false positives. In general, if estimates of the PPV of the diagnoses are available, the estimated IR from databases can provide a quantitative estimate of under ascertainment and under notification in TESSy. Vice

versa, if under notification to TESSy is known to be small, estimates of the PPV for the algorithm can be obtained.

Results from seroprevalence surveys have provided estimates for the incidence of *Bordetella pertussis* infection [21-23]. These have provided prevalence estimates beyond those of the surveillance systems, partly as they also capture asymptomatic or mildly symptomatic infections. On the contrary, in this study, we observed that estimates of incidence obtained from databases are roughly comparable with those of TESSy.

4.3. Scope of the component strategy

The scope of this component strategy goes beyond ADVANCE and has the potential of being a comprehensive tool to address heterogeneity and disease misclassification in databases, particularly in multi-database pharmacoepidemiology studies, when the characteristics of the databases affect the operational definition of the outcomes and benchmarking.

Inspection of components can provide knowledge that can inform the interpretation of the heterogeneity of the study results. The component strategy can support quantitative bias analysis. In this study, we first developed a set of components with increasing sensitivity and decreasing PPVs. We explored several scenarios for possible PPVs of the components, but in many European databases, estimating directly the PPV of simple algorithms such as components is feasible in a relatively timely and inexpensive way [24-26]. If this is possible, then a consequence of our formulas in Table 3 is that the only value needed to obtain a complete estimate of validity is the sensitivity of the composition of the algorithms, as the rest can be analytically derived. In many cases, sensitivity of the composition could be argued to be very high. In the case of pertussis, we can speculate for instance that cases that were missed from SIDIAP were either seen in a hospital outside of the network that transmits their data to the database, or were very mild and did not require medical attention. The percentage of cases with those characteristics may be estimated from external sources. If estimating this quantity is not possible, then the formulas of Table 3 can still be applied, and they can provide an upper limit for the sensitivity of all the components, that is, the maximum possible sensitivity: to obtain this, it is sufficient to make an assumption on the sensitivity of the *composite* algorithm.

If the validity of the variables that enter the analysis can be convincingly proven to be high, this analysis provides evidence that the study results are robust to misclassification. If not, comparing the distribution of components across exposure strata can indicate if differential misclassification is to be suspected. If it is suspected, it can be an important source of bias, as shown by the simulations we report here, as well as in other studies [5, 6]. Components with different validity can, thus, be used to design sensitivity analyses of the study results, applying repeated adjustments for validity to check if the result is robust. If both the PPV and sensitivity are suspected to be non-differential, then the estimate may be unbiased, but the confidence intervals of the estimate need to be adjusted for validity [8]. In future work, the estimates provided by the component strategy could be validated against actual validation studies. Moreover, the components could be analysed using latent class modelling, which enables to estimate the validity conditional on various covariates, e.g., age [27].

4.4. Strengths and limitations

In this study, we used standardised component algorithms as a transparent way of documenting the data extraction process across multiple databases. At the same time, we could also perform a qualitative evaluation of the expected validity of each component Bordetella pertussis, based on its specified semantics and setting. Quantitative scenarios for the validity of each component can also be made using the same approach. We showed that estimates of the validity of various composite algorithms can then be derived in a purely algebraic manner. We could use the incidence estimates based on data from the TESSy surveillance system, which is where European Union member states are required to report pertussis cases, as a reference value, although we cannot exclude the possibility that they may also be subject to under ascertainment and underreporting.

The estimates of sensitivity that we obtained for SIDIAP cannot be generalised to the other PC databases. The sensitivity of the PC databases depends on how often a person with the disease symptoms would seek attention in a PC practice. Although in all the databases, the PC physicians have a gatekeeper role, emergency care can be sought without PC referral, and PC practices may not be accessible at night or weekends. Referrals from other settings may be recorded in the PC practice, but no automatic mechanism is in place. In the absence of a database-specific estimate, estimates from another database are a realistic alternative to assuming that sensitivity is 100%.

5. Conclusions

This study demonstrated the value of quantifying the impact of different event-finding algorithms across databases and the possibility of benchmarking with disease surveillance data as well as assessing validity estimates when data from different settings can be linked. The validity parameters could be used to correct disease IR estimates from healthcare databases.

Acknowledgements

The authors medical writing and editorial assistance from Margaret Haugh (MediCom Consult, Villeurbanne, France).

Disclaimer

The results described in this publication are from the proof of concept studies conducted as part of the IMI ADVANCE project with the aim of testing the methodological aspects of the design, conduct and reporting of studies for vaccine benefit-risk monitoring activities. The results presented relate solely to the methodological testing and are not intended to inform regulatory or clinical decisions on the benefits and risks of the exposures under investigation. This warning should accompany any use of the results from these studies and they should be used accordingly. The views expressed in this article are the personal views of the authors and should not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organisations with which the authors are affiliated.

Funding source

The Innovative Medicines Initiative Joint Undertaking funded this project under ADVANCE grant agreement n° 115557, resources of which were composed of a financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and in kind contributions from EFPIA member companies.

Declaration of potential conflicts of interest

Rosa Gini, Caitlin Dodd, Kaatje Bollaerts, Claudia Bartolini, Giuseppe Roberto, Consuelo Huerta-Alvarez, Elisa Martín-Merino, Talita Duarte-Salles, Gino Picelli, Lara Tramontan, Giorgia Danieli, Ana Correa, Chris McGee, Benedikt Becker, Charlotte Switzer, Jorgen Bauwens, Nicoline van der Maas, Gianfranco Spiteri, Emmanouela Sdona declared no conflicts of interest. Sonja Gandhi-Banga declared that she works for Sanofi Pasteur and holds company shares. Daniel Weibel declared that he has received personal fees from GSK for work unrelated to the submitted work. Miriam Sturkenboom declared that she has received grants from Novartis, CDC and Bill & Melinda Gates Foundation, for work unrelated to the submitted work.

References

[1] Sturkenboom M, van der Aa L, Bollaerts K, Emborg HD, Ferreira G, Gino R, et al. The ADVANCE distributed network system for evidence generation on vaccines coverage, benefits and risks based on electronic health care data. Vaccine. 2018;Paper 2 in supplement.

[2] Sturkenboom M, Weibel D, van der Aa L, Braeye T, Gheorge M, Becker B, et al. ADVANCE database characterization and fit for purpose assessment for multi-country studies on the coverage, benefits and risks of vaccinations. Vaccine. 2018;Paper 3 in Supplement.

[3] Gini R, Schuemie M, Brown J, Ryan P, Vacchi E, Coppola M, et al. Data extraction and management in networks of observational health care databases for scientific research: a comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE strategies. EGEMS (Washington, DC). 2016;4:1189.

[4] Avillach P, Coloma PM, Gini R, Schuemie M, Mougin F, Dufour JC, et al. Harmonization process for the identification of medical events in eight European healthcare databases: the experience from the EU-ADR project. J Am Med Inform Assoc. 2013;20:184-92.

[5] Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. Curr Epidemiol Rep. 2014;1:175-85.

[6] Hofler M. The effect of misclassification on the estimation of association: a review. Int J Methods Psychiatr Res. 2005;14:92-101.

[7] De Smedt T, Merrall E, Macina D, Perez-Vilar S, Andrews N, Bollaerts K. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness. PloS One. 2018;13:e0199180.

[8] Brenner H, Gefeller O. Use of the positive predictive value to correct for disease misclassification in epidemiologic studies. Am J Epidemiol. 1993;138:1007-15.

[9] Roberto G, Leal I, Sattar N, Loomis AK, Avillach P, Egger P, et al. Identifying cases of type 2 diabetes in heterogeneous data sources: strategy from the EMIF Project. PloS One. 2016;11:e0160648.

[10] Barlow RS, Reynolds LE, Cieslak PR, Sullivan AD. Vaccinated children and adolescents with pertussis infections experience reduced illness severity and duration, Oregon, 2010-2012. Clin Infect Dis. 2014;58:1523-9.

[11] McNamara LA, Skoff T, Faulkner A, Miller L, Kudish K, Kenyon C, et al. Reduced severity of pertussis in persons with age-appropriate pertussis vaccination-United States, 2010-2012. Clin Infect Dis. 2017;65:811-8.

[12] European Parliament and of the Council. (Decision EU 2012) Commission implementing decision of 8 August 2012 amending Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision No 2119/98/EC of the European Parliament and of the Council. Annex to L 262. Official Journal of the European Union 27/9/2012. Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32012D0506&qid=1428573336660&from=EN#page=22. Accessed on: 9 November 2018.

[13] European Centre for Disease Prevention and Control. The European Surveillance System (TESSy). Available at: https://ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy. Accessed on: 9 November 2018.

[14] Becker BFH, Avillach P, Romio S, van Mulligen EM, Weibel D, Sturkenboom M, et al. CodeMapper: semiautomatic coding of case definitions. A contribution from the ADVANCE project. Pharmacoepidemiol Drug Saf. 2017;26(8):998-1005.

[15] Bellettini CV, de Oliveira AW, Tusset C, Baethgen LF, Amantea SL, Motta F, et al. [Clinical, laboratorial and radiographic predictors of Bordetella pertussis infection]. Revista paulista de pediatria : orgao oficial da Sociedade de Pediatria de Sao Paulo. 2014;32:292-8.

[16] Hurtado-Mingo A, Mayoral-Cortes JM, Falcon-Neyra D, Merino-Diaz L, Sanchez-Aguera M, Obando I. [Clinical and epidemiological features of pertussis among hospitalized infants in Seville during 2007-2011]. Enferm Infecc Microbiol Clin. 2013;31:437-41.

[17] Eurostat. Population data. Available at: https://ec.europa.eu/eurostat/web/population-demography-migration-projections/population-data/database. Accessed on: 12 November 2018.

[18] Ulm K. A simple method to calculate the confidence interval of a standardized mortality ratio (SMR). Am J Epidemiol. 1990;131:373-5.

[19] McDonald SA, Teunis P, van der Maas N, de Greeff S, de Melker H, Kretzschmar ME. An evidence synthesis approach to estimating the incidence of symptomatic pertussis infection in the Netherlands, 2005-2011. BMC Infect Dis. 2015;15:588.

[20] Schielke A, Takla A, von Kries R, Wichmann O, Hellenbrand W. Marked underreporting of pertussis requiring hospitalization in infants as estimated by capture-recapture methodology, Germany, 2013-2015. Pediatr Infect Dis J. 2018;37:119-25.

[21] Barkoff AM, Grondahl-Yli-Hannuksela K, He Q. Seroprevalence studies of pertussis: what have we learned from different immunized populations. Pathog Dis. 2015;73.

[22] de Greeff SC, de Melker HE, van Gageldonk PG, Schellekens JF, van der Klis FR, Mollema L, et al. Seroprevalence of pertussis in The Netherlands: evidence for increased circulation of Bordetella pertussis. PloS One. 2010;5:e14183.

[23] de Melker HE, Versteegh FG, Schellekens JF, Teunis PF, Kretzschmar M. The incidence of Bordetella pertussis infections estimated in the population from a combination of serological surveys. J Infect. 2006;53:106-13.

[24] Coloma PM, Valkhoff VE, Mazzaglia G, Nielsson MS, Pedersen L, Molokhia M, et al. Identification of acute myocardial infarction from electronic healthcare records using different disease coding systems: a validation study in three European countries. BMJ Open. 2013;3.

[25] Valkhoff VE, Coloma PM, Masclee GM, Gini R, Innocenti F, Lapi F, et al. Validation study in four health-care databases: upper gastrointestinal bleeding misclassification affects precision but not magnitude of drug-related upper gastrointestinal bleeding risk. J Clin Epidemiol. 2014;67:921-31.

[26] Gini R, Schuemie MJ, Mazzaglia G, Lapi F, Francesconi P, Pasqua A, et al. Automatic identification of type 2 diabetes, hypertension, ischaemic heart disease, heart failure and their levels of severity from Italian General Practitioners' electronic medical records: a validation study. BMJ Open. 2016;6:e012413.

[27] Hui SL, Walter SD. Estimating the error rates of diagnostic tests. Biometrics. 1980;36:167-71.

Supplementary material

Supplementary File 1

Incidence rate as a proportion and sensitivity analysis

The incidence of an algorithm in the age band 0-14 is the number if cases $\#C_{0-14}$ in persons who are aged 0-14 at algorithm date, divided by the number of person years PY_{0-14} aged 0-14 in the study population. In this study this number we considered the persons who were 0-14 at the end of each follow-up year (that is, those who were 0-13 at index date, plus those who were born during the follow-up year), and counted the cases recorded during the year in this population. This way four numbers were changed: a whole year of follow-up was computed for persons who were in fact *born* during the year, and therefore contributed on average a *half* person-year; second, the person-time of those who were 15 at the end of follow-up, but had been 14 for, on average, half a year, was excluded from the computation of follow-up; third, cases in persons who were 15 at the end of follow-up year were assumed to were assumed to have all taken place after birthday; last, the person-time of those who were cases, and therefore not at risk anymore, was not discarded from the denominator. The assumptions underlying the approximation are: first, the population was quite stable, therefore the number of persons aged 0 and aged 15 at the end of each follow-up year, respectively N_{0_e} and N_{15_e}, were similar, so the balance including half a year for the former and excluding half a year for the latter was approximately even. As a consequence the number of persons aged 0-14 at the end of follow-up, N_{0-14_e}, was similar to the number of person-years observed. Second, the number of cases in persons who turned 15 during the year, that should have been included while they are still 14, and that were nevertheless excluded, was very low compared to the number of cases in younger subjects, so the number of cases $\#C_{0-14}$ was very similar to the number of cases $\#C_{0-14_e}$ observed in persons aged 0-14 at the end of the follow-up time. Last, the number of cases was so small that including person time not at risk did not affect the estimate of incidence rate. In formulas

$$IR_{0-14} = \frac{\#C_{0-14}}{PY_{0-14}} \sim \frac{\#C_{0-14}}{N_{1-14_e} + .5 \times N_{0_e} + .5 \times N_{15_e}} \sim \frac{\#C_{0-14_e}}{N_{0-14_e}}$$

To quantify the impact of those assumptions we performed a sensitivity analysis. We measured the difference between number of persons who were born and the number of persons who turned 15 during the follow-up years as a percentage of the study population and we compared the number of cases in persons aged 0-14, which actually entered the analysis, with half of the number of cases in persons aged 15 at the end of each follow-up year (the average number of additional cases that would have been included, had they been correctly classified with their calendar age).

The difference between the number of children born during each year and the number of children who were 15 at the end of each year was less than 2% of the overall population in each database, except in PEDIANET, where it was 5%, The assumption on stability of population may have had a overestimated the IR in PEDIANET, although by a negligible amount. In BIFAP and PEDIANET there were no cases in children aged 15 at the end of follow-up. In the other databases, adding half of such cases would have increased the number of cases by less than 5%, except one the case of THIN, where they would have added 7% to the PC specific diagnosis and one case out of four to the positive laboratory value component. This would not have affected our analysis.

Supplementary Table 1. Concepts selected for the component algorithms, projected to free text strings and codes in the coding systems in use in the participating databases. The concepts belonging to the data domain of 'laboratory test' and of 'Results from laboratory test' were all projected to local coding systems

Label of the set of concepts	Description of each included concept	Concept Unique Identifier	Languages		Coding systems					
			Spanish	Italian	ICD10	ICD9	ICPC or BIFAP ICPC-based	Read-CTv3	Read-v2	ATC
(Bordetella pertussis)	Bordetella pertussis	C0043167	Bordetella pertussis	Bordetella pertosse	A37.0	033.0	R71001		A330.	
	Whooping cough due to Bordetella pertussis without pneumonia	C2887068								
	Whooping cough due to Bordetella pertussis with pneumonia	C2887069								
(Pertussis unspecified)	Whooping cough due to unspecified organism	C0043168	'ferina' No como *parapertussis* or *pertusis*	'Pertoss' 'Pertuss'	A37 other than A37.0, A37.1, A37.8	033 other than 033.0, 033.1, 033.8	R71 other than R71001 484.3-Neumonia en tosferina	A33z., Ayu3A, X76Hf	A33.. other than A330., Ayu3A	
	Bordetella Infections	C0006015			-	-	-	X70I8, XE0Qw		
	Whooping cough-like syndrome	C0343485						XM00D	XM00D	
	Notification of whooping cough							65VA.	65VA.	
	Pneumonia in pertussis	C0155865				484.3		H243.	H243.	
(Symptoms compatible with pertussis)	Concept of 'tos pertusoide' in Spanish general practice		'Tos pertusoide' or 'Tos pertuss' in diagnosis description			786.2.3 Tos pertusoide	R05.10 Tos pertussoide			

(Symtoms in infants)	Apnea	C0003578				786.03	R04012	X76Gw	R0604	
	Cyanosis	C0010520			R23.0	782.5	K29005, S08011	XM07N	R025.	
	Post-tussive vomiting	C1740793								
	Paroxysms of coughing	C0231911								
(Macrolides)	Macrolides									J01FA*
(Bordetella pertussis test)	polymerase chain reaction test									
	culture or serology									
	isolation of Bordetella pertussis from a clinical specimen									
(Positive result from a Bordetella pertussis test)	positive polymerase chain reaction test									
	positive culture or serology									
	positive isolation of Bordetella pertussis from a clinical specimen									

Supplementary Table 2. Percentage of children aged 0 or 1 in the population aged 0-14, from the EUROSTAT source and from the study population.

% of infants in the population 0-14	SIDIAP (ES)	BIFAP (ES)	PEDIANET (IT)	THIN (UK)	RCGP (UK)
EUROSTAT population 2012 and 2014	13.5	13.5	12.9	14.3	14.3
Study population	12.1	16.1	13.0	14.8	15.1

Formulas linking true proportion of cases of a disease in the study population, and sensitivity, positive predictive value, and observed proportion of one or two component algorithms for that disease

Rosa Gini, Caitlin Dodd

1 Definitions

In this document, we consider a study population composed if N subjects, and a disease with C cases. We denote by π the proportion C/N.

When an algorithm is available on the study population detecting the disease, the following quantities are defined

TP	number of true positives detected by the algorithm
FP	number of false positives detected by the algorithm
FN	number of false negatives detected by the algorithm
P	observed frequency of the algorithm in the study population
PPV	positive predictive value of the algorithm
SE	sensitivity of the algorithm

By definition, the following formulas hold

$$C = \mathbf{TP} + \mathbf{FN}$$

$$\pi = \frac{C}{N} = \frac{\mathbf{TP} + \mathbf{FN}}{N}$$

$$P = \frac{\mathbf{TP} + \mathrm{FP}}{N}$$

$$SE = \frac{\mathbf{TP}}{C} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}}$$

$$PPV = \frac{\mathbf{TP}}{\mathbf{TP} + \mathrm{FP}}$$

It is important to note that P can be observed from the data.

2 Formulas for one algorithm

From the definitions, the following sequence of equations hold

$$
\begin{aligned}
\pi &= \frac{\mathbf{TP}+\mathbf{FN}}{N} \\
&= \frac{1}{N}(\mathbf{TP}+\mathbf{FN}) \times \frac{\mathbf{TP}+\mathrm{FP}}{\mathbf{TP}} \times \frac{\mathbf{TP}}{\mathbf{TP}+\mathrm{FP}} \\
&= \frac{1}{N}(\mathbf{TP}+\mathrm{FP}) \times \frac{\mathbf{TP}+\mathbf{FN}}{\mathbf{TP}} \times \frac{\mathbf{TP}}{\mathbf{TP}+\mathrm{FP}} \\
&= P \times \frac{1}{SE} \times PPV
\end{aligned}
$$

This proves the following formulas

$$\pi = \frac{P \times PPV}{SE} \tag{1}$$

$$SE = \frac{P \times PPV}{\pi} \tag{2}$$

$$PPV = \frac{SE \times \pi}{P} \tag{3}$$

3 Formulas for two algorithms

Let A and B be two algorithms. It is important to remark that N, C and π do not depend on A or B, while sensitivity and PPV do. Moreover, P_A, P_B, $P_{A\,\mathbf{OR}\,B}$ and $P_{A\,\mathbf{AND}\,B}$ can be observed from the data.

Remark the persons who are true positives for A **OR** B are those who are true positives for either of the two. To count the true positives for A **OR** B we need to add the true positives from the two algoritms, but, in order not to count the same persons twice, we need to subtract from the total those who are positive for both A and B, that is, we need to subtract those who are true positives for A **AND** B. In formulas this is the following

$$\mathbf{TP}_{A\,\mathbf{OR}\,B} = \mathbf{TP}_A + \mathbf{TP}_B - \mathbf{TP}_{A\,\mathbf{AND}\,B}$$

By dividing the previous formulas by C, and remembering the definition of SE from Section 1,

$$
\begin{aligned}
SE_{A\,\mathbf{OR}\,B} &= \frac{\mathbf{TP}_{A\,\mathbf{OR}\,B}}{C} \\
&= \frac{\mathbf{TP}_A + \mathbf{TP}_B - \mathbf{TP}_{A\,\mathbf{AND}\,B}}{C} \\
&= \frac{\mathbf{TP}_A}{C} + \frac{\mathbf{TP}_B}{C} - \frac{\mathbf{TP}_{A\,\mathbf{AND}\,B}}{C}
\end{aligned}
$$

hence the following formula hold

$$SE_{A\,\mathbf{OR}\,B} = SE_A + SE_B - SE_{A\,\mathbf{AND}\,B} \tag{4}$$

Now, applying formula 2 to the three elements on the right of the previous equation generated the following formula

$$SE_{A\,\mathbf{OR}\,B} = \frac{P_A \times PPV_A}{\pi} + \frac{P_B \times PPV_B}{\pi} - \frac{P_{A\,\mathbf{AND}\,B} \times PPV_{A\,\mathbf{AND}\,B}}{\pi} \tag{5}$$

which in turn, from multiplying both sides for $\frac{\pi}{SE_{A\,\mathbf{OR}\,B}}$, implies this formula

$$\pi = \frac{P_A \times PPV_A + P_B \times PPV_B - P_{A \text{ AND } B} \times PPV_{A \text{ AND } B}}{SE_{A \text{ OR} B}} \tag{6}$$

Finally, formula 3 can be applyed to the algorithm A **OR** B using the previous two formulas. Since $SE_{A \text{ OR} B}$ cancels out, this becomes

$$PPV_{A \text{ OR } B} = \frac{P_A \times PPV_A + P_B \times PPV_B - P_{A \text{ AND } B} \times PPV_{A \text{ AND } B}}{P_{A \text{ OR } B}} \tag{7}$$

6.3 ESTIMATING THE INCIDENCE OF ADVERSE EVENTS FOLLOWING VACCINATION IN OBSERVATIONAL DATABASES WITH INCOMPLETE EXPOSURE DATA

Caitlin Dodd

Kaatje Bollaerts

Daniel Weibel

Maria de Ridder

Olivia Mahaux

Francois Haguinet

Tom de Smedt

Simon de Lusignan

Chris McGee

Talita Duarte-Salles

Hanne-Dorthe Emborg

Consuelo Huerta

Elisa Martín-Merino

Gino Picelli

Klara Berencsi

Giorgia Danieli

Miriam Sturkenboom

Submitted

Abstract

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid vaccine benefit-risk monitoring using existing European healthcare databases. Incidence rate (IR) estimates of vaccination-associated adverse events that are needed to model vaccination risks can be calculated from existing healthcare databases when vaccination (exposure) data are available. We assessed different methods to derive IRs when data are missing in one database, using estimated IRs from other databases for febrile seizures, fever and persistent crying. IRs were estimated for children aged 0-5 years in outcome-specific risk and non-risk periods following the first dose of acellular pertussis (aP) vaccination in four primary care databases and one hospital database. We compared derived and observed IRs in each database using three methods: 1) multiplication of non-risk period IR for database *i* by IR ratio (IRR) obtained from meta-analysis of IRRs estimated using the self-controlled case-series method, from databases other than *i*; 2) same method as 1, but multiplying with background IR; and 3) meta-analyses of observed IRs from databases other than *i*. IRs for febrile seizures were lower in primary care databases than the hospital database. The derived IR for febrile seizures using data from primary care databases was lower than that observed in the hospital database, and using data from the hospital database gave a higher derived IR than that observed in the primary care database. For fever and persistent crying the opposite was observed. We demonstrated that missing IRs for a post-vaccination period can be derived but that the type of database and the method of event data capture can have an impact on potential bias. We recommend IRs are derived using data from similar database types (hospital or primary care) with caution as even this can give heterogeneous results.

1. Introduction

The Accelerated Development of VAccine beNefit-risk Collaboration in Europe (ADVANCE) is a public-private collaboration aiming to develop and test a system for rapid benefit-risk (B/R) monitoring of vaccines using existing healthcare databases in Europe (see Appendix for consortium members). A series of proof of concept (POC) studies were designed to assess the processes and system proposed for generating the required data to generate evidence on coverage, risks and benefits of vaccines as well as benefit-risk analyses.

Modelling is one method that is widely used to analyse vaccine benefit-risk, to understand the impacts of diseases, interventions, and environmental exposures deterministically or in simulated populations [1]. Valid estimates of incidence rates (IRs) for vaccine-preventable disease and adverse event following immunisation, and vaccination coverage are needed to model the benefit-risk of vaccination. These input parameters may be obtained from the literature or by using data from available healthcare databases. However, when using healthcare databases, their heterogeneity and potentially important missing information on vaccinations need to be taken into consideration [2].

The first vaccines developed against *Bordetella pertussis* contained whole killed organisms [3]. Due to the reactogenicity of this vaccine, Between 2004 and 2015 several countries switched from whole-cell pertussis (wP) to acellular pertussis (aP) vaccines for infants and children due to the reactogenicity of the wP vaccine [4]. In the ADVANCE POC studies the benefits and risks of wP and aP vaccines in children were compared as an example. For this, IRs of known benefits and adverse events in outcome-specific risk periods following each dose of wP and aP vaccine were required. Since we used existing healthcare databases that collected data for purposes other than for research, we were faced with the problem of comparing the effects of exposure which occurred in distinct time periods, often with missing exposure data for the period before the switch from wP to aP. To compare the B/R for the wP and aP vaccines, we attempted to estimate IRs for various outcomes following wP vaccination in some databases that were established too recently to contain wP exposure data. In this paper we compared different methods for deriving IRs in the risk period following vaccination. To test these methods we limited the analysis to aP exposure, assuming that the aP exposure data were missing, which allowed us to compare the observed and derived IRs.

2. Methods

2.1 Data sources and population

This study was conducted with data generated for the ADVANCE proof of concept risk study that included seven population-based healthcare databases from Denmark, Spain, UK and Italy (**Table 1**) [5, 6]. Two databases were excluded in this methods study: AUH because it is a subset of the national SSI database in Denmark, and PEDIANET from Italy, in which vaccination data was linked only for the 2006 and 2007 birth cohorts. We excluded data from the SSI database, which is a hospital database, in sensitivity analyses to study the impact of hospital data on the results.

Table 1: Databases providing data for the ADVANCE POC safety study [6]

Country	Database	Geographic coverage	Type of data	Years with available data	Switch from wP to aP	Size (N persons)	Children exposed to aP	Primary care diagnoses	Hospital discharge diagnoses
Denmark	SSI	National	National claims data record linkage	2000 - 2014	1997	7.5 million	980,843	No	Yes (ICD-10)
Spain	BIFAP	Multi regional sample	GP medical records	2002 - 2013	2000-2004	4.8 million	320,638	Yes (ICPC-based codes + free text)	Limited to free text comments recorded by the GP
Spain	SIDIAP	Regional (Cataluña)	GP medical records & partial linkage to hospital	2005-2014	2000-2004	5.8 million	570,225	Yes (ICD-10)	Yes (ICD-9)
United Kingdom	RCGP RSC	National sample	GP medical records	2003 - 2014	2004	2.0 million	152,784	Yes (READ)	Yes (READ)
United Kingdom	THIN	National sample	GP medical records	1996-2013	2004	8.3 million	576,151	Yes (READ)	Yes (READ)

AUH = Aarhus University Hospital, SSI = Statens Serum Institute, BIFAP = Base de datos para la Investigación Famacoepidemiológica en Atención Primaria, SIDIAP = Information System for Research in Primary Care, RCGP RSC = Royal College of General Practitioners Research and Surveillance Centre, THIN = The Health Information Network, GP = General Practitioner, ICD = International Classification of Diseases

The study population comprised all children aged <6 years registered in any of the participating databases during the study period, who had received at least the first dose of aP vaccine. For the calculation of background rates, children were followed from start of the study period (1 January 1990), one month after their date of birth (to allow for pre-vaccination person time and to avoid pre-term related or birth-induced increase in IR), or date of valid data in the database, whichever occurred the latest. For the calculation of baseline rates and incidence rate ratios, children were followed from 31 days before their first dose of aP vaccine. All children were followed until the end of study period (31 December 2015), until they received their pertussis booster dose, transferring out of the database, death, reaching age 6 years, or end of data availability in the database, whichever occurred first. Children with missing date of birth or sex were excluded.

Data from each participating database was extracted locally and transformed into a common data model, comprising vaccination, event, and population files [7].

2.2 Outcomes

To test the methodology we selected three outcomes from the risk study that have different patterns of care: febrile convulsions, fever, and persistent crying. Febrile convulsions are rare and are usually considered to be serious clinical events requiring presentation to the emergency room. Fever is common

but does not often require hospitalisation. Persistent crying is non-specific and often lacks a specific diagnosis code even in primary care. Definitions, codes and methods for data extraction and harmonisation can be found in other papers in this supplement [6, 7].

2.3 Definition of exposure

Data on aP vaccination were obtained from the healthcare databases [6]. Although our study was driven by the need to estimate IRs during the wP risk period, we limited our methodological study to aP risk period since the IRs could be estimated in all participating databases, therefore we could compare the IRs derived using different methods with the estimated IRs.

Outcome-specific risk windows were defined as day 0 to 3 for febrile convulsions and fever and day 0 to 1 for persistent crying, with day 0 being the day of vaccination. Baseline periods were defined as 31 to 8 days before dose one and the interval from the last day of the risk window to 31 days after the dose. The week prior to vaccination was excluded from the baseline period to avoid the 'healthy vaccinee effect', i.e. vaccine avoidance by subjects experiencing an illness (**Figure 1**) [8]. The pertussis vaccination schedules were 3, 5 and 12 months, 2, 4 and 11 months and 2, 3 and 4 months for Denmark, Spain and UK, respectively.

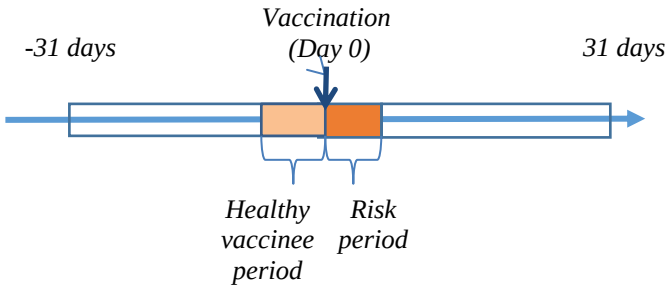


Figure 1: Schematic representation of the timeline of a typical observation period for dose 1.

2.4 Statistical methods

IRs were calculated by age in months and in the aP vaccination risk and non-risk period for each outcome. We conducted self-controlled case series (SCCS) analyses for each of the outcomes to obtain IRRs, comparing risk to non-risk periods for the first dose of aP vaccination [9]. The study population for each outcome-specific SCCS analysis included children who experienced the event at least once during their follow-up.

For each database *i* and event, a leave-one-out (L-O-O) random effects IRR was estimated using a meta-analysis of the IRRs from all databases other than database *i,* independent of the type of data source [10]. The result is referred to as L-O-O_IRR _ma. IRs in the risk period following vaccination were derived using three methods (**Box**). In the first method, we multiplied the baseline IR calculated in non-risk periods around aP vaccination in database *i* by the L-O-O_IRR_ma that excluded database *i* (IR_bl) (**Figure 2**). In the second method, we multiplied the background IR that was calculated in the month of age at the recommended first dose by the L-O-O_IRR_ma that excluded database *i* (IR-bg). In the third method, we derived a pooled risk period IR using a meta-analysis of the IRs for the observed risk period for all databases other than *i* (IR_ma). We then assessed the agreement between observed and derived risk period IRs.

Box: Methods used to derive incidence rates in risk period following vaccination

Derived from baseline IR (IR_bl):
The baseline IR in database *i* was multiplied by the L-O-O_IRR_ma calculated excluding database *i*. Confidence intervals (CIs) were obtained by calculating the standard error of the log IR_bl as follows:
The standard error of the sum of the log IR and the log L-O-O_IRR_ma was calculated as:

$$\sqrt{se(\log(IR))^2 + se(\log(L-O-O\_IRR\_ma))^2} \quad (1)$$

where

$$se(\log(IR)) = \frac{1}{\sqrt{N\_events}} \quad (2)$$

Derived from background IR (IR_bg):
The background IR of each outcome in the month of age when the first dose was recommended in the country of database, *i*, was multiplied by the L-O-O_IRR_ma calculated excluding database *i*. CIs were obtained by calculating the standard error of the log IR_bg as in equations (1) and (2).
Derived via meta-analysis of risk period IRs (IR_ma):
The log-transformed risk period IRs of all databases except database *i* were meta-analysed, providing IR_ma. CIs were obtained using the DerSimonian and Laird method for random effects meta-analysis [10].

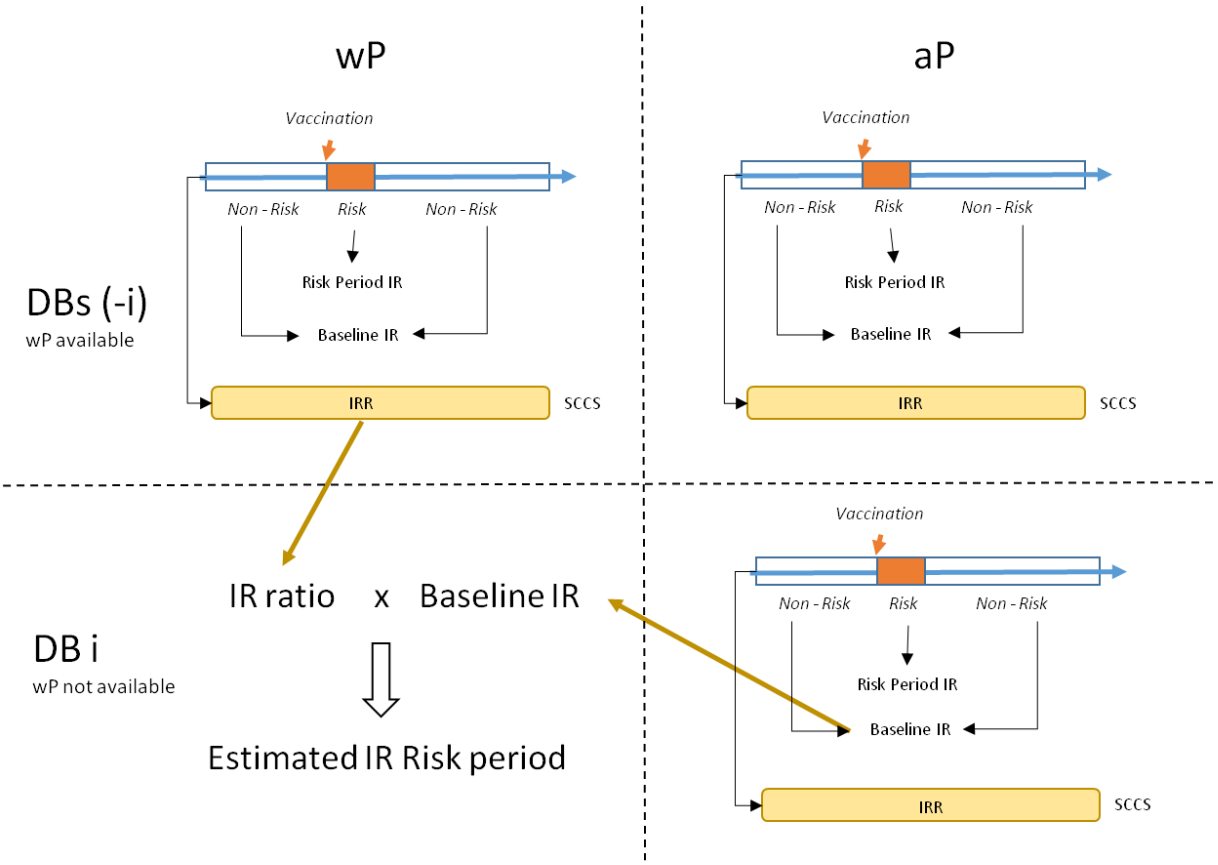


Figure 2: Approach for calculating risk window specific incidence rates in databases when wP exposure is missing or under the assumption of missing aP exposure [5]

3. Results

The study population comprised 2.6 million children aged <6 years who had received at least one dose of aP-containing vaccine. The database-specific sample sizes varied from 152,784 (RCGP RSC) to 980,843 (SSI) (**Table 1**). Over 400,000 children experienced at least one of the three events of interest during the study period.

The overall background IR (per 1,000 person-years) in this paediatric population for febrile convulsion ranged between 3 (BIFAP) to 11 (SSI; hospitalization). The age-stratified IRs peaked between 1 and 2 years of age in all databases (**Figure 3**). For fever, the overall IR (per 1,000 person-years) varied between the databases from 8 (SSI) to 184 (BIFAP). The age-stratified IRs for fever were high up to 18 months of age in most of the databases (**Figure 3**).

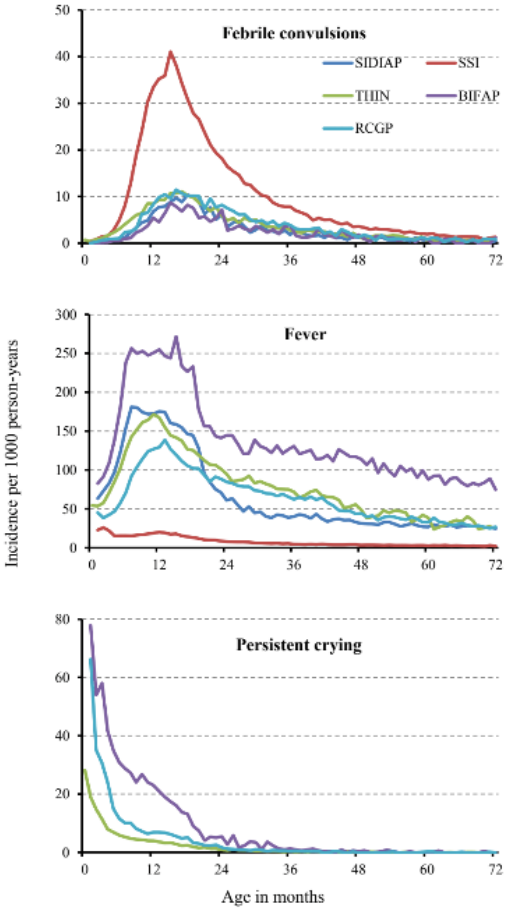


Figure 3: Background incidence of events of interest per 1,000 person years by age in months and database (NB: the y-axes are not the same scale)

SSI = Statens Serum Institute, BIFAP = Base de datos para la Investigación Famacoepidemiológica en Atención Primaria, SIDIAP = Information System for Research in Primary Care, RCGP RSC = Royal College of General Practitioners Research and Surveillance Centre, THIN = The Health Information Network

The overall IRs (per 1,000 person-years) of persistent crying ranged from 2 (THIN) to 22 (BIFAP). The age-stratified IRs peaked in the first months of life and then declined rapidly (**Figure 3**). No data for persistent crying were available in the SIDIAP and SSI databases since there are no specific ICD-9 or ICD-10 codes for this event. The event was identified using BIFAP-specific-ICPC or ICD-9 codes as well as free-text in the BIFAP database. .

IRRs for adverse events following vaccination which compared the IRs in risk periods after aP vaccination with those at baseline, as estimated via SCCS analyses, varied between databases. For febrile convulsions, no significant association after dose one of aP vaccine was seen in the BIFAP and RCGP RSC databases, while the risk was significantly lower in the SSI and THIN databases. L-O-O IRR_ma estimates were closer to 1 than those estimated in the SCCS in all databases. Statistically significant protective effects observed in the SSI and THIN databases were no longer present in the L-O-O_IRR_ma estimates. When the estimates from the SSI database were excluded, the L-O-

O_IRR_ma estimates increased slightly closer 1 due to removal of the significantly protective IRR in the SSI database (**Table 2**).

Table 2: Self-controlled case series (SCCS) and leave-one-out (L-O-O) incidence rate ratios (IRRs) following dose one of acellular pertussis vaccine

Event	Database	SCCS IRR (95% CI)	L-O-O IRR (95% CI)	L-O-O IRR without SSI (95% CI)
Febrile convulsions	SSI	0.24 (0.18; 0.31)	0.88 (0.32; 2.39)	NA
	BIFAP	2.23 (0.77; 6.47)	0.46 (0.18; 1.18)	0.63 (0.20; 1.98)
	SIDIAP	0.40 (0.13; 1.27)	0.72 (0.20; 2.57)	1.12 (0.33; 3.77)
	RCGP RSC	1.93 (0.66; 5.65)	0.48 (0.18; 1.32)	0.67 (0.19; 2.30)
	THIN	0.31 (0.10; 0.98)	0.76 (0.21; 2.74)	1.23 (0.43; 3.50)
Fever	SSI	1.33 (1.21; 1.47)	0.83 (0.62; 1.11)	NA
	BIFAP	0.72 (0.67; 0.78)	0.96 (0.65; 1.43)	0.87 (0.56; 1.33)
	SIDIAP	0.58 (0.54; 0.62)	1.02 (0.78; 1.33)	0.93 (0.72; 1.21)
	RCGP RSC	1.12 (0.96; 1.30)	0.87 (0.61; 1.22)	0.75 (0.54; 1.04)
	THIN	1.01 (0.94; 1.08)	0.89 (0.60; 1.31)	0.77 (0.57; 1.04)
Persistent crying	SSI	NA	2.38 (1.55; 3.64)	NA
	BIFAP	1.60 (1.34; 1.91)	2.95 (2.56; 3.39)	2.95 (2.56; 3.39)
	SIDIAP	NA	2.38 (1.55; 3.64)	2.38 (1.55; 3.64)
	RCGP RSC	2.83 (2.18; 3.66)	2.19 (1.18; 4.06)	2.19 (1.18; 4.06)
	THIN	3.00 (2.54; 3.54)	2.11 (1.20; 3.68)	2.11 (1.20; 3.68)

BIFAP = Base de datos para la Investigación Famacoepidemiológica en Atención Primaria, RCGP RSC = Royal College of General Practitioners Research and Surveillance Centre, SIDIAP = Information System for Research in Primary Care, SSI = Statens Serum Institute, THIN = The Health Information Network, SCCS = Self Controlled Case Series. L-O-O = Leave-one-out

IRRs for fever showed a significant protective effect in the BIFAP and SIDIAP databases whereas the risk was increased in the SSI database and no association was observed in the THIN and RCGP RSC databases. Again, L-O-O meta-analysis removed much of the heterogeneity in these results. All L-O-O_IRR_ma estimates had confidence intervals including one (**Table 2**).

Persistent crying was significantly elevated in all databases that provided data for this event. L-O-O_IRR_ma results were consistent across databases and remained significantly greater than one. Because SSI did not contribute persistent crying cases, removal of SSI had no impact on L-O-O_IRR_ma estimates (**Table 2**).

The IR_bl and IR_bg methods performed similarly for febrile convulsions, tending to underestimate observed risk period IRs. In the primary care databases, with the exception of RCGP RSC, the derived IR_ma tended to be higher than the observed IR, because of the impact of the elevated incidence from the hospital database, SSI. For the SSI database, the observed risk period IR was higher than the derived IR_ma as this was based on the risk period IRs of the primary care databases. In analyses excluding SSI, IR_bl and IR_bg performed similarly and were in agreement with the observed risk period IR except in the RCGP RSC database. The IR_ma method produced higher estimates with wider confidence intervals than IR_bl and IR_bg in all scenarios (**Figure 4**).

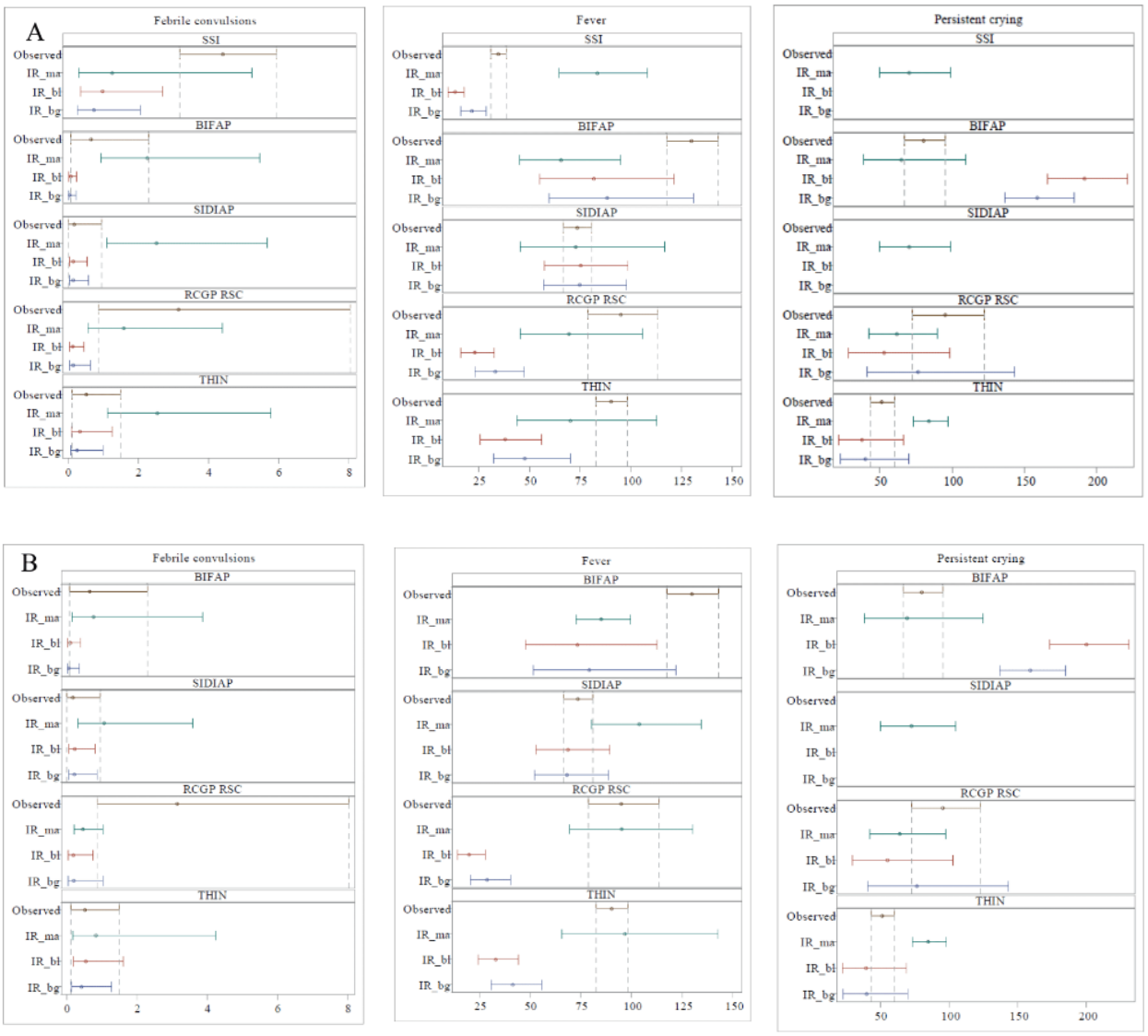


Figure 4: Comparison of results from the three methods for calculating incidence rates (IRs) for febrile convulsions, fever and persistent crying following aP vaccination (A) in all databases and (B) in primary care databases (excluding SSI)

BIFAP = Base de datos para la Investigación Famacoepidemiológica en Atención Primaria, RCGP RSC = Royal College of General Practitioners Research and Surveillance Centre, SIDIAP = Information System for Research in Primary Care, SSI = Statens Serum Institute, THIN = The Health Information Network, aPE = acellular pertussis vaccine

For fever the IR_bl and IR_bg methods gave similar results, i.e., derived IR estimates that were generally lower than the observed estimates. The derived IR_ma estimates were similar across databases. In the BIFAP database where the background IR for fever was highest, the IR_ma underestimated the observed IR for the risk period while in the SSI database, where the background rate of fever was the lowest, the IR_ma overestimated the IR for the risk period compared with the observed IR. (**Figure 3, Figure 4**). In analyses excluding SSI, IR_bl and IR_bg significantly underestimated the observed risk period IRs in all databases except for the BIFAP and SIDIAP databases, while the IRs from IR_ma were similar across databases and produced an underestimation of observed risk period IR in BIFAP.

For persistent crying, the results from the IR_bl and IR_bg approaches were similar. In the UK databases, the IRs derived by both methods were slightly lower than the observed risk period IRs, but not statistically significantly lower, whereas the IRs derived by both IR_bl and IR_bg were higher than those observed for the BIFAP database. The risk period IRs derived by the IR_ma method were similar across databases but they were underestimated compared with the observed risk period IRs in the BIFAP and RCGP RSC databases, and overestimated compared with the observed risk period IRs in the THIN database. Since no data for persistent crying events were available from the SSI database, its removal had no impact on the estimated IRs.

4. Discussion

The results from this study demonstrate that it is possible to obtain estimates for event-specific IRs occurring during risk windows after vaccination in a certain database using incidence rate ratios and incidence rates from other data sources, even if the data on the type of vaccination (for the IR_bl method) or the occurrence of vaccination (for the IR_bg and IR_ma methods) are not available in that database. The results also demonstrate that use of IR estimates from other data sources may not always be valid, since the type of data source (e.g. primary care setting versus hospital setting) has a major impact, which differs by type of event and the care pattern for that event.

Febrile convulsions are acute and can lead to emergency room visits and, therefore, primarily appear in hospital records [11, 12]. Since the SSI database contains only hospital-derived data, this might explain why the background, baseline and risk period incidence rates are higher in the SSI database than in the other databases which contain primary care-derived data (SIDIAP, BIFAP, THIN, and RCGP RSC). The observed IRs for febrile convulsions and their peak at around 15-16 months of age, especially in the SSI database, are consistent with those in the literature that reports a peak incidence at around 18 months old [13, 14]. The derived estimates for febrile convulsions IRs were in much better agreement with observed risk period IRs when the SSI hospital-based database was removed because of the difference in background incidence between primary care and hospital databases.

The post-vaccination IRs for fever derived using baseline or background rates produced estimates that were lower than the observed IRs in the risk window. Fever had a very low background incidence in the SSI database because it is a symptom and is unlikely be recorded as a hospital discharge diagnosis. The IRs derived using meta-analysis also tended to be lower that the observed risk period IRs except in the SSI database where the observed risk period IR was low. Removal of SSI did not improve the agreement between the derived IRs and observed risk period IRs due to its small contribution and therefore minimal changes to the L-O-O estimates.

Persistent crying is a non-specific condition that is not easy to record using medical coding systems and only the BIFAP database had specific codes for this event. Agreement was good for the methods in all databases, except BIFAP where the derived IRs using baseline and background rates were over-estimates compared with the observed risk period IRs, due to the higher baseline and background rates of persistent crying in BIFAP. The usefulness of the IR_ma estimates for the BIFAP, RCGP RSC and THIN databases is uncertain as they are derived from the meta-analysis of data from the other two databases while for the SIDIAP and SSI databases the IR_ma is the only estimate available due to the absence of persistent crying events in these databases.

In general, IR_ma estimates produced wider CIs due to our use of a random-effects meta-analysis and therefore, the 95% CIs for the IR_ma estimates were more likely to contain the observed IR. The L-O-O_IRR_ma estimates were similar across databases for each event, irrespective of which database was left out, suggesting that any differences in the resulting IR_bl or IR_bg estimates were due to difference in underlying baseline or background rates.

The aim of this study was to assess methods to fill gaps in information in one database using estimates from other databases. We demonstrated that this is possible, but that how data for each event are captured should be taken into consideration, as this may have a greater impact on the absolute IRs than on the IRRs. If an event, such as fever or persistent crying, is not captured in a database, we recommend that the pooled IRs (IR_ma) from databases which were able to capture the event of interest in similar settings are used. For example, the incidence of febrile convulsions was lower in the primary care databases than in the hospital database, but the IR_ma method produced derived IRs that were more in line with those observed in the hospital database. This method may be preferable if observed IRs in primary care databases are assumed to be underestimated.

Although the type of event type may have an important impact on the performance of methods for derivation, we demonstrated that the IR_bl and IR_bg methods provided very similar results for the events we used, which means that the approach using the background IRs (which does not require vaccine exposure time) can be used. This may be because the risk periods represent a very small period in comparison with the total follow-up period, and the risk increase was small during the risk period. These methods may be preferable if background and baseline IRs are assumed to be accurate, and the IR_bg method may be preferable if the risk period is short or cannot be observed due to missing exposure data.

5. Conclusions

Although we were able to compare derived and observed IRs for aP exposure, we did not have the estimates of the true incidence of each event in the post-wP vaccination risk period in all databases. We cannot draw general conclusions regarding which method provides the best estimates of the true incidence, but we can conclude that, in case of short risk windows and small increases in IRRs, the IR_bl and IR_bg methods provide similar estimates. Additionally, the IR_ma method may provide derived IRs that are closer to the observed IRs when these latter come from a similar type of database. However, it is important to note that this method is sensitive to heterogeneity in baseline incidence in each of the database as it uses absolute measures of incidence, [15, 16].

We demonstrated that the type of events and databases have a large impact and it is important to distinguish if the events are diagnosed in primary care, hospital or both, and perform stratified analyses for the type of events the databases capture. It is important to have a clear understanding of the external and internal validation of the databases as well as the heterogeneity of the studied databases and those used for deriving the parameters before proceeding to parameter derivation. We conclude that derived IRs for events following vaccination in the absence of specific vaccine exposure data in a specific database is possible if the background IRs can be calculated and IRRs are available from a similar type of database.

Acknowledgements

The authors acknowledge input and guidance from Heather Whitaker (Open University, United Kingdom) and Paddy Farrington (Open University, United Kingdom). They acknowledge Bart Spiessens (Johnson & Johnson, Belgium) who authored the SAS code used to conduct the self-controlled case series. They also acknowledge medical writing and editorial assistance from Margaret Haugh (MediCom Consult, Villeurbanne, France).

Disclaimer

The results described in this publication are from the proof of concept studies conducted as part of the IMI ADVANCE project with the aim of testing the methodological aspects of the design, conduct and reporting of studies for vaccine benefit-risk monitoring activities. The results presented relate solely to the methodological testing and are not intended to inform regulatory or clinical decisions on the benefits and risks of the exposures under investigation. This warning should accompany any use of the results from these studies and they should be used accordingly. The views expressed in this article are the personal views of the authors and should not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organisations with which the authors are affiliated.

Funding source

The Innovative Medicines Initiative Joint Undertaking funded this project under ADVANCE grant agreement n° 115557, resources of which were composed of a financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and in kind contributions from EFPIA member companies.

Declaration of potential conflicts of interest

Caitlin Dodd, Kaatje Bollaerts, Maria de Ridder, Tom de Smedt, Chris McGee, Talita Duarte-Salles, Hanne-Dorthe Emborg, Consuelo Huerta, Elisa Martín-Merino, Gino Picelli, Klara Berencsi, Giorgia Danieli declared that they have no potential conflicts of interest. Daniel Weibel declared that he has received personal fees from GSK for work unrelated to the submitted work. Olivia Mahaux and Francois Haguinet declared that they are employed by GSK and hold company shares. Simon de Lusignan declared that he has received grants from GSK, Takeda, and Seqirus / JSS, and also personal fees from Sequirus and Sanofi, for work unrelated to the submitted work. Miriam Sturkenboom declared that she has received grants from Novartis, CDC and Bill & Melinda Gates Foundation, for work unrelated to the submitted work.

References

[1] Funk S, Salathe M, Jansen VA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J R Soc Interface. 2010;7:1247-56.

[2] Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. Value Health. 2003;6:9-17.

[3] Plotkin S OW, Offit PA. Vaccines. 6th edn. ed. Philadelphia: Elsevier Saunders; 2012.

[4] Zhang L, Prietsch SO, Axelsson I, Halperin SA. Acellular vaccines for preventing whooping cough in children. Cochrane Database Syst Rev. 2014:CD001478.

[5] Weibel D. ADVANCE POC I Risk pillar - Testing new approaches to monitoring benefit/risk with pertussis vaccines as test case: Incidence rates of safety outcomes of whole-cell pertussis and acellular pertussis vaccines in pre-school children: European Network of Centres for Pharmacoepidemiology and Pharmacovigilance; 2017 [Last accessed 28 May 2018]. Available from: http://www.encepp.eu/encepp/viewResource.htm?id=13813.

[6] Weibel D, Dodd C, Mahaux O, Haguinet F, de Smedt T, Duarte-Salles T, et al. ADVANCE system testing: can safety studies be conducted using electronic health data: an example with pertussis vaccination? Vaccine. 2018;Paper 6 in Supplement.

[7] Sturkenboom M, van der Aa L, Bollaerts K, Emborg HD, Ferreira G, Gino R, et al. The ADVANCE distributed network system for evidence generation on vaccines coverage, benefits and risks based on electronic health care data. Vaccine. 2018;Paper 2 in supplement.

[8] Virtanen M, Peltola H, Paunio M, Heinonen OP. Day-to-day reactogenicity and the healthy vaccinee effect of measles-mumps-rubella vaccination. Pediatrics. 2000;106:E62.

[9] Whitaker HJ, Farrington CP, Spiessens B, Musonda P. Tutorial in biostatistics: the self-controlled case series method. Stat Med. 2006;25:1768-97.

[10] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177-88.

[11] Le Saux N, Barrowman NJ, Moore DL, Whiting S, Scheifele D, Halperin S. Decrease in hospital admissions for febrile seizures and reports of hypotonic-hyporesponsive episodes presenting to hospital emergency departments since switching to acellular pertussis vaccine in Canada: a report from IMPACT. Pediatrics. 2003;112:e348.

[12] Quantin C, Benzenine E, Velten M, Huet F, Farrington CP, Tubert-Bitter P. Self-controlled case series and misclassification bias induced by case selection from administrative hospital databases: application to febrile convulsions in pediatric vaccine pharmacoepidemiology. Am J Epidemiol. 2013;178:1731-9.

[13] Hauser WA. The prevalence and incidence of convulsive disorders in children. Epilepsia. 1994;35 Suppl 2:S1-6.

[14] Sammon CJ, Charlton RA, Snowball J, Weil JG. The incidence of childhood and adolescent seizures in the UK from 1999 to 2011: A retrospective cohort study using the Clinical Practice Research Datalink. Vaccine. 2015;33:7364-9.

[15] Engels EA, Schmid CH, Terrin N, Olkin I, Lau J. Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses. Stat Med. 2000;19:1707-28.

[16] Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. Stat Med. 2002;21:1575-600.

CHAPTER 7. GENERAL DISCUSSION

GENERAL DISCUSSION

Active surveillance of drugs is often divided into different stages: signal detection, rapid assessment of risk, thorough evaluation of risk, and near real time monitoring. Each of these stages applies different methods and uses different types of databases, although recently a blending of methods and data sources is occurring. The development of methods and findings reported in this thesis are discussed according to two dimensions. The first dimension focuses on the type of method with respect to the hierarchy of drug evaluations: signal detection, rapid safety assessment, and thorough evaluation. The second dimension is the modality of evaluation methods, specifically focusing on the methods around multi-site studies and the use of a common data models and common analytics. These two dimensions are discussed together in each of the methods sections.

Summary

Methods for signal detection in children

Comparison of signal detection methods applied to spontaneous reporting databases (SRDs) have found that different disproportionality methods perform similarly and that the choice of algorithm must be made on a per-purpose basis (4-8). In contrast to the performance of methods used in spontaneous reporting databases, methods used to conduct signal detection on electronic health care databases which include traditional disproportionality analyses as well as epidemiological designs, differ in performance and not all have been quantified (1).

Performance assessment in both types of databases presents challenges since a reference set of negative drug-event associations and positive drug-event associations is needed. The negative controls present a greater challenge as absence of evidence does not necessarily mean evidence of absence. The choice and composition of reference set is an important factor for performance assessment (2, 3).

Methods for signal detection in SRD have been studied extensively and researchers are closer to a consensus on how they should best be used. However, debate remains regarding the impact of masking (meaning the diminished ability to detect signals due to preponderance of commonly reported drugs and events) and the impact of stratification. Some prior studies have produced conflicting results, revealing that the decision to account for a masking effect, effect modification by age, or some other variable must be made on a situational basis(4-8). In this thesis we added to existing methods knowledge by investigating the impact on performance of age stratification and the extent of masking by vaccines as we focused on children, where age stratification is relevant. When children grow, there is a rapid change in organ system and function, therefore the effect of drugs may differ substantially between ages (9). For this methods on SDR work we evaluated performance of different stratification approaches, and the reference sets for therapeutics and vaccines produced for the GRiP project (10, 11).

We created a common data model (CDM) for spontaneous reporting databases, as described in chapter 2, because multiple SRDs exist (e.g. FAERS, VAERS, Eudravigilance, WHO-Vigibase), and especially when looking at children only, power is limited. Conversion of the EudraVigilance and VAERS databases to a common data model allowed for direct comparison of methods applied to the databases separately as well as to their combination. It also highlighted deficiencies in each database by revealing those variables which were unavailable for conversion to CDM or, if available, with insufficient frequency to make application of methods feasible.

Our methodological studies using the FAERS and Eudravigilance databases demonstrate that stratification by age groups in children reveals some signals while masking others and that this is dependent upon the specific age group. We also demonstrated that in a combined therapeutic and vaccine database, including reports concerning vaccine exposures when assessing safety signals related to drugs can produce more robust estimates because of

the larger denominators, as compared to estimates produced when vaccine exposures are excluded. However, masking may be present and may differ by the type of association being investigated, particularly upon whether the event of interest is commonly reported in association with vaccines, and upon the age group under study.

In comparing methods across different data sources we found that traditional disproportionality methods, referred to as signal detection algorithms (SDAs) when paired with standard thresholds for defining a signal, do not perform as well as vaccine-specific methods in signal detection for vaccine exposures. Linearly combining a traditional SDA and a vaccine-specific method only marginally improved performance over the vaccine-specific method alone. Although we used large publicly available datasets from the USA and EU, for many true positive associations in the reference set, few events were reported, which limited our ability to accurately assess sensitivity and positive predictive value. Based on this experience and the lack of other pediatric specific reference sets, we have suggested methods for assessing performance in the absence of a reference set and advocate that reference set-independent methods be further developed. In addition, because the group of people typically exposed to vaccines differs from those prescribed drugs in terms of age and underlying health status, we advise that methods for vaccine safety signal detection be refined and developed since standard disproportionality methods do not appear to perform well.

Methods for rapid safety signal assessment

Rapid assessment of a safety concern is usually conducted to look quickly at population impact, which is of relevance to regulators and public health organizations for initial decision making. Most rapid assessment methods rely on ecological methods, which utilize easily accessible population level data in which groups or periods have been defined by a natural experiment such as a targeted vaccination campaign, enabling the study of time trends on a population level. In chapter 3 we have investigated the utility of ecological methods for vaccine safety assessment and found that, while ecological methods are useful for rapid assessment, the power of such studies to detect risk when the time from onset to diagnosis is long or when vaccine coverage is low, is insufficient. In the ecological study of incidence rates of narcolepsy before, during and after the 2009 H1N1 pandemic and vaccination campaign, we demonstrated that rapid assessment is possible across four continents using electronic health care data sources; results were generated with a common protocol, common data model and analytics in a distributed manner. However, based upon simulations conducted within the same study, estimates of relative incidence was underestimated by the ecological methods used. The utility of ecological methods for rapid assessment therefore relies on the type of association that is of interest.

Vaccine evaluation studies in distributed networks

Beyond the Vaccine Safety Datalink (VSD) which was initiated in 1990, single database studies were the common rule until the 2009 H1N1 pandemic for the evaluation of vaccine safety. A single site study may be more efficient when a study can be addressed fully and reliably in one data source, although (as is done in genetics) replication would be recommendable. In the single database studies we conducted within the THIN general practice database in chapter 2, we found that after application of inclusion criteria, sample sizes are often too limited to detect rare safety events. This provides grounds to advocate for the standard use of multiple data sources when safety issues need to be addressed properly.

Multi-database studies are necessary to study drug effects for many reasons. Primarily because of increased power to detect effects with increased study population, secondly to profit from differences in drug use in different countries and thirdly because of the ability to look at consistency across sites. The Rofecoxib scandal in 2004 and the H1N1 pandemic in 2009 gave a strong boost to the analysis of different data sources in parallel (12, 13). Within Europe several projects demonstrated the ability to conduct distributed studies using a common protocol, common

data model and common analytics (14). VAESCO, GRiP and ADVANCE leveraged the expertise that was built up in these initial projects and applied it to the pediatric (GRiP) and vaccine area (VAESCO, ADVANCE) (15-17). Several studies in this thesis were conducted as part of these projects, and aimed to generate and test methods for signal detection, rapid assessment and evaluation using a distributed multi-database study. The work of the VSD in the United States was pioneering in that the system was the first to establish a distributed data model of multiple stakeholders with common standards for data quality and sharing, allowing for rapid assessment of vaccine safety concerns(18). Based upon the success of the VSD, similar collaborations have been initiated in North America (Sentinel in the United States and the Canadian Immunization Research Network in Canada) (19, 20). In Europe, however, post-marketing surveillance for vaccines continues to be conducted primarily on a per-country and per-vaccine basis rather than through an established network with fixed funding (17). One of the key decisions in these collaborations is the choice of a/the common data model (CDM) to allow for common analytics on the electronic health care data. Sentinel and CNODES are using the Sentinel CDM, VSD has its own CDM, OHDSI uses the OMOP CDM and ADVANCE uses its own CDM, which is very similar to the CDM from prior EU projects, allowing for re-use of analytical tools.

European collaborative studies

Transformation of data to a CDM is a complex issue in Europe, because different countries use different coding, have different health care structures and use different languages. To improve consistence between different databases and transparency, Gini et al have worked toward defining methods for creating and quantifying the impact of different case finding algorithms using components from various database domains such as diagnoses, prescriptions, laboratory tests, and procedures (21). In addition, for the ADVANCE project, Becker et al have developed a tool (Codemapper) which maps from text definitions to Unified Medical Language System (UMLS) concepts and subsequently to codes in diverse coding systems (22). This tool has been used in studies conducted within the ADVANCE project, described in chapter 6 of this thesis.

Although one may use different case finding algorithms, it would be good to validate these against a gold standard and get performance measures. This is not always possible. To support validity assessment we developed a method for deriving validity indices (positive predictive value, sensitivity, specificity, etc.) from a subset of indices and have shown that with estimates of population prevalence of the event, prevalence of the composite of the components of the algorithm to identify events in the data source, and one other validity measure, all others may be derived. This approach was applied to algorithms for detection of pertussis infection in chapter 6 of this thesis.

Global collaborative studies

With increasing globalization comes the threat of a rapidly spreading zoonosis or viral mutation leading to increased virulence. This requires global collaboration and flexibility to use primary and secondary data collection methods, while using a common protocol, common data model and analytics. We demonstrated that using a common data model and common analytics, signal detection and rapid assessment is possible. However, the evaluation of detected signals presents a greater challenge, particularly in lower and middle income countries which may have less robust infrastructure in place.

The globally collaborative safety evaluation studies in this thesis illustrate the challenges inherent in using existing data for vaccine safety assessment; in many occasions no electronic systems are available and data needs to be collected in a dedicated manner. This limitation resulted in the fact that in the study of Guillain Barré Syndrome and narcolepsy and H1N1 vaccine in chapter 4 of this thesis, no low income countries were included. The Global Vaccine Safety Multi country study aimed at including both low and middle income countries as well as high income countries. Through a proof of concept study assessing the risk of thrombocytopenic purpura and aseptic meningitis

following measles and mumps containing vaccines, it was demonstrated that it is feasible to conduct vaccine safety assessment in low and middle income countries using a common protocol and common analytics. This was feasible with secondary use of electronic health data but also with ad hoc data collection.

Methodological considerations

Across the different studies, methods and data sources, we encountered methodological issues and would like to provide recommendations.

Vaccines cannot be treated like other drugs when conducting signal detection

The safety profile demanded by the public for vaccines due to their use in healthy subjects and their ubiquity in populations lead to different patterns of reporting for adverse events experienced following immunization (AEFI) as compared to adverse events following drug exposure. Extra vigilance is needed for vaccines to be able to quickly detect and assess vaccine safety concerns and to address vaccine hesitancy adequately. To ensure adequate detection of signals and avoid masking, it may be inappropriate to apply standard signal detection algorithms to vaccine reports. We have shown that the vaccine specific time-to-onset (TTO) method outperforms the empirical Bayes geometric mean (EBGM). Additionally, we have shown that a linear combination of the TTO method and EBGM improves performance over either method alone but that the improvement over TTO alone was not significant. We recommend that vaccine-specific SDAs be employed when AEFIs and that spontaneous reporting databases containing both drug and vaccine exposures be analyzed with and without stratification by report type, i.e. drug related and vaccine related, where possible.

Comparison of signal detection methods should be independent of a reference set

In the studies which make use of the vaccine and drug-related GRiP pediatric reference sets to compare the performance of methods to detect signals in pediatrics, we found that reference set size and low numbers of positive control associations in the data set to be tested made assessment of performance difficult and reduced the precision of performance measures. This was compounded by the fact that many drugs and/or vaccines included in each reference set were used in specific pediatric subpopulations while many events occurred only in other pediatric subpopulations. For example, the pediatric drug reference set used in chapters 2.1 and 2.2 included psychosis and suicide as events, which are unlikely to occur in neonates and infants. Similarly, the GRIP vaccine reference set used in chapter 2.3 includes influenza vaccines which are less likely to be used in infants and neonates as well as the MMR vaccine which is unlikely to be used in adolescents except as a catch-up vaccine. The vaccine schedule is specific to pediatric age groups, with most childhood vaccines received in infancy (23). We recommend use of reference set independent performance assessments such as those presented in chapter 2.2 in the study of masking by vaccines in EudraVigilance. In this study, we were able to show, in the absence of a reference set, that the magnitude of estimates generated by disproportionality analysis generally decreased with the removal of vaccine exposures. A similar approach has been used by Zeinoun et al (5). We also recommend that reference sets be regularly updated and expanded based on the most current evidence. Additionally, because of differences in drug/vaccine use and in events incidences, age group-specific reference sets may be called for when performing signal detection.

Rapid assessment performance depends on onset to diagnosis lag time of the event as well as vaccine coverage

In chapter 3, which describes the use of ecological data to assess the incidence of narcolepsy before, during, and after the H1N1 pandemic and vaccination campaign, we showed that rapid assessment of adverse events following vaccination requires an understanding of the onset to diagnosis interval (lag time) as well as accurate recording of

onset. Long lag times and use of diagnosis dates instead of onset, will attenuate potential increases in risk. While we were able to analyze incidence rates of narcolepsy in periods prior to H1N1 virus circulation, during virus circulation, and following the vaccination campaign, these analyses were hampered by availability of diagnosis dates only, which may occur long after disease onset. Beyond impact of long lag times, coverage has an impact on the ability to assess risks. We demonstrated that ecological methods are likely to underestimate risk when coverage is low. Reliance on rapid assessment methods in triage or regulatory decision making should take these factors into account.

Vaccine safety evaluation requires collaborative studies to increase sample size and leverage exposure differences

Power can be limited in single database studies, especially when stringent inclusion and exclusion criteria are applied. If limited to a single database, conclusions may not be generalizable and consistency/replication cannot be tested. Even large databases may not provide adequate power to study rare exposures and outcomes or to produce estimates in sparse strata. As shown in chapter 4.1, which describes results of a single database study assessing the risk of Bell's Palsy following influenza vaccination in the THIN database, there was sufficient power for non-stratified analyses. However, power was reduced when analyses were stratified by year and age group, meaning that the increased risk of diagnosis on the day of vaccination was no longer detected within strata. Although over 6,000 cases of Bell's Palsy were found in this large database of over 8 million subjects, addressing some age, sex, and vaccine year strata-specific questions were not possible and would necessitate collaboration with other databases. Similarly, in the study of measles-induced immune suppression conducted in THIN and described in chapter 4.2, the a priori planned outcome of death was dropped from the final analysis due to an insufficient number of cases, and incidence rates for hospitalizations were unstable as compared to those for other outcomes due to low case counts. Both studies, due to their limitation to a single population, were able to exploit population-specific features such as reduced MMR uptake (for the measles study) and influenza vaccination recommendations for the elderly and pregnant women (for the Bell's palsy study) in the UK population. However, limitation to a single database also limits the extent to which effects due to exposure can be distinguished from effects due to underlying population characteristics. For example, in chapter 5.1, which describes a collaborative study of Guillain Barré Syndrome following p-H1N1 vaccination, an analysis limited to the Mexican data set would have indicated an increased risk of Guillain-Barré Syndrome following p-H1N1 vaccination while analysis including other databases attenuated this finding.

Common definitions and quantitative assessment of differences in outcomes is necessary in multi-database studies

In each of the multi-database studies in chapters 5 and 6, common protocols with common definitions were used and approaches were taken to harmonize extraction of events in disparate data sources. In the study of Guillain Barré Syndrome following p-H1N1 vaccination, data was retained locally in its original form but cases were classified according to Brighton Collaboration criteria while vaccine exposures were classified by adjuvant (24). In the two other studies included in chapter 5, one of which aimed to assess the risk of narcolepsy following p-H1N1 vaccination and the other which aimed to assess known associations following measles and mumps containing vaccines, data were collected de novo using a common protocol, common data entry tools including an online case report form, and harmonized event definitions. Data were locally transformed using the CHAMELEON (Chameleon, Erasmus Medical Center, Rotterdam, the Netherlands) tool and shared, allowing for a common analysis. In the manuscripts which originated from the ADVANCE project in chapter 6, data access providers of electronic health data were using a common data model. Each data provider extracted events and vaccines and converted these data into the CDM using harmonized code lists generated by Codemapper (22). In the subsequent harmonization process, differences between data sources in codes and incidence rates were discussed and algorithms for case detection were iteratively harmonized. Study teams then assessed whether a data source was fit for purpose to

participate in studies. Subsequently, data was analyzed using common analysis scripts in SAS or R. In each of these projects, heterogeneity was reduced largely due to the common definitions/analysis but by no means removed, because of different provenance of the data (hospital versus primary care) and granularity of the coding schemes (READ vs. ICPC). Meta-data on provenance, such as healthcare delivery and coding practices for a database, are important factors in interpreting data from disparate data sources, and these pieces of information are available when there is close collaboration among data custodians and investigators. Additionally, studies reporting relative rather than absolute rates can avoid some of these issues by comparing cases to controls or at-risk person time to control time within one data source, which may produce more consistent estimates across heterogeneous data sources.

Electronic health care databases differ in their coding systems, coding practices, and incorporation of data from different domains such as hospitalizations, diagnoses, laboratory findings, prescriptions, procedures and dispensing. Because of these differences, case finding algorithms must be developed per database and, in multi-database studies, harmonization must be conducted on a study by study basis. In addition, validation may be undertaken to measure the predictive value of the resulting algorithms. In each of the collaborative studies included in this thesis, this process of mapping to various coding systems and exploiting data originating from different domains was undertaken. In the global proof of concept studies included in chapter 5, which relied on collection of clinical cases, the Brighton Collaboration case definitions for aseptic meningitis, immune thrombocytopenic purpura, encephalitis, narcolepsy, and Guillain Barré Syndrome were used to ensure comparability and consistency. In the study describing and testing case finding algorithms for identification of pertussis infections based on secondary use of health care databases throughout Europe, the interrelations among validity measures were exploited to estimate positive predictive value of a set of algorithms.

Electronic health care databases have been set up at different times in European countries and the data therefore cover different time periods. We tested methods for deriving incidence of events following an exposure which has not been observed in a certain databases (because of lack of the relevant calendar time period) using estimates from databases in which the exposure has been observed. While the exercise revealed that this type of derivation is possible, the results highlighted the importance of taking into account the types of electronic health care databases being used. For example, febrile convulsions, a serious event for which emergency care is often sought, had a higher incidence in the databases that captured hospitalization as outcomes, compared to general practice databases. We could not use the hospital based rates to derive incidence rates in general practice databases. The use of absolute rates to derive incidence produced rates similar to those observed if the databases used in the derivation were of similar type. Incidence rates derived using a combination of absolute and relative rates were influenced by underlying absolute rates, revealing that a thorough understanding of database provenance is of particular importance when using absolute rates.

Through the processes of consensus reaching and case validation, heterogeneity in case extractions from primary and secondary data collection can be limited and even quantified. In the absence of or in addition to time and resource expensive processes of chart validations, the impact of using different case finding algorithms can be assessed. The heterogeneity in data sources allows for the use of different case finding algorithms and their comparison to one another. Relationships among validity indices can then be exploited to measure performance of each algorithm. The utility of using data from one data source to derive missing data in another data source, however, can only be done with similar type of underlying data (e.g. hospitalizations or primary care)

Data pooling in multi-database studies can align with privacy constraints

Global collaboration means dealing with different legislations and regulations on data sharing. Some countries may share individual level de-identified data, whereas other countries cannot share data based on less than a pre-specified number of individuals. We demonstrated through different studies that, in spite of these restrictions, it is still possible to pool information for vaccine safety studies.

For example, the study of Guillain Barré Syndrome following pH1N1 vaccine used a hybrid of individual level anonomized data as well as a meta-analysis of estimates by site. The study of narcolepsy following p-H1N1 vaccines used a hybrid model due to the restriction that a subset of sites could not share individual level-data, requiring that results from these sites be meta-analyzed with the results of the pooled individual level analysis in other databases. In the global proof of concept study assessing known adverse events following measles and mumps containing vaccines, all individual level data was available but due to differences in case and exposure ascertainment, primary analyses were restricted to the subset of sites which were able to ascertain exact dates of exposure. These examples highlight that in a distributed global collaboration that uses a common protocol, standard case definitions and analytics, the governance of databases to share data remains and must be incorporated at the analysis stage using one or two stage pooling or a hybrid approach.

Assessment of vaccine safety requires consideration of AEFI-specific sources of bias

While all epidemiological studies are prone to biases that researchers try to limit through design, vaccine safety concerns in the population have the potential to make suspected associations even more difficult to study due to changes in behavior of patients and healthcare providers. In chapter 6 of this thesis, which focused on the association between p-H1N1 vaccination and narcolepsy, we conducted a simulation study to determine the impact of differential exposure classification and detection bias when studying narcolepsy in the presence of awareness about the association. We found that reduction in the time from onset to diagnosis and misattribution of onset dates to the period following vaccination in the presence of awareness interacted to inflate risk estimates. With a limited study period, recall bias (attributing onset to the period following vaccination) more significantly increased risk estimates than a reduction in the onset to diagnosis interval. We found that analysis using a case-control design rather than designs with person time offsets resulted in estimates with less extreme bias, possibly because exposed and unexposed person-time become irrelevant. However, the analysis was conducted in a simulated population with perfect exposure classification so this result may not be generalizable to real world settings. Accurate reporting of onset dates, blinded review of cases, and possibly use of dates less prone to bias (i.e. date of first contact with a physician) could reduce bias due to detection bias. The reduction in the onset to diagnosis interval is a more difficult issue as subjects must be diagnosed to enter a study as cases. When we extended our study period in simulation up to 25 years, we were able to obtain an unbiased estimate in the presence of a reduced onset to diagnosis interval. This of course is not an option when conducting rapid assessment of a vaccine safety signal. Quantitative bias analyses should be employed to assess the impact of increased diagnosis and reporting.

Future perspectives

Methods for signal detection

While much effort has been put into the testing of methods for detection of safety signals in adults, little has been done for children. Assessment of performance for drug and especially for vaccine safety signal detection in children may be hampered by small number of reported cases from reference sets. Development of broader vaccine safety reference set, perhaps with gradation of certainty regarding positive controls, could allow for more flexible methods testing. Current methods testing makes use only of the extremes: multiply replicated true

positive associations and negative associations with complete absence of reports. Further methods should be developed to test performance of methods when uncertainty in the veracity of an association is incorporated.

Established associations will have different reporting patterns than emerging associations and performance assessment limited to these associations may not give a clear indication of how well methods perform. Solutions such as limiting performance testing to the period before an association was added to the summary of product characteristics have been advocated, but including emerging and possible signals in a subset of analyses may give a clearer picture of method and database performance (reference). Additionally, rapid detection of safety signals is easier to conduct when data is complete and accurate. Spontaneous reporting systems should encourage completion of all fields and follow up with reporters to complete fields when possible. For signal detection in pediatrics in general and in vaccines specifically, it is of particular importance that fields related to age, vaccine components, and time to onset be completed. We noted that age was often missing and, as described in the manuscript on vaccine signal detection in EudraVigilance and VAERS, time to onset was missing in 7% of VAERS reports and in over 50% of EudraVigilance reports.

Due to developmental differences, differing drug use patterns, and age-dependency of adverse drug and vaccine reactions, stratification by age should be a standard part of all signal detection in spontaneous reporting databases as this may serve to effectively highlight true associations which would otherwise go undetected in the full database. Masking should be taken into account when investigating drug-related signals involving events which are known or suspected AEFIs.

Another area of signal detection which has not been fully developed is that of drug-drug interactions. Some work has been done in this area using machine learning (25-27). However, this has not yet been extended to vaccines where it is of particular importance due to co-administration. With better data on vaccine exposures, signal detection could move beyond the limited vaccine-AEFI model to one in which signal detection is conducted at the level of antigen, adjuvant, preservative, manufacturing process, and the interaction of these features of vaccines.

The majority of signal detection algorithms currently in use have been developed for drugs. The utility of the time-to-onset method, developed specifically for vaccines, has been displayed in this thesis and elsewhere. However, time-to-onset may be missing, uniform across outcomes, or unknown due to nonspecific symptoms at onset. Other methods which take into account the features unique to vaccine exposures (transience, co-administration, high coverage, etc.) should be developed to better conduct safety signal detection specifically for vaccines. Additionally, many recent signals, especially those associated with Human Papilloma Virus vaccines, are non-specific and are characterized by a cluster of symptoms rather than a specific diagnostic code. For vaccine safety signal detection, development of algorithms to detect unusually frequent clusters of symptoms in SDR and EHR could lead to earlier detection of unexpected adverse events (28).

In addition to spontaneous reporting databases which are highlighted in this thesis, electronic health record data is being actively explored as another data source for detection of signals. These data sources show great promise for detection and assessment of safety signals. The availability of longitudinal data as well as denominator data allows for the use not only of traditional disproportionality measures but also pharmacoepidemiological designs, sequential testing, machine learning, and scan statistics (1). Methods for signal detection in electronic health care databases have been studied intensely in recent years, especially within the Observational Medical Outcomes Partnership (OMOP), Pharmacoepidemiological Research on Outcomes of Therapeutics by a European ConsorTium (PROTECT), Exploring and Understanding Adverse Drug Reactions (EU-ADR) and Sentinel projects, which aimed to investigate whether electronic health care databases might be used for signal detection (29-31). Continued research on the utility of EHR for signal detection and validation is necessary. Exciting developments are also

possible when combining these data sources with other sources such as those which may be searchable through natural language processing such as scientific literature and social media, patient-reported outcomes in SRD and other sources, and genomic data. In the era of precision medicine we are now entering, EHR databases will serve as a resource for determining the outcomes which other sources of data will be employed to predict. For these reasons, it is necessary to understand and harmonize available EHR data sources.

Finally, drugs and vaccines come to market following accumulation of data on safety and effectiveness in clinical trials. Unfortunately, this data generally remains unavailable to researchers seeking to understand the safety profile of a vaccine post-marketing. With the collaboration of vaccine manufacturers and researchers, it may be possible to use data generated pre-marketing to supplement post-marketing studies, either as Bayesian priors in analyses to sequentially update risks and benefits or as indicators for predictive features to weigh more heavily in post-marketing surveillance.

Methods for rapid Assessment of safety signals

The United States Sentinel system has made significant progress in the area of rapid assessment. Using their distributed data model, it is possible to quickly deploy queries against participating databases. The majority of the queries deployed in Sentinel are simple: background rates within strata, incidence rate ratios, and incidence of concurrent exposures (19). While rapid assessment of population impact using ecological methods may provide information to regulators rapidly, we have found that its utility is limited for rare events with a long onset to diagnosis interval.

As illustrated by the success of the Sentinel project , simple analyses as well as more complex analyses can be conducted in a network of databases when the data has been previously converted to a CDM. An additional solution to rapid assessment which has not been fully explored is a sequential updating of benefits and risks in a sequential or Bayesian sequential benefit risk assessment. This has been explored in the ADVANCE project and elsewhere but remains an area with potential for significant advancements, which will depend upon availability of timely data (32-35). As evidenced by collaborations such as the VSD and Sentinel, availability of timely data is best facilitated through conversion of databases to a common data model. Again, data from clinical trials could be made available to researchers conducting post-marketing surveillance. Mapping data from the Clinical Data Interchange Standards Consortium (CDISC) standards used in clinical trials to ontologies used in EHR or to a CDM would accelerate this process.

Models such as VSD and Sentinel which have been successful in rapidly assessing signals rely upon individual level data. Globally, as new vaccines are introduced in developing countries and globalization facilitates spread of infectious agents, it will no longer be sufficient to focus rapid assessments on the United States population covered by VSD and Sentinel. Instead, it is necessary to move toward a global vaccine data network, building upon infrastructure which has been built up in developed countries through previous collaborative studies and in developing countries through clinical trials and health surveillance systems.

Collaborative Studies

As illustrated in chapter 4 of this thesis, collaborative studies provide researchers with the opportunity to increase power to study rare events, to exploit differences in underlying populations, timing of exposures, and healthcare systems as well as myriad other as yet unexploited differences. However, we have also illustrated that collaborative studies such as those presented in this thesis require tools for translatability of coding systems, common tools for extraction, transformation, and loading of data files, and support for data recording and retention in diverse sites. Each of the studies presented here provides solutions for dealing with these obstacles

but much work remains to be done. For example, the Codemapper tool developed by Becker et al and used in this thesis increases translatability of coding systems and provides a platform through which investigators can reach consensus. The component algorithm approach first described by Gini et al and applied in ADVANCE also serves to reduce heterogeneity among databases by providing measures for determining the best case finding algorithm per database. This field of research is new and much progress remains to be made. Many advocate the use of common data models to allow for rapid deployment of common programs across a large network of databases. While conversion to common data models has obvious benefits – including evolution of databases toward complete CDM data (ie databases recognizing and rectifying areas of incompleteness during and following conversion to a CDM), the potential to conduct large pooled analyses, and the collaboration inherent in use of a common data model, - there are potential negative consequences. As we have shown in this thesis, database heterogeneity is a hindrance but we should carefully consider whether we want to significantly reduce database heterogeneity in pursuit of a universal CDM. Those differences which hinder rapidity and collaboration may be the same differences which could help to disentangle future epidemiological questions.

The future of vaccine safety surveillance is in the use of big data and common analytics and in order to smooth this transition to the use of big data, it is important to retain the lessons learned. As common data models become the norm, researchers will need to retain a focus on data quality and completeness rather than assuming that the size of the available data set will correct for misclassifications. Additionally, as shared and especially open-source analytics become more widespread, care should be taken that these tools are consistently updated and do not become black boxes which are applied without proper understanding of the underlying theory. Finally, specific to vaccines, it is necessary to create a vaccine ontology which will allow for common analytics on vaccines with different antigen combinations, adjuvants, and additives. While ontologies such as RxNorm and Anatomical Therapeutic Chemical classification systems exist for drugs (and include vaccines, although in insufficient detail), progress is currently being made to develop a more detailed ontology for vaccines (36, 37). This has been spurred by the development by the article 57 database of the EMA and further progress has been made within the ADVANCE consortium (38, 39). This will continue to be an area of development, especially as many of the new vaccines currently in development come to market.

Dealing with Heterogeneity and Bias

Within collaborative research networks, data still typically resides with the data provider. This limits pooling of data but also limits how much data from one source can be informed by data from another source, as typically only coefficients from statistical analysis or limited analysis sets are shared. With secure environments in which to share data or even IT tools to allow a hub to 'look' at data without transmitting it, it may be possible to borrow information from one data source where it is lacking in another. For example, common imputation methods use non-missing data to impute missing variables. Additionally, results generated in a full data set could be applied to analysis in a limited data set. For example, prediction models could be iteratively updated and calibrated using data from each data source, using the subset of variables available in that data source. Finally, if data is shared in a secure fashion, it may become possible to move beyond the vaccine-AEFI association model and to analyze associations at the level of antigen, strain match (for influenza vaccines), manufacturing process, etc. while avoiding the collinearity these features typically have with data provider when individual level data cannot be shared.

Finally, conducting studies globally reveals differences in data quality and access in diverse locations. In a rapidly globalizing world, if we want to rapidly conduct high quality studies of emerging epidemics or vaccine safety concerns, support for data recording and retention in lower income settings is needed. The first step toward this is epidemiological training and funding for database infrastructure. Another important step to assuring vaccine safety

in diverse locations is to conduct validation which is consistently the most time and resource intensive. This process could be facilitated and made less expensive by employing natural language processing to the validation of cases (40, 41).

Outcome and exposure misclassification will continue to hamper studies conducted using observational databases. Solutions have been proposed and tested within this thesis. However, assessment of the positive predictive value either using a gold standard such as manual review or by exploiting the interrelations of validity measures should become common practice in all database studies. In addition, reporting of these measures will help to inform future studies and provide input parameters for sensitivity analyses. Finally, simulation is a useful tool which should be applied more often not only in the development and testing of methods but in investigation into unmeasured confounding and biases and their potential impact on empirical estimates.

References

1. Arnaud M, Begaud B, Thurin N, Moore N, Pariente A, Salvo F. Methods for safety signal detection in healthcare databases: a literature review. Expert opinion on drug safety. 2017;16(6):721-32. Epub 2017/05/12.

2. Sridhar G, Tian F, Forshee R, Kulldorff M, Selvam N, Sutherland A, et al. Evaluation of optic neuritis following human papillomavirus vaccination. Human vaccines & immunotherapeutics. 2017;13(7):1705-13. Epub 2017/05/04.

3. Greene SK, Peterson ER, Kapell D, Fine AD, Kulldorff M. Daily Reportable Disease Spatiotemporal Cluster Detection, New York City, New York, USA, 2014-2015. Emerging infectious diseases. 2016;22(10):1808-12. Epub 2016/09/21.

4. Wang HW, Hochberg AM, Pearson RK, Hauben M. An experimental investigation of masking in the US FDA adverse event reporting system database. Drug safety. 2010;33(12):1117-33. Epub 2010/11/17.

5. Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. Human vaccines. 2009;5(9):599-607. Epub 2009/07/21.

6. Maignen F, Hauben M, Hung E, Van Holle L, Dogne JM. Assessing the extent and impact of the masking effect of disproportionality analyses on two spontaneous reporting systems databases. Pharmacoepidemiology and drug safety. 2014;23(2):195-207. Epub 2013/11/19.

7. Hopstadius J, Norén GN, Bate A, Edwards IR. Impact of stratification on adverse drug reaction surveillance. Drug safety. 2008;31(11):1035-48.

8. Woo EJ, Ball R, Burwen DR, Braun MM. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug safety. 2008;31:667-74.

9. Duncan DT, Rienti M, Jr., Kulldorff M, Aldstadt J, Castro MC, Frounfelker R, et al. Local spatial clustering in youths' use of tobacco, alcohol, and marijuana in Boston. The American journal of drug and alcohol abuse. 2016;42(4):412-21. Epub 2016/04/21.

10. Park R, O'Brien TF, Huang SS, Baker MA, Yokoe DS, Kulldorff M, et al. Statistical detection of geographic clusters of resistant Escherichia coli in a regional network with WHONET and SaTScan. Expert review of anti-infective therapy. 2016;14(11):1097-107. Epub 2016/08/18.

11. Osokogu OU, Fregonese F, Ferrajolo C, Verhamme K, de Bie S, t Jong G, et al. Pediatric drug safety signal detection: a new drug-event reference set for performance testing of data-mining methods and systems. Drug safety. 2015;38(2):207-17. Epub 2015/02/11.

12. Kulldorff M, Kleinman K. Comments on 'a critical look at prospective surveillance using a scan statistic' by T. Correa, M. Costa, and R. Assuncao. Stat Med. 2015;34(7):1094-5. Epub 2015/03/11.

13. Trifiro G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, et al. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? Journal of internal medicine. 2014;275(6):551-61. Epub 2014/03/19.

14. Yih WK, Lieu TA, Kulldorff M, Martin D, McMahill-Walraven CN, Platt R, et al. Intussusception risk after rotavirus vaccination in U.S. infants. N Engl J Med. 2014;370(6):503-12. Epub 2014/01/16.

15. Dieleman J, Romio S, Johansen K, Weibel D, Bonhoeffer J, Sturkenboom M, et al. Guillain-Barre syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe. BMJ. 2011;343:d3908. Epub 2011/07/14.

16. Joseph Sheehan T, DeChello LM, Kulldorff M, Gregorio DI, Gershman S, Mroszczyk M. The geographic distribution of breast cancer incidence in Massachusetts 1988 to 1997, adjusted for covariates. International journal of health geographics. 2004;3(1):17. Epub 2004/08/05.

17. Peters U, Sinha R, Bell DA, Rothman N, Grant DJ, Watson MA, et al. Urinary mutagenesis and fried red meat intake: influence of cooking temperature, phenotype, and genotype of metabolizing enzymes in a controlled feeding study. Environmental and molecular mutagenesis. 2004;43(1):53-74. Epub 2004/01/27.

18. Mathew A, Peters U, Chatterjee N, Kulldorff M, Sinha R. Fat, fiber, fruits, vegetables, and risk of colorectal adenomas. International journal of cancer. 2004;108(2):287-92. Epub 2003/11/26.

19. The Sentinel Initiative: National Strategy for Monitoring Medical Product Safety. Programs OoCP; 2008.

20. Prestwood KM, Kenny AM, Kleppinger A, Kulldorff M. Ultralow-dose micronized 17beta-estradiol and bone density and bone metabolism in older women: a randomized controlled trial. JAMA. 2003;290(8):1042-8. Epub 2003/08/28.

21. Roberto G, Leal I, Sattar N, Loomis AK, Avillach P, Egger P, et al. Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. PLoS One. 2016;11(8):e0160648. Epub 2016/09/01.

22. Becker BFH, Avillach P, Romio S, van Mulligen EM, Weibel D, Sturkenboom M, et al. CodeMapper: semiautomatic coding of case definitions. A contribution from the ADVANCE project. Pharmacoepidemiology and drug safety. 2017;26(8):998-1005. Epub 2017/06/29.

23. Control ECfDPa. [27 December 2018]; Available from: https://vaccine-schedule.ecdc.europa.eu/.

24. Sejvar JJ, Kohl KS, Gidudu J, Amato A, Bakshi N, Baxter R, et al. Guillain-Barre syndrome and Fisher syndrome: case definitions and guidelines for collection, analysis, and presentation of immunization safety data. Vaccine. 2011;29(3):599-612. Epub 2010/07/06.

25. Ayvaz S, Horn J, Hassanzadeh O, Zhu Q, Stan J, Tatonetti NP, et al. Toward a complete dataset of drug-drug interaction information from publicly available sources. Journal of biomedical informatics. 2015;55:206-17. Epub 2015/04/29.

26. Vilar S, Uriarte E, Santana L, Tatonetti NP, Friedman C. Detection of drug-drug interactions by modeling interaction profile fingerprints. PloS One. 2013;8(3):e58321. Epub 2013/03/23.

27. Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther. 2011;90(1):133-42. Epub 2011/05/27.

28. Chandler RE, Juhlin K, Fransson J, Caster O, Edwards IR, Noren GN. Current Safety Concerns with Human Papillomavirus Vaccine: A Cluster Analysis of Reports in VigiBase((R)). Drug safety. 2017;40(1):81-90. Epub 2016/09/18.

29. Leis AA, Stokic DS. Zika Virus and Guillain-Barre Syndrome: Is There Sufficient Evidence for Causality? Frontiers in neurology. 2016;7:170. Epub 2016/10/18.

30. Islam MB, Islam Z, Farzana KS, Sarker SK, Endtz HP, Mohammad QD, et al. Guillain-Barre syndrome in Bangladesh: validation of Brighton criteria. Journal of the peripheral nervous system : JPNS. 2016;21(4):345-51. Epub 2016/09/13.

31. Alcalde-Cabero E, Almazan-Isla J, Garcia Lopez FJ, Ara-Callizo JR, Avellanal F, Casasnovas C, et al. Guillain-Barre syndrome following the 2009 pandemic monovalent and seasonal trivalent influenza vaccination campaigns

in Spain from 2009 to 2011: outcomes from active surveillance by a neurologist network, and records from a country-wide hospital discharge database. BMC neurology. 2016;16:75. Epub 2016/05/22.

32. Greene SK, Kulldorff M, Yin R, Yih WK, Lieu TA, Weintraub ES, et al. Near real-time vaccine safety surveillance with partially accrued data. Pharmacoepidemiology and drug safety. 2011;20(6):583-90. Epub 2011/05/04.

33. Li L, Kulldorff M. A conditional maximized sequential probability ratio test for pharmacovigilance. Stat Med. 2010;29(2):284-95. Epub 2009/11/27.

34. Bollaerts K, De Smedt T, Donegan K, Titievsky L, Bauchau V. Benefit-Risk Monitoring of Vaccines Using an Interactive Dashboard: A Methodological Proposal from the ADVANCE Project. Drug safety. 2018;41(8):775-86. Epub 2018/03/28.

35. Leite A, Andrews NJ, Thomas SL. Near real-time vaccine safety surveillance using electronic health records-a systematic review of the application of statistical methods. Pharmacoepidemiology and drug safety. 2016;25(3):225-37. Epub 2016/01/29.

36. Nelson SJ, Zeng K, Kilbourne J, Powell T, Moore R. Normalized names for clinical drugs: RxNorm at 6 years. Journal of the American Medical Informatics Association : JAMIA. 2011;18(4):441-8. Epub 2011/04/26.

37. Organization WH, Organization WH. Anatomical therapeutic chemical (ATC) classification index with defined daily doses (DDDs). Oslo: WHO Collaborating Centre for Drug Statistics Methodology. 2000.

38. Borg JJ, Tanti A, Kouvelas D, Lungu C, Pirozynski M, Serracino-Inglott A, et al. European Union pharmacovigilance capabilities: potential for the new legislation. Therapeutic advances in drug safety. 2015;6(4):120-40. Epub 2015/08/25.

39. Becker B. Vaccine semantics: Automatic methods for recognizing, representing, and reasoning about vaccine-related information. Rotterdam: Erasmus University; 2019.

40. Love TJ, Cai T, Karlson EW. Validation of psoriatic arthritis diagnoses in electronic medical records using natural language processing. Seminars in arthritis and rheumatism. 2011;40(5):413-20. Epub 2010/08/13.

41. Sada Y, Hou J, Richardson P, El-Serag H, Davila J. Validation of Case Finding Algorithms for Hepatocellular Cancer From Administrative Data and Electronic Health Records Using Natural Language Processing. Medical care. 2016;54(2):e9-14. Epub 2013/08/10.

SUMMARY

Since their inception, vaccines have prevented illness, disability, and death throughout the populations in which they have been administered. However, as is true of any intervention, vaccines are not without risk. Assessing vaccine risks and balancing them with measures of effectiveness and benefit in a timely manner is necessary to maintain public trust in vaccine programs and ensure continued protection against vaccine preventable diseases.

Following licensure, as large numbers of people are exposed to a vaccine, the processes of passive and active signal detection along with assessment and evaluation of potential safety signals must begin in earnest. The past 30 years have seen a swift move from passive reporting systems to use of large linked databases and novel statistical methods to achieve the goals of detection, assessment, and evaluation. With globalization and the collaborations it both enables and requires, evaluation methods have become increasingly collaborative through multisite studies. In this thesis, I follow this trajectory from signal detection in passive reporting systems through to methodological issues in multi-database investigations using a set of test cases which are timely and of public health importance.

In the general introduction, I describe the importance of vaccines as a public health intervention and the processes by which vaccines are assessed following licensure through the analysis of spontaneous reports, studies in observational databases, and collaborative multi-database studies. I outline methodological challenges to the assessment of vaccine safety including the rarity and non-specificity of many adverse events following immunization (AEFI), patterns in reporting and ascertainment of AEFI that may lead to bias, and the difficulty in differentiating vaccine effects from effects due to other causes such as infection or underlying health status. I also outline challenges inherent to the use of existing data sources such as limited sample size, incompleteness, and heterogeneity.

Spontaneous reporting system databases, which contain only data on suspected adverse reactions to drug and vaccine exposures, provide the potential to identify safety signals quickly. However, these databases do not contain denominator data and included reports are a non-representative sample of all adverse reactions. Conduct of the studies in chapter 2 was motivated by the continued use of spontaneous reporting systems for vaccine safety signal detection and their importance, particularly in settings in which active surveillance is less feasible due to resource constraints. The study described in 2.1 focuses on the impact of age stratification in the United States Food and Drug Administration Adverse Event Reporting System (FDA FAERS) database. We determine that age stratification reveals some true associations while masking others but that overall age adjustment does not improve performance of signal detection algorithms. While age-adjustment may be called for in some circumstances, we argue that age strata-specific estimates may reveal some true safety signals while masking others. In chapter 2.2 we investigate the extent of masking (reduced likelihood of detecting safety signals due to excess reports related to the same event and/or exposure) of true signals by vaccine reports in EudraVigilance, the mixed drug and vaccine database from the European Medicines Agency. We find that the impact of removal of vaccine reports in a mixed vaccine-drug database is dependent upon how often each event is reported in association with vaccines. Finally, in 2.3 we use the vaccine specific subset of EudraVigilance together with the vaccine-only Vaccine Adverse Event Reporting System (VAERS) database to compare a traditional signal detection algorithm to one developed specifically for vaccines and to assess the potential benefit of combining thresholds of both methods. The vaccine specific signal detection algorithm, which compares the distribution of time-to-onset (days from the vaccination of interest to onset of the event of interest) to time-to-onset distributions for other vaccines and events, provides superior performance, but only when data on time-to-onset relatively complete. In the same study, we convert both the VAERS and EudraVigilance databases to the same common data model and determine whether combining the two data sources leads to improved performance of signal detection algorithms.

We find that combining a traditional signal detection algorithm with a vaccine-specific leads to improved performance but conducting signal detection in the union of two spontaneous reporting databases does not. Overall, the analysis choices in a spontaneous reporting database should be informed by knowledge about the association(s) of interest and their biological plausibility in specific age groups and following specific drugs or vaccines as well as the quality and completeness of the available data.

Ecological studies are those which measure exposures and outcomes at the population rather than at the individual level. The study included in chapter 3 is inspired by the wide availability of population level data and an interest in its utility to rapidly assess adverse events following vaccination. We compare incidence rates of narcolepsy before, during, and after targeted pandemic H1N1campaigns, following which we simulate data with a known relative incidence to better understand the utility of ecological methods. Results from population-level data are shown to be influenced by delays between onset and diagnosis as well as low vaccination coverage, leading us to conclude that ecological methods are suitable for signal generation but not for risk assessment.

Studies conducted in a single database have long been the standard for assessment of vaccine effects. Limiting a study to a single database means that many issues of underlying population heterogeneity as well as heterogeneity in coding systems, healthcare delivery, and data structure are avoided. In chapter 4 we present two single database studies: one assessing risk of Bell 's palsy following influenza vaccine and one measuring the impact of hypothesized measles-induced immune suppression on infections, anti-infective prescriptions, and hospitalizations in the years following infection. In chapter 4.1, we use the self-controlled case series method and show that the risk of Bell 's palsy was not increased following either pandemic or seasonal influenza vaccines in The Health Improvement Network (THIN) database. In chapter 4.2,we illustrate that rates of infections and prescriptions for anti-infective medications are increased in the years following measles infection in measles-infected children as compared uninfected children who have received a measles vaccine. These studies are facilitated by the use of one coding system and consistency of health care delivery within the system which generated the data. However, in both studies planned analyses (stratification by year in the study of Bell 's palsy following influenza vaccination, incidence of death in the study of measles-induced immune suppression) are underpowered. While the studies could be implemented without the need to address the heterogeneity that is present in multi-database studies, they also illustrate the shortcomings of single database studies and the need for collaboration.

Single database studies may lack power to measure associations in which the AEFI is rare. Additionally, studies limited to a single database are also limited in the diversity of populations and vaccine exposures which can be included. In chapter 5 we describe three internationally collaborative vaccine safety studies, each of which was conducted with the motivation to prove the utility of international collaboration as well as its feasibility. We aim, in each study, to use as far as possible a common data model, protocol, analytics, and case definitions and to apply each of these in a collaboration which has global scope. The first, in which we assess risk of Guillain-Barré Syndrome following pandemic H1N1 vaccines, we employ common analytics as well as common case definitions to detect an increased risk following vaccination in line with previous studies. The study described in chapter 5.2, in which we analyze the association between pandemic H1N1 vaccines and narcolepsy, makes use of a common protocol and common analytics, to find no evidence of increased risk of narcolepsy following pandemic H1N1 vaccination in either children or adults. Finally, in chapter 5.3 we describe a study in which we are able to detect known associations following measles and mumps containing vaccines using a common protocol and common analytics, and including data from low and middle income countries. Each of these studies represents progress toward standardization to reduce heterogeneity between data sources but none of them achieves use of a common data model, protocol, analytics, and case definitions on a global scale. However, using knowledge through the conduct of each of these collaborative studies, we have developed significant capacity and gained knowledge which can be applied to the conduct of future global assessments of vaccine safety.

Multi-database and especially multi-country studies inevitably require that heterogeneity in data sources be addressed. Additionally, as is true of any epidemiological study, both single and multi-database studies are prone to information and selection biases. In the studies included in chapter 6, we address heterogeneity and bias in observational database studies of vaccine safety. Two of the studies, originating from the ADVANCE consortium, address heterogeneity in multi-database studies while in the third we employ simulation to understand the impact of bias. In chapter 6.1 we describe the results of a simulation study to assess the impact of detection bias and differential exposure misclassification on estimates of the risk of narcolepsy following pandemic H1N1 vaccine and show that, in the absence of an association, these sources of bias can produce estimates suggesting increased risk. In chapter 6.2 we apply a method to identify events of interest in a database using components from domains such as diagnosis, laboratory tests, and prescribing, and show how the validity of these case finding algorithms can be assessed. In the final manuscript of this chapter, we investigate derivation of post-vaccine exposure incidence rates in a multi-database study in which exposure data is missing in a subset of databases.

In the general discussion I again focus on the vaccine life cycle post-licensure, from signal detection in spontaneous reporting databases through rapid assessment, single database studies, collaborative studies, and assessment of heterogeneity and bias. I provide methodological considerations for studies of vaccine effects using existing data sources and future perspectives for the study of vaccine safety in the era of globalization and real world evidence.

SAMENVATTING IN HET NEDERLANDS

In alle bevolkingsgroepen waarin ze toegediend zijn, hebben vaccins sinds hun begin ziektes, handicaps en overlijden voorkomen. Desondanks zijn vaccins niet zonder risico's, hetgeen waar is voor alle medische interventies. Het tijdig beoordelen van de risico's van vaccineren en het balanceren met maatstaven van baten en effectiviteit is noodzakelijk om het vertrouwen van het algemene publiek in vaccinatieprogramma's te behouden en om voortdurende bescherming tegen door vaccinatie te voorkomen ziektes te garanderen.

Na goedkeuring, wanneer grote aantallen mensen aan het vaccin blootgesteld worden, moet er serieus begonnen worden met actieve en passieve signaaldetectie samen met het beoordelen en evalueren van potentiele veiligheidssignalen. In de afgelopen dertig jaar was er een snelle beweging van passieve rapportagesystemen naar het gebruik van grote gelinkte databases en nieuwe methoden in de statistiek om de doelen van detectie, beoordeling en evaluatie te bereiken. Door globalisering en de samenwerking die daardoor zowel mogelijk als noodzakelijk wordt, zijn de evaluatiemethoden in toenemende mate collaboratief geworden met multi-site studies. In dit proefschrift volg ik het traject van signaaldetectie in passieve rapportagesystemen tot methodologische kwesties in multi-databaseonderzoeken, gebruikmakend van een aantal proefgevallen die zowel tijdig als van belang voor de volksgezondheid zijn.

In de algemene introductie beschrijf ik het belang van vaccinatie als een interventie in te volksgezondheid en beschrijf ik de processen waarmee vaccins na goedkeuring beoordeeld worden doormiddel van de analyse van spontane rapportages, studies in observationele databases en collaboratieve multidatabasestudies. Ik geef een schets van de methodologische uitdagingen in het beoordelen van vaccinatieveiligheid, inclusief de zeldzaamheid en de niet-specifieke aard van veel ongewenste voorvallen na immunisatie (*adverse events following immunization*, AEFI), patronen in het rapporteren en vaststellen van AEFI's die tot vertekening kunnen leiden en de moeilijkheid in het onderscheid maken tussen vaccineffecten en effecten van andere oorzaken zoals een infectie of de onderliggende gezondheidssituatie. Ik schets ook de uitdagingen die inherent zijn aan het gebruik van bestaande gegevensbronnen, zoals beperkte steekproefomvang, onvolledigheid en heterogeniteit.

Databases van spontane rapportagesystemen bevatten alleen gegevens van ongewenste reacties op blootstelling aan geneesmiddelen en vaccins, en geven de mogelijkheid om risicosignalen snel te identificeren. Desondanks bevatten deze databases geen informatie over de noemer, en de geregistreerde rapportages zijn geen representatieve steekproef van alle ongewenste bijwerkingen. Het uitvoeren van hoofdstuk 2 was gemotiveerd door het aanhoudende gebruik van spontane rapportagesystemen voor het opvangen van risicosignalen van vaccins en hun belang met name in situaties waarin actieve surveillance minder haalbaar is vanwege beperkte middelen.

De studie in hoofdstuk 2.1 richt zich op de impact van stratificatie op leeftijd in de United States Food and Drug Administration Adverse Event Reporting System (FDA FAERS). We stelden vast dat stratificeren op leeftijd de prestaties van signaaldetectiealgoritmen niet verbetert. Hoewel het in bepaalde situaties nodig kan zijn op leeftijd te adjusteren, beargumenteren wij dat strata-specifieke schattingen sommige veiligheidssignalen kan onthullen terwijl het andere kan maskeren. In hoofdstuk 2.2 onderzoeken we de mate waarin ware signalen gemaskeerd worden (een afname in de waarschijnlijkheid van het waarnemen van veiligheidssignalen door overmatige rapportage gerelateerd aan hetzelfde voorval en/of dezelfde blootstelling) in EudraVigilance, de gemengde geneesmiddelen en vaccins database van de European Medicines Agency. Wij zien dat de impact van het verwijderen van vaccinatierapportages uit een gemengde database met vaccins en geneesmiddelen afhankelijk is

van hoe vaak elk voorval gerapporteerd is in samenhang met vaccins. Tenslotte gebruiken we in 2.3 het vaccinatie specifieke deel van EudraVigilance samen met de vaccinatie specifieke VAERS database om een traditioneel signaaldetectiealgoritmen te vergelijken met een algoritme dat specifiek ontwikkeld is voor vaccins, en om te beoordelen welk voordeel er mogelijk ligt in het combineren van de drempels van beide methoden.
Het signaaldetectiealgoritme dat specifiek voor vaccins ontwikkeld, en die de verdeling van tijdsduur-tot-aanvang (het aantal dagen vanaf de onderzochte vaccinatie tot het begin van het onderzochte voorval) vergelijkt met de tijdsduur-tot-aanvang van andere vaccins en andere voorvallen, geeft betere prestaties, maar alleen in die gevallen wanneer de gegevens over de tijdsduur-tot-aanvang relatief volledig is. In dezelfde studie zetten we zowel de VAERS als de EudraVigilance databases om in hetzelfde algemene datamodel, en bepalen we of het combineren van de beide gegevensbronnen leidt tot verbeterde prestaties van signaaldetectiealgoritmen. We zien dat het combineren van een traditioneel signaaldetectiealgoritme met een vaccin specifiek signaaldetectiealgoritme leidt tot verbeterde prestaties maar dat dit niet het geval is voor het uitvoeren van signaaldetectie in de vereniging van beide databases van spontane rapportages. In het algemeen kan gesteld worden dat de keuzes in het analyseren van spontane rapportagedatabases geïnformeerd moeten worden door kennis van de onderzochte associaties en hun biologische plausibiliteit in specifieke groepen en volgende op specifieke geneesmiddelen of vaccinaties, als ook de kwaliteit en volledigheid van de beschikbare gegevens.

Ecologische studies zijn die studies waarin blootstellingen en uitkomsten gemeten worden op het niveau van de bevolking en niet op het niveau van het individu. De studie in hoofdstuk 3 is geïnspireerd door de wijdverspreide beschikbaarheid van gegevens op bevolkingsniveau en interesse in het nut ervan om ongewenste voorvallen na vaccinatie snel te kunnen bestuderen. We vergelijken incidentiecijfers van narcolepsie voor, tijdens en na gerichte pandemische H1N1 campagnes, waarna we gegevens simuleerden met een bekende relatieve incidentie om beter te begrijpen wat het nut is van ecologische methoden. Het is aangetoond dat de resultaten op bevolkingsniveau beïnvloed worden door vertraging tussen aanvang en diagnose, en ook door lage vaccinatiedekking, hetgeen ons leidde tot de conclusie dat ecologische methoden gepast zijn voor het genereren van risicosignalen, maar niet voor het beoordelen van de risico's.
Studies die uitgevoerd worden in één enkele database zijn lang de standaard geweest voor het beoordelen van de effecten van vaccins. Het beperken van een studie tot één database betekent dat een groot aantal kwesties van de onderliggende bevolkingsheterogeniteit en ook heterogeniteit in coderingssystemen, verzorging van de gezondheidszorg en datastructuur worden vermeden. In hoofdstuk 4 presenteren we twee studies met één database: een studie die het risico beoordeelt op Bellse parese na griepvaccinatie en een studie die de impact meet van een hypothetische door mazelen veroorzaakte immunosuppressie op infecties, voorschriften van anti-infectieuze middelen, en hospitalisaties in de jaren na infectie. In hoofdstuk 4.1 gebruiken we de methode van de *self-controlled case series* en laten we zien de het risico op Bellse parese niet verhoogd was na zowel een pandemische als een seizoensgriepvaccinatie in The Health Improvement Network (THIN) database. In hoofdstuk 4.2 illustreren we het feit dat er in verhouding meer voorschriften van anti-infectieuze middelen gegeven worden in de jaren na mazeleninfectie in kinderen die geïnfecteerd zijn met mazelen, vergeleken met kinderen die een mazelenvaccin ontvangen hebben. Deze studies worden gefaciliteerd door het gebruik van één coderingssysteem en consistente verzorging van gezondheidszorg binnen het systeem waaruit de gegevens voorgekomen zijn. Toch hebben in beide studies geplande analyses (stratificatie op jaartal in de studie naar Bellse parese na griepvaccinatie, en de incidentie van overlijden in de studie naar door mazelen veroorzaakte immuunsupressie) te weinig onderscheidingsvermogen. Hoewel de studies geïmplementeerd konden worden zonder de noodzaak zich te richten op de heterogeniteit die aanwezig is in multidatabasestudies, illustreren ze ook de tekortkomingen van studies van één database en de noodzaak van samenwerking.

Studies van één enkele database kunnen mogelijk te weinig onderscheidingsvermogen hebben om relaties te meten waarin de AEFI zeldzaam is. Daarnaast zijn studies beperkt tot een enkele database ook beperkt in de diversiteit van de bevolkingsgroepen en blootstelling aan vaccins die geïncludeerd kunnen zijn. In hoofdstuk 5 beschrijven we drie internationaal samenwerkende vaccinatieveiligheidsstudies, die allemaal uitgevoerd zijn met de motivatie om nut en haalbaarheid van internationale samenwerking te bewijzen. Wij stellen in elk van die studies ons tot doel om zo ver als het kan een gemeenschappelijk data model, protocol, analyses en gevalsdefinities te gebruiken en om elk van deze toe te passen in een samenwerkingsverband met mondiaal bereik. In de eerste studie, waarin we het risico op het syndroom van Guillain-Barré na H1N1 vaccinatie beoordelen, passen we algemene analyses en ook algemene gevalsdefinities toe om een toegenomen risico na vaccinatie te vinden, hetgeen in overeenstemming is met voorgaande studies. De studie van hoofdstuk 5.2, waarin we de relatie tussen narcolepsie en pandemische H1N1 vaccins analyseren, maakt gebruik van een gemeenschappelijk protocol en gemeenschappelijke analyses, om geen bewijs te vinden voor een verhoogd risico op narcolepsie na vaccinatie voor pandemische H1N1, noch in kinderen, noch in volwassen. Tenslotte beschrijven we in hoofdstuk 5.3 een studie waarin we in staat waren reeds bekende relaties te ontdekken na gebruik van vaccins die mazelen en bof bevatten, gebruikmakend van een gemeenschappelijk protocol en gemeenschappelijke analyses, en ook met de inclusie van gegevens uit landen met lage- en middeninkomens. Deze studies staan allemaal voor vooruitgang in de standaardisatie om heterogeniteit tussen gegevensbronnen te verminderen, maar geen van deze studies verwezenlijkt een algemeen datamodel, protocol, analyses en gevalsdefinities op mondiaal niveau. Desondanks hebben we, gebruikmakend van de kennis opgedaan door het uitvoeren van deze collaboratieve studies, een significante capaciteit ontwikkeld en kennis verworven die toegepast kan worden in het uitvoeren van toekomstige mondiale evaluaties van vaccinveiligheid. Multidatabase studies, in het bijzonder die met databases uit verschillende landen, hebben de onvermijdelijke noodzaak te letten op de heterogeniteit van hun gegevensbronnen. Bovendien zijn zowel enkele als multidatabasestudies, net als elke epidemiologische studie, geneigd tot vertekeningen door informatie en selectie. In de studies van hoofdstuk 6 bespreken we de vertekening door heterogeniteit in observationele databasestudies naar vaccinveiligheid. Twee van de studies, die voortkomen uit het ADVANCE consortium, bespreken heterogeniteit in multidatabasestudies, terwijl we in de derde studie een simulatie gebruiken om de invloed van vertekening te begrijpen. In hoofdstuk 6.1 beschrijven we de resultaten van een simulatiestudie die opgezet is om de impact van detectievertekening en differentiële blootstellingsmisclassificatie te evalueren op schattingen op het risico op narcolepsie na het pandemische H1N1 vaccin, en laten we zien dat in afwezigheid van een relatie, deze oorzaken van vertekening een schatting kunnen produceren die een verhoogd risico suggereert. In hoofdstuk 6.2 passen we een methode toe om belangrijke voorvallen te identificeren in een database, gebruikmakend van componenten van domeinen zoals diagnoses, laboratoriumonderzoeken en geneesmiddelenprescripties, en laten we zien hoe de validiteit van deze voorval opsporende algoritmes geëvalueerd kunnen worden. In het laatste manuscript van dit hoofdstuk onderzoeken we het afleiden van incidentiecijfers na blootstelling aan een vaccin in een multidatabasestudie waarin gegevens over de blootstelling ontbreekt in een deel van de databases.

In de algemene discussie richt ik me wederom op de levenscyclus van vaccins na goedkeuring, van signaaldetectie in spontane rapportagedatabases tot snelle beoordeling, studies met één database, collaboratieve studies, en het beoordelen van heterogeniteit en vertekening. Ik verzorg methodologische overwegingen voor studies naar de effecten van vaccins die gebruik maken van bestaande gegevensbronnen en ik geef perspectieven voor de toekomst van het bestuderen van vaccinatieveiligheid in het tijdperk van globalisering en werkelijk bewijs.

LIST OF PUBLICATIONS

Manuscripts Included in this Thesis					
Title	Authors	Journal	Volume	Number	Year
Estimating risk of adverse events following vaccination in observational databases with incomplete exposure data	Caitlin Dodd, Francois Haguinet, Olivia Mahaux, Daniel Weibel, Kaat Bollaerts	*Submitted: Vaccine*			
Pediatric Signal Detection in the VAERS and EudraVigilance databases	Caitlin Dodd, Lionel van Holle, Maria de Ridder, Daniel Weibel, Miriam Sturkenboom	*In Preparation*			
Quantifying outcome misclassification in multi-database studies: the case study of pertussis in the ADVANCE project	Rosa Gini, Caitlin Dodd, Kaat Bollearts, Claudia Bartolini, Giuseppe Roberto	*Submitted: Vaccine*			
Masking by vaccines in pediatric signal detection: A study in the EudraVigilance database	Caitlin Dodd, Alexandra Pacurariu, Osemeke U. Osokogu, Daniel Weibel, Carmen Ferrajolo, Dang H. Vo, Benedickt Becker, Jan A. Kors, Miriam Sturkenboom	Pharmaco-epidemiology and Drug Safety	27	11	2018
Incidence rates of narcolepsy diagnoses in Taiwan, Canada, and Europe: a method for rapid assessment of potential safety issues on a population level: the SOMNIA study	Caitlin Dodd, Maria de Ridder, Wan-Ting Huang, Daniel Weibel, Maria Giner-Soriano, Silvia Perez-Vilar, Javier Diez-Domingo, Larry Svenson, Salah Mahmud, Bruce Carleton, Monika Naus, Jeffrey C. Kwong, Brian Murray, Lisen Arnheim-Dahlstrom, Lars Pedersen, Rosa Morros, Steven Black, Miriam Sturkenboom	PLOS ONE	13	10	2018
Narcolepsy and adjuvanted pandemic influenza A (H1N1) 2009 vaccines: a global investigation	Daniel Weibel, Miriam Sturkenboom, Steven Black, M.D., Maria de Ridder, Caitlin Dodd, Jan Bonhoeffer, Ann Vanrolleghem, Nicoline van der Maas, Gert Jan Lammers, Sebastiaan Overeem, Angela Gentile, Norberto Giglio, Vanesa Castellano, Jeffrey C. Kwong, Brian J. Murray, Karen Cauch-Dudek, Diana Juhasz, Michael Campitelli, Alexandre N. Datta, Ulf	Vaccine	36	41	2018

	Kallweit [&], Wan-Ting Huang, Yu-Shu Huang, Chung-Yao Hsu, Hsi-Chung Chen, Maria Giner-Soriano, Rosa Morros, Carles Gaig, Ester Tió, Silvia Perez-Vilar [&], Javier Diez-Domingo, Francisco Javier Puertas, Lawrence W. Svenson, PhD, Salah Mahmud, Bruce Carleton, Monika Naus, Lisen Arnheim-Dahlström, Lars Pedersen, Frank DeStefano, MD, MPH, Tom T. Shimabukuro, MD, MPH, MBA				
The impact and longevity of measles-associated immune suppression: a population based matched cohort study	Caitlin N Dodd, Kartini Gadroen, Gwen MC Masclee, Maria AJ de Ridder, Daniel Weibel, Michael J. Mina, Bryan T. Grenfell, Miriam CJM Sturkenboom, David A.M.C. van de Vijver, Rik L de Swart	BMJ Open	8	11	2018
Bell's palsy and influenza(H1N1)pdm09 containing vaccines: a self-controlled case series	Wijnans, Leonoor; Dodd, Caitlin; Weibel, Daniel; Sturkenboom, Miriam;	PLOSONE	12	5	2017
Enhancing vaccine pharmacovigilance in low and middle-income countries: Proof-of-concept study on aseptic meningitis and immune thrombocytopenic purpura following measles-mumps containing vaccination	Silvia Perez-Vilar, Daniel Weibel, Miriam Sturkenboom, Steven Black, Christine Maure, Jose Luis Castro, Pamela Bravo-Alcántara, Caitlin N. Dodd, Silvana A. Romio, Maria de Ridder, Swabra Nakato, Helvert Felipe Molina-León, Varalakshmi Elango, Patrick L.F. Zuber and the WHO Global Vaccine Safety-Multi Country Collaboration	Vaccine	35	32	2017
Pandemic influenza vaccine & narcolepsy: simulations on the potential impact of bias	Dodd, Caitlin; Wijnans, Leonoor; de Ridder, Maria; Romio, Silvana; Weibel, Daniel; Overeem, Sebastiaan; Lammers, Gert Jan; Bonhoeffer, Jan; Black, Steve; Sturkenboom, Miriam;	Expert review of vaccines	15	5	2016
Drug Safety Monitoring in Children: Performance of Signal Detection Algorithms and Impact of Age	Dodd, Caitlin; Osokogu, Osemeke U; Pacurariu, Alexandra; Kaguelidou,	Drug safety	39	9	2016

Stratification	Florentia; Weibel, Daniel; Sturkenboom, Miriam CJM;				
International collaboration to assess the risk of Guillain Barré Syndrome following Influenza A (H1N1) 2009 monovalent vaccines	Dodd, Caitlin N; Romio, Silvana A; Black, Steven; Vellozzi, Claudia; Andrews, Nick; Sturkenboom, Miriam; Zuber, Patrick; Hua, Wei; Bonhoeffer, Jan; Buttery, Jim;	Vaccine	31	40	2013
Other Manuscripts					
Pharmacoepidemiological safety studies in children: a systematic review	Osokogu, Osemeke U; Dukanovic, Julijana; Ferrajolo, Carmen; Dodd, Caitlin; Pacurariu, Alexandra C; Bramer, Wichor M; tJong, Geert; Weibel, Daniel; Sturkenboom, Miriam CJM; Kaguelidou, Florentia;	Pharmaco-epidemiology and drug safety	25	8	2016
Vitamin D status and hospitalisation for childhood acute lower respiratory tract infections in Nigeria	Ahmed, Patience; Babaniyi, IB; Yusuf, KK; Dodd, Caitlin; Langdon, Gretchen; Steinhoff, Mark; Dawodu, Adekunle;	Paediatrics and international child health	35	2	2015
Comparison of a new transport medium with universal transport medium at a tropical field site	Schlaudecker, Elizabeth P; Heck, Joan P; MacIntyre, Elizabeth T; Martinez, Ruben; Dodd, Caitlin N; McNeal, Monica M; Staat, Mary A; Heck, Jeffery E; Steinhoff, Mark C;	Diagnostic microbiology and infectious disease	80	2	2014
Dyadic adjustment and spiritual activities in parents of children with cystic fibrosis	Grossoehme, Daniel H; Szczesniak, Rhonda; Dodd, Caitlin; Opipari-Arrigan, Lisa;	Religions	5	2	2014
Administrative data misclassifies and fails to identify nephrotoxin-associated acute kidney injury in hospitalized children.	Schaffzin JK, Dodd CN, Nguyen H, Schondelmeyer A, Campanella S, Goldstein SL.;	Hospital Pediatrics	4	3	2014
A simulation study to compare three self-controlled case series approaches: correction for violation of assumption and evaluation of bias	Hua, Wei; Sun, Guoying; Dodd, Caitlin N; Romio, Silvana A; Whitaker, Heather J; Izurieta, Hector S; Black, Steven; Sturkenboom, Miriam CJM; Davis, Robert L; Deceuninck, Genevieve;	Pharmacoepide miology and drug safety	22	8	2013
Parental Reporting of Allergenic Food Introduction: A Study of Urban and Suburban Populations	Hartman, Heather N; Dodd, Caitlin; Assa'ad, Amal H;	Journal of Allergy and Clinical Immunology	131	2	2013
Neonatal outcomes after influenza immunization during pregnancy: a randomized controlled trial	Steinhoff, Mark C; Omer, Saad B; Roy, Eliza; El Arifeen, Shams; Raqib, Rubhana; Dodd, Caitlin; Breiman,	Canadian Medical Association Journal	184	6	2012

	Robert F; Zaman, K;				
Etiology and seasonality of viral respiratory infections in rural Honduran children	Schlaudecker, Elizabeth P; Heck, Joan P; MacIntyre, Elizabeth T; Martinez, Ruben; Dodd, Caitlin N; McNeal, Monica M; Staat, Mary A; Heck, Jeffery E; Steinhoff, Mark C;	The Pediatric infectious disease journal	31	11	2012
Pregnancy modifies the antibody response to trivalent influenza immunization	Schlaudecker, Elizabeth P; McNeal, Monica M; Dodd, Caitlin N; Ranz, Joseph B; Steinhoff, Mark C;	Journal of Infectious Diseases	206	11	2012
Antibody persistence in mothers one year after pneumococcal immunization in pregnancy	Schlaudecker, Elizabeth P; Steinhoff, Mark C; Omer, Saad B; Roy, Eliza; Arifeen, Shams E; Dodd, Caitlin N; Altaye, Mekibib; Raqib, Rubhana; Breiman, Robert F; Zaman, K;	Vaccine	30	34	2012
Low prevalence of Chlamydia trachomatis infection in non-urban pregnant women in Vellore, S. India	Vidwan, Navjyot K; Regi, Annie; Steinhoff, Mark; Huppert, Jill S; Staat, Mary Allen; Dodd, Caitlin; Nongrum, Rida; Anandan, Shalini; Verghese, Valsan;	PLOSONE	7	5	2012

ACKNOWLEDGEMENTS

My mentors: Miriam, Steve, Maria, Daniel – for asking me difficult questions and trusting me to answer them. My life would look so very different without you. Thanks for seeing the challenges and inviting me to take them on. Thanks especially to Miriam and Steve for setting me on this path and helping me along at every step, to Maria for her patience and diligence, and to Daniel for helping to keep it all in perspective.

The mathematicians, statisticians and epidemiologists I admire most: Paddy, Heather, Nick, Kaat, Rosa, Hector, Bruce, Maria, Miriam, Steve – for encouraging me, fielding my questions, exposing me to new knowledge via new experiences, and collaborating with me. You are some of the most creative researchers I have had the good fortune to know. Thank you for being great role models.

The teachers I've been lucky enough to have: Mom/Mrs. Crowley, who encouraged my curiosity and especially my love of literature both in and outside the classroom. Dave Aggen, who first recognized and encouraged my interest in math. David Wetta, Jared Gardner, and Michael Cavanagh, who taught me to take nothing for granted when approaching a problem critically. Although they are professors of Literature, their instruction still influences the way I think about problems, whether literary, artistic, philosophical or statistical. Bob Downer, who urged me to join the inaugural class of his Masters in Biostatistics program, mentored me through the next 2 years, and managed to convince me that I could become a decent statistician. Thanks to each of you for challenging me and teaching me how to learn.

My former colleagues in the US: Matt, Mekibib, Rhonda, Mark, Elizabeth, Josh, Adekunle and Gretchen – You helped to turn an inexperienced but enthusiastic young biostatistician into a scientist with her own research interests. Thank you for taking me seriously and sharing your knowledge with me. Thanks especially to Gretchen for being a friend, mentor, and reliable source of refreshingly acerbic wit.

My colleagues and coauthors: Alexandra, Swabra, Osemeke, Rosa, Kaat, John, Kartini, Leonoor, Silvia, Silvana, Rik, David, Wan-Ting, Lionel, Francois, Klara, Daniel, Marius – Working with each of you has shown me a new facet of talent and intelligence. Thank you for being a reminder of how many varied and wonderful ways there are to be impressed by others.

Andreas and Solange – I can't imagine working and living here without your help. Thank you for listening and clueing me in to the inner workings of the Dutch bureaucracy, and especially for the open friendliness you displayed every time I stepped into your office.

Mees and Kris – Thank you for being patient with my questions and rescuing me from forgotten passwords and sluggish servers. You are more than your job titles would have anyone believe. Special thanks to Mees for always graciously agreeing to take time away from his important tasks to help me with mine.

The staff and families of Eduard van Beinum school: Especially Judith and Onno, Marjolein and Rolf, Floor and Jasper, Elma and Jos, Barbara, Sharon, Judith, and Martijn. Thank you for welcoming my daughters into your lives and the community.

Donna, Kassi, Hilary – For encouragement, support, and the love you show to Nora and Millie. Thank you.

Sheila and Larry – Thank you for always being interested in me and the girls, for making time to see us, and for telling us how proud you are. It means a lot to hear it from you.

The Cincinnati Rollergirls – Thank you for providing a community, gut-busting laughter, and the gains in confidence without which I may not have decided to give this PhD a shot. Special thanks to GiGi GoGlitter, Hot Slice, Railroad, Buckhead Betty, Sista Shovechild, Sk8 Crime, Jocelyn Bones, Jungle Lacy, and coach Quad Almighty.

My NL friends: Marius, Daniel, Kartini, Marten – for showing me perspectives I would not otherwise have seen, for reminding me of what I've accomplished and what I'm capable of, and for laughs. You've made these years brighter. Special thanks to Marius for pizza lunches and Vagabonding, to Kartini for therapeutic venting sessions and a healthy dose of skepticism, to Marten for an introduction to Bible scholarship and an endless stream of Big Lebowski references, and to Daniel for invaluable life advice.

Jenny – Best friend a lady could ask for, constantly knocking me out with the originality of your thought and the steadfastness of your love.

Brendan – Thank you for being an unusually kind, supportive, and protective big brother. I have always been able count on you to make me laugh, tell me I'm the cleverer and better looking sibling (some may disagree), and reassure me that I am, in fact, doing alright.

Mom and Dad – Thank you for encouraging my curiosity and providing me with endless opportunities to satisfy it, helping me gain independence, and supporting me in my decisions even when it meant you had to worry about me (again). I have had every advantage and opportunity. Thank you for the sacrifices you made to provide them.

Richard – I've learned more from knowing and loving you than I could ever learn from a silly old PhD. Thanks for always reflecting back to me the best parts of myself, for being the place I belong, and for loving me even when it's not easy. Thank you for sharing yourself and your life so generously with me. You're my dude, dude.

Nora and Millie – The world's greatest two reasons to get out of bed in the morning and come back home at night. I am so grateful for your bravery and resilience here in a foreign country far from your family and your culture. I am beyond proud to have watched you grow and change during our time in The Netherlands. Thank you for being the people I want most to make proud, and thank you most of all for the laughs and hugs and adventures we share. I love you two doodles more than I can say.

About the author

Caitlin Dodd was born in Duluth, Minnesota in 1980 before proceeding to live an itinerant life in places as far-flung as Salida, Colorado; New York, New York; Frankfort, Kentucky; Chicago, Illinois; Grinnell, Iowa; Cincinnati, Ohio; and 4 different cities in Michigan. She completed a bachelors degree in literature with a minor in film in 2001 at Grinnell College, followed by 5 more itinerant years of restaurant and factory work – manning assembly lines and slinging hash from Texas to Michigan – before the arrival of her first daughter, Nora, in 2006. It was this arrival that spurred her to go back to school at Grand Valley State University where she was persuaded to join a fledgling Masters in Biostatistics program led by professor Bob Downer. She completed her masters in 2009 following the birth of her second daughter, Millie.

Following completion of her masters she joined the Global Health Center at Cincinnati Children's Hospital as a biostatistician where she was introduced to the study of vaccines through the work of Steve Black and Mark Steinhoff. From there, she moved on to the biostatistical consulting unit within Cincinnati Children's where she worked as a statistical consultant on projects as varied as identification of acute kidney injury in medical records, religiosity and adherence to cystic fibrosis treatment, and introduction of allergenic foods in urban vs. suburban infants, all while continuing to study vaccine safety with Steve Black. It was through Steve that Caitlin was introduced to Miriam Sturkenboom with whom she collaborated, together with a global network of researchers, on a study of Guillain-Barré Syndrome following pandemic influenza vaccines. Through her work on this study, Caitlin made her first trips to Rotterdam where she was offered a PhD position.

In 2013, Caitlin accepted and moved to Rotterdam with her two daughters to begin her PhD studies under the promotorship of Miriam Sturkenboom and the supervision of Maria de Ridder and Daniel Weibel.

Caitlin spends her time outside of work trying and failing and trying again to be a decent piano player, making surreal drawings for all those she knows and loves, writing short stories very, very slowly, reading as much and as widely as she can, traveling the world with her partner in search of sunshine and live music, washing and folding copious quantities of laundry, swatting mosquitoes, feeding the cat, unclogging drains, and enjoying *almost* every moment she gets with her children.

Caitlin currently works as a statistician and pharmacoepidemiologist in the Global Health group in the Julius Center at Utrecht University Medical Center. She lives in Amsterdam with her two daughters, Nora and Millie, and their cat, Moonshadow.

PHD PORTFOLIO

Oral Presentations

Narcolepsy and Pandemic H1N1 Vaccine: A Simulation Study to Explore the Effect of Bias, International Society for Pharmacoepidemiology Annual Conference, 2014, Taipei.

*Narcolepsy incidence rates in the SOMNIA (Systematic observational method for narcolepsy and influenza immunization) stud*y, International Society for Pharmacoepidemiology Annual Conference, 2016, Dublin.

The Impact and Longevity of Measles-Associated Immune Suppression, International Society for Pharmacoepidemiology Annual Conference, 2016, Dublin.

Poster Presentations

International Collaborative Case Series Safety Monitoring for Pandemic 2009 H1n1 Vaccines: Estimation of the Risk of Guillain-Barré Syndrome, International Society for Pharmacoepidemiology Annual Conference, 2012, Barcelona.

Estimating incidence of adverse events following vaccination in observational databases when exposure information is unavailable: A contribution from the advance project, International Society for Pharmacoepidemiology Annual Conference, 2018, Prague

Signal detection in VAERS and EudraVigilance using disproportionality and time to onset method and their combination, International Society for Pharmacoepidemiology Annual Conference, 2018, Prague

Teaching

EU2P: European Training Program in Pharmocvigilance and Pharmacoepidemiology, Instructor Study design, Statistical Analysis, 2018-2019