Use of routinely collected electronic healthcare data for postlicensure vaccine safety signal detection: a systematic review

Yonatan Moges Mesfin, Allen Cheng, Jock Lawrie, Jim Buttery

ABSTRACT

Background Concerns regarding adverse events following vaccination (AEFIs) are a key challenge for public confidence in vaccination. Robust postlicensure vaccine safety monitoring remains critical to detect adverse events, including those not identified in prelicensure studies, and to ensure public safety and public confidence in vaccination. We summarise the literature examined AEFI signal detection using electronic healthcare data, regarding data sources, methodological approach and statistical analysis techniques used.

Methods We performed a systematic review using the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines. Five databases (PubMed/Medline, EMBASE, CINAHL, the Cochrane Library and Web of Science) were searched for studies on AEFIs monitoring published up to 25 September 2017. Studies were appraised for methodological quality, and results were synthesised narratively.

Result We included 47 articles describing AEFI signal detection using electronic healthcare data. All studies involved linked diagnostic healthcare data, from the emergency department, inpatient and outpatient setting and immunisation records. Statistical analysis methodologies used included non-sequential analysis in 33 studies, group sequential analysis in two studies and 12 studies used continuous sequential analysis. Partially elapsed risk window and data accrual lags were the most cited barriers to monitor AEFIs in near real-time.

Conclusion Routinely collected electronic healthcare data are increasingly used to detect AEFI signals in near real-time. Further research is required to check the utility of non-coded complaints and encounters, such as telephone medical helpline calls, to enhance AEFI signal detection.

Trial registration number CRD42017072741

INTRODUCTION

Vaccination is one of the most effective public health interventions. Current immunisation programmes provide protection against up to 26 diseases and prevent an estimated 2–3 million deaths every year. It is estimated that 1.5 million more deaths could be saved through further increasing vaccination coverage of existing vaccines. However, this remarkable success has been challenged due to vaccine safety concerns and increasing vaccine hesitancy, largely due to fear of adverse event following immunisation (AEFIs). Notably, following the sharp reduction of incidence of vaccine-preventable diseases the public attention to AEFI has increased. This can result in loss of confidence in vaccination, a resultant drop in vaccine coverage and eventually lead to a re-emergence of controlled disease (figure 1). Hence, timely detection of potentially causally related adverse events (AEs) and more rapidly refute spurious claims regarding AEs using real-world data is critical to maintain the community and providers confidence in vaccine programmes. Nevertheless, recent analysis of global AEFI reporting found that more than 36% of WHO

Key questions

What is already known?

► Adverse event(s) following immunisation (AEFI) signal detection has primarily relied on passive surveillance reporting.

What are the new findings?

► AEFIs signal monitoring using population-based electronic health records (EHRs) is increasing, but has been primarily limited to diagnostic data from hospital settings.

► Continuous sequential (rapid cycle) analysis method allows AEFIs signal monitoring in near real-time.

► Data delays (data accrual lags) are the key challenges to perform near real-time AEFI monitoring using EHRs.

What do the new findings imply?

► A complementary and efficient AEFI signal monitoring system is feasible using EHRs.

► Further research is required to evaluate the utility of syndromic data/proxy measures to enhance the timeliness of monitoring AEFIs.
member countries do not have a functional postlicensure safety monitoring system for vaccines.5

Postlicensure AEFIs monitoring is often classified into three stages: signal detection, signal refinement and signal confirmation. A vaccine safety signal is defined as 'reported information on a possible causal relationship between an adverse event and a vaccine, the relationship being unknown or incompletely documented previously'.6 Generally, AEFI signal detection has been undertaken using passive surveillance or active surveillance system. Passive surveillance systems, the prevailing AEFI monitoring system, monitor reports of AEs that are spontaneously submitted by healthcare providers, vaccinated individuals/their caregivers or others. Its wide population coverage allows for detection of new and unanticipated AEs but has limitations of under-reporting and imprecise risk estimates due to lack of appropriate denominator data.7 According to the 2015 Global Vaccine Safety Initiative meeting report, low passive AEFI reporting rates are a significant barrier to detect vaccine safety signal timely.8 In contrast, active surveillance systems are resource intensive and takes substantial time to achieve the required sample size to study rare AEs. Hence, their use in many settings are largely limited to investigate signals detected from the passive surveillance systems, literature review or possible preclicensure trial safety questions.7 9 10

Encouragingly, in recent years, new studies have shown that routinely collected electronic health records (EHRs) can be used as an alternative data source to monitor for AEFI signals in near real-time.11 12 For example, in the USA, newly marketed vaccines are monitored for potential AEFIs weekly using the Vaccine Safety Datalink (VSD) collaboration between the US Centre for Disease Control and eight healthcare organisations. In the VSD, patient

Figure 1  Potential stage in the evolution of an immunisation programme, vaccine safety. Diagram adapted from Chen et al. The Vaccine Adverse Effect Reporting System (VAERS). Vaccine 1994:12(6):542–50.

Considering the increasing availability of EHRs and the necessity of further improving the capacity of vaccine safety monitoring, particularly in low-income and middle-income countries, EHRs can offer an alternative data source to establish complementary active AEFI surveillance systems. By systematically summarising these literature, we intend to provide valuable information for countries considering establishing AEFI signal detection system based on EHRs. Therefore, we aimed to: (1) describe the features of postlicensure vaccine safety studies employing EHRs primarily for safety signal detection and (2) catalogue the nature of data sources, methodological approaches and analysis techniques applied.

METHODS

Search strategy

A systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines,16 as provided in online supplementary file 1. The protocol was registered at the international prospective register of systematic reviews (registration number CRD42017072741). We searched OVID Medline (1946 to September week 3 2017), OVID Embase (1974 to 2017 September 10), the Cochrane Library, Scopus and Web of Science. Comprehensive search terms for all databases were developed in consultation with a medical librarian to identify all potentially relevant studies. A combination of keywords and Medical Subject Headings (MeSH) were used in each database with appropriate adjustment. Final searches were performed on 25 September 2017. An example of the search strategy used in Ovid MEDLINE is shown in online supplementary file 1. In addition, bibliographies of relevant studies, conference papers/proceedings and grey literature databases, such as who.int and greylit.org, were searched to identify further important and unpublished studies.

Studies selection criteria and screening

We included studies primarily focussing on AEFI signal detection using EHRs. Studies were included regardless of vaccine type, population group studied, study setting and methodology used. However, studies based on passive pharmacovigilance data or administrative (claim) data; studies conducted solely to test or verify the previously identified signals and feasibility studies or studies conducted to evaluate methodologies were excluded from the review. We also excluded non-English records and conference abstracts.
Search results were downloaded and managed in EndNote X8. Articles were screened in three stages (titles alone, abstracts and then full-text review) based on the PRISMA flow of information (figure 2). At the initial stage, titles and abstracts were screened to remove duplicate records and studies clearly outside the scope of the review. Then, two reviewers conducted a full-text review to assess the eligibility based on the inclusion criteria. Study screening stages and the reasons for articles exclusion during full-text review are described in figure 2.

Quality assessment and data extraction

We used a checklist adapted from the Food and Drug Authority (Best Practices for Conducting and Reporting Pharmaco-epidemiologic Safety Studies Using EHR). Many of the critical appraisal tools extensively used to appraise observational studies, such as Ottawa-Newcastle tool and strengthening the reporting of observational studies in epidemiology (STROBE), are not suitable for evaluating pharmaco-epidemiological studies and public health surveillance as they are reasonably different from the standard epidemiological studies. The lead author (YMM) assessed risk of bias of all the included studies, and the second independent reviewer (TK) evaluated 25% of the studies randomly for verification. As there was no substantial risk of bias identified, we considered all appraised studies for the final review. The methodological quality and risk of bias assessment criteria were:

- Well defined research questions.
- Sample representativeness.
- Clear inclusion and exclusion criteria.
- Appropriateness of study design and comparison groups.
- Follow-up (risk interval) long enough for the events to occur.
- Appropriateness of data integration method, when relevant.
- Adjustment of confounders.

Employed appropriate statistical analyses method.
- Used objective criteria to measure outcomes.

The lead author consistently extracted the required data using pretested data abstraction template. The following information were extracted across the included studies:

- Study author.
- Publication year.
- Study setting and period.
- Data source(s) and nature of the data (diagnostic vs prediagnostic).
- Study design(s) employed.
- Studied population.
- Vaccine(s) and AE(s) studied.
- Statistical analysis approaches and signal detection method used.
- Frequency of assessment.
- Method(s) of controlling confounders reported and challenges reported.
- Main findings (signal(s) identified or not).

Data analysis

Key features of the studies are described quantitatively. Results from the selected studies are synthesised in a narrative analysis. The structure of the detailed review includes: vaccines monitored; AEs studied; study design(s) used; data analysis approach and signal detection method employed.

Patient and public involvement statement

No patient data were considered in this study.

RESULT

Studies identified and characteristics

After removal of duplicate articles, we screened the titles and abstracts of 606 articles and excluded articles clearly out of the scope of this review. Then, we screened the remaining 235 full-text articles according to the exclusion criteria (figure 2). Studies could be excluded for more than one reason. Forty-seven articles, conducted between 2002 and 2017, were included in the final synthesis. No studies were excluded based on quality or bias.

Almost all studies included in this review were conducted in the USA (n=45). Two additional studies were conducted in the UK and Taiwan. A considerable number of studies (n=13, 28%) assessed the safety of vaccines administered to high-risk groups (pregnant women or elderly subjects). Fourteen (30%) studies assessed the AEFIs in near real-time (table 1).

Vaccines studied

Multiple types of vaccines, including live, inactivated, monovalent and combined, were monitored after licensure for potential AEFI. Seasonal influenza vaccines (trivalent inactivated influenza vaccines (TIIV), live attenuated influenza vaccines, monovalent influenza vaccines and live attenuated monovalent influenza vaccines) were most frequently studied (n=17), followed by combined...
Table 1  Summary characteristics of selected studies

<table>
<thead>
<tr>
<th>Study characteristics</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
</tr>
<tr>
<td>Retrospective</td>
<td>37</td>
</tr>
<tr>
<td>Prospective</td>
<td>10</td>
</tr>
<tr>
<td>Data source</td>
<td></td>
</tr>
<tr>
<td>Immunisation record linked with:</td>
<td></td>
</tr>
<tr>
<td>Outpatient, emergency department and inpatient data</td>
<td>35</td>
</tr>
<tr>
<td>Emergency department and Inpatient data</td>
<td>8</td>
</tr>
<tr>
<td>Outpatient and inpatient data</td>
<td>3</td>
</tr>
<tr>
<td>Outpatient (general practice) data</td>
<td>1</td>
</tr>
<tr>
<td>Study type</td>
<td></td>
</tr>
<tr>
<td>Near real-time surveillance</td>
<td>14</td>
</tr>
<tr>
<td>Phase IV observation study</td>
<td>33</td>
</tr>
<tr>
<td>Study design</td>
<td></td>
</tr>
<tr>
<td>Self-controlled study</td>
<td></td>
</tr>
<tr>
<td>Self-controlled risk interval</td>
<td>22</td>
</tr>
<tr>
<td>Self-controlled case series</td>
<td>4</td>
</tr>
<tr>
<td>Cohort study</td>
<td></td>
</tr>
<tr>
<td>Historical comparison (current vs historical design)</td>
<td>20</td>
</tr>
<tr>
<td>Concurrent/Parallel comparison group</td>
<td>9</td>
</tr>
<tr>
<td>Case-crossover study</td>
<td>2</td>
</tr>
<tr>
<td>Studied outcomes of interest</td>
<td></td>
</tr>
<tr>
<td>Preselected adverse events</td>
<td>35</td>
</tr>
<tr>
<td>All medically attended events</td>
<td>12</td>
</tr>
<tr>
<td>Analysis method</td>
<td></td>
</tr>
<tr>
<td>Non-sequential analysis</td>
<td>33</td>
</tr>
<tr>
<td>Group sequential analysis</td>
<td>2</td>
</tr>
<tr>
<td>Continuous sequential (rapid cycle) analysis</td>
<td>12</td>
</tr>
</tbody>
</table>

AEIFs studied and data source

Most of the reviewed studies (n=35) studied preidentified AEs using a fixed postvaccination risk interval. AEs were selected based on the safety concerns from passive surveillance reports and prelicensure clinical trials. Frequently studied AEs were Guillain-Barré syndrome, febrile convulsions, seizures, anaphylaxis, meningitis/encephalitis and local reactions. Potential maternal and infant outcome (AEFIs), such as pre-eclampsia/eclampsia, maternal death, small for gestational age, preterm birth, stillbirth and neonatal death were also evaluated. Studied AEFIs were mainly identified using International Classification of Diseases (ICD) Clinical Modification codes as well as relevant ICD-9 or ICD-10 codes from electronic records (outpatient, inpatient and emergency department settings). In some studies, patients’ charts/medical records were manually reviewed to verify the AEs.

In this review, 14 statistically elevated vaccine-AE pairs (signals) were detected, and 6 were confirmed. These were measles, mumps, rubella and varicella vaccine and seizure/febrile convolution.38–43 2010–2011 TIIV and febrile seizure,57 monovalent rotavirus vaccine and intussusception,61 2014–2015 TIIV and febrile seizures48 and Tdap vaccine and chorioamnionitis.41

Figure 3  Type of vaccines studied by the selected studies.

Study designs employed

Self-controlled design was the most frequently used study design (n=22),18–21 25 27 28 30–34 36 38 39 44 46–48 53 57–59 62 63 followed by cohort design with historical comparison (also called observed vs expected analysis) (n=20).18 22–26 29 34 38 39 43 45 47–49 57 60 61 63 64 Self-controlled design can be self-controlled risk interval (SCRI) or self-controlled case series (SCCS). Cohort design with concurrent/parallel comparison group,19 20 29 40–42 50–52 mostly to examine vaccines administered to pregnant women, and case-crossover study designs were also employed.28 32 Of note, 18 studies (38.3%) employed more than one study design; of these, SCRI and current versus historical designs were often used together.23 34 38 39 47 48 57 63

Statistical analysis and signal detection method

Two broad data analytic approaches, non-sequential analysis and sequential analysis, were employed to identify elevated risk of AEs associated with a given vaccine. In studies that employ a non-sequential analysis approach (n=33), statistical tests are performed after all the data are collected/accumulated. Detailed description of these studies and their analytic approaches are provided in online supplementary file 2. The sequential analysis approach allows repeated examination of data to check for AEFI increased occurrence. This was implemented in two different ways in the included studies: (i) as group sequential analysis (n=2), which involved a periodic statistical test and limited number of statistical tests over
time and (ii) as continuous sequential analyses (n=12), also called ‘rapid cycle analysis’, which involved a weekly statistical test until the end of the study period (table 2).

The choice of specific statistical tests was guided by the data analysis approach used. Standard analytic tests, such as logistic and Cox regression, were used to examine the data at the end of the study period (end-of-study analysis). A sequential hypothesis test statistic, the sequential probability ratio test (SPRT), was used to examine data for an elevated risk of AEFI continually over time. In particular, maximised sequential probability ratio test (MaxSPRT) was the most frequently applied sequential hypothesis test statistic. It has different versions: Poisson MaxSPRT, Binomial MaxSPRT and Conditional MaxSPRT (table 2). Further, supplementary analyses were performed to verify the detected signals and instances of elevated risks. These included temporal scan statistics, to evaluate clustering of events after vaccination, and case-centred regression and logistic regression.

**Confounder adjustment and potential challenges**

Many different potential confounders were measured including age, gender, chronic conditions, site, seasonality, trend, concomitant vaccines and delay in the arrival of patient data. Generally, studies adjusted confounding variables in three ways: using data restriction, matching and stratification (alone or in combination). Strategies chosen were often design-based and included the following: (i) using a matched control design to adjust baseline confounders and seasonal trends; (ii) using self-controlled design, which automatically addresses time-invariant confounders and (iii) adjusting the expected rate calculated from historical data. Interestingly, during analysis, MaxSPRT inherently allows controlling bias due to repeated tests. In this review, the most cited challenges, particularly in the case of continuous sequential analysis, were uncertainty in estimating background rates, outcome misclassification, partially elapsed risk window and late-arriving data (data accrual lags).

**DISCUSSION**

Routinely collected EHRs are increasingly used for the detection of AEFI signal besides for testing hypothesis based on known signals. Evidence from this review suggests that electronic healthcare data have a significant potential to establish a near real-time AEFI surveillance systems. All the included studies used coded diagnostic medical data to get information about the studied AEs. Further, non-pharmacovigilance studies have also suggested that alternative non-coded medical information, such as telephone triage data and ambulance data, have potential for near real-time syndromic surveillance and rapidly detection of outbreak signal.66-67

A near real-time surveillance systems involves continuous checking (rapid cycle analysis (RCA)) of the EHRs for an elevated occurrence of AEs as the new data are added over the study period. It was first used to evaluate the safety of meningococcal conjugate vaccine using electronic healthcare data from the VSD in the USA,14 though Davis et al established its feasibility by replicating the previously recognised rotavirus-intussusception signal.68 Since then, we identified 12 studies that examined AEFI signal using RCA method.14 22 24 29 39 43 47 48 57 61 62 64 The RCA method has been also used based on an alternative data sources other than EHRs. For example, in the UK, H1N1 vaccine was monitored using passive surveillance data,66 and in Australia seasonal influenza vaccines have been monitored since 2015, based on data collected directly from consumers using SMS-messaging and email (AusVaxSafety).70

The near real-time AEFI surveillance systems use sequential analysis approach, primarily MaxSPRT, to continuously evaluate data for signals while adjusting bias due to multiple testing. MaxSPRT is an improved type of the classical SPRT, which uses a two-sided alternative hypothesis and a predefined relative risk (RR) value usually other than 1. MaxSPRT uses one-sided composite alternative hypothesis by defining the RR usually as >1 to declare statistically significant risk.71 The key advantage of MaxSPRT over the classical SPRT is that it helps to minimise the risk of late detection of AEs due to an incorrect choice of RR and make it suitable for data monitoring more frequently.14 Indications, advantages and weakness of both classical and MaxSPRT, including the three variants of MaxSPRT, are provided in table 3.24 47

As vaccines are often recommended for all persons in a given age group, traditional epidemiological cohort and case-control designs are usually not suitable to study vaccines AEs after licensure. The main reasons include an inadequate number of comparison groups (unvaccinated individuals), concern regarding comparability of the vaccinated to unvaccinated groups (selection bias), insufficient power and timeliness.72 Rather, self-controlled design (SCRI and SCCS) and cohort design, with a historical comparison, are the preferred design choice in postlicensure vaccine safety studies (table 4). In self-controlled design, comparisons are made with individuals in two different periods, vaccination risk period and control period. The incidence of AEFI is compared between prespecified postvaccination risk period and control period (unexposed period).73 Studies showed that including a prevaccination control period is essential to facilitate timely data analysis for vaccines administered in a short period, mostly in case of seasonal influenza vaccine. However, if there are clinical confounders that are a contraindication for vaccination (eg, allergic reaction) or indications for vaccination (eg, seizure disorder), a prevaccination control period is not recommended.39 47 48 57 74 75
<table>
<thead>
<tr>
<th>Study, Country</th>
<th>Data sources and period</th>
<th>Study design</th>
<th>Study subject</th>
<th>Vaccine studied</th>
<th>Adverse events studied</th>
<th>Analysis approach and signal detection method</th>
<th>Main finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yih, 2009 USA</td>
<td>VSD (from August 2005 to May 2008)</td>
<td>Prospective active surveillance with observed vs expected analysis</td>
<td>10–64 years of age</td>
<td>Tdap (new vaccine)</td>
<td>Encephalopathy-encephalitis-meningitis, paralytic syndromes, seizures, cranial nerve disorders (including Bell's palsy) and GBS.</td>
<td>Weekly sequential analysis, PMaxSPRT</td>
<td>No increased risks were identified for any of the outcomes over the course of 145 weeks surveillance</td>
</tr>
<tr>
<td>Klein, 2010 USA</td>
<td>VSD (from January 2006 to October 2008)</td>
<td>Prospective active surveillance with observed vs expected analysis</td>
<td>Children age 12–23 months</td>
<td>MMRV</td>
<td>Seizures and fever</td>
<td>Weekly sequential analysis, PMaxSPRT</td>
<td>Signal for seizure during days 7–10 was identified and confirmed, IRR=1.98 (1.43–2.73)</td>
</tr>
<tr>
<td>Huang, 2010 Taiwan</td>
<td>(2009/2010 season)</td>
<td>Prospective active surveillance with SCRI and observed vs expected analysis</td>
<td>≥6 months old age</td>
<td>H1N1 vaccine (LAMV and MIV)</td>
<td>Neurological, allergic and haematological AEs.</td>
<td>Weekly sequential analysis, BMaxSPRT and PMaxSPRT</td>
<td>No increased risks were identified for any of the outcomes over the course of 22 weeks follow-up</td>
</tr>
<tr>
<td>Belongia, 2010 USA</td>
<td>VSD (May 2006 to May 2008)</td>
<td>Prospective active surveillance with observed vs expected analysis</td>
<td>Infant aged 4–48 weeks</td>
<td>Penta-Valente rotavirus (new vaccine)</td>
<td>Intussusception and other (seizures, meningitis/encephalitis, myocarditis, Gram negative sepsis, gastrointestinal bleeding and Kawasaki syndrome)</td>
<td>Weekly sequential analysis, PMaxSPRT</td>
<td>No increased risks were identified for any of the outcomes over the course of 164 weeks surveillance</td>
</tr>
<tr>
<td>Lee, 2011 USA</td>
<td>VSD (November 2009 to April 2010)</td>
<td>Prospective active surveillance with SCRI and observed vs expected analysis</td>
<td>≥6 months old age</td>
<td>H1N1 and seasonal influenza vaccine</td>
<td>11 potential neurological, allergic and cardiac AEs.</td>
<td>Weekly sequential analysis, BMaxSPRT and PMaxSPRT</td>
<td>Signal was observed for Bell's palsy at week 20, but not confirmed after further analysis</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study, Country</th>
<th>Data sources and period</th>
<th>Study design</th>
<th>Study subject</th>
<th>Vaccine studied</th>
<th>Adverse events studied</th>
<th>Analysis approach and signal detection method</th>
<th>Main finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gee, 2011 USA&lt;sup&gt;29&lt;/sup&gt;</td>
<td>VSD (August 2006 to October 2009)</td>
<td>► observed vs expected and ► Cohort with concurrent comparison</td>
<td>9–26-year-old female</td>
<td>Quadrivalent human papillomavirus vaccine (HPV4)</td>
<td>GBS, stroke, venous thromboembolism (VTE), appendicitis, seizures, syncope, allergic reactions, and anaphylaxis</td>
<td>► Weekly sequential analysis, PMaxSPR and exact sequential analysis ► Supplementary analysis: Temporal cluster analysis and Case-centred and logistic regression analysis</td>
<td>Excess risk for appendicitis was identified but not confirmed</td>
</tr>
<tr>
<td>Tse, 2012 USA&lt;sup&gt;37&lt;/sup&gt;</td>
<td>VSD (August 2010 to February 2011)</td>
<td>observed vs expected analysis and SCRI designs</td>
<td>Children ages 6–59 months</td>
<td>TIV</td>
<td>Febrile seizures in the 0–1 days following first dose TIV</td>
<td>Weekly analysis, both BMaxSPRT and PMaxSPRT</td>
<td>Excess risk for seizures identified and confirmed, IRR=2.4 (1.3–4.7)</td>
</tr>
<tr>
<td>Wise, 2012 USA&lt;sup&gt;60&lt;/sup&gt;</td>
<td>Emerging Infections Programme (EIP) (October 2009 to May 2010)</td>
<td>Retrospective active surveillance with SCRI design</td>
<td>All individuals who received the vaccine</td>
<td>Influenza A (H1N1) monovalent vaccines</td>
<td>GBS during the 42 days following vaccination</td>
<td>► Weekly sequential analysis, PMaxSPRT ► Sensitivity analysis ► Temporal cluster analysis</td>
<td>Excess risk for GBS was identified, not confirmed</td>
</tr>
<tr>
<td>Tseng, 2013 USA&lt;sup&gt;60&lt;/sup&gt;</td>
<td>VSD (from April 2010 to January 2012)</td>
<td>Observed vs expected analysis</td>
<td>1 month to 2 years</td>
<td>PCV13 Vaccine</td>
<td>Febrile seizures, encephalopathy, urticaria and angioedema, asthma, anaphylaxis, thrombocytopenia, Kawasaki disease</td>
<td>► Group sequential analysis (12 repeated tests were performed)</td>
<td>Excess risks for encephalopathy and Kawasaki disease identified but not confirmed</td>
</tr>
<tr>
<td>Nelson, 2013 USA&lt;sup&gt;60&lt;/sup&gt;</td>
<td>VSD (September 2008 to January 2011)</td>
<td>Prospective active surveillance with observed vs expected analysis</td>
<td>children aged 6 weeks to 2 years</td>
<td>DTaP-IPV-Hib (combination)</td>
<td>MAF, seizure, meningitis/encephalitis/myelitis, series no anaphylactic allergic reaction; not formally tested—anaphylaxis, GBS, invasive Hib disease, all hospitalisation</td>
<td>► Group sequential testing, Poisson MaxSPRT (11 repeated tests were conducted) ► Sub group end-study-analysis</td>
<td>No increased risks were identified</td>
</tr>
<tr>
<td>Weintraub, 2014 USA&lt;sup&gt;61&lt;/sup&gt;</td>
<td>VSD (April 2008 to March 2013)</td>
<td>Retrospective active surveillance with observed vs expected analysis</td>
<td>Infants ages of 4 and 34 weeks</td>
<td>Monovalent Rotavirus vaccine</td>
<td>Intussusception within 7 days following vaccination</td>
<td>► Weekly sequential analysis, PMaxSPRT ► Temporal cluster analysis ► Exact logistic regression</td>
<td>Increased risk identified and confirmed, IRR=9.4 (1.4–103.8)</td>
</tr>
</tbody>
</table>

Table 2 Continued
<table>
<thead>
<tr>
<th>Study, Country</th>
<th>Data sources and period</th>
<th>Study design</th>
<th>Study subject</th>
<th>Vaccine studied</th>
<th>Adverse events studied</th>
<th>Analysis approach and signal detection method</th>
<th>Main finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kawai, 2014 USA</td>
<td>VSD (September 2012 to February 2013)</td>
<td>Retrospective active surveillance with SCRI and observed vs expected analysis</td>
<td>6 months to 17 years and 2–49 years</td>
<td>TIV, LAIV (first-dose vaccine)</td>
<td>Seizures, GBS, encephalitis and anaphylaxis</td>
<td>Weekly sequential analysis, BMaxSPRT and PMaxSPRT End of surveillance Logistic regression</td>
<td>No increased risks for any of the outcomes were identified</td>
</tr>
<tr>
<td>Daley, 2014 USA</td>
<td>VSD (January 2009 to September 2012)</td>
<td>Prospective active surveillance</td>
<td>4–6 years old DTaP-IPV combination</td>
<td>Meningitis/encephalitis, seizures, stroke, GBS, Stevens–Johnson syndrome and anaphylaxis</td>
<td>Weekly sequential analysis, PMaxSPRT and conditional MaxSPRT Posthoc analysis</td>
<td>No increased risks for any of the outcomes were identified</td>
<td></td>
</tr>
<tr>
<td>Li, 2016 USA</td>
<td>VSD (June 2013 to April 2015)</td>
<td>Retrospective active surveillance with SCRI and observed vs expected analysis</td>
<td>≥6 months old age</td>
<td>First dose of IIV3, IIV4 and LAIV4</td>
<td>Acute disseminated encephalomyelitis, anaphylaxis, Bell’s palsy, encephalitis, GBS, febrile seizures and transverse myelitis</td>
<td>Weekly sequential analysis using both BMaxSPRT and PMaxSPRT End of surveillance analysis after all the data have been collected</td>
<td>Excess risks for febrile seizure were identified and confirmed after vaccination of—IIV3: IRR=5.25 (1.57–1.75) and IIV4: IRR=12.3 (2.5–58.9)</td>
</tr>
</tbody>
</table>

BMaxSPRT, binomial-based maximised sequential probability; CMaxSPRT, conditional maximised sequential probability; GBS, Guillain-Barré syndrome; IIV3, trivalent inactivated influenza vaccine; IIV4, quadrivalent inactivated influenza vaccine; IRR, incidence rate ratio; LAIV, live attenuated influenza vaccine; LAMIV, live attenuated monovalent influenza vaccine; MIV, monovalent influenza vaccine; MMRV, measles, mumps, rubella and varicella vaccine; PCV13, 13-valent pneumococcal vaccine; PMaxSPRT, Poisson-based maximised sequential probability test; TIV, trivalent influenza vaccine; VSD, vaccine safety data-link.
### Table 3
Sequential statistical approaches for postlicensure vaccine safety surveillance (description, indication and challenges)

<table>
<thead>
<tr>
<th>Statistical approaches</th>
<th>General description</th>
<th>Advantage/indication</th>
<th>Challenges/weakness</th>
</tr>
</thead>
</table>
| Group sequential analysis               | ▶ Involves repeated (periodic) analyses overtime as data accumulate, at regular or irregular interval.  
▶ Compares the test statistic to a prespecified signalling threshold, and stops if the observed test statistic is more extreme than the threshold | ▶ Commonly used in clinical trials  
▶ More appropriate when data updates are less frequent  
▶ Yield increased study power for a given sample size | ▶ Does not allow to capture the safety problems as soon as possible  
▶ Very complex to compute  
▶ Limited ability to control potential confounders |
| Continuous sequential analysis (rapid cycle analysis) | ▶ Allows examination of data frequently (as often as desired) over time.  
▶ Surveillance starts as soon as uptake of the vaccine starts or delayed until a pre-set number of events occur | ▶ Allows to monitor the vaccine safety problems in real-time  
▶ Suitable to identify true safety signals sooner. This method can signal after single AEs, if that event occurs sufficiently early.  
▶ Require updated data in a real-time or in a continuous fashion | ▶ All data related to vaccinations and AEFIs may not be available timely for analysis (data accrual lags)  
▶ The risk windows might be not fully elapsed for some AEFIs at the time of each analysis (partially elapsed window), particularly in case of influenza vaccine  
▶ Inherently reduces statistical power |

### Signal detection method/ statistical test

| Binomial-based MaxSPRT | Based on the binomial distribution  
▶ Events occurring among vaccine exposed individuals or time periods compared with the number of events among unexposed individuals to the studied vaccine/matched periods | Best fit for self-controlled designs  
▶ More suitable when the AEs are relatively common  
▶ Account bias due to multiple looks at a data | Limited ability to control potential confounders |
| Poisson-based MaxSPRT | Assumes a Poisson distribution  
▶ Compare the observed number of events in a given preidentified risk period with a historical data or the scientific literature  
▶ Does not depends on choice of RR, it uses a one-sided composite alternative hypothesis of RR>1 | More suitable when AEFIs are very rare  
▶ Minimise the risk of late detection of AEFIs due to an incorrect choice of RR  
▶ Adjust for multiple looks at a data | Relies on having accurate background rate of the outcomes for comparison  
▶ Does not consider uncertainty in the estimation of expected rates, if the data are limited  
▶ Limited ability to control potential confounders |
| Conditional-based MaxSPRT | Assumes a Poisson process for the cumulative person-time to observe a number of AEFIs | Accounts for uncertainty in historical data  
▶ Adjust for multiple looks at a data | Assumes constant event rates are in historical and surveillance data  
▶ Limited ability to control potential confounders |

AE, adverse event; AEFI, adverse events following immunisation; MaxSPRT, maximised sequential probability ratio test; RR, relative risk.

A cohort study design with a historical comparison is used frequently for detecting AEFI signals. This design compares the observed incidence of AEFI in the risk period after vaccination of the studied vaccine(s) against the expected incidence of AEFI projected based on the historical data. It helps to improve the timeliness of detecting the AEFI signal because only data for the risk window is collected rather than waiting for data for
<table>
<thead>
<tr>
<th>Study design</th>
<th>Population</th>
<th>Comparison groups</th>
<th>Strength and preference</th>
<th>Weakness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cohort study design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study with historical comparison group, also called current vs historical design</td>
<td>Individual vaccinated with the vaccine of interest</td>
<td>Historical incidence rate of AEFIs calculated from historical data on individuals that have not been exposed to the vaccine of interest vs Incidence of AEFIs in the prespecified risk period/window following vaccination</td>
<td>It has greater statistical power to detect rare AEFIs signal earlier</td>
<td>Highly dependent on accurate estimation of background incidence rates of the AEFIs for comparison</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>It is less affected by data lags as it only collects for the risk window, rather than both for risk and comparison windows</td>
<td>It may be subjected to difference in confounders between current and historical vaccinees, seasonality and secular trends in AEFIs, diagnostic or coding criteria</td>
</tr>
<tr>
<td>Cohort study with concurrent/parallel comparison group (matched or not)</td>
<td>Individual vaccinated with the vaccine of interest</td>
<td>Incidence of AEFIs in the prespecified risk period/window following vaccination vs Incidence of AEFIs among individuals without the vaccine of interest or totally non vaccinated individuals</td>
<td>Reduce the likelihood of false or missed signals due to secular trends in disease, diagnostic patterns or coding criteria</td>
<td>Difficult to get adequate number of unvaccinated control group, in case of studying routinely administered vaccines</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>May be subjected to bias due to difference in characteristics of vaccinated and unvaccinated groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>For a rare AE, it may not provide the earliest possible signal</td>
</tr>
<tr>
<td><strong>Self-controlled design</strong></td>
<td>Vaccinated cases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCRI</td>
<td></td>
<td>Within subject comparison</td>
<td>Automatically control time-invariant confounders that vary between individual, such as sex, socioeconomic status</td>
<td>Has less statistical power due to fewer events occurring in the shorter control interval</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incidence of AEs in the predefined risk period following vaccinations vs Incidence of AEFIs in the predefined control/non risk period</td>
<td>Less prone to miscategorization of exposure, vaccinated cases are considered to collect information</td>
<td>Less suitable to capture subacute or chronic AEFIs for example, autoimmune disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Confounded by indication</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>It is not free from bias due to time-varying confounders for example, Age and seasonality</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Selecting a risk interval that is too wide or too short may bias the risk estimate relative to the true risk window.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitive to indication bias</td>
</tr>
<tr>
<td>Study design</td>
<td>Population</td>
<td>Comparison groups</td>
<td>Strength and preference</td>
<td>Weakness</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SCCS</td>
<td>Primarily vaccinated persons, but unvaccinated persons and experienced the AEs can be considered</td>
<td>▶ Within subject comparison ▶ Incidence of AEFIs in the predefined risk period following vaccinations vs Incidence of AEFIs in the control period (time period before or after vaccination)</td>
<td>▶ Inherently control time-invariant confounders ▶ Can be advantageous when identification of a vaccinated group is challenging and the outcome is rare</td>
<td>▶ Less suitable to capture subacute or chronic AEFIs ▶ More susceptible to bias because of time-varying confounders, as the observation period is often longer than SCRI ▶ Problem with defining risk interval (selecting a risk interval that is too wide or too short may bias the risk estimate relative to the true risk window). ▶ Sensitive to indication bias</td>
</tr>
<tr>
<td>Case-crossover design</td>
<td>All individuals who are vaccinated and cases</td>
<td>Subjects serve as their own matched controls with defined by prior time periods in the same subject</td>
<td>▶ Preferred method for studying risk of acute AEFIs ▶ Robust to time-invariant confounders by making within-person comparisons</td>
<td>▶ Does not address confounders that vary over time ▶ Susceptible to exposure trend bias for example, due to change in policy for a vaccine ▶ Sensitive to indication bias</td>
</tr>
</tbody>
</table>

AE, adverse event; AEFI, adverse event following immunisation; SCCS, self-controlled case series; SCRI, self-controlled risk interval.
the comparison window. However, studies showed that accurate baseline risk estimation is a very challenging task, and it may introduce bias if the historical population are considerably different from the studied population. Nevertheless, this problem can be minimised through simultaneous use of the self-controlled design as they have complementary strengths (table 4).

The essential requirement to conduct a near real-time AEFI surveillance based on EHRs is the availability of timely data. Both data accrual lag and partially elapsed risk window, the risk windows might not be fully elapsed for some AEs at the time of each analysis, can deter performing RCA. Data accrual lag in EHRs can occur due to several reasons and the level of delay may vary depending on the outcomes studied. A study from UK showed that up to 30 days or more are required to completely record AEFI diagnoses at general practice level. Two studies were included in this review, and methodological evaluation studies suggested that various design-based measures can be taken for adjusting partially elapsed risk window and data accrual lags. These include: (i) calculating the expected counts of AEFIs comparable to the elapsed risk window length; (ii) restricting comparison periods proportional to the elapsed risk period or (iii) AEFIs occurring in later weeks in the risk window can be ignored if the matching weeks in the control period have not elapsed.

CONCLUSION
The utility of routinely collected EHRs for AEFI monitoring globally has been demonstrated, with most published experience drawn from US literature. In addition, the advancement of statistical analysis techniques and RCA provide a significant potential to detect AEFI signal in near real-time.

To date, AEFI monitoring based on EHRs use is limited to diagnostic medical information. Potential incorporation of other electronic health information, including non-coded complaints and encounters, offers further opportunities to improve AEFI real-time surveillance systems to help maintain safe immunisation programmes and maximise confidence in those programmes.

REFERENCES


