Methodological frontiers in vaccine safety: Qualifying available evidence for rare events, use of distributed data networks to monitor vaccine safety issues, and monitoring the safety of pregnancy interventions

Situation Paper

Prepared by:

- Caitlin Dodd, Miriam Sturkenboom (UMC Utrecht)
- Nick Andrews (Public Health England)
- Steve Black (Global Vaccine Safety Data Network)
- Saad B. Omer (Yale University)
- Helen Petousis-Harris (University of Auckland)

This draft situation paper will be finalized after the GACVS anniversary symposium.

The authors alone are responsible for the views expressed in this draft situation paper and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.
Executive Summary

This situation paper describes the current state of the art in methodology for assessing vaccine safety in observational data sources with a focus on rare events, assessment of pregnancy exposures, and criteria by which to grade existing evidence.

While vaccines are rigorously tested for safety and efficacy in clinical trials, these trials do not include enough subjects to detect rare adverse events, and they generally exclude special populations such as pregnant women. For these reasons, it is necessary to conduct post-marketing safety assessments using observational data sources.

The study of rare events has been enabled in recent decades through the increasing availability of large linked databases in combination with development of case-only, case-coverage and case-centered methods such as the self-controlled case series. The study of rare events has been further bolstered by the use of distributed data networks which serve to increase power and sample size. These distributed data networks necessitate common protocols, definitions, data models, and analytics and the processes of developing and employing these tools are rapidly evolving. Particularly, a variety of models for analysis in distributed data networks, including privacy-preserving methods, have been proposed and utilized.

Although pregnant women have typically been excluded from clinical trials, vaccines are increasingly indicated for use in pregnancy, and additional vaccines are undergoing pregnancy trials. Assessment of vaccine safety in pregnancy is complicated by the physiological changes a woman undergoes during pregnancy, the challenges of mother-child linkage in databases, and the need for long-term infant follow-up. Potential sources of bias including differential access to and utilization of antenatal care, immortal time bias, seasonal timing of pregnancy, and unmeasured determinants of adverse pregnancy outcomes have yet to be fully explored.

Assessment of evidence generated in post-marketing studies must be evaluated prior to making recommendations. Typical tools for conducting this assessment downgrade evidence from observational data and prioritize evidence from randomized controlled trials. However, inclusion and exclusion criteria from trials limit the generalizability of results. For this reason, real world evidence based upon real world data is increasingly being used for safety assessments, and new tools for evaluating real world evidence have been developed.

The future of vaccine safety surveillance, particularly for rare events and in special populations, comprises the use of big data in single countries as well as in collaborative networks. This move
toward the use of real world data requires continued development of methodologies to generate and assess real world evidence.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADVANCE</td>
<td>Accelerated Development of Vaccine Benefit-risk Collaboration in Europe</td>
</tr>
<tr>
<td>AEFI</td>
<td>Adverse Event Following Immunization</td>
</tr>
<tr>
<td>CDC</td>
<td>United States Centers for Disease Control</td>
</tr>
<tr>
<td>CDM</td>
<td>Common Data Model</td>
</tr>
<tr>
<td>EMA</td>
<td>European Medicines Agency</td>
</tr>
<tr>
<td>FDA</td>
<td>United States Food and Drug Administration</td>
</tr>
<tr>
<td>GBS</td>
<td>Guillain-Barré Syndrome</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>GVDN</td>
<td>Global Vaccine Data Network</td>
</tr>
<tr>
<td>GVS-MCC</td>
<td>Global Vaccine Safety Multi-Country Collaboration</td>
</tr>
<tr>
<td>HPV</td>
<td>Human Papillomavirus</td>
</tr>
<tr>
<td>LMIC</td>
<td>Low and Middle Income Country</td>
</tr>
<tr>
<td>MCO</td>
<td>Managed Care Organization</td>
</tr>
<tr>
<td>MMR</td>
<td>Measles, Mumps, and Rubella</td>
</tr>
<tr>
<td>OHDSI</td>
<td>Observational Health Data Sciences and Informatics</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality Adjusted Life Year</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomized Controlled Trial</td>
</tr>
<tr>
<td>ROBINS-I</td>
<td>Risk Of Bias In Non-Randomized Studies - of Interventions</td>
</tr>
<tr>
<td>RWE</td>
<td>Real World Evidence</td>
</tr>
<tr>
<td>SAGE</td>
<td>Strategic Advisory Group of Experts</td>
</tr>
<tr>
<td>SCCS</td>
<td>Self-Controlled Case Series</td>
</tr>
<tr>
<td>SOMNIA</td>
<td>Systematic Observational Method for Narcolepsy and Influenza Immunization Assessment</td>
</tr>
<tr>
<td>VAC4EU</td>
<td>Vaccine monitoring Collaboration for Europe</td>
</tr>
<tr>
<td>VAESCO</td>
<td>Vaccine Adverse Events Monitoring and Communication</td>
</tr>
<tr>
<td>VSD</td>
<td>Vaccine Safety Datalink</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Introduction and Rationale

The awareness that medicines can cause harm has been known since ancient times. However, it was the thalidomide incident of the late 1950’s and early 1960’s that stimulated the monumental shift to proactively regulate drugs internationally. While new mandates around drug registration were implemented, there were also new regulations for post marketing surveillance. Developments in ways to code adverse events emerged through the 60s and the 70s and continue today, and drug adverse event reporting systems have been established. In 1978 a feasibility study to evaluate standardizing spontaneous reports funded by the United States Food and Drug Administration (FDA) evolved to what is now known as the International Pharmacovigilance Monitoring Centre based in Uppsala, Sweden. However, while modern pharmacovigilance incorporates spontaneous reporting data and experience from over 128 countries exists to detect signals, the data are not suitable to estimate risk [1].

At the time of their licensure, vaccines have been assessed for efficacy, quality and safety in randomized controlled trials. While randomized studies are traditionally viewed as the gold standard for vaccine outcomes, they are usually limited by their sample size when it comes to detecting very rare events or longer-term outcomes [2]. This is where post marketing surveillance, based on observational studies, is required.

Pharmacoepidemiology is the study of the use and effects of drugs and other medicinal products in populations. It uses epidemiological methods to study effects in large numbers of people and can contribute important information about the effectiveness and safety of a vaccine that is not available from a pre-licensure clinical evaluation. The importance of observational studies in complementing clinical trials of drugs is a well-established cornerstone of pharmacoepidemiology. Observational studies in vaccine safety include a range of methodologies that essentially compare the occurrence of outcomes of interest among people exposed and unexposed to a vaccine. These include the cohort, case-referent (i.e. case-control) studies, self-controlled studies, analyses of secular trends (ecological studies), case series, and case reports, in descending order of levels of evidence [3]. Vaccine safety studies rely on the first three approaches to assess risk, while the others serve to generate hypotheses. In recent years, sequential analyses have also been used to prospectively monitor suspected vaccine safety concerns [4].

Over the past three decades, the ability to conduct observational studies has evolved to a point that extremely large, heterogeneous populations can be used to assess vaccine safety outcomes. Traditional methods have used cohort and case-control designs but more recently newer designs
such as the self-controlled case series have found favor in overcoming potential confounder and selection bias [5] (Table 1).

While monitoring the safety of vaccines to the highest standard achievable using the tools available is fundamentally ‘right’ and scientifically important, there is also a need to address perceptions about vaccine safety. In part due to the well-coordinated international anti-vaccine movement, vaccine hesitancy has risen to become one of the top ten threats to global public health. Achieving and maintaining public confidence in vaccines needs rapid and credible responses to vaccine safety concerns, real or perceived. It is concerns about vaccine safety that contribute the most to vaccine hesitancy and being able to offer reassurance and effective communications supported by high quality data is essential [6].

Another reason for conducting large observational post-licensure vaccine safety studies is to examine potential associations with rare events, differences between vaccine brands and formulations, and possible risks in special sub-populations. Rare events such as Guillain-Barre Syndrome pose challenges with respect to study power. For example, to detect a 2-fold increased relative risk with a background incidence of 1/100,000 and assuming a 1:1 ratio of vaccinated to non-vaccinated subjects, a study population of over 4.7 million is required [3]. Such risks cannot be assessed in clinical trials, nor among smaller populations. Only through large administrative data collections for populations in the millions can a vaccine-associated risk for rare events be assessed. In addition, the globalization of vaccine manufacturing increases the need to be able to compare brands, as illustrated by measles vaccines and infant mortality [7], mumps vaccines and aseptic meningitis [8], and influenza vaccines in association with febrile convulsions [9] or narcolepsy [10, 11]. Finally, there is also the need to assess the safety of vaccines in sub-populations such as pregnant women and their infants [12], people with certain health conditions such as immune suppression [13], and among different ethnic groups [14] and indigenous peoples [15]. These sub-populations are often excluded from RCTs. To these ends, linkable data for large and diverse populations is required.

In recent years there have been major advances in the tools and methodologies required to undertake the types of large, robust observational studies needed to address these diverse challenges. In 1990 the United States Centers for Disease Control and Prevention (CDC) in collaboration with large managed care organizations (MCOs) created the Vaccine Safety Datalink (VSD) [2]. This project linked medical events and demographic information with vaccine exposure information. Since then the system has evolved to include more populations, changes in the way data is collected, the development of near real-time data files, and the ability to monitor new vaccines or changes in vaccine use in near real-time [16]. Since the early 1990’s the number of
publications internationally referring to multiple observational databases has exploded thanks to increased existence of electronic databases, the appreciation of the importance of data-linkage, and information technology for analyzing multiple databases [17]. Recent collaborations such as VAESCO, ADVANCE, SOMNIA, and the WHO Global Vaccine Safety-Multi Country Collaboration have successfully incorporated data from multiple countries to produce estimates of vaccine safety and effectiveness [11, 18-21]. The purpose of this paper is to appraise current methodologies in post-marketing vaccine safety in terms of how these can be used to assess rare events and special populations, and how evidence derived from these studies can be evaluated.

Studying rare events

*Self-controlled case-only designs and other methods*

Over the past 20 years there have been an increasing number of epidemiological studies assessing vaccine safety for rare events which cannot be easily assessed in trials. This has been made possible through the availability of large linked databases and innovative tools for analysis. Whilst traditional case-control and cohort studies continue to be used, novel designs such as case-only, case-coverage and case-centered methods have been developed and evaluated for use with these databases [22-26]. Improved understanding of the benefits and limitations of the methods and data sources in terms of potential bias and confounding helps determine which are most suitable for addressing a specific adverse event and vaccine. Case-only designs, which only require an unbiased set of cases, are particularly useful for assessing rare events due to their efficiency in only needing cases and their elimination of time-invariant confounding factors. Such factors may be multifactorial and difficult to measure and could include frailty, socio-economic status, and levels of health care usage. The most commonly used case-only method is self-controlled case-series (SCCS) and its derivatives (e.g. self-controlled risk-interval) [22,24]. SCCS requires definition of a follow-up period in which cases and their vaccine history are obtained, along with risk intervals post vaccination and baseline intervals for comparison. Another method, developed prior to SCCS and applied in the GVSI-MCC study of measles and mumps containing vaccines, is the case crossover method in which vaccine exposure in risk windows just prior to the event is compared to earlier control windows within cases [20,23]. Case-coverage and referent designs use data on cases and their vaccination status and compare this to vaccination uptake in a referent population or cases which may be from a different data source [25, 26]. These novel tools have proved highly useful in recent decades for assessing rare events and with more availability of linked databases should continue to be used and developed.
Detection, power, specificity, bias and delayed onset

To study an adverse event, it is first necessary to identify it as a possible adverse reaction. Events likely to be missed and therefore requiring methods for identification are those that are rare, new, unexpected and with delayed onset after vaccination. Even when there is a signal, or other reason to study a rare event, it is often the case that this cannot be done rapidly due to the rarity of the event. For new vaccines this may mean that the full safety profile will remain unknown for a long time and this limitation should be documented and solutions developed. Large-linked databases and collaborative studies will enhance power, but many countries, including many LMIC, do not have such databases. This means that, for vaccines used only in these countries, enhanced safety assessment with epidemiological studies is limited. Case-only methods offer a good solution due to their efficiency but still require the ability to identify an unbiased set of cases with linked vaccination records. Whilst some adverse events are well diagnosed and coded in records, others may have non-specific symptoms with unclear onset meaning coded events have a low positive predictive value. Validation, whilst desirable, may not be possible or expensive. It is therefore important to understand methodologically the implications of assessing events where specificity may be lacking. Further work on bias analysis would be beneficial in this area as well as the merits of validation on subsets of cases. Finally dealing with confounding, particularly in studies where case-only methods are not appropriate remains a challenge, particularly with events with onset many months or years after vaccination.

Collaborative studies

Although most epidemiological studies of rare events to date have been done within countries, this requires the availability of data from large study populations with high vaccine exposure or over long periods of time. Collaborative studies in which data or results are combined between countries or regions within a country offer the opportunity to assess rare and very rare adverse events where numbers within one site are insufficient to address a question. They may also enable assessment within sub-populations such as pregnant women or using data from several smaller populations such as those obtained through sentinel hospitals. Collaborative studies of vaccine safety are complex due to the variety of potential signals, available data sources, and possible study designs. A solution that has been successfully employed within the US is the vaccine safety datalink (VSD) (see distributed data network section) where a common methodology and analysis is possible when using a network of healthcare organizations [27]. There have also been several international collaborative studies in the past decade. This was first done in Europe through the VAESCO consortium for Pandemic influenza and Guillain-Barre syndrome (GBS) (7 countries) and then internationally through the Global H1N1 GBS consortium (10 countries), the SOMNIA pandemic vaccine narcolepsy study (6
countries), and the proof-of-concept WHO Global vaccine Safety Multi Country Collaboration in LMICs assessing measles-containing vaccines association with thrombocytopenic purpura and aseptic meningitis (16 countries) [11, 18, 28-30]. There have also been bipartite collaborations such as Denmark and Sweden for HPV and autoimmune conditions [31,32]. Most recently the ADVANCE project created a system for monitoring vaccine benefits and risks and is now continuing as the Vaccine Monitoring Collaboration for Europe [19, 33-38]. Most of these studies have involved data sharing to a hub and whilst some have involved a single methodology, others have allowed different study designs. Key issues to have arisen are the cost and timeliness of such studies and in LMIC having sentinel hospitals that code discharges and link to vaccination data.

Distributed data networks

Common data models, protocols, analytics, definitions

Currently, distributed data networks can be described along two axes: first, whether they employ an ad-hoc or dedicated network and second, the extent to which they utilize common standards. Along the ad-hoc to dedicated network axis, collaborative studies can be implemented either through a group of data sources assembled to address a specific study question or through a sustained database network via which studies can be deployed rapidly. Many examples of distributed analyses conducted within ad-hoc networks exist (see Collaborative Studies section) while the only examples of dedicated networks specific to the study of vaccines are the VSD [2] and the ADVANCE/VAC4EU in Europe. Other dedicated networks of observational data sources include the FDA Sentinel System and the Observational Health Data Sciences and Informatics (OHDSI) community, of which only Sentinel has been employed in the assessment of vaccine safety [39-41].

These data networks also vary in the extent to which they employ common standards such as common protocols, definitions, data models, and analytics. Common study protocols are written and agreed upon by the network and implemented either centrally by distributing analysis scripts which can be run locally against data in a common data model or implemented locally by each database. Common definitions used to harmonize data from heterogeneous sources include those for AEFI developed by the Brighton Collaboration which require medical record review, or harmonization of codes from various coding systems both for events and vaccine exposures in retrospective database analyses [42-44]. Common data models harmonize the structure and content of the participating data sources to enable the use of common analytics developed centrally to run against all data sources. Finally, common analytics include data management, data quality assessment, and analysis
scripts which can be deployed against data in a common data model and, in a dedicated database network, can be modularized for re-use.

**Models for analysis**

Several models exist to analyze data within a distributed data network, the choice of which is dependent upon both the privacy restrictions at each contributing data site and the extent to which the network has utilized common standards. The models vary in how much individual-level information is shared, how analysis results are combined, and the autonomy of each site [45]. At the extreme of local autonomy and privacy preservation, a common protocol is implemented locally against data in its original format and estimates from all sites are meta-analyzed. An alternative which reduces local autonomy is one in which data is converted to a common data model and analysis is conducted locally using a program written centrally, after which results are meta-analyzed. This model can be adapted to allow for pooling of individual-level data by deploying centrally written scripts which generate aggregated or de-identified data sets which can be pooled centrally and analyzed. At the extreme of centralization and data sharing, all individual-level data are pooled and centrally analyzed.

**Data privacy and implications for 1 vs 2-stage meta-analysis**

Typically, studies are conducted in individual sites and synthesis of evidence is conducted through systematic reviews or meta-analysis. In distributed data networks, depending upon the model for analysis chosen, data may be retained locally at each site or shared at the aggregate or individual level. This choice depends primarily on privacy concerns and the desire for local control of the data. The approach of conducting analysis locally using only local data has the drawbacks that some data sources fall out of the analysis due to paucity of cases and resulting non-convergence, and that naturally occurring heterogeneity in vaccine formulations and schedules cannot be exploited. In the SOMNIA study, a hybrid approach was used, allowing for pooling of individual-level data from sites for which sharing was possible followed by meta-analytical pooling with results from sites for which sharing of individual-level data was not possible [11].

Approaches to facilitate 1-stage pooling have been discussed and include sharing of aggregate data sets or sharing of individual-level data sets. Currently, a lot of research is being conducted on approaches to privacy protecting distributed analysis methods. This includes development of DataSHIELD, which exploits the iterative fitting strategy of generalized linear models to produce coefficients and estimates from data processed in parallel identical to those from data pooled at the individual level [46]. More recently, simulation as well as empirical studies have shown that privacy-
protecting methods, such as sharing of aggregated data sets, perform similarly to analysis in individual level pooled data [47-49].

**Studying exposure in pregnancy**

Currently, vaccines are increasingly indicated for use in pregnancy, and additional vaccines are undergoing pregnancy trials. Vaccine safety assessment in pregnancy presents unique challenges as pregnancy is a physiologically dynamic state and safety considerations involve both the mother and the fetus. Levels of sex hormones estradiol and progesterone increase during the course of pregnancy and changes in sex hormones are associated with changes in the immune system. For example, as the pregnancy progresses, there is an increase in type 2 helper T-cell (Th2) responses and attenuation of type 1 helper T-cell (Th1); hence a Th1-to-Th2 shift in pregnancy [50]. However, the increases in sex hormones and associated immunological changes are not perfectly linear. Moreover, these changes occur relatively quickly. Therefore, exposure assessment in pregnancy should account for the gestational age at vaccination. Often the gestational age at vaccination is grouped by trimester for statistical analyses. However, given the pace of change in pregnancy, trimester-based classification of gestational age might be too crude. If there is sufficient power, week of gestation is a better measure while assessing exposure time in pregnancy.

Considerations of toxic and teratogenic effects (i.e. permanent anatomical, functional or developmental disruptions) on the embryo or fetus are dependent on the stage of development. Prior to implantation, adverse events can be a result of injury to a large proportion cells resulting in spontaneous abortion. If there is injury to a small number of cells, then often there is survival without abnormalities. In the embryonic period, i.e. 2-9 weeks of embryonic gestational age, there is substantial organogenesis. Therefore, for exposures in this period, outcomes of concern include malformations and altered function. After 9 weeks of gestation through term, there is a period of fetal growth, differentiation, and maturation; hence, outcomes of concern include preterm and small for gestational age birth, fetal death, minor malformations, and altered function.

Exposure in pregnancy also warrants infant follow up. While the duration of follow up depends on the vaccine and the outcome of interest, there is increasing consensus on following the infant for at least one year after birth -particularly in maternal vaccine trials.

Because conduct of trials in pregnant women has implications for the health and safety of both mother and infant, their conduct is limited, increasing the need for observational studies [51,52]. However, only a limited number of observational studies of vaccine exposure in pregnancy have been conducted and include assessments of influenza, tetanus, meningococcal, measles, mumps,
oral polio, pertussis, and yellow fever vaccines among others [53,54]. This limitation to the number of studies conducted in large observational databases may be related both to the difficulty of conducting mother-child linkage and the range of adverse pregnancy and neonatal outcomes which should be assessed in association with vaccine exposure. Additionally, observational studies of vaccination in pregnancy are prone to sources of biases including differential access to and utilization of antenatal care, immortal time bias related especially to late-term exposures, seasonal timing of pregnancy, and covariates related to adverse pregnancy outcomes such as nutrition, smoking, and alcohol use which are typically not captured in observational data sources [55]. Adverse pregnancy and birth outcomes must be assessed objectively using standardized definitions and, as is the case for all AEFI, availability of background rates is vital. Toward these ends, the GAIA Consortium has developed and tested case definitions for a set of maternal and neonatal AEFI [56]. The US Sentinel system has developed and included a dedicated mother-child linkage table in its common data model [57,58].

**Options for evaluating existing evidence**

When public health policy decisions are made regarding recommendations for vaccine usage, a potential safety issue, or regarding the cost effectiveness of a vaccine, decision makers often use criteria to evaluate the quality of the evidence that is available. The most common assessment tool is the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) criteria [59]. GRADE classifies studies into four levels of evidence – very low, low, moderate, and high – where evidence from randomized controlled trials (RCTs) is rated as high quality and evidence from observational data is rated as low quality.

Despite widespread acceptance by recommending bodies including the WHO SAGE committee, the assumption that observational data is routinely of low quality and randomized trials are always the best evidence source for recommendations is itself based upon low quality evidence and is increasingly being challenged [60, 61]. Frieden pointed out that while RCT data have good internal validity and assess efficacy in perfect situations, they have very low external validity, which is now generally accepted. RCTs have in-and exclusion criteria that select certain populations and follow-up that does not reflect reality. In fact, the FDA, EMA and other regulators and health technology organizations are now embracing real world data and real-world evidence (RWE) as the state of the art for safety assessments, because of the limited size and follow-up for RCTs. The 21st century Act places additional focus on the use of RWE for regulatory decision making [62]. The European Medicines Agency has developed the OPTIMAL framework and is increasingly using RWE for decision making [63].
While treatment assignment is not randomized in observational studies (leaving potential for confounding) and data is not always source verified (leaving space for bias), they still can offer a robust “real world” picture of the effectiveness and safety of a vaccine in the target populations that it is being used, including people who might not have been eligible to participate in a clinical trial. Study designs and methods have been developed to deal with confounding such as the self-controlled case series, and the use and matching on propensity scores emulates a clinical trial. Assessment of the correctness of the outcome can be validated and exposure data can be benchmarked [64]. The availability of electronic health records in many countries has provided the opportunity to assess vaccine coverage, safety, effectiveness and impact on large populations in a retrospective fashion. The Vaccine Safety Datalink in the US (8 million source population) and the ADVANCE/VAC4EU infrastructure in Europe (40 million source population) are able to rapidly assess the safety, coverage, effectiveness and benefit-risk of vaccines post marketing and the same infrastructures can been used to evaluate vaccine impact [65-67].

To illustrate the poor predictive value of some RCT data, when results of a model based upon real world data for effectiveness for pneumococcal conjugate vaccine were compared to the pre-licensure randomized trial for PCV7, it was evident that the pre-licensure RCT may have underestimated vaccine impact for pneumonia and otitis media and thus the data from the RCT found the vaccine cost $50,000 per QALY when in fact it was likely to be cost saving [68]. This disparity was likely due to the inability to assess herd effects in the clinical trial setting. In a study assessing potentially reduced efficacy of herpes zoster vaccine when administered concomitantly with 23-valent pneumococcal vaccine, no increased risk of herpes zoster was detected in subjects co-administered both vaccines. This evidence contradicts findings of reduced efficacy from clinical trials based upon measured antibody levels as a measure of efficacy [69]. Similarly, for rotavirus vaccines, while pre-licensure RCT trials with more than 100,000 individuals found no evidence of an increased risk of intussusception following two different RCTs, large studies using real world evidence post licensure found such a risk [70].

**Future Perspectives**

Current methodologies for post-marketing vaccine safety are sufficient, in general, to assess rare events and safety concerns in general populations. There needs to be further effort in identifying and applying these methods to special populations such as pregnant women, persons with co-morbidities, and diverse ethnicities. Moreover there is a need to look at combination schedules rather than separate vaccines.
**Studying rare events**

While the study of events too rare to assess within a single population is made possible through collaborative studies at larger scale, these studies also require tools for harmonization: common protocols, data models, definitions, and analytics. The future of vaccine safety surveillance comprises the use of big data in single countries as well as in collaborative networks and requires that networks have sustainable funding rather than being assembled ad-hoc. Progress toward this goal is evident in the VAC4EU consortium, which aims to develop a sustained vaccine-benefit risk network for Europe, and the Global Vaccine Data Network currently in development, which aims to maintain a global network of vaccine experts and data owners capable of quickly deploying vaccine safety studies in populations beyond North America and Europe [71].

While each of these approaches will lead to increases in the availability of real world evidence on vaccine safety, each requires access to large volumes of harmonized data, which in turn requires a sustained data network, alongside global capacity building in this area.

**Exposure in pregnancy**

Due to barriers to participation in trials for pregnant and lactating women, observational studies are of particular importance in the assessment of vaccine safety in pregnancy. This requires sustainable networks employing common data models that prioritize mother-child linkage and recording of variables which allow for accurate ascertainment of the timing of exposure during pregnancy. Additionally, there is much room for methodological development in addressing bias and confounding in studies of vaccine exposure during pregnancy.

**Grading existing evidence**

Our ability to use real world evidence is growing exponentially. Methods to deal with confounding and bias have matured and guidance documents on the transparency of reporting of such studies and new tools such as ROBINS-I have been developed to evaluate the potential for bias in observational studies [72,73].

It is time to reevaluate the GRADE criteria themselves and provide more balance in our evaluation of available evidence. It would seem more reasonable to argue, as Frieden has done, that there is no one perfect gold standard for evidence. Rather, RCTs and observational studies each have their own strengths and weaknesses and can complement one another as we make decisions regarding the vaccines that we use.
References


10. Sarkanen TO, Alakuijala AP, Dauvilliers YA, Partinen MM. Incidence of narcolepsy after H1N1 influenza and vaccinations: Systematic review and meta-analysis. 2018;38:177-86.


(H1N1) 2009 vaccine: multinational case-control study in Europe. BMJ. 2011 Jul 12;343:d3908. doi: 10.1136/bmj.d3908.


60. https://www.who.int/immunization/sage/Guidelines_development_recommendations.pdf
67. https://vac4eu.org


### Table 1. Post-marketing evidence generation in vaccine safety [74]

<table>
<thead>
<tr>
<th>Level of evidence</th>
<th>Designs</th>
<th>Data Sources</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal Detection/</td>
<td>Observed vs. Expected analyses</td>
<td>Spontaneous reports, Observational databases*</td>
<td>Binary signal/non-signal based upon pre-defined thresholds and observed values</td>
</tr>
<tr>
<td>Hypothesis Generation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scan statistics</td>
<td></td>
<td>Observational databases</td>
<td>Binary signal/non-signal based upon event clustering</td>
</tr>
<tr>
<td>Self-controlled case series</td>
<td></td>
<td>Spontaneous reports, Observational databases</td>
<td>Incidence rate ratio</td>
</tr>
<tr>
<td>Time-to-onset analyses</td>
<td></td>
<td>Spontaneous reports</td>
<td>Binary signal/non-signal based upon pre-defined thresholds and p-values</td>
</tr>
<tr>
<td>Sequential methods</td>
<td></td>
<td>Prospective cohorts</td>
<td>Binary risk/non-risk based upon pre-defined stopping rule</td>
</tr>
<tr>
<td>Hypothesis Strengthening</td>
<td>Ecological methods including Interrupted Time Series</td>
<td>Observational databases, Surveillance data</td>
<td>Incidence rate ratio, p-values for slope or level change (in ITS)</td>
</tr>
<tr>
<td>Hypothesis Testing</td>
<td>Cohort-based studies</td>
<td>Prospective cohorts, Observational databases</td>
<td>Absolute risk, hazard &amp; survival function</td>
</tr>
<tr>
<td></td>
<td>Case-referent studies</td>
<td>Observational databases</td>
<td>Odds ratio, hazard ratio</td>
</tr>
<tr>
<td></td>
<td>Self-controlled methods</td>
<td>Observational databases</td>
<td>Incidence rate ratio</td>
</tr>
</tbody>
</table>

*Observational databases include population-based health database such as administrative billing and electronic health record databases