www.advance-vaccines.eu

Accelerated Development of VAccine beNefit-risk Collaboration in Europe

Grant Agreement nº115557

D4.2 Report on appraisal of vaccine safety methods

WP4 – Methods for burden of disease, vaccination coverage, vaccine safety and effectiveness, impact and benefit-risk monitoring

V2.0
Final

Lead beneficiary: P95
Date: 24/10/2014
Nature: Report
Dissemination level: PU

TABLE OF CONTENTS

DOCUMENT INFORMATION

Grant Agreement Number	115557	**Acronym**	ADVANCE
Full title	Accelerated Development of VAccine beNefit-risk Collaboration in Europe		
Project URL	http://www.advance-vaccines.eu		
IMI Project officer	Angela Wittelsberger (angela.wittelsberger@imi.europa.eu)		

Deliverable	**Number**	4.2	**Title**	Report on appraisal of vaccine safety methods
Work package	**Number**	4	**Title**	Methods for burden of disease, vaccination coverage, vaccine safety and effectiveness, impact and benefit-risk monitoring

Delivery date	**Contractual**	Month 12	**Actual**	Month 9
Status	Current version / V2.0		Draft ☐ Final X	
Nature	Report × Prototype ☐ Other ☐			
Dissemination Level	Public ☐ Confidential ×			

Authors (Partner)	Nick Andrews (PHE), Kaat Bollaerts (P95), Catherine Cohet (GSK), Caitlin Dodd (EMC), Paddy Farrington (OU), Elisa Martin Merino (AEMPS), Angela Noufaily (OU), Catherine Panozzo (SP), Dominique Rosillon (GSK), Christel Saussier (ANSM), Suzie Seabroke (MHRA), Lionel Van Holle (GSK), Heather Whitaker (OU)			
Responsible Author	Paddy Farrington		**Email**	c.p.farrington@open.ac.uk
	Partner	OU	**Phone**	+44 1908 654840

Description of the deliverable	Broad view of methodological requirements for vaccine risk assessment, from the contrasting perspectives of the public, public health agencies, regulatory authorities, vaccine manufacturers, and academia.
Key words	Adverse event, Epidemiology, Risk, Statistical methods, Surveillance, Vaccine

DOCUMENT HISTORY

NAME	DATE	VERSION	DESCRIPTION
Nick Andrews (PHE), Kaat Bollaerts (P95), Catherine Cohet (GSK), Caitlin Dodd (EMC), Paddy Farrington (OU), Elisa Martin Merino (AEMPS), Angela Noufaily (OU), Catherine Panozzo (SP), Dominique Rosillon (GSK), Christel Saussier (ANSM), Suzie Seabroke (MHRA), Lionel Van Holle (GSK), Heather Whitaker (OU)	08/06/2014	1.0	First draft
Kaat Bollaerts (P95), Vincent Bauchau (GSK), Anna Cantarutti (ASL Cremona), Catherine Cohet (GSK), Stephen Evans (LSHTM), Dan Fife (JRDUS), Angela Noufaily (OU), Dominique Rosillon (GSK), Suzie Seabroke, (MHRA), Miriam Sturkenboom (EMC), Lionel Van Holle (GSK), Heather Whitaker (OU)	20/08/2014	1.0	Comments
Paddy Farrington (OU)	25/08/2014	1.1	Revised draft
Steven Bailey (Pfizer), Katherine Donegan (MHRA), Marianne Van Der Sande (RIVM)	15/10/2014	1.1	Internal review
Paddy Farrington (OU)	21/10/2014	2.0	Final version

DEFINITIONS AND ABBREVIATIONS

ABBREVIATIONS

- Participants of the ADVANCE Consortium are referred to herein according to the following codes:
  - **EMC**. Erasmus Universitair Medisch Centrum Rotterdam (Netherlands) - Coordinator
  - **UNIBAS**. Universitaet Basel (Switzerland) - **Managing entity of the IMI JU funding**
  - **EMA.** European Medicines Agency (United Kingdom)
  - **ECDC.** European Centre for Disease Prevention and Control (Sweden)
  - **SURREY.** The University of Surrey (United Kingdom)
  - **P95.** P95 (Belgium)
  - **SYNAPSE.** Synapse Research Management Partners, S.L. (Spain)
  - **OU.** The Open University (United Kingdom)
  - **LSHTM.** London School of Hygiene and Tropical Medicine (United Kingdom)
  - **PEDIANET.** Società Servizi Telematici SRL (Italy)
  - **KI.** Karolinska Institutet (Sweden)
  - **ASLCR.** Azienda Sanitaria Locale della Provincia di Cremona (Italy)
  - **AEMPS.** Agencia Española de Medicamentos y Productos Sanitarios (Spain)
  - **AUH.** Aarhus Universitetshospital (Denmark)
  - **UTA.** Tampereen Yliopisto (Finland)
  - **WIV-ISP.** Institut Scientifique de Santé Publique (Belgium)
  - **MHRA.** Medicines and Healthcare products Regulatory Agency (United Kingdom)
  - **SSI.** Statens Serum Institut (Denmark)
  - **RCGP.** Royal College of General Practitioners (United Kingdom)
  - **RIVM.** Rijksinstituut voor Volksgezondheid en Milieu * National Institute for Public Health and the Environment (Netherlands)
  - **GSK.** GlaxoSmithKline Biologicals, S.A. (Belgium) – EFPIA Coordinator
  - **SP.** Sanofi Pasteur (France)
  - **NOVARTIS.** Novartis Pharma AG (Switzerland)

- **SP MSD.** Sanofi Pasteur MSD (France)
- **CRX**. Crucell Holland BV (Netherlands)
- **PFIZER**. Pfizer Limited (United Kingdom)
- **TAKEDA**. Takeda Pharmaceuticals International GmbH (Switzerland)

- **ADE** Adverse Drug Event
- **AEFI** Adverse Event Following immunisation
- **AF** Attributable Fraction
- **AR** Attributable Risk
- **ARE** Asymptotic Relative Efficiency
- **ARL** Average Run Length
- **CDC** Centers for Disease Control and Prevention
- **CPRD** Clinical Practice Research Database
- **CUSUM** Cumulative Sum
- **DALY** Disability-Adjusted Life Year
- **DT** Diphtheria and Tetanus
- **DTaP** Diphtheria, Tetanus and acellular Pertussis
- **DTP** Diphtheria, Tetanus and Pertussis
- **ECDC** European Centre for Disease Prevention and Control
- **FDA** Food and Drug Administration
- **FDAAA** Food and Drug Administration Amendments Act
- **GP** General Practitioner
- **GRIP** Global Research in Paediatrics
- **HPV** Human Papilloma Virus
- **OHDSI** Observational Health Data Sciences and Informatics
- **O-E** Observed - Expected
- **OR** OddsRatio
- **MaxSPRT** Maximised Sequential Probability Ratio Test
- **MMR** Measles, Mumps and Rubella
- **NNH** Number Needed to Harm
- **OMOP** Observational Medical Outcomes Partnership
- **OPV** Oral Polio Virus
- **POC** Proof Of Concept
- **RI** Relative Incidence
- **RR** Relative Risk
- **PRISM** Post-Licensure Rapid Immunization Safety Measurement
- **RD** Risk Difference
- **SCCS** Self-Controlled Case Series
- **SPRT** Sequential Probability Ratio Test
- **UK** United Kingdom
- **US** United States
- **USA** United States of America
- **VAERS** Vaccine AdverseEvents Reporting System
- **VAESCO** Vaccine AdverseEvents Surveillance and Communication
- **VSD** Vaccine Safety Datalink

- **Assessment Criteria**

- M1: Statement of the measure of effect directly available from the method.
- S1: Asymptotic relative efficiency and power of the estimator.
- S2: Finite sample performance: bias and coverage probabilities.
- T1: Effectiveness for sequential and/or routine monitoring.
- T2: Operational sample size required for specified power.

- R1: Assumptions (statistical and epidemiological) of the method.
- R2: Robustness to failure of the assumptions.
- O1: Data requirements.
- O2: Complexity of implementation.

DEFINITIONS

- **Grant Agreement.** The agreement signed between the beneficiaries and the IMI JU for the undertaking of the ADVANCE project (115557).
- **Project.** The sum of all activities carried out in the framework of the Grant Agreement.
- **Work plan.** Schedule of tasks, deliverables, efforts, dates and responsibilities corresponding to the work to be carried out, as specified in Annex I to the Grant Agreement.
- **Consortium.** The ADVANCE Consortium, comprising the above-mentioned legal entities.
- **Project Agreement.** Agreement concluded amongst ADVANCE participants for the implementation of the Grant Agreement. Such an agreement shall not affect the parties' obligations to the Community and/or to one another arising from the Grant Agreement.

EXECUTIVE SUMMARY

We provide a broad view of methodological requirements for vaccine risk assessment, from the contrasting perspectives of the public, public health agencies, regulatory authorities, vaccine manufacturers, and academia. We focus on methods for signal strengthening and signal confirmation, rather than signal detection. These varied perspectives lead to the formulation of nine assessment criteria within five categories: measures of effect, temporality, statistical validity, robustness, and operational practicability.

We undertake an extensive review of study designs used for vaccine risk assessment, or potentially useable for this purpose. This covers designed vaccine introduction, cohort methods, case referent methods and self-controlled methods. We also consider sequential methods, methods for signal strengthening and other analytical features such as control for confounding and meta-analysis techniques.

We then assess these designs in the light of our criteria, focusing on the estimation of direct effects. This exercise is supplemented by three further components. First, we consider aspects of the interplay between signal detection and signal confirmation. Second, we rapidly review other projects of relevance to this component of ADVANCE, including the methodological investigations of the Observational Medical Outcomes Partnership (OMOP) and the practical implementations in the Vaccine Safety Datalink (VSD). Third, we describe some of the methods used to evaluate indirect adverse effects of vaccination.

In the final section, we discuss implications for integrating risk assessments into risk – benefit evaluations, focusing on: study designs, effect measures, and types of effect. We identify ten topics meriting further research. We recommend that four research components be included in the proof of concept studies to be undertaken as part of Work Package 5:

- Study of heterogeneity of vaccine risk between databases
- Evaluation of sequential methods for new vaccines
- Comparative evaluation of standard methods for vaccines in routine use
- Signal detection, strengthening and confirmation within a single database.

1. INTRODUCTION

1.1 Background on the ADVANCE project and the risk assessment task

The ADVANCE project seeks to develop '*a validated and tested best practice framework that could rapidly provide robust data on vaccine benefits and risks to support accelerated decision making*' (ADVANCE Description of Work, section 2.1.3). The present review of methods for vaccine safety assessment is undertaken to further this goal.

The evaluation of vaccine safety is the central component of any risk assessment for vaccines. The present review focuses on the methods used to quantify the risk of direct and indirect (that is, population level) adverse health effects that may be associated with vaccination. Other types of risk may be associated with vaccination programmes – for example, their wider social and political impact or the opportunity cost incurred by not undertaking other interventions perceived to be of lesser priority. These types of risks are not usually included in formal risk – benefit analyses, and are not considered here. We focus entirely on adverse health events in individuals associated, directly or indirectly, with vaccination, but to the exclusion of those associated with efficacy (for example, the perverse individual effects that vaccination can have on risk-related behaviour), to be dealt with in a separate report.

The present report is focused primarily on methods for vaccine safety assessment, rather than data or outcome events of interest, though the issue of data sources will arise in relation to multiple uses of data, and the evaluation of indirect risks is too bound up with the specificities of outcome events to ignore these completely. It is informed by the expertise of practitioners drawn from diverse stakeholder groups (clinicians, public health bodies, regulators, manufacturers, and academics), and by the findings of other projects relating to vaccine safety and to drug safety more widely.

It is appropriate at this stage to emphasize the specificities of safety assessments for vaccines, and how they differ from safety assessments for other pharmaceutical drugs. There are three main differences. First, vaccines are administered with the purpose of preventing ill-health, rather than treating illness, and are administered primarily, frequently on a very large scale, to healthy individuals, very often children. This profoundly affects how risks and benefits are weighed, and how the public perceives vaccine risks. It also imposes a challenging ethical imperative on health providers to monitor vaccination programmes so as to detect problems quickly if any are suspected, and to collect accumulating evidence of safety if there are none. In methodological terms, it implies that methods to assess vaccine safety must be powered to assess rare events, typically occurring at a rate well under 1 in 1000 doses, and to do so rapidly. Second, vaccinations are often administered with high coverage, according to age-related or seasonal schedules. From a methodological point of view, this implies that finding appropriate control groups may be difficult, and that confounding by age or season need to be taken into account in assessing possible associations between vaccination and health events, if these are also age or season related. In addition, vaccines are often co-administered, which may make assessing causality with any one particular vaccine difficult. Third, vaccination programmes are typically large-scale ecological interventions that perturb the host-pathogen relationship.

This can have indirect consequences, which may be deleterious to health on a population level, and therefore qualify as vaccine-associated risks (there can also be indirect benefits, notably through the protection afforded by herd immunity, to be considered in a separate report on the benefits of vaccination). The magnitude of these indirect risks, if there are any, typically depends in a complex way on the coverage and the efficacy of the vaccine. Thus, indirect risks are very dependent on the specific context. In addition, since the indirect risks associated with vaccination programmes are the consequence of shifts in infection dynamics, they may only emerge many years after the vaccination programme has started. Partly for these reasons, the methods used to investigate them cannot so easily be described in a general framework.

1.2 Scope of the report

The evaluation of risks associated with vaccination takes place throughout a vaccine's development, licensure, incorporation into a vaccination programme, and routine use. However, in line with the project proposal (ADVANCE Description of Work, section 2.1.3), we will not include either approaches to pre-licensure safety monitoring or safety signal detection methods using data mining techniques in the present review. We will focus primarily on methods used post-marketing for the rigorous evaluation of possible associations based on prior hypotheses. Such prior hypotheses can be motivated by signal generation systems, by other epidemiological studies, by prior knowledge of the biological mechanisms involved, or by public concern.

Recent developments in pharmaco-epidemiology, including vaccines, impelled by the wide availability of large databases, have motivated the conduct of 'signal strengthening' analyses. We will also review these methods which occupy the increasingly populated grey area between signal detection and confirmation, and involve an initial verification of signals or potential problems which can be conducted rapidly but falls short of more rigorous confirmation.

The bulk of the present report will focus on assessing the potential for direct risks of vaccines, since most quantitative methods of vaccine safety assessment relate to these. However, it is essential that risk – benefit assessments of vaccination programmes should also consider indirect risks, if only in qualitative terms, and for this reason we shall also discuss the evaluation of such indirect risks, though in considerably less detail than for direct risks for the reasons set out in Section 1.1.

1.3 Method and structure of the report

In line with the project's conceptual vision encompassing all stakeholders (ADVANCE Description of Work, section 2.1.3), we obtained scoping documents from key participants in order to inform both the criteria for evaluation and the methods to be evaluated. These documents are summarized in Section 2, which describes the different perspectives on risk assessment.

These perspectives informed the list of assessment criteria, which are set out in Section 3, and the list of methods to be reviewed, each of which is briefly described in Section 4. The methods

to be reviewed were selected based on expert opinion within the ADVANCE team and a search of the literature. As well as methods that have been used to evaluate vaccine safety, we decided to include methods that could be used for this purpose but which, to our knowledge, have not so far been used.

In Section 5 the criteria of Section 3 are applied to the methods of Section 4. In section 6 we discuss the interplay between signal detection methods and signal confirmations methods in the context of an overall framework for risk assessment.

In Section 7, we review existing systems for rapid risk assessment, as well as other projects related to ADVANCE. This section is based on briefing documents drawn up by members of the ADVANCE team with expertise in these systems and projects. Section 8 contains a brief review of the methods used for quantifying the indirect risks associated with vaccination.

Finally in Section 9 we discuss the findings, focus on some of the methodological problems of integrating risk assessments in risk – benefit evaluations, suggest areas where new research may be required, and recommend some issues for investigation in proof of concept studies.

2. PERSPECTIVES ON RISK ASSESSMENT

Different stakeholders will necessarily have different priorities, and hence the criteria by which they evaluate risk assessment methods may differ. In addition, vaccination carries risks and generates benefits at both the individual and the societal level. Consequently, how risks and benefits are balanced is likely to vary according to what perspective and in what context they are considered. Furthermore, the balance struck is likely to vary over time as the vaccination programme matures. In this section, we briefly review the different perspectives from which vaccine risks might be evaluated, and the criteria for risk assessment which are of key importance from each perspective. Our aim is not to provide a detailed description of the risk assessment issues or activities within each perspective, but to focus on those aspects with direct implications for assessing methods. We consider only post-marketing risk assessment.

Many safety monitoring activities are undertaken in consultation between public health authorities, regulators and manufacturers, and so do not uniquely fit within a single perspective. In order to avoid too much repetition, these joint activities are described in the first section in which they are relevant (so, for example, post-marketing surveillance of vaccine safety, which is important to public health agencies, regulators and manufacturers, is described under the public health perspective). We also take it for granted that all perspectives include a commitment to ethical methods and good statistical and epidemiological practice.

2.1 The public's perspective

Public perceptions of vaccine-associated risks can profoundly affect vaccination programmes, as public health authorities and vaccine manufacturers discovered to their cost in the mid-1970s, when concerns about the safety of the whole cell DTP vaccine led to the momentary collapse of the pertussis vaccination programmes in several countries, notably Sweden, the UK and Japan, and lasting impacts in several others (Gangarosa 1998). While in some countries (eg Sweden) concern had been expressed about the efficacy of the vaccine, in most this was not a major issue – indeed, the DTP vaccine was to some extent a victim of its success, pertussis incidence having been reduced to historically low levels, with the effect of also reducing public concern about the risk associated with whooping cough.

Public responses to the complex interactions between vaccine efficacy, disease incidence, vaccine coverage and adverse event are discussed by Chen (1999), and a simplified overview is represented in Figure 1. As a vaccination programme matures, coverage increases, causing a reduction in disease incidence and a possible increase in vaccine-related adverse events. The vaccination programme becomes more vulnerable to safety scares, perhaps resulting in loss of public confidence which causes a drop in coverage and an outbreak of disease. However, single case reports can also have a large impact very early on in a new vaccination programme.

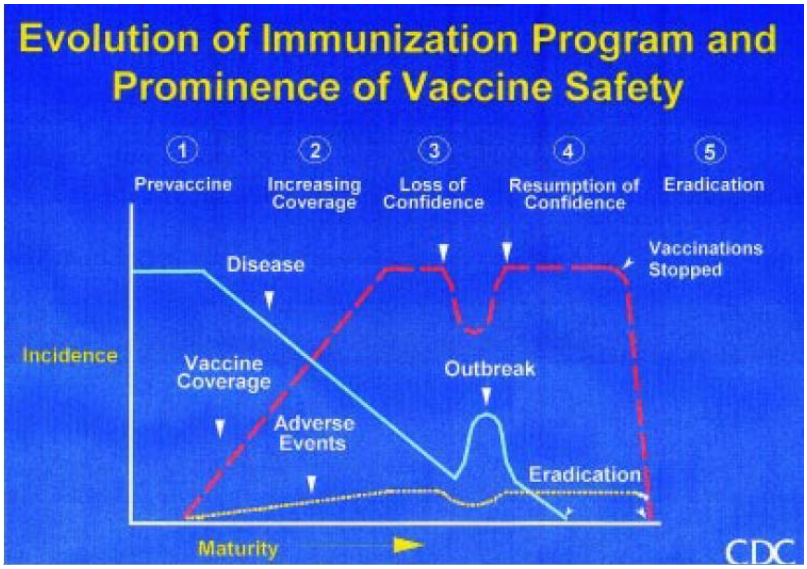

Figure 1 *Potential stages in the evolution of an immunization programme, showing the interaction between vaccine coverage, incidence of disease and incidence of vaccine-related adverse events (reproduced from Chen 1999)*

In addition, individual perceptions of vaccine-related risks may be structured by different priorities from those governing the provision of public health. The discrepancy arises because vaccination programmes impart indirect protection to unvaccinated individuals: thus, when vaccine coverage with an effective vaccine is high, there is lesser risk of disease for an unvaccinated person. In contrast, the public health priority is to maintain high coverage to ensure that low risks to unvaccinated individuals are maintained (Fine & Clarkson 1986). Related tensions between individual and public priorities, and the risk assessments involved, may also arise when individuals sharing responsibility for disease transmission are not those

potentially most affected by disease, as with boys and HPV vaccination (Georgousakis 2012), or children and influenza vaccination (Galvani 2007).

Public perceptions of vaccine risk can be affected by actual or perceived conflicts of interest on the part of those undertaking risk assessment studies. Studies sponsored by vaccine manufacturers are particularly prone to such suspicions. However, the independence of public health bodies and health professionals, especially those with close links to governments, may also be queried, especially when individual and collective health priorities are perceived as being not wholly synonymous. The media, and movements or pressure groups questioning vaccine policy, can help give voice to concerns about vaccine safety, which may rapidly be amplified and politicised, and may impact upon the benefit-risk balance.

Implications for methods to assess vaccine safety

The safety issues that may suddenly threaten a vaccination programme owing to loss of public confidence are likely to be unpredictable, and may not be based on good prior evidence or biological plausibility (such was the case with autism and MMR vaccine). They are likely to be rare, serious events (as was the case with encephalopathy and DTP vaccine). Often, such events will arise naturally at or soon after the recommended age at vaccination, the temporal association thus reinforcing the appearance of causality (a feature of both autism and MMR, and encephalopathy and DTP). Sometimes, as with concerns over thiomersal-containing vaccines and developmental disorders, the suspected adverse events will be ill-defined.

Effective methods for safety risk assessment need to be able to respond to such public concerns by providing compelling evidence of safety, or an accurate quantification of the risk. The methods employed must therefore be able to cope with the following features: studies need to be powered for several sub-analyses (with different design criteria such as risk periods, observation period, index date) or for several endpoints; they need to be able to handle rare events, and to allow for confounding (notably by age). Furthermore, they need to be able to produce results quickly, and these results need to be communicated effectively.

Finally, risk assessment studies for the purpose of benefit – risk evaluations should be, and be seen to be, undertaken independently of any vested interests involved.

2.2 The perspective of public health agencies[1]

Public health agencies, whether located within government structures or independent of these, carry an overall responsibility for improving and safeguarding the public health. Vaccines are one of the great success stories of public health, and in consequence the efficacy, coverage and safety of vaccination programmes are central to the mission of public health agencies. As far as vaccine safety is concerned, there are four broad areas of particular relevance to the present project.

[1] This section was informed by a document drawn up by Nick Andrews (Public Health England) which is reproduced as Annex 1.

The first is to have access to post-marketing surveillance of vaccine safety. This is undertaken routinely by regulatory authorities and manufacturers to identify temporal or spatial clusters of safety signals with a vaccine or its delivery, or with a particular batch or make of vaccine. When required, measures will be agreed in collaboration with the regulatory authority and the manufacturer (an intervention could include, for example, recommendations to alter the vaccination schedule, add booster doses, undertake special vaccination campaigns, replace vaccines of certain types, change vaccine delivery methods, alter indications or contraindications to vaccination, or halt vaccination). Such surveillance systems can be passive or (preferably, to avoid problems associated with low sensitivity) active, and will typically span a wide range of events.

Second, public health agencies, sometimes along with regulatory authorities or vaccine manufacturers, carry out controlled epidemiological studies to assess possible associations. A key aspect is to have access to suitable data sources, with the ability and permissions to link them if needed; these may include primary health care data, hospital data, vaccine registries, and disease registers. In some cases, studies requiring the proactive collection of new data not available in existing databases may be needed. Access to, or the development of, special registers, such as pregnancy registers, may also be useful in some circumstances. Such studies will be rigorous epidemiological investigations, undertaken using a pre-established protocol with clearly stated hypotheses, and validated methods. In the case of studies conducted by other parties, access to relevant results is important.

The third general requirement is access to the required expertise. Expert advice will often be required on the adverse event of interest (for example in order to specify relevant case definitions and code lists for data extraction). In addition, availability of epidemiological, statistical and data processing expertise and resources, both in-house and external, are essential to undertake the necessary investigations, or to develop new methods if required.

The fourth and final area of special importance is communication. This involves both educating the public on how vaccine safety is monitored and assessed, and communicating complex messages about safety and the benefit – risk balance in specific instances. Such communication requires striking a careful balance between scientific accuracy and direct relevance to individuals and families facing the decision to vaccinate. Effective communication benefits from close collaboration with regulatory authorities and between public health agencies in different countries.

Implications for methods to assess vaccine safety

Surveillance systems to detect new signals, and carry out an initial assessment of them (eg using observed vs. expected analyses), must be sensitive and timely, without generating large numbers of false positives, which would rapidly undermine the practical usefulness of such systems. Active surveillance systems are particularly valuable in being able to provide accumulating evidence of safety. Such systems must be adapted to the planned distribution of vaccines.

The epidemiological studies undertaken (which may be studies requiring the collection of new data, as well as studies in existing databases) must be rigorous in methodological terms, in both their statistical (eg hypotheses, power, bias, control of confounding) and clinical components (eg case definitions, case and exposure ascertainment). It is useful to have access to a broad range of methods, with different strengths and weaknesses, but the appropriateness of these methods in different settings must have been validated. Resources (including expertise) and data must be mobilised sufficiently rapidly to answer questions of interest authoritatively yet without undue delay. Effective methods for communicating risk and safety messages are needed, including graphical techniques.

The availability of brand-specific and/or batch-specific data is also important for public health agencies, notably when vaccines from several manufacturers are in contention within the same study. Similarly, comparisons between studies using different vaccines are likely to be of interest. Formal comparisons between vaccines may require non-standard null hypotheses, as in the case of tests of non-inferiority or equivalence.

2.3 The perspective of regulatory authorities[2]

One of the primary responsibilities of regulatory authorities in relation to licensed vaccines is to monitor their safety after licensure, in order to protect the public against possible adverse events associated with them. At the time of licensure, safety information is available, obtained from laboratory experiments and clinical trials. However, such information is inevitably constrained by two main factors: first, the scale and duration of the pre-licensure trials is necessarily limited (phase 3 vaccine efficacy trials seldom involve more than some thousands or tens of thousands of participants, followed for months rather than years); second, these trials are undertaken in controlled conditions which usually differ in some respects from those that apply in the clinical settings in which and in the populations to which vaccines are dispensed.

Regulatory authorities conduct vaccine safety surveillance activities including routine signal detection based on spontaneous reports (submitted directly through national schemes or submitted by manufacturers), signal strengthening activities (such as observed versus expected analyses) and confirmatory epidemiological studies. Additionally, regulatory authorities assess data submitted from manufacturers on all the above activities.

Routine signal detection based on spontaneous reports allows signals to be generated with no prior hypothesis (a process known as data mining). The pre-licensure information available on vaccines, along with knowledge gained cumulatively from other vaccine programmes, generates essential prior information upon which additional post-marketing surveillance is built. Thus, the events monitored are potential adverse events identified from clinical trial data, from past experience with similar vaccines, or by potential events suggested by the biological mechanisms of vaccine action. In addition, it might be decided to monitor temporally

[2] This section was informed by a document drawn up by Suzie Seabroke (MHRA) which is reproduced as Annex 2.

coincidental events that are prevalent in the target population, and which are therefore likely to be reported as potential adverse reactions.

Regulatory authorities may also carry out epidemiological studies to assess possible associations. These may be conducted independently or in collaboration with public health agencies. The same requirements for data sources therefore exist for regulatory authorities as for public health agencies, and in particular access to brand and batch level data, as regulatory action can be taken at both levels. Timely access to the results from studies conducted by other parties is also critical.

The regulatory perspective on safety assessment features two main dimensions: monitoring of new vaccines using methods suitable for an initial rapid assessment of safety, followed by extensive confirmation studies to investigate the signals identified in rapid assessment monitoring, or by other means for mature vaccination programmes.

From a regulatory perspective it is critical to obtain information on potential adverse reactions very rapidly after the introduction of a new vaccine, in order that prompt action can be taken to protect the public, or that communications regarding vaccine safety supported by evidence can be made if necessary. Thus, rapid safety signalling and strengthening methods are of key importance. At this stage, speed is essential: rapidity of the evaluation is more important than its definitiveness, and hence methods can be cruder than those used for signal verification.

If these rapid assessment methods indicate that there may be a problem with a particular vaccine, the focus then shifts to undertaking or commissioning confirmatory analyses, in order to investigate the issue further, and quantify the risk involved. This will often involve several different studies, typically undertaken in different contexts (possibly different countries) and based on different data sources. The key requirement at this stage is to obtain a definitive safety assessment with clear policy direction, without undue delay. Therefore, as for public health authorities, access to the required expertise is also necessary.

Implications for methods to assess vaccine safety

Rapid assessment methods must be implemented in as close to real time as is possible, especially for seasonal vaccines, in order for the results to inform ongoing vaccination campaigns. For this reason, it is desirable to make maximum use of sources of data, such as spontaneous reports, that do not suffer from long reporting delays. Such data, however, need to be suitably contextualised (for example, by comparing observed with expected values) by data on background rates, if available. The methods will typically be sequential. Ecological analyses may also be informative for high-uptake vaccines. While grossly biased methods are clearly inappropriate, the methods used for rapid assessment can be relatively crude, as there will seldom be time or power fully to control for confounders.

Confirmatory methods, on the other hand, need to be rigorously controlled, and undertaken in the relevant target population. There is particular benefit in replicating results in different databases: robust methods are therefore particularly valuable. Whereas the emphasis in rapid assessment methods is on signal strengthening (usually within a limited range of potential

signals of prior interest), a further consideration for confirmatory methods is to quantify the effect, if there is one. Thus, avoidance of bias is important, and to this end information from a variety of data sources and statistical methods is preferred. Finally, while rapidity is not the primary consideration, as it is for rapid assessment, it is desirable to obtain definitive answers as promptly as possible, within the limits of the power available.

2.4 The perspective of vaccine manufacturers[3]

Vaccine manufacturers have a special responsibility regarding the safety of their products, which they monitor continuously for safety throughout the life cycle of the vaccine, within a highly regulated framework governed by stringent quality procedures. They systematically collect and periodically review data on adverse events, following a pre-established risk management plan which sets out the potential and identified risks and outlines risk minimisation activities. The potential and identified risks to be monitored are determined using clinical trial data, past experience with similar vaccines, and are updated based on information from accumulating post-marketing surveillance. Signals are evaluated and investigated according to manufacturers' internal procedures, aligned with regulations and guidance on pharmacovigilance and pharmacoepidemiology. Increasingly, manufacturers are assessing vaccine safety in the context of benefits through formal periodic benefit – risk evaluation reports and other means.

The safety surveillance activities undertaken by vaccine manufacturers span a wide range, from signal detection based on spontaneous reports, through signal evaluation, to confirmatory epidemiological studies. These activities have already been evoked above. While these activities are informed by the wider context, notably the experience of and literature on vaccines from other manufacturers, they are necessarily focused entirely on the manufacturer's own products. Manufacturers are also in a special position in that they have full access to the manufacturing data and to pre-licensure data from clinical trials and other studies, which may sometimes be of use (though usually of low power) in assessing risks of events that occur unexpectedly after the vaccine is licensed.

The decision to undertake a fully-fledged epidemiological study can be initiated either as a company decision, or following a requirement from a regulator or other competent authority. Such studies tend to be most frequently undertaken close to licensure or the launch of the vaccination programme, or to support variations in labelling, indication or formulation. While vaccine manufacturers have considerable in-house expertise to undertake risk evaluations, studies are often undertaken in collaboration with other researchers.

Implications for methods to assess vaccine safety

Manufacturers are particularly keen to ensure that evaluations of their product are undertaken fairly, using methods that have been independently validated. The methods used may be

[3] This section was informed by a document drawn up by Catherine Cohet and Dominique Rosillon (GlaxoSmithKline Vaccines) which is reproduced as Annex 3.

expected to conform to relevant statutory requirements and, to the extent possible, to accepted codes of conduct and guidelines, such as the ENCePP methodological guide (http://www.encepp.eu/standards_and_guidances/methodologicalGuide9_2_1_4.shtml) . The key methodological issues are: choice of a suitable study design; control of bias from confounding, control selection or other sources; accuracy and completeness of exposure information, including dose and risk period; appropriate definition of the adverse event and valid case ascertainment procedures; sample size and power.

2.5 An academic perspective

The term risk is often used synonymously for probability. However, in the context of benefit – risk assessments, risk is more appropriately understood as a combination of probability and consequence. In relation to adverse events caused by vaccines, these are the probability that the vaccine causes an adverse event, and the loss incurred from the occurrence of the adverse event. If there are several possible independent adverse events $A_1,\ldots,A_k$, the overall risk takes the form

$$\text{Risk} = p(A_1)L(A_1) + \ldots + p(A_k)L(A_k)$$

where p(A) is the probability that event A occurs, and L(A) is the loss incurred from event A. Implicit in this equation is the key assumption that it makes sense to add these different components: this will only be the case if the losses are expressed on a common scale. These quantities may also vary according to covariates, notably age.

The issues involved in risk evaluation stretch far beyond statistics, and are relevant to very many academic disciplines, from climate change to politics and finance. The subject is highly multi-disciplinary, encompassing the complexities involved in understanding risk perception (Slovic 2000, 2010), and the contrasting perspectives on risk of social and cultural theory described by Lupton (2008) stemming from the work of anthropologist Mary Douglas, sociologist Anthony Giddens and philosopher Michel Foucault. Closer to ADVANCE, in the context of pharmacoepidemiology, some of the issues involved have been touched upon as part of the iMi PROTECT project (www.imi-protect.eu), and will be considered in greater detail as they relate to vaccines in a separate document on benefit – risk evaluation.

Implications for methods to assess vaccine safety

Epidemiological methods employed to evaluate the safety of vaccines focus primarily on evaluating the probabilities p(A). This is typically achieved by estimating them from epidemiological studies. The statistical issues involved are the standard ones of statistical inference (Cox & Hinckley 1974): estimators for these probabilities must be consistent (that is, they should converge to the true value as the sample size increases), should be nearly unbiased in large samples (that is, their expected value should be close to the truth), and should preferably be robust to failure of model assumptions. In addition, they must be adequately adjusted for the presence of confounders, which may or may not be known; this is probably the single greatest challenge. Finally, there must be a degree of consensus that the probability thus estimated

represents a causal relation between vaccination and the adverse event. This final requirement is not achievable via the methods of statistical inference: evidence for causality is seldom directly available through experiments, but must usually be accrued incrementally by marshalling other information, for example using the criteria proposed by Bradford Hill (1965) applied to vaccines (WHO 2001). More recently, a new field known as causal inference has emerged, though the methods typically rely on strong and often untestable assumptions (Pearl 2009). Causal inference methods include structural equation modelling, propensity scores, and instrumental variable analysis, some of which will be reviewed in Section 4. Since the 1990s, more direct methods to eliminate confounding have been developed, known as self-controlled methods, also to be reviewed in Section 4.

The major problem in conducting risk assessments in the context of benefit-risk evaluations, however, is in the evaluation of the losses L(A). As previously noted, these must be measured on some common scale by which they can be compared (and offset against benefits, also to be expressed on this scale). In medicine, much effort has been devoted to defining appropriate scales to quantify impact, for example QALYs (quality-adjusted life years) and DALYs (disability-adjusted life years) (see Murray 1994, Thacker 2006). These measures, however, have not routinely been used with vaccines. Other, simpler, measures of impact can be used: for example, a recent study of rotavirus vaccine used both numbers of hospital admissions and deaths as the scale on which to measure both risks and benefits (Clark 2014).

The choice of scales, and more generally the choice of loss functions, implies a choice of perspective. For example, choosing hospital admissions as the scale on which to measure incurred losses represents largely an institutional healthcare perspective. Some Bayesian perspectives on decision theory, notably that of Savage (1954), promote an approach based on individual subjective loss functions and probabilities. The diversity of the contending approaches ensures that the topic of decision-making in the presence of uncertainty remains hotly debated among academics and practitioners.

3. CRITERIA FOR RISK ASSESSMENT METHODS

The criteria set out below will be used to evaluate the different study designs available, to be described in Section 4. They are chosen so as to capture the key features required of risk estimates to be used in benefit – risk assessments. The criteria are grouped under five headings, and discussed in the next five sections. In each, we first give the rationale for this set of criteria, then list the specific criteria with some contextualisation and discussion.

Some effort has been made to keep the number of criteria down to a manageable number – we propose nine criteria in all. Section 3.6 contains the full list of the nine criteria for ease of reference.

3.1 Effect measure criteria (M)

Rationale

Benefit – risk assessments typically require absolute measures of risk for the populations in which the assessments are undertaken. Such measures might include, for example, the expected number of adverse events attributable to the vaccine, by age group and time period, which will then be weighed against the expected benefit for that same age group and time period (or other groupings). For adverse events with prolonged sequelae, the expected number might be combined with a measure of disease burden by converting it to DALYs (disability-adjusted life years; Murray 1994, Thacker 2006). The definition of what events are included (for example, it might be decided to consider only events leading to hospitalisation), and the weightings used (via DALYs or otherwise), are study-specific and will not be considered here.

The measures of risk required in any given benefit – risk assessment are seldom likely to be available directly, either because the studies available have been undertaken in different populations from those for which assessments are sought, or because the measures of risk differ from those required in such assessments. In both cases, the available measure of risk must be converted into the measure desired, and this will usually need some extra information pertaining to the population for which the assessment is sought – often, a measure of the baseline rate of events of interest, and a measure of vaccine coverage. For example, the measure available might be a post-vaccination relative incidence R associated with a risk period T. Given a per-capita baseline rate B per unit time, and a number V of persons newly vaccinated, the absolute number of vaccine-associated events expected within the risk period T is V x B x (R-1) x T. In this calculation, the quantity V depends on the population and the vaccine considered, B depends on the population and the event of interest, while T and R are generally assumed to be specific to the vaccine-event pair under study (the latter assumption may not necessarily be valid). The issue of context-specificity and vaccine-specificity of different quantities, and its relevance to benefit – risk assessments, is discussed further in Section 9.

Criterion

In choosing a method for risk assessment, it is important to know what effect measure it will generate, and to know that the information will be available to convert this effect measure into the risk measure required. The criterion could therefore be stated as: '*Can the effect measure be converted into the required risk measure?*' However, since the answer to this question will depend on what risk measure is to be used in the benefit – risk assessment, the criterion will instead be stated as follows:

Criterion M1: Statement of the measure of effect directly available from the method.

Such measures could include:

Absolute risks $p_v \; in \; vaccinated, \; p_u \; in \; unvaccinated$

Absolute rates $\lambda_v \; in \; vaccinated, \; \lambda_u \; in \; unvaccinated$

Relative risk $R = \frac{p_v}{p_u}$

Relative incidence $RI = \frac{\lambda_v}{\lambda_u}$

(or relative rate, or hazard ratio)

Odds ratio $OR = (\frac{p_v}{1-p_v})/(\frac{p_u}{1-p_u})$

Attributable fraction $AF = 1 - \frac{1}{R}$ *or* $AF = 1 - \frac{1}{RI}$

Attributable risk $AR = p_v - p_u$

Attributable rate $RD = \lambda_v - \lambda_u$

Number needed to harm $NNH = \frac{1}{AR}$

In the above, 'risk' denotes a dimensionless probability, while 'rate', 'incidence' and 'hazard' are per unit time; 'fraction' is a ratio of either. To be meaningful, all types of measure usually also require specification of the 'risk period' T during which the potential for vaccine-associated adverse effects is assessed.

These measures of effect are standard in epidemiology; see Schechtman (2002) for a discussion of some of them. The terms attributable fraction and attributable risk can denote different measures from those defined here: see Benichou (2001), who also discusses several other related measures of attribution including the prevented fraction, preventable fraction, and generalized attributable risk or generalized impact fraction. These are seldom, if ever, used in connection with vaccines and will not be reviewed here.

The absolute risks and rates in the above formulas usually relate to a defined post-vaccination risk period of duration τ, which may be indefinite; they may be dose-specific or relate to completed courses of vaccination. For rare events, the odds ratio is numerically close to the relative risk. Rates are commonly used for potentially recurrent events; when the events are not recurrent one might use hazards, or risks. For short risk periods τ and/or rare events, rates and risks are approximately related by
$p_v = \lambda_v \tau, \quad p_u = \lambda_u \tau.$

The relative risk, relative incidence, odds ratio, and attributable fraction are relative effect measures, expressed as dimensionless ratios. The attributable fraction, attributable rate, and number needed to harm (NNH) are absolute measures, which may be derived by combining a ratio measure and an absolute risk or rate for the unvaccinated population.

Some study designs yield only relative effect measures, but because they are applied to an entire database, or to an entire population, can be used to obtain absolute effect measures. For example, the National Childhood Encephalopathy Study (Alderslade 1981) was a case-control study, thus yielding odds ratios which, since the events were rare, approximate relative risks. However, because the study was a national one and included all reported cases, it was possible to derive absolute measures of effect. Similarly, the self-controlled case series method produces

relative incidences. But applied to a population in which the number of vaccine doses administered is known, it has been used to calculate attributable risks, expressed as 1 event per x thousand doses (Farrington 1995b). Similarly, conversion of relative incidences into NNH is discussed by Wilson (2013).

3.2 Statistical criteria (S)

Rationale
The statistical criteria relate to the formal properties of the effect estimator for each method, assuming that all assumptions required by the method are satisfied. All estimators considered in this report are assumed, unless otherwise specified (there are some exceptions), to be obtained by maximum likelihood in regular settings, and are therefore (in a technical statistical sense) asymptotically unbiased and consistent. The main issues of interest are therefore how the estimators compare in terms of asymptotic efficiency and power, and how good the point and interval estimators are in finite samples.

Criteria
The statistical criteria are:

Criterion S1: Asymptotic relative efficiency and power of the estimator.
Criterion S2: Finite sample performance: bias and coverage probabilities.

The asymptotic relative efficiency (ARE) of two estimators T_{1n} and T_{2n} of a parameter θ based on a sample size n is the limit, as the sample size tends to infinity, of the ratio of mean squared errors:

$$ARE = \lim_{n\to\infty} \frac{E(T_{1n}-\theta)^2}{E(T_{2n}-\theta)^2}.$$

The estimator T_{1n} is usually taken to be a 'gold standard' estimator, whereas T_{2n} might be some other estimator. Usually, the 'gold standard' will be based on a full cohort analysis of the data, against which another method (case-control or self-controlled case series, for example) using only some of the information in the sample is to be evaluated. High asymptotic efficiencies are desirable.

The power is the probability of rejecting the null hypothesis, for alternative hypothesis values of θ. Power and efficiency are related, and for that reason are put together under a single criterion; the difference is that efficiency relates to estimation and power to hypothesis testing. For risk assessment, the alternative hypothesis values of θ of primary interest are those corresponding to an increased risk (eg, relative risk R > 1, attributable risk AR > 0).

Finite sample performance relates to how the estimator performs for small or moderate sample sizes. The statistical theory of maximum likelihood estimation guarantees that maximum likelihood estimators perform well in large samples, but this does not tell us what happens in other situations. These usually need to be investigated by simulation. Interest focuses, typically,

on the bias $E(T_{2n}) - \theta$ of the estimator, which ought to be small, and how close the empirical coverage probability of 95% confidence intervals for θ (that is, the long-run proportion of intervals containing the true parameter value in independent replications of the experiment) is to the theoretical value 0.95.

In the present context, these statistical criteria are not of predominant practical importance: most standard methods have estimators with reasonable properties, and methods can sometimes be adjusted to improve their performance (eg by choosing more controls per case in a case-control study). However, there are situations (rare adverse events, for example, and hence small sample sizes) where they can become an issue.

3.3 Timeliness criteria (T)

Rationale
This set of criteria relates to the timeliness with which the results of risk assessments may be obtained, an early assessment clearly being desirable, though not at the cost of gross inaccuracy.

The issue of timeliness applies rather differently to new and established vaccines. For the purpose of this section, new vaccines are taken to include influenza vaccines, as these are targeted at specific strains of influenza which change on an annual basis. New vaccines also include those newly introduced, variations on existing vaccines, and existing vaccines targeted at different populations (for example, pertussis vaccination in pregnancy). Established vaccines include those used in mature vaccination programmes. The distinction is essential because for new vaccines the numbers exposed will tend to be low initially. Thus, the power to investigate a specified potential association will necessarily be low to start with, and will build incrementally. In such situations, sequential designs are called for, as these minimise the expected sample size (and hence, usually, the waiting time) needed to reach a decision (Wald 1947), which in the present context could be to flag up a potential association, or to fail to identify a problem and thereby provide evidence of safety.

For mature vaccination programmes, on the other hand, there is lesser need for sequential methods, as substantial power is usually available from accumulated data, if necessary by combining databases. One exception where sequential methods might be relevant for established vaccines, is long-term, routine monitoring with the aim of detecting departures from a steady state (which in this case is an acceptable safety profile), as done for some industrial processes – using the methodological framework of Statistical Process Control first developed by Shewhart (1931) and later elaborated by Page (1954). A second exception may be for very rare events, where the accumulated data are insufficient or incomplete.

A second factor influencing the speed with which results can be obtained is the extent of data checking and validation that needs to be undertaken, for example, by returning to case notes and immunisation records. This will generally depend on the size of the study and the quality of the data. If the data are of high quality, it might be decided not to undertake any checking of exposures or events. If the data cannot be assumed to be wholly trustworthy, some checking

might be necessary. In this case, study designs that, for a specified power, require lower numbers of participants may possess a temporal advantage.

Criteria
In line with the discussion above, we propose two timeliness criteria:

Criterion T1: Effectiveness for sequential and/or routine monitoring.
Criterion T2: Operational sample size required for specified power.

The first criterion is primarily relevant to the sequential evaluation of new vaccines, typically using the sequential probability ratio test, but also perhaps for routine monitoring of established vaccines, for example using cumulative sums. Other issues covered under this criterion include the statistical properties of the sequential implementation, notably whether the sequential design is efficient (in that it minimises the expected time to decision, for sequential hypothesis tests, or the ratio of average run lengths ARL_1/ARL_0, for routine monitoring), and whether unbiased estimates of effect can be obtained (estimates obtained at stopping by the SPRT, for example, being biased away from the null).

The second criterion relates to the number of subjects needed to reach a specified power. We refer to this as the 'operational sample size' to distinguish it from the population size required to generate sufficient cases. In a case-control study, the operational sample size will be the number of cases plus the number of controls; in an SCCS study it will be the number of cases. This criterion presumes that the time to check the data (or in some instances, to obtain the data) grows proportionately to the number of subjects included in the study.

3.4 Restrictions and Robustness criteria (R)

Rationale
This set of criteria relates to how a method performs when reality does not conform to model assumptions.

All statistical methods require assumptions. Some of these are shared by all methods (such as the need for exposure data to be collected independently of outcomes). Others are specific to particular methods. Traditionally, statisticians have tended to think of assumptions as being limited to those relating to the underlying stochastic process generating the data (for example: '*events arise in a Poisson process*') and to the model structure (for example: '*the log of the Poisson rate is a linear function of the regression parameters*'), and not to our knowledge of the data – the latter being perhaps described as epidemiological assumptions (for example: '*all confounders are included in the model*'). The distinction is not very meaningful (for example, when missing data are involved assumptions need to be made about the missingness process) as both relate to our knowledge of the state of Nature. Both types of assumptions will be considered under the same heading.

Assumptions are important, particularly, insofar as violation of assumptions may lead to bias. In what follows, we say a method is robust if it is not overly dependent on assumptions (note:

this is not the same as the technical statistical meaning of robustness, which relates to estimators, not methods). Even if the bias can be substantial, it may help to know the direction of this bias.

Finally, some methods may be inapplicable in certain settings, in the sense that they cannot be used at all (rather than being useable, but biased). Such restrictions need to be made explicit in evaluating methods. We consider them under this heading alongside issues of robustness, as they also, in a more extreme way, relate to assumptions.

Criteria
Based on the above discussion, we propose the following criteria:

Criterion R1: Assumptions (statistical and epidemiological) of the method.
Criterion R2: Robustness to failure of the assumptions.

The first of these criteria will involve listing the key assumptions required for the method to yield good (that is, efficient in the technical statistical sense) estimates. Note that assumptions common to all methods will not be listed, for example that exposure and outcome data must be ascertained independently.

The second criterion consists of evaluating the impact of failure of these assumptions. This will include itemising those circumstances, if any, in which the method cannot be applied, and any modifications that might be required as a result.

3.5 Operational criteria (O)

Rationale
This final set of criteria concern the ease of implementation of the method, including the data requirements and data availability, the complexity of implementation, and any other issues relating to use of the method, such as financial cost.

Criteria
We propose the following two criteria under this heading:

Criterion O1: Data requirements.
Criterion O2: Complexity of implementation.

The first of these criteria sets out the data requirements for each method, to include the data that must be assembled for the method to be applied, and any issues related to particular sensitivity to imperfections in the data not mentioned elsewhere. Note that the data requirements include what is usually understood informally under this heading, rather than a more technical specification of the minimal sufficient statistics for the relevant parameters of the likelihood function[4].

[4] The detailed data requirements for some commonly used models are in Annex 4.

The second criteria relates to implementation issues, to include technical issues (such as availability of standard software) but also any special permissions or consent issues that may be required as standard, which may affect the choice of approach (as is the case with randomised controlled trials).

3.6 The assessment criteria at a glance

The full list of the nine criteria discussed above is as follows:

Criterion M1: Statement of the measure of effect directly available from the method.
Criterion S1: Asymptotic relative efficiency and power of the estimator.
Criterion S2: Finite sample performance: bias and coverage probabilities.
Criterion T1: Effectiveness for sequential and/or routine monitoring.
Criterion T2: Operational sample size required for specified power.
Criterion R1: Assumptions (statistical and epidemiological) of the method.
Criterion R2: Robustness to failure of the assumptions.
Criterion O1: Data requirements.
Criterion O2: Complexity of implementation.

4. METHODS OF RIKS ASSESSMENT FOR VACCINES

A large number of study designs and statistical methods have been used to evaluate the safety of vaccines. Others have not been used for this purpose, but could be used in principle. In this review of methods for risk assessment we have cast our net widely, and sought to include most types of approach.

Our survey of methods is organised so as to begin with the more formal approaches (namely designed experiments), which are typically the most data intensive, and to work towards those which are least demanding in terms of data (for example, ecological methods, which do not use individual level data). We give a description of each method in non-technical terms, with important variants listed and described separately, in order to emphasize the richness of the methodological toolbox available.

We do not seek to give any detailed account of how each method has been used for vaccine risk assessment, that is, we do not list all the vaccine – event pairs for which the method has been applied. We do, however, give a few examples of use for vaccine safety assessment, if available, or for other purposes, if not, and note in passing any major methodological issues involved. For less well-known methods, technical references are provided as appropriate. At this stage, no attempt is made to evaluate the methods using the criteria set out in Section 3, as for this purpose we will generally collapse sub-headings and consider methods and their variants together.

4.1 Designed vaccine introduction

The introduction of a new routine vaccination programme offers unique opportunities for risk (and effectiveness) evaluation, because the population is unvaccinated. These opportunities are seldom exploited. While there are probably good reasons for this (logistical problems, ethical objections, cost), and also less good ones (inertia, bureaucracy), perhaps studies at introduction merit greater attention, especially within the context of an overall risk-benefit evaluation. In particular they offer a unique opportunity for post-licensure randomised studies of vaccination against no vaccination on a very large scale. Cluster-randomized introductions of this sort also provide a unique way of assessing the vaccine's indirect risks (and indeed benefits) via herd immunity effects.

4.1.1 Randomised trials at vaccine introduction

The vast 1954 field trial of the Salk inactivated polio vaccine combined a large randomised component of over 600,000 children with an even larger open study of over one million, effectively heralding the introduction of mass vaccination against polio. The controversies surrounding the trial are reviewed by Meldrum (1998). The Salk vaccine story is relevant also from a vaccine risk perspective in view of the Cutter incident (Nathanson 1963, Offit 2005).

4.1.2 Stepped wedge designs

The stepped wedge design exploits the phased introduction of a new vaccination programme, which allows rates in the vaccinated group to be compared with rates in the contemporaneous unvaccinated group during the introduction of the vaccine. The design was first used in The Gambia to assess the effectiveness of HBV vaccine (Gambia Hepatitis Study Group 1987). There are cluster randomised and non-randomised versions of the design, which has been the subject of a systematic review (Brown 2006). Stepped wedge introductions of mass vaccination programmes could be used to assess vaccine safety, and indeed risk – benefit more widely, cluster randomized designs perhaps offering an opportunity for assessing short-term indirect effects.

4.2 Cohort studies

Cohort studies are in a sense the benchmark method of evaluating risks associated with vaccines. We include randomised controlled trials as particular types of cohort studies.

4.2.1 Randomised controlled trials

Randomised controlled trials are generally considered to be a 'gold standard' as, in theory, they control for confounders. However, for risk assessment they are seldom large enough to provide acceptable power for evaluating rare events. For vaccines, randomized trials are usually parallel group designs (i.e. not crossover trials) as they are most frequently designed to assess efficacy, safety often being a secondary endpoint. For ethical reasons, randomised trials with an unvaccinated (or placebo) group are seldom if ever used in post-licensure surveillance.

Post-licensure randomised vaccine-vaccine comparisons may present fewer ethical issues, but perhaps owing to logistical difficulties have not been used that much on a large scale. Examples include the Swedish DTP vaccine trials of the 1980s and 1990s comprising both acellular and whole cell vaccine arms (as well as a DT arm in some trials) (Ad Hoc Group for the Study of Pertussis Vaccines 1988, Gustaffson 1996, Olin 1997). The Swedish studies are a little unusual in that no DTP vaccination was used in Sweden at the time, thus making it possible to use a placebo.

A key trial in safety terms was the trial of high titre measles vaccine versus standard measles vaccine, which revealed higher all-cause mortality in girls in the high-titre arm (Garenne 1991, Aaby 1994). This led to the abandonment of such vaccines, and much further work on non-specific effects of vaccines, notably from the Aaby group, which presents big methodological challenges (see Section 4.2.3). The impact of different vaccination schedules (notably the age at vaccination and the order in which different vaccines are administered) on vaccine safety could perhaps also be studied by randomised comparisons without insuperable ethical issues. The feasibility of randomized trials to examine the issue of non-specific effects is discussed in Shann (2010).

4.2.2 Post-unblinding surveillance of randomized cohorts

Randomised controlled vaccine trials can be large but are usually of relatively short duration. However once a trial has been unblinded it may sometimes be possible to undertake long-term follow-up of the randomised groups in order to study long-term risks associated with the vaccines. This is being done for HPV vaccines, owing to the long lag time between exposure and outcome. Information bias and selection bias owing to dropouts may be an issue, however, as may be catch-up vaccination. Examples include the long-term follow-up of some of the Swedish DTP vaccine trials (Olin 2003, Gustafsson 2006), and the very long-term follow up of the 1964 (systematically allocated) measles vaccine trial in the UK (Ramsay 1994). These long-term follow-up studies of large trials have been used primarily to assess vaccine efficacy and evidence of waning, but could perhaps be used for safety.

4.2.3 Parallel group non-randomized cohort analyses

Comparisons of event rates in vaccinated and unvaccinated will often be subject to selection biases related to the non-random allocation of vaccines. However comparisons restricted to vaccinees are likely to be less problematic, and can be undertaken retrospectively in large databases. For example, such methods have been used to investigate the effect of thiomersal, a preservative made from ethyl mercury included in some vaccines, on the incidence of developmental disorders in later life. The exposure here can be quantified according to the cumulative dose of thiomersal received, and is nil for non-thiomersal-containing vaccines. See for example Verstraeten (2003) and Andrews (2004). Methods of analysis include Cox regression (Cox 1972) and Poisson regression (Frome 1983).

A special mention is perhaps needed of the methods used to study possible non-specific effects of vaccines, notably their impact on all-cause mortality. These methods are primarily relevant to low-income, high child mortality countries, and so perhaps of less relevance for the present project – though these issues might eventually emerge in Europe and affect the relative timing

of childhood vaccinations. Essentially, they are cohort methods with various adjustments for selection bias and/or informative missingness. There is a growing, complex and evolving literature on non-specific effects, starting with Kristensen (2000) but traceable back to the high-titre measles vaccine story (see Section 4.2.1). The key statistical issues are (a) selection biases which are likely to affect the comparison between vaccinated and unvaccinated, and (b) biases due to informative missingness, notably what has been called 'survival bias'. Attempts have been made, with varying degree of success, to address the first problem by using propensity scores, and the second by landmark analyses. These and other methodological issues are discussed in detail in Fine (2009) and Farrington (2009).

4.2.4 Risk interval cohort studies

Risk interval cohort studies differ from parallel group cohort studies in that exposures vary over time, so that the same individual can be exposed and unexposed. Such designs are particularly relevant for vaccine safety studies since many adverse events, if they occur at all, tend to occur soon after vaccination. One big advantage of risk interval studies is that they can be undertaken in vaccinated populations, thus getting round some of the selection biases associated with vaccination. A disadvantage is that they require pre-specification of the post-vaccination risk interval. This is not generally a big problem in practice and can be addressed by using several such intervals.

As the next sections demonstrate, there is a burgeoning nomenclature surrounding sub-versions of the risk-interval cohort study, a similarity between some designs, and in some cases possible links with self-controlled methods. A review and more formal statistical description of these designs might be useful in order to clarify connections and differences between them.

4.2.4.1 Standard risk interval cohort studies

Standard risk interval studies are among the most commonly used designs for investigating vaccine safety. Typically, they use all (or most) of the relevant person-time available in the database. So for a study of primary MMR vaccination it makes sense to use everyone during the time period in which the first dose of MMR vaccine is typically given – say the second year of life. Such studies were described by Ray (1989) (he described control in such studies as 'self-control') and are now very commonly used, for example for studies in the Vaccine Safety Datalink (VSD) in the USA and in the Clinical Practice Research Database (CPRD) in the UK. See for example Baggs (2011) for a general description of the VSD, and Barlow (2001) for a description of an early risk interval cohort study from the VSD. Risk interval cohort studies can also be used for indefinite risk intervals, as used by Madsen (2002) for the study of autism and MMR vaccine in a Danish database.

4.2.4.2 Sequence Symmetry Analysis (SSA)

Also called Prescription Symmetry Analysis, this is a cohort method developed by Hallas (1996). The population studied comprises all individuals who have been prescribed a given drug and have also had the event of interest. The risk interval is the (indefinite) post-exposure time, the control interval is the pre-exposure time. The quantity estimated is the number of events following exposure, divided by the number preceding exposure, adjusted for time trends. There is some conditioning involved, since each individual has an event, so the method bears

some similarity, yet to be elucidated, with the SCCS method. However, it's not clear that the SSA method of analysis takes account of this conditioning. The method has been used to quantify relative risks in pharmacoepidemiology (eg van Boven 2013) but not, apparently, in vaccine studies. There may be an issue of bias in such an application owing to the healthy vaccinee effect, which could perhaps be mitigated by excluding some immediate pre-exposure time.

4.2.4.3. Other variants on the standard risk interval cohort study design

A further variant has recently been proposed which the authors have called 'self-control cohort' method (Ryan 2013). This is a risk interval study starting at some determinate time before exposure and ending at some determinate time after exposure. Self-control appears to be at the group level rather than the individual level. Applied to vaccines, the analysis may need to be adjusted for age effects since the vaccine risk period always follows the control period. Like the SSA method (Section 4.2.4.2) it may also be necessary to treat separately the immediate pre-vaccination period owing to the healthy vaccinee effect, vaccination being postponed if the child is ill, which would tend to favour the vaccine in such a design.

4.3 Case referent studies

These designs compare exposures in cases to exposures in separate but 'comparable' controls. The big advantage over cohort methods is the reduction in operational sample size, though choosing suitable controls can be tricky and may introduce a selection bias.

4.3.1 Case-control studies

This very popular design has been much used in vaccine studies. Exposure may be defined as 'ever-never', as in the study by Smeeth (2004) of autism and MMR vaccine, or using risk intervals, as in the big pre-databases National Childhood Encephalopathy Study (Alderslade 1981) of brain damage and DTP vaccine. Most case-control studies of vaccines are matched; in the case of risk-interval case-control studies, this is necessary from an operational perspective in order to define the index date for the controls. Analysis of matched case-control studies is by conditional logistic regression. A review of the methodology as it applies to vaccines is available (Rodrigues 1999).

Case-control studies are still done within databases (for example Black 2003), presumably because they are so much more economical in data handling terms for rare outcomes, and therefore quick to do. One disadvantage of individually matched case-control studies compared to the unmatched design is that cases that are not matched can't be used (Hocine 2007); looser forms of matching, such as frequency matching, may circumvent this problem.

4.3.2 Variants on the case-control design

4.3.2.1 Nested case-control method

Nested case-control studies are case-control studies undertaken within cohort studies. Each incident case is matched with a (usually fixed) number of controls sampled from the risk set for that case, namely individuals who have not experienced the outcome event at the time of occurrence of the case. Nested case-control studies, which were introduced by Liddell (1977)

and formalised by Breslow (1983), offer an attractive alternative to full cohort analysis when obtaining covariate information is expensive, or when it is desired to simplify the analysis (notably when there are few cases in a very large cohort). The analysis is as for matched case-control studies. The design has been used frequently for assessing vaccine safety, even within databases where there is no extra burden of data collection on covariates; see Hernan (2004) and Jick (2008) for examples relating to hepatitis B vaccination and multiple sclerosis, and oral polio vaccine and intussusception, respectively, both undertaken within the UK's General Practice Research Database.

4.3.2.2 Counter matched nested case-control method

Counter matching was introduced by Langholz (1995) as a way of improving the efficiency of nested case-control studies, when information on a covariate (the counter-matching variable) correlated with the exposure is available for all members of the cohort. The risk set for each incident case is stratified on the counter matching variable, and controls are sampled within the strata. A gain in efficiency is obtained when the counter matching variable is correlated with exposure, and the main effect of the counter-matching variable is estimable. The analysis proceeds as for nested case-control methods, with an adjustment for the stratified sampling scheme. The method has been used to good effect when a correlate of exposure is available, but more detailed exposure information is costly to obtain. We are not aware of applications to vaccine safety.

4.3.2.3 Case-cohort method

The case-cohort design (also called case-base design) is another instance of a case-referent design undertaken within a cohort study. The case-cohort design was proposed by Prentice (1986) as a way of reducing the burden of data collection on covariates. In this design, a subcohort is sampled from the original cohort, and combined with all the cases from the original cohort not already sampled within the subcohort. The underpinning statistical theory, which requires the construction of a pseudolikelihood, is more complex than for the nested case-control design (Self 1988) owing to the possible overlap between cases and controls. To our knowledge, the design has not been used in vaccine safety studies; one potential advantage over the nested case-control method is that it can more readily be used to analyse multiple outcomes. Moulton (1995) describes an application of case-cohort methodology to the estimation of vaccine efficacy. The method is also called the case-coverage method.

4.3.2.4 Case-coverage method with external coverage cohort

Case-coverage designs can also use a coverage cohort distinct from that in which the cases arise. This greatly simplifies the analysis, since it avoids the possible overlap between cases and controls in the case-cohort method, which necessitates more complex analytical techniques. The technique has been used by Miller (2013) to assess the association between pandemic influenza vaccine and narcolepsy. The cases were ascertained from sleep centres and paediatric neurology centres; the coverage cohort, matched for age and other variables, was drawn from the weekly returns of the Royal College of General Practitioners. The analysis method was as for the screening method for estimating vaccine efficacy (Farrington 2002), the key difference being that the expected odds are based on a sample rather than a census. The method does not allow for uncertainty in the expected odds. This is immaterial when the coverage cohort is large,

the major issue being the possible bias if cases are drawn from a population with a different vaccination profile from that of the coverage cohort.

4.3.2.5 The case-centred method

The case-centred method is similar to a nested case-control study, using all rather than a sample of relevant controls within each risk set. Originally developed by Fireman (2009) to measure influenza vaccine effectiveness against death, the method has also been applied to evaluate vaccine safety (Rowhani-Rahbar 2012). In this method, the vaccination odds of each case is compared to the expected odds in at-risk individuals in view at the time of occurrence of the case, and matched to the case on pre-specified stratifying covariates. The analysis is via the same statistical model as the screening method for estimating vaccine efficacy (Farrington 1992), the main difference being that each data point corresponds to a single case and that the exposure odds are obtained from the same database as the cases, rather than from an independent census. Fireman (2009) states that the method is based on the same likelihood function as a stratified Cox regression model. However, it does not appear to allow for uncertainty in estimation of the expected odds.

4.3.2.6 The test-negative case-control method

The test-negative case-control design is a special case of the matched case-control design, in which participants are selected among individuals satisfying some common clinical criterion of 'caseness', who are subsequently subjected to confirmatory testing. The cases are then chosen among the test positives and the controls among the test negatives. This design is used to evaluate vaccine effectiveness (Orenstein 2007, De Serres 2013) but has not been used for evaluating adverse events. In this context, the method may be useful to minimise selection and ascertainment biases, especially for complex conditions such as auto-immune or neurological diseases. However, care is required to avoid bias in situations where the test used to distinguish between cases and controls involves an assessment of causality.

4.4 Self-controlled methods

Self-controlled methods have been increasingly used since the 1990s. The term 'self-controlled' here means that each individual is controlled by her/himself; note that this use differs from that of Ray (1989) described in Section 4.2.4.1. In consequence, time-invariant confounders acting multiplicatively on the baseline incidence are automatically controlled for, even if unmeasured. A further consequence of self-control at the individual level is that non-cases contribute no information.

There are two basic self-controlled study designs, each with variants: case-crossover, and self-controlled case series (SCCS). There has been some confusion in the terminology used to describe these various designs, notably in the literature of environmental epidemiology where some versions of the case-crossover design are actually SCCS. The key distinction is that the case-crossover method is based on case-control logic, the event being regarded as fixed and exposures random, whereas the SCCS method is based on cohort logic, the time of occurrence of the event being considered to be random and exposures to be fixed.

4.4.1 The case-crossover method

This is a type of risk-interval case-control study in which one or more control intervals are taken from the pre-event history of each case. The method was developed by Maclure (1991). In vaccine safety research it has been used by Confavreux (2001) and Ki (2003). The key assumption is that exposure risk is time invariant, which is generally not valid for paediatric vaccines. Actually, the required assumption is a little stronger than this: exposures must be exchangeable across time periods (Vines 2001).

4.4.2 Variants of the case-crossover method

Several attempts have been made to weaken the assumption of exposure time-invariance required by the case-crossover method. So far as we know, these variants have not been used in vaccine safety studies.

4.4.2.1 The case-time-control method

The case-time-control method (Suissa 1995) supplements the case-crossover design with a correction for trends in exposure. The method has been further extended by Jensen (2014) to model exposure trends with splines. The idea is to supplement the case-crossover analysis by a second case-crossover analysis in controls using the same exposure. This second case-crossover analysis estimates the effect of trends in exposure, for which the original analysis can then be corrected.

4.4.2.2 The case-case-time-control method

The case-time-control method can potentially be biased through an inappropriate choice of external controls to assess trends in exposures. Wang (2011) proposed a 'case-case-time-control' method to reduce such bias. In this method, the supplementary case-crossover analysis is undertaken not in separate controls, but in future cases. Thus, the impact of exposure time trends is estimated within the same population as the impact of exposures on outcomes.

4.4.2.3 Other choices of referent windows in the case-crossover method

A large literature on the choice of referent widows for case-crossover designs has emerged in environmental epidemiology. Much of this literature seeks to mitigate biases associated with failure of assumptions, notably non-constancy (and non-exchangeability) of exposures. The bidirectional case-crossover method uses control widows after the event as well as before it (Navidi 1998). Some versions of this method (in which the event is considered to be random) are identical to SCCS. The time-stratified case-crossover design (Lumley 2000) method is also a form of SCCS. Vines (2001) and Whitaker (2007) discuss these and other case-crossover designs and their relationship with SCCS.

4.4.3 The self-controlled case series (SCCS) method

This was developed specifically for vaccine safety evaluation (Farrington 1995, 2006) and has been widely applied in this field (see Weldeselassie 2011 for a recent review). The method is not affected by trends in exposure risks, but is prone to confounding by time-varying covariates, which can however be controlled explicitly; two important time-varying confounders for

vaccine studies are age and (particularly for seasonal vaccines) season. The key assumptions of the method are that events influence neither the observation period nor future exposures.

Various designs have appeared under different names, which turn out to be versions of the basic SCCS model. A recent such example in vaccine safety evaluation is the 'self-controlled risk interval' design (Greene 2012), which is an SCCS design with short observation periods and generally without adjustment for age. In this specific application, the observation period was defined to be 1 to 84 or 127 days from vaccination, the first 42 days constituting the risk period.

4.4.4 Variants of the SCCS method

The assumptions that events influence neither the observation period nor future exposures may fail if, for example, the event is death (in which case the observation period ends at the event) or the event is a contra-indication to vaccination (in which case the event influences future vaccinations).

4.4.4.1 Variants to handle event-dependent observation periods

For single dose vaccines, deaths can be handled by starting the observation period at vaccination and using a nominal end of observation corresponding to the end of data capture. For multi-dose vaccines, a simple method has been proposed by Kuhnert (2010) and applied to sudden infant deaths after hexavalent vaccine. This method works for multi-dose vaccines provided there is a minimum time separation between doses. It can also be applied to single-dose vaccines, in which case it is identical to the self-controlled risk interval method of Greene (2012).

Sometimes, the event of interest is not death itself, but may increase short-term mortality. Thus, the event is still correlated with the observation period. An SCCS method to cope with this has been developed (Farrington 2011), which involves modelling the time from event to end of observation using a mixture model. This variant has not been applied to vaccines, as events associated with high short-term mortality (other than death itself) are seldom of concern.

4.4.4.2 Variants to handle event-dependent exposures

A simple adjustment to the standard SCCS method if the event has only a short-term impact on subsequent vaccination is to include a pre-vaccination dummy 'risk' period (Farrington 2006). This has been used, for example, in studies of OPV and intussusception (Galindo Sardinas 2001), to allow for likely delay in vaccination after a hospitalisation.

Longer term effects (as arise if the event is a contra-indication to vaccination) can be handled by a further variant of the SCCS method (Farrington 2009). Applications of this latter variant SCCS method include rotavirus vaccine and intussusception, and influenza vaccine and GBS (Andrews 2009). This further variant is considerably more complex than the standard SCCS method, and involves construction of a pseudolikelihood function.

4.5 Sequential methods

The methods described so far have deterministic sample sizes and study times. Sequential methods, on the other hand, involve accumulating evidence until a threshold is reached at which

point a decision is reached. Sequential methods are particularly useful for monitoring recently introduced vaccines (or new schedules), for which extensive past data are unavailable, yet early results are needed.

Sequential methods can be continuous (in which case the test statistic is updated when each new case arises) or grouped. Continuous updating methods are commonly used in sequential clinical trials (Whitead 1997); group sequential methods may be more readily applicable in observational epidemiology; Nelson (2013) provides an example of group sequential monitoring of the safety of a pentavalent DTaP-IPV-Hep B vaccine.

Several sequential methods specifically designed for vaccine safety surveillance have been proposed in recent years. Most are based on the sequential probability ratio test (SPRT), originally proposed and studied by Wald (1947). Alternatively, routine surveillance of established vaccination programmes can be undertaken using statistical process control methods originally developed by Shewhart (1931) and Page (1954).

The literature on sequential methods is vast; all the applications in vaccine safety monitoring are group sequential methods. Note that these applications are formulated within a hypothesis testing framework: the estimates of vaccine effect obtained when an alert is triggered may therefore be biased.

4.5.1 Sequential probability ratio test (SPRT) methods

These methods involve monitoring the likelihood ratio which is updated at regular intervals, and result in a decision to accept or reject the null hypothesis of no association, once the likelihood ratio reaches a pre-determined boundary.

4.5.1.1 The MaxSPRT method

This is the approach used in the VSD as part of its Rapid Cycle Analysis (Yih 2011, Davis 2013), which runs weekly. The statistical method it is based on is the maximised SPRT (Brown 2007, Kulldorff 2011) which the authors have called MaxSPRT. This is a special case of a sequential generalised likelihood ratio test. There are two versions of this MaxSPRT, Poisson and binomial. The Poisson MaxSPRT contrasts the current post-vaccination event rate with the expected rate in the absence of vaccination, calculated using historical data or a control vaccine, whereas the binomial MaxSPRT uses concurrent controls (Lieu 2007). There is also a 'self-control period' version, the post-vaccination incidence being compared to the immediate pre-vaccination incidence (Yih 2011).

In these various implementations, the critical thresholds are computed numerically, based on the desired type I error probability and the maximum duration of surveillance.

The MaxSPRT method has also been used for 'signal strengthening' in the UK, in relation to influenza vaccine (Bryan 2010) and HPV vaccine (Donegan 2013); see Section 4.6.

4.5.1.2 The conditional MaxSPRT method

A drawback of the MaxSPRT method is that the expected values are assumed to be perfectly estimated. When based on small samples, this may bias the procedure towards signalling. To

this end, a conditional Poisson maximised SPRT method was developed (Li 2009). This adjusts for the estimation of the historical rates, but requires a different perspective from MaxSPRT, the threshold values depending on the expected numbers of historical as well as current events.

4.5.1.3 Other sequential generalised likelihood ratio methods

Shih (2009) criticises the MaxSPRT approach on various counts, notably because it is not the most efficient (and hence powerful) generalisation of the SPRT for composite alternative hypotheses. They propose an alternative method, along with various applications to vaccine safety monitoring in clinical trials and surveillance settings.

4.5.1.4 Sequential case series analysis

The historical contemporary control based MaxSPRT methods are prone to some degree of confounding, as both rely on between-individual comparisons. Hocine (2009) proposed an SPRT version of the self-controlled case series method, which removes such confounding. Rather than go down the generalised likelihood ratio route, Hocine (2009) exploited the adaptive scheme of Huang (2004) to bound the type I error probability while avoiding specification of an alternative hypothesis. Various applications to vaccine safety monitoring are discussed. Unlike MaxSPRT, no pre-specified study duration is required with this adaptive scheme, and the null hypothesis (that there is no association) is never accepted. A major disadvantage of the SCCS-based SPRT is that it could not be run weekly, as the self-controlled case series is retrospective and requires sufficient time to have accrued to define disjoint observation periods. Hocine (2009) suggests a 6-monthly run.

4.5.2 CUSUM methods

An alternative to the SPRT is the cumulative sum, or CUSUM, method, originally proposed by Page (1954). Like the SPRT, increments based on the log likelihood are summed. Also like the SPRT, the CUSUM signals if it exceeds a pre-specified boundary. Unlike the SPRT, the CUSUM signals (eventually) with probability 1 even if there is no association. For this reason, its operational characteristics are described in terms of average run lengths (ARLs) rather than probabilities: a high ARL is required if there is no association, a short ARL if there is.

CUSUM methods have not received much interest in vaccine surveillance, unlike other areas of surveillance where they are very commonly used. A CUSUM method for sequential surveillance based on the self-controlled case series method has been proposed by Musonda (2008), where it is argued that while the SPRT is best suited to monitor a newly introduced vaccine, the CUSUM is best suited for long-term monitoring of an existing vaccine which is presumed safe, in case the safety profile deteriorates for some reason. As with the SCCS-based SPRT, the SCCS-based CUSUM cannot be incremented in short intervals.

4.6 Signal strengthening methods

There is a well-understood distinction between signal generating methods, and signal confirmation methods. Signal generating methods are data mining techniques to uncover possible associations, without any prior hypotheses. Signal confirmation techniques, on the other hand, aim to evaluate the strength of evidence for a specific signal (that is, a hypothesized association between a vaccine and an adverse event). So-called ‘signal strengthening’ methods

occupy the ground between these two types of approaches. In broad terms, a signal strengthening method is used to assess the plausibility of a signal, yet without achieving the rigour required of signal confirmation .

Signal strengthening lacks clear definition; the fact that a method does not achieve the rigour associated with signal confirmation may be due, either, to limitations of the data, or to limitations of the methods, or both. Nevertheless, such methods have been found useful as a way of screening signals (or potential associations), prior to undertaking more formal confirmatory analyses.

4.6.1 Observed – Expected analyses

Observed – Expected (or O-E) analyses seek to compare rates of adverse events obtained through surveillance, with the rates expected, as obtained from national statistics, administrative databases, or publications, for a non-vaccinated population as similar as possible in its demographic and other relevant characteristics to the vaccinated population. Typically, in such analyses, control for confounders is limited, and the populations from which the observed and the expected values are obtained may differ. However, such analyses are valuable in helping to contextualise the frequencies of adverse events obtained in surveillance systems.

Nazareth (2013) undertook O-E analyses of a pandemic H1N1-2009 influenza vaccine, using observed rates obtained prospectively from general practitioners, and observed rates from a variety of sources including publications and national statistics. Bryan (2010) and Donegan (2013) are further examples of O-E analyses, for H1N1 influenza vaccine and HPV vaccine respectively. These analyses were implemented sequentially using the MaxSPRT (see Section 4.5.1.1), using spontaneous reports to obtain observed rates (adjusted for under-reporting using a range of assumptions), and data from the Clinical Practice Research Database to obtain expected values.

4.6.2 Self-controlled analysis of spontaneous reports

This method seeks to apply the SCCS method to spontaneous reports data (Escolano 2013). It works non-parametrically for 2-dose vaccines provided there is a different risk profile for each dose, or parametrically (with an assumption about the reporting rate and how it evolves with time since vaccination) for single doses. It has been applied to rotavirus vaccine safety (Escolano 2011).

The method cannot be relied upon to confirm (or dismiss) a signal, as it is based on some strong assumptions regarding the rate at which adverse events are reported following vaccination in spontaneous reports databases. Rather, it aims to control for some fixed confounders *à la* SCCS, and thus strengthen the interpretation of signals generated from spontaneous reports. In this sense, it may be regarded as a 'signal strengthening' method.

4.7 Ecological Methods

Ecological methods are those that do not involve individual exposure or outcome data. Generally, they are deemed to provide less compelling evidence than studies based on individual data, though this will depend on the precise circumstances of their implementation.

It is difficult to describe ecological methods in any generality, as they are very dependent on specific circumstances. However, they are generally relatively easy to implement.

4.7.1 Before and after comparisons

A simple ecological analysis is to compare rates of a specified event in a given population after the introduction of routine vaccination, with rates in a comparable population prior to this. This method was used, for example, by Donegan (2013) to study chronic fatigue syndrome in girls aged 12 – 20 years, before and after the introduction of HPV vaccine in the UK.

Before and after comparisons, however, are confounded by changes in reporting and diagnostic practice, and by changes in the natural incidence of some events. For example, the proponents of a link between MMR vaccine and autism made much of the apparent temporal correlation between increased vaccination and increased autism, the latter most likely being due to improving ascertainment and changing case definitions for autism spectrum disorders (Taylor 1999).

4.7.2 Natural experiments based on vaccination schedule changes

If a vaccination schedule is suddenly shifted from age A to age B, then a drop in event incidence in age group A and an increase in age group B might be supportive of a causal relationship. See Shields (1988) for an example related to pertussis vaccination and febrile convulsions in Denmark, where such a phenomenon was indeed observed.

4.7.3 Short vaccination campaigns or pulse vaccination

In some countries, vaccines are administered in pulse vaccination campaigns that last only a few days or weeks. Documenting the changes in event incidence before, during and after the campaign can help throw light on a causal association, or lack of one: if a marked peak in incidence is observed shortly after the campaign, rates subsequently returning to pre-campaign levels, this may be suggestive of an association. Examples include an application to measles vaccination and Guillain-Barré syndrome in several Latin American countries, which did not support an association (da Silveira 1997), and an application to MMR vaccine and aseptic meningitis in Brazil, which did (Dourado 2000).

4.8 Other statistical considerations

The choice of design does not just involve the type of study to use, but also how to implement that study type. In this section three such generic issues are dealt with in a little more detail, though not exhaustively.

4.8.1 Confounder control

Confounding is a major issue in all pharmacoepidemiology, and in vaccine safety investigations in particular, owing to the fact that vaccines are not administered randomly. Accordingly, several approaches to controlling confounding have been developed which are briefly reviewed below. We also discuss adjustment for time-varying confounders.

4.8.1.1 Matching and direct adjustment

These are the most commonly used methods for confounder control. Matching is common in case-referent designs (see Section 4.3), but can also be used in cohort designs: see Klein (2010) for an example relating to DTaP vaccine. Direct adjustment in the analysis is used in both cohort and case-referent designs, sometimes in addition to matching.

These methods require that the confounders are (a) known and (b) measured. Generally, it is not possible to match on more than a few confounders (but see Section 4.8.1.2 on propensity scores). Matching or adjusting on variables that are not confounders can incur a cost in reduced efficiency, a phenomenon known as 'overmatching'. Matching and adjusting for variables that are on the causal pathway can produce biased results.

4.8.1.2 Propensity scores

Propensity scores can be used to correct for differences in the propensity for individuals to be vaccinated. Briefly, a separate analysis is carried out to model the probability that an individual is vaccinated. This model produces a score, which may then be incorporated into the model for the outcome of interest, or used for matching.

It would appear that propensity scores are only to be recommended when there are a lot of covariates to adjust and not very many events (Cepeda 2003). Recent work suggests that propensity scores seldom produce different answers from direct adjustment (Shah 2005, Sturmer 2006). They also need to be used with caution to avoid over-adjustment (Senn 2007): propensity scores are concerned about reducing bias, sometimes at the cost of efficiency. And of course, propensity scores cannot adjust for unmeasured or unknown confounders.

Propensity scores have been used in vaccine safety studies, notably in relation to non-specific effects, but in this context as more generally, often produce rather little difference from standard methods of adjustment (see references in Farrington 2009). Possible reasons for their lack of effectiveness are discussed in Farrington (2009), where it is suggested that time-varying propensity scores might work better, since vaccination is age-dependent.

Studies may incorporate propensity scores by matching or adjustment in a regression model, or both. Whether the analysis should then be adjusted for the matching has been the subject of some debate: see Austin (2008) and subsequent discussion.

4.8.1.3 High dimensional propensity scores

High dimensional propensity scores are propensity scores constructed from hundreds of covariates. This approach has been proposed as a way of adjusting for measured and unmeasured confounders, through their correlation with other variables (Schneeweiss 2009, Garbe 2013). The method has not been fully evaluated theoretically or by simulation. A note of caution is in order until such a formal evaluation has been undertaken, in view of the potential for over-adjustment. To our knowledge, high dimensional propensity scores have not been used in vaccine safety assessments.

4.8.1.4 Instrumental variables[5]

Instrumental variable analysis provides one of the very few ways of allowing for unmeasured confounders, as well as measured confounders. An instrumental variable must satisfy three conditions: it must be related to exposure, preferably strongly so; it must be independent of the confounders; and it must be independent of the outcome, conditionally on the exposure and the confounders (Greenland 2000). A review of the uses of instrumental variables in pharmacoepidemiology has been undertaken by Chen (2011).

We are unaware of the explicit use of instrumental variable analyses for vaccine safety assessment, however the approach has been used to assess the efficacy of influenza vaccines (Yoo 2006, Groenwold 2010, Trogdon 2010, Wong 2012). The critique of some of these studies undertaken by Chen (2011) emphasizes the difficulties involved in instrumental variable analysis, notably choosing a sufficiently strong instrument (that is, one strongly associated with exposure), while meeting the stringent independence requirements. The instruments used in these studies include: history of arthritis; history of gout; antacid medication; GP-specific vaccination rates; locality-specific vaccination rates; history of influenza; vaccination in the previous year. Further work on the application of instrumental variables to vaccine safety studies is clearly required.

4.8.1.5 Self-controlled methods

The self-controlled methods reviewed in Section 4.4 control for all time-invariant confounders, whether known and measured or not, provided that these act multiplicatively on the baseline incidence. These methods are particularly useful for use in administrative databases, where information of confounders is missing or incomplete. In particular, self-controlled methods can sometimes more effectively adjust for confounding by indication than standard methods. For an example, see Kramarz (2000) on influenza vaccine and asthma exacerbations. How sensitive self-controlled methods are to failure of the multiplicative assumption remains to be determined.

4.8.1.6 Exogenous time-varying confounders: age and season

In vaccine studies in children, it is essential to adjust for age as paediatric vaccines are given according to very prescriptive age-dependent schedules. In addition, many adverse events are age-dependent as well. Similarly, studies of seasonal vaccines (like influenza vaccine) and seasonal events should incorporate an adjustment for season: often the study period is stratified by influenza season, as in Kramarz (2000). Sometimes both age and season adjustment is

[5] This section was contributed by Elisa Martin Merino (AEMPS).

necessary; see Galindo Sardinas (2001) for an example relating to OPV vaccination and intussusception. Some designs, notably nested case-control studies (see Section 4.3.2.1) and related designs, lend themselves to matching on age and/or season.

4.8.1.7 Endogenous time-varying confounders: the healthy vaccinee effect

This is perhaps the single most important yet incompletely addressed issue in statistical methods for vaccine safety evaluation. The magnitude of the bias has been described in detail by Fine (1992), and it can be considerable. Briefly, children who are unwell will have their vaccinations deferred, and this can bias the estimated effect to favour the vaccine (that is, the relative risk will be underestimated). Fine (1992) provides extensive evidence for this effect in studies relating to vaccinations in infancy and Sudden Infant Death Syndrome.

4.8.2 Independence of detection and confirmation studies

The availability of large administrative databases with clinical data and information on vaccinations (eg VSD, CPRD, Mini-Sentinel) has led to a great increase in numbers of statistical analyses from such databases, and to a new impetus to develop new methodologies. Such developments are generally to be welcomed. However, care is required to maintain the key distinction between signal generation and signal confirmation. If a database is used to generate a signal, it cannot validly be used to confirm that signal. One way to ensure degree of independence between detection and confirmatory investigations is to split the data into two sets, one to be used as a training set for signal detection, and the other as a test set for signal confirmation. 'Signal strengthening' analyses (see Section 4.6) may further complicate the picture if undertaken in the same database as the confirmatory analyses.

4.8.3 Meta-analyses of vaccine studies

So far, attention has been focused entirely on single studies. For rare adverse events, these may lack power. It is also relevant, and important, to consider between-study variability, even when the power of individual studies is good. The obvious way to do this is via meta-analysis.

Meta-analyses and systematic reviews of vaccine safety have been done within the Cochrane Collaboration (Jefferson 2003a, 2003b, 2004). Some of these have included criticism of the quality of vaccine safety studies. At the same time, the methodological criteria used in systematic reviews may themselves need to be updated to keep up with new study designs. Meta-analyses involving new methods of assessment have been undertaken to good effect, notably Dodd (2013) and Salmon (2013) which both investigated the association between H1N1 (2009) influenza vaccine and Guillain Barré syndrome.

5. EVALUATION OF THE AVAILABLE DESIGNS

The following evaluations are for groups of study designs, rather than individual variants. They comprise two parts: first, an overview in tabular form against the criteria set out in Section 3; second, a discussion of the evaluation, evidence-based where appropriate.

As previously noted, some generic issues common to all methods are not systematically commented upon under each heading, in order to avoid repetition. These issues include:

- The need for events and exposures to be ascertained independently, failure of which may result in bias.
- The impact of misclassification of the disease outcome or the vaccine exposure: if misclassification is at random, this will have the effect of biasing effect estimates towards the null (i.e. absence of an effect).
- The impact of errors in the date of vaccination or in the length of the risk window, which will also induce misclassification and, if it occurs at random, will bias effect estimates towards the null.
- The possible bias arising from counting multiple instances of the same event (for example when several hospital admissions arise as part of the same illness episode).

The assessments do not cover ecological methods (reviewed in Section 4.7) or other statistical considerations (reviewed in Section 4.8), as these do not comprise generic designs. Thus, ecological methods are very dependent on circumstances, while other statistical considerations can typically be relevant to several different methods.

5.1 Designed vaccine introduction

Designed vaccine introduction was considered in Section 4.1. This evaluation is focused on the stepped-wedge design.

Criterion[6]	Overview
M1	Absolute risks in individually or cluster randomised unvaccinated and vaccinated groups; indirect effects are estimable in cluster-randomised stepped wedge designs
S1	Good though not generally as high as for non-stepped studies; power increases with the number of clusters and the number of steps, depending on the intra-cluster correlation
S2	Not at issue, as sample sizes are intentionally large
T1	Suitable for sequential monitoring of short-term adverse effects
T2	Intentionally large: aims to cover the entire target population
R1	Exchangeability of the cluster-randomised sites (if cluster randomised)
R2	Randomised and population-based, hence robust to confounding; information bias possible if vaccination status is not concealed
O1	Effective population-based data tracking systems required
O2	Very complex, novel, requiring very substantial resources and commitment; possible ethical issues; new methodology and piloting required

[6] The criteria are summarized in Section 3.6.

The stepped wedge, cluster-randomized introduction of a new vaccination programme is not commonly used, though there has been some methodological work on the analysis of stepped wedge cluster-randomised trial by Hussey(2007). These authors conclude:

'The stepped wedge design provides an innovative choice for a cluster randomized crossover trial that is subject to constraints that limit the use of more conventional designs. The stepped wedge seems particularly suited to investigations of community level public health interventions that have been proven effective in individual level trials and so-called "phase IV" effectiveness trials.'

The analysis of such trials proceeds using random effects models (via GLMM or GEE) to estimate the within-cluster and between cluster variances. The methodology is very similar to that of cluster-randomised trials, and the above evaluation is largely based on our understanding of such analyses. Specific issues relating to the use of the stepped wedge design for vaccine introduction include the likely inability to conceal vaccination status from participants and vaccinators (Brown 2006). Designed vaccine introductions are likely to be most useful for adverse events that occur shortly after vaccinations, though long-term follow up could be planned in principle.

The major obstacle to the use of stepped-wedge cluster randomised vaccine introduction is likely to be logistical. A rationale for the stepped wedge design is usually that it is not possible to introduce a new intervention at all locations at the same time – this is seldom an issue with routine vaccination programmes in Europe. The added complications of a stepped wedge introduction may be met by resistance from health authorities. Since the design involves withholding the intervention from some individuals for a time, it may also raise ethical issues. To our knowledge, an evaluation of the practicality of such an approach to vaccine introduction in Europe has so far not been undertaken.

On the other hand, the stepped wedge introduction of a new vaccine appears to meet many of the requirements of benefit – risk monitoring, including sequential evaluation and evaluation of indirect effects. Further methodological work specifically on the application of this approach to mass vaccination and the estimation of direct and indirect effects may be warranted.

5.2 Cohort studies

Cohort studies were considered in Section 4.2. The present evaluation is focused on non-randomised studies; some comments on randomised comparisons, and on other variants of the cohort design, are made in the discussion.

Criterion	Overview
M1	Absolute risks in unvaccinated and vaccinated groups
S1	Usually taken as 'gold standard'
S2	Only an issue for rare events
T1	Prospective cohort studies can be implemented in sequential mode, provided the data on events and denominators are rapidly updated

Criterion	Overview
T2	Operational sample size is usually large: for database studies, typically the entire database
R1	Statistical assumptions vary according to analysis method; in the absence of randomisation the key epidemiological assumption is that all confounders have been included in the model
R2	Statistical assumptions are usually verifiable and seldom pose a serious issue; the assumption about confounders cannot be checked without external information, and results may be sensitive to its failure; the direction of bias is usually unknown
O1	Counts or times of events and denominators, with covariate information
O2	Easily implemented in administrative databases; onerous if data checking is required

Cohort studies are sometimes considered to be the 'gold standard' in observational studies, perhaps because of their superficial similarity to randomised controlled trials (the key difference being of course that essential ingredient, randomisation). For this reason, they are usually the benchmark against which other methods are evaluated – certainly in terms of asymptotic relative efficiency and power. Absolute risks, hazards or survival functions are estimable from cohort studies, so both absolute and relative effect measures can be derived.

Cohort studies can be undertaken retrospectively in pre-existing databases, or prospectively in which case they lend themselves to sequential analysis using methods reviewed in Section 4.5. The main limitation on sequential implementations for observational, rather than clinical trial settings, is the rapid updating of data, and suitable control for confounding. The very numerous methods for group sequential monitoring of clinical trials can, in principle, be applied to observational cohorts. The issues are considered in greater detail in Section 5.5.

The assumptions required in the analysis of cohort studies depend on the method used. For Cox regression, for example, a proportional hazards assumption may be needed (Cox 1972); for generalised linear modelling approaches, specific distributional assumptions are required, along with the choice of link function (McCullagh 1989). These assumptions and modelling choices can generally be checked, using methods that are now standard in many areas of epidemiology (Breslow 1987).

If data are missing or censored in cohort studies, some assumption on the mechanism giving rise to the missing data (and particular, whether the data are missing at random or not) is usually required – which may not be verifiable. However, the issue of missing data is generally less problematic for cohort studies of vaccines, where post-vaccination risk periods are typically brief and events of interest are often acute and do not involve repeated measurements. Two exceptions are: cohort studies of child survival, where informatively missing vaccination has been an issue, though exclusively in very specific income-poor settings (Farrington 2009); and long-term follow-up of randomised cohorts, described in Section 4.2.2. Generally, technical modelling issues seldom present critical problems of robustness for cohort studies of vaccine safety.

The major modelling assumption in all cohort studies is that all important confounders have been included in the model. This assumption is untestable within the study, since an important confounder may not have been measured. External information on confounders is needed to provide reassurance that the study avoids major bias. The problem of confounding is compounded by the fact that it is not usually possible to say in which direction any bias would lie. Commonly used methods of confounder control for non-randomised cohort studies were reviewed in Section 4.8.1; with the exception of instrumental variable approaches, which rely on untestable assumptions and may incur a severe loss of efficiency, all such methods require the confounders, or at least correlates of them, to have been measured – a further untestable assumption. The only sure way to remove confounding is randomisation, considered in Section 4.2.1, which is likely to be limited to vaccine – vaccine comparisons.

Cohort designs are very versatile. Provided the data are available, they lend themselves to sophisticated analyses, for example competing risks (Andersen 2012, Fine 1999), recurrent or multiple events (Wei 1997), and modelling of dose-response relationships (Abrahamovicz 2012). We are not aware of any major restrictions on their use. Standard risk-interval cohort studies, reviewed in Section 4.2.4.1, and which are the norm in vaccine safety investigations, do not present any additional challenges over parallel group designs, other than correct specification of the vaccine-associated risk period. Sequence Symmetry Analysis, reviewed in Section 4.2.4.2, would benefit from some methodological work to clarify its properties and the assumptions upon which it rests. This design, however, has not so far been used in vaccine safety studies.

5.3 Case referent studies

This evaluation is focused on the standard case-control design reviewed in Section 4.3.1. Its many variants, reviewed in Section 4.3.2, are briefly evaluated in the subsequent discussion.

Criterion	Overview
M1	Odds ratios of exposure in cases compared to controls; hazard ratio in nested case-control and case-cohort designs; absolute hazards estimable from nested case-control studies
S1	Efficiency and power close to that of cohort studies when there are 4 or more controls per case
S2	Only an issue when there are few cases
T1	Nested case-control studies could in principle be implemented in sequential mode, provided the data on events and risk sets are rapidly updated
T2	Operational sample size much smaller than for cohort studies
R1	The model structure relies on statistical assumptions; controls must be sampled from the same population as the cases; the key epidemiological assumption is that all confounders have been included in the model
R2	Statistical assumptions are usually verifiable and seldom pose a serious issue; the appropriateness of the controls may not be easy to check and an inappropriate choice may induce bias; the assumption about confounders cannot be checked without external information, and results may be sensitive to its failure; the direction of bias is usually unknown

Criterion	Overview
O1	Case-control sets (matched if required) with covariates
O2	Easily implemented; less onerous than cohort studies if data checking is required

Standard case-control studies are retrospective, and do not permit the estimation of absolute risks or rates: only the odds ratio may be estimated, which for rare events approximates the relative risk.

However, if the study includes all cases in a defined population of known size, then absolute rates can be estimated, at least approximately, as was done in the NCES (Alderslade 1981). In contrast, nested case-control studies are essentially prospective designs, and relative hazards may be estimated from them. In addition, provided that all cases in the underlying cohort are included, and the cohort size and risk set sampling fractions are known, then good estimates of absolute risks may be obtained (Langholz 1997). Similarly, absolute risks may also be estimated from case-cohort studies.

The efficiency of case-control studies relative to cohort studies with the same cases has been studied by Ury (1975), who showed that the asymptotic relative efficiency is k/(k+1) where k is the number of controls per case. Thus if 4 or more controls are used per case, the asymptotic relative efficiency is in excess of 80%. The relative efficiency and power of matched versus unmatched case-control studies, and looser matching such as frequency matching, has been investigated by Sturmer (2001). The relative efficiency of case-cohort and nested case-control studies has been investigated by Langholz (1990), who conclude that little efficiency is to be lost, and some might be gained, from nested case-control studies.

Nested case-control studies lend themselves to sequential analysis, including continuous monitoring, under the same proviso as cohort studies (namely rapid data updating).

The assumptions required in case-control studies are similar to those of cohort studies, which they often inherit: for example, nested case-control methods are based on a similar partial likelihood as used in the Cox model. One important requirement for individually matched case-control studies is that matching should be taken into account in the analysis; this method of analysis is often described as ‘conditional logistic regression’. Failure to do this can produce bias, as the asymptotic theory upon which maximum likelihood estimation is based may not be valid.

A key additional requirement is that the controls must be selected from the same population of the cases – or at least, that any differences can be controlled by including suitable variables in the model. The other key requirement, shared with cohort studies, is that all confounders have been included in the model, or eliminated by matching or stratification. Like cohort studies, omission of a confounder may produce biased results, and the direction and magnitude of the bias is usually unknown.

Case-control studies, like cohort studies, are extremely versatile, and we are not aware of any important restrictions on their applicability in principle. For example, they can be used in

conjunction with both indefinite and time-limited risk periods. Their main advantage over cohort studies is their greatly reduced operational sample size, especially for rare events.

5.4 Self-controlled methods

Self-controlled methods are, to our knowledge, the only so far available that control for unknown as well as known (time-invariant) confounders, whether measured or not; the only close contender in this respect is analysis using instrumental variables (reviewed in Section 4.8.1.4), where the assumptions involve the confounders, which must therefore be known, but can in principle be unmeasured. This section is split into two subsections, as the two basic self-controlled designs differ in important respects.

5.4.1 Case-crossover designs

This evaluation is focused on the standard case-crossover design, reviewed in Section 4.4.1; the variations on the standard design, reviewed in Section 4.4.2, are considered in the subsequent discussion.

The case-crossover method is a particular instance of a matched case-control design with exposures in specified risk intervals; controls are not separate individuals but control periods within the past history of each case, prior to the case period. Thus the method yields odds ratios, as in a case-control study, and has similar efficiency and power properties as case-control studies.

In principle, a case-crossover design could be readily implemented in a sequential framework, the likelihood ratio being incremented by the contribution of each new case as it arises. However, we are not aware of such an implementation ever having been attempted.

Criterion	Overview
M1	Odds ratios of exposure in the case period compared to control periods
S1	Efficiency and power close to that of cohort studies when there are 4 or more control periods per case
S2	Only an issue when there are few cases
T1	Case-crossover designs could in principle be implemented in sequential mode, with continuous monitoring if exposure information were readily available
T2	Operational sample size smaller than for cohort or case-control studies; for the case-time-control method it is the same size as for a case-control study
R1	The method is not applicable with indefinite periods; the key assumption is that exposures are globally exchangeable, which implies that the probability of exposure is constant within individuals at all times; this assumption can be relaxed in case-time-control and case-case-time-control designs
R2	Results are sensitive to failure of the exchangeability (and hence constant exposure probability) assumption; the direction of bias can be deduced if there is a monotone trend in exposure
O1	Cases (and controls for the case-time-control method) with exposure information at selected times

Criterion	Overview
O2	Easily implemented; less onerous than cohort or case-control studies if data checking is required

The standard case-crossover method (and its case-case-time-control variant) uses only cases, and so its operational sample size is much less than other designs. The case-time-control method, on the other hand, uses separate controls, so has an operational sample size similar to that of a case-control study.

Originally, the case-crossover method was developed for brief exposures (Maclure 1991), such as triggers for myocardial infarction. It is thus applicable for vaccines for which the risk period is short. Suissa (1995) discusses its application to longer exposures, for which a correction for exposure trends becomes more important. However, the method is not applicable to very long and potentially indefinite exposures (such as indefinite post-vaccination risk periods). The major advantage of the case-crossover method is that all time-invariant, multiplicative confounders are implicitly controlled. Time-varying confounders are not allowed for.

The key assumption of the standard case-crossover method is that exposures within the set of time periods used (comprising a case period and one or more control periods) are exchangeable (Vines 2001); intuitively, this mimics the tacit assumption that, in a matched case-control study, the labelling of the controls is immaterial. This assumption fails when the probability of exposure varies between time periods (for example if there is a trend in exposure), or when the exposure process is autocorrelated (which can occur even when there is no trend in exposure). In consequence, it is assumed that there are no time-varying confounders (since such confounders are associated with exposure, which would then be time-dependent).

If there is a monotone increasing trend in exposure, exposure is more likely to occur in the case period, and hence the odds ratio will be overestimated; similarly, if there is a decreasing trend, the odds ratio will be underestimated. In the case of vaccine studies, there is typically a strong temporal dependence of vaccination, owing to vaccination schedules (for paediatric vaccines) or seasonal vaccination (for influenza vaccine). However, the trend is most likely non-monotone. The case-time-control method allows for non-constancy of exposure, provided that a suitable set of externals controls (with the same exposure trend as the cases) can be found; Jensen (2014) has shown that the case-time-control method gives good results even in the presence of autocorrelation. The case-case-time-control method (Wang 2011) provides a further option, obviating the need for separate controls. However, using cases twice as proposed and evaluating trends in coverage in cases may perhaps require more sophisticated analysis methods than suggested. These variants have not been evaluated in the setting of vaccine studies.

The case-crossover methods discussed so far are applicable to unique events; Luo (2008) discusses extensions of the method to accommodate recurrent events, taking into account the possible dependence between matched sets within individuals.

5.4.2 Self-controlled case series designs

This evaluation is focused on the standard self-controlled case series method, reviewed in Section 4.4.3; the variations on this method, reviewed in Section 4.4.4, are considered in the subsequent discussion.

Criterion	Overview
M1	Relative incidence (or relative hazard) in the risk period compared to control periods
S1	Efficiency and power close to that of cohort studies when the risk period is short in relation to the observation period; relative efficiency declines as risk period duration increases and true relative incidence reduces
S2	Only an issue when there are few cases
T1	Self-controlled case series method with short observation periods can be implemented as group sequential procedures
T2	Operational sample size smaller than for cohort or case-control studies, and identical to standard case-crossover studies
R1	Events must be independently recurrent or rare; the key assumption is that events are do not affect subsequent exposures (or observation); time-varying confounders must be allowed for explicitly
R2	Results are sensitive to failure of the exogenous exposure assumption, though the direction of bias can be deduced; failure to allow for time-varying confounders may induce bias
O1	Cases with exposure information throughout a pre-defined observation period
O2	Easily implemented; less onerous than cohort or case-control studies if data checking is required

The self-controlled case series (SCCS for short) method yields an estimate of the relative incidence (for independently recurrent events) or relative hazard (for rare unique events). The likelihood, which is conditional Poisson or equivalently product multinomial, is obtained by conditioning on the number of events experienced by each individual within a cohort, over a pre-specified observation period. This determines most of the properties of the method.

The efficiency of the SCCS method relative to the full cohort method declines as the proportion of time included within the post-vaccination risk period increases; the method is thus most efficient for short risk periods, though it can be used for long and indeed indefinite risk periods (see Farrington 2001, for an example with MMR vaccine and autism). The relative efficiency also declines as the relative incidence reduces. For short risk periods, the relative efficiency is close to that of a full cohort design.

The method uses only cases, and thus benefits from a small operational sample size, equivalent to that of a case-crossover analysis (though more exposure information is required). Small sample performance has been investigated by Musonda (2008a), who conclude that asymptotic results are valid for sample sizes in excess of 20 – 50 cases, depending on the risk period and relative incidence.

The use of the SCCS method for sequential analysis has been considered by Hocine (2009) and Musonda (2008), within SPRT and CUSUM implementations respectively. The method is only appropriate for group sequential monitoring owing to its retrospective features.

There are three modelling assumptions: the event of interest must be recurrent or rare (that is, with frequency of less than 10% during the period of observation, see Farrington 2011); occurrence of an event must not affect the subsequent probability of exposure (an assumption equivalent to requiring that exposure may be considered exogenous); and occurrence of an event must not affect the observation period. For vaccine safety studies, the events of interest usually are rare but do not increase short-term mortality; thus only the second assumption is critical.

Failure of the assumption that events do not affect subsequent exposures (to the same vaccine) produces bias, the direction of which can sometimes be predicted: if the event reduces the chance of subsequent vaccination, then the relative incidence will be biased upwards; if the event increases the chance of subsequent vaccination, the relative incidence will be biased downwards. Variants on the basic SCCS method, reviewed in Section 4.4, have been proposed to circumvent the assumptions. These variants are not applicable to indefinite risk periods. Further SCCS methods have been proposed to cater for recurrent but non-independent events (Farrington 2010, Simpson 2013).

The SCCS model is also prone to bias if time-varying confounders are not included in the model. Adjustment for age and/or season is usually essential, unless the observation period is very short. This can be done parametrically, as originally described in Farrington (1995), or semi-parametrically (Farrington 2006). The impact of mis-specification of the risk period and misclassification have been considered, respectively, by Mohammed (2012, 2013) and Quantin (2013).

The standard SCCS method is easy to implement; a key (but unusual) data requirement is that information on post-event exposures is required. Some of the variants reviewed in Section 4.4 are considerably more challenging to apply than the standard SCCS method.

5.5 Sequential Methods[7]

Sequential methods were discussed in Section 4.5. This evaluation is focused on the sequential probability ratio test and its variants, with brief mention of sequential case series and CUSUM methods.

Criterion	Overview
M1	No effect measures are directly available from sequential methods, but require correction for stopping bias
S1	Power is dependent upon pre-specified null and alternative relative risks, together with time allowed to reach an acceptance or rejection threshold, i.e. power increases with increasing time

[7] This section was contributed by Caitlin Dodd, Erasmus University Medical Centre.

Criterion	Overview
S2	Cannot be conducted in the finite sample setting, requires a large and continually updating database
T1	N/A
T2	Dependent upon accuracy of null and alternative relative risks, has been described as optimal , meaning that the minimum number of cases is required
R1	In non-self-controlled variations, assumes that expected rates can be calculated from a control population or obtained from the literature
R2	Very sensitive to misspecification of expected rates and stopping rules
O1	Requires a large continually updated database except in self-controlled sequential methods
O2	Moderately easy to implement given a sufficient database; methods not available in most standard software except R package Sequential (maxSPRT) and SAS (group sequential designs)

The main limitation of the SPRT method is that it requires, as an alternative hypothesis, a specific relative risk, above which the association will be designated as a 'signal'. Failure to choose the correct RR could result in a delay in signalling even when the RR differs from 1 (Lieu 2007).

The MaxSPRT method (see Section 4.3.1) was designed to address this problem by testing the alternative hypothesis that the RR is greater than 1, meaning that only the alpha level and length of surveillance need be specified a priori. The Max SPRT requires either concurrent controls, leading to the same matching issues described in Section 5.3, or historical data, implying that results are partially dependent on the quality of this historical data for use as a reference, notably the presence of secular trends. Asymptotic approximations are not needed as critical values are calculated using iterative calculations. It has been pointed out by Shih (2011) that the MaxSPRT is not the most efficient approach among sequential generalised likelihood ratio tests; its relative efficiency remains to be determined.

A limitation shared by all sequential methods is the requirement for continually updated data and for rapid linkage between exposures and events, two requirements which may be impossible or impractical in many data sources.

In an attempt to account for the uncertainty in expected counts required for each sequential method, the Conditional MaxSPRT was developed (Li 2010) and purportedly preserves the type I error rate even when the sample size of the historical data is small. It is conditioned on the number of adverse events in the historical and surveillance populations while the person time cumulated while observing these counts is the random variable. This means that event rates which differ within subpopulations could be problematic if this difference is unknown.

CUSUM methods (see Section 4.5.2) were designed to detect a change in a process from a state of equilibrium rather than to monitor the rate of a known safety problem (Kulldorff 2010). One potential application is to detect a sudden deterioration in a safe product, for example owing to occurrence of a fault in the manufacturing process.

Group sequential methods developed for use in clinical trials require less frequent testing and should therefore lead to fewer costly false positive signals. This approach may be preferable when detection of false positives due to repeated testing is a concern and when time to detection is of lower priority.

Self-controlled methods lend themselves to sequential analysis, but, like group sequential methods, are better suited to situations when time to detection is not the priority. These methods have the advantage over other sequential methods that they control for time-invariant confounding, making stratification or matching unnecessary.

In general, sequential methods can be limited by the requirement to specify expected event counts, the impracticality of controlling for confounders by means other than stratification or matching, and the need for access to continually updated data sources. In these circumstances they are perhaps most useful as signal detection or strengthening methods, rather than for use as signal confirmation and causality assessment.

5.6 Signal strengthening methods

We briefly assess the two signal-strengthening methods reviewed in Section 4.6. In view of their very different features, we consider them separately. In evaluating these methods, it is important to bear in mind that they are not intended as signal confirmation methods – but rather to provide rapid contextualisation of a signal.

5.6.1 Observed – Expected methods

This method was reviewed in Section 4.6.1.

Criterion	Overview
M1	Relative risk of observed compared to expected
S1	Similar to that of a cohort study
S2	Only an issue when there are few observed cases, or when expected values are based on few cases, or there is limited background incidence data
T1	Can be implemented in sequential mode
T2	Operational sample size depends on implementation but can be low if limited to the observed cases
R1	The key assumption is that observed and expected quantities relate to the same populations; reporting or case ascertainment is assumed identical for observed and expected quantities
R2	Results are sensitive to failure of the assumptions – but then this is a hypothesis strengthening method
O1	Expected counts and baseline numbers expected
O2	Easily implemented; provides a rapid check on a hypothesized association

The Observed-Expected method produces an O/E ratio, which given suitable assumptions about the relevant denominators can be thought of as a relative risk or rate (when divided by a

population or person-time denominator, respectively). The O and E values can be expressed as absolute risks or rates.

Formally, O-E methods are cohort designs, possibly with a determinate comparator (i.e. the expected rates can be derived separately and regarded as fixed), and thus have similar statistical properties. These methods can be implemented sequentially, and indeed have been used in this way (see Section 4.6.1). The operational sample size depends on the details of the design, but in its most economical form (using fixed expected rates obtained separately) only requires counts of observed cases.

The key assumption is that the observed and expected values relate to the same population – and thus that all confounders have been controlled for. This is most unlikely to be the case in practice – nor is detailed control expected in a hypothesis strengthening method. Rather, steps need to be taken to avoid gross bias, for example control for broad age groups. A further assumption is that case ascertainment or reporting is equally sensitive in the two populations. If this is likely to be untrue, the results can be supplemented by a sensitivity analysis, as was done by Bryan (2010) and Donegan (2013).

The primary purpose of the method is to provide rapid contextualisation of a signal, and hence the implementation of the method is geared to meeting that objective.

5.6.2 Self-controlled analysis of spontaneous reports

This method was reviewed in Section 4.6.3.

Criterion	Overview
M1	Ratio of relative incidences (non-parametric version), or relative incidence (parametric version)
S1	Similar to that of a self-controlled case series method with short observation period
S2	Only an issue when there are few observed cases
T1	Can in principle be implemented in sequential mode
T2	Operational sample size is that of a self-controlled case series study
R1	Applicable only to short risk periods; the nonparametric version is applicable only to multi-dose vaccines; the nonparametric version relies on the assumption that the reporting rates after different doses are proportional; the parametric version makes an explicit assumption about the reporting rate
R2	Results are sensitive to failure of the assumptions – but then this is a hypothesis strengthening method
O1	Spontaneously reported post-vaccination cases with number of last dose
O2	Easily implemented; provides a rapid check on a hypothesized association

The effect measure for the nonparametric version of the method is a ratio of relative incidences (that is, relative incidence associated with dose 1, divided by relative incidence associated with dose 2). For the parametric method, the effect measure is a relative incidence.

The method is a special case of the SCCS method (see Section 5.4.2), with further assumptions, and so shares its statistical properties, and its ease of implementation.

The method uses a short observation period, and so can easily be implemented in sequential mode. The method only applies if the risk period is short. The nonparametric method essentially produces a disproportionality measure between vaccine doses – and so is applicable only when the vaccine has 2 or more doses. In this case the key assumption is that the reporting rate as a function of time since last dose is proportional between doses. The parametric version can be applied to a single dose vaccine, but makes the stronger assumption that the reporting rate varies as a specified function of time since vaccination (typically exponential).

These assumptions are unverifiable, and for this reason the method cannot provide definitive evidence of association or lack of it. However, because the method is self-matched, it removes confounding due to fixed confounders, and in this sense is a 'signal strengthening' method.

6. THE IMPACT OF SIGNAL DETECTION ON VACCINE RISK ASSESSMENT[8]

Although it is not part of the ADVANCE remit to review and assess signal detection methods, these methods are nevertheless relevant to risk assessment and benefit – risk evaluation. In this section, we consider three ways in which the methods used to undertake signal detection can help, or indeed hinder, risk assessments. Throughout, signal detection is taken to mean a data mining technique applied with no prior hypothesis to identify possible associations between vaccination and an adverse event.

6.1 Sensitivity of signal detection methods for vaccines

Many of the data mining techniques used, primarily with databases of spontaneous reports, are based on different types of disproportionality analyses, which essentially involve identifying drug – event pairs that appear with higher than expected frequency. These methods include proportional reporting ratios (Evans 2001), reporting odds ratios (van Puijenbroek 2002), information components (Bate 1998), and empirical Bayes methods (Dumouchel 1999); the first two are formulated in a classical framework, the latter two in a Bayesian framework. These measures were designed to be applied to the wide variety of drugs and events that are typically found in databases of spontaneous reports used by regulatory agencies.

In contrast, databases of spontaneous reports associated with vaccines, especially those available to manufacturers, typically comprise a much smaller number of different products, which are often used in very specific populations. It has been shown that the application of some disproportionality techniques in these circumstances can lead to a loss of sensitivity (Van Holle 2014). For the detection of vaccine-related signals, disproportionality analyses may

[8] This section is heavily based on a document produced by Lionel Van Holle (GSK).

usefully be supplemented by other methods, such as those based on the distribution of times from vaccination to event (Van Holle 2012).

Clearly, for the purpose of assessing vaccine risks, it is important that methods appropriate for vaccines are used to detect possible signals.

6.2 Types of signals and their impact on benefit-risk evaluations

The safety problems associated with vaccines which are picked up by signal detection may be very different in nature, and may thus require very different responses. Thus, a safety problem might be related to the quality of the manufacturing process, programmatic errors, or an inherent feature of the vaccine product itself (Autran 2009). Problems associated with the manufacturing process may be limited to a particular batch, without calling into question the overall risk profile of the vaccine, or necessitating a re-evaluation of the vaccine's benefit – risk profile.

Different types of problems are likely to require different responses from manufacturers, regulatory agencies and public health authorities. For this reason, it is important at an early stage to seek to throw as much light as possible on the likely nature of a new signal – for example, by investigating whether it is associated with a particular batch or type of vaccine. Such information is required as soon as possible after signal detection – and usually as a prelude to undertaking signal confirmation investigations.

6.3 The distinction between signal generation and signal confirmation

The ready availability of large databases has revolutionised observational pharmacoepidemiology, making it relatively easy and cheap to undertake large-scale investigations of putative associations between vaccines and adverse events. There is also a trend towards using such data for signal generation using data mining techniques (see Section 7.2), or for undertaking signal strengthening analyses (see Section 7.1).

These developments present a new methodological challenge, resulting from the blurring of the distinction between signal generation and signal confirmation. This distinction has traditionally been ensured by confining signal generation investigations to spontaneous reports data, while undertaking signal confirmation investigations on more robust, population-based data or specially designed studies. There is a danger that the ability to confirm associations will be lost, if the same data have been used to generate the signal; the issue appears as yet to attract little attention in the published literature (Harpaz 2012), although early discussions are now starting to occur.

Currently, the only method available to undertake signal generation and signal confirmation within the same database is to split the data into a training set (for detection purposes) and a test set (for confirmation). Further work on – and awareness of – this issue is needed.

7. EXPERIENCE FROM OTHER PROJECTS AND SYSTEMS[9]

In this section we will review (a) existing systems for rapid risk assessment, and (b) other projects on vaccine safety assessment that are closely or distantly related to ADVANCE. In each case we will briefly review the system or project, and seek to summarise its conclusions and lessons to be learned insofar as they relate to ADVANCE.

7.1 The Vaccine Safety Datalink[10]

The Vaccine Safety Datalink (VSD) was set up in 1990 as a partnership between the Centers for Disease Control and Prevention, and several Health Maintenance Organisations (HMOs) in the United States of America, in order to create an active surveillance system specifically designed to study adverse events in relation to vaccines (Chen 1997). The system has since been expanded substantially to include other HMOs, increase the age range of the population monitored. In 2011, data for over 18 million people of all ages and spanning 16 years were available for research (Baggs 2011).

The VSD has made use of a wide range of study designs, including retrospective cohort, case-control and self-controlled methods, the latter including self-controlled case series methods and other new case-only methods. Although the majority of VSD studies relate to vaccine safety, studies have also been undertaken on coverage, burden of disease, and cost-effectiveness (Baggs 2011).

Since 2005, Rapid Cycle Analysis (RCA) has been implemented for newly introduced vaccines, to provide early evidence of adverse events (Davis 2005). A range of sequential procedures have been developed for this purpose, notably the MaxSPRT method (see Section 4. 5.1.1). The genesis and development of the approach is described by Davis (2013). Yih (2011) reviews the operation of the RCA system between 2006 and 2009: 30 vaccine-event pairs were monitored, and there were 10 signals. Only one of these was deemed to be genuine, relating to risk of seizures 7 to 10 days after MMRV vaccination.

Yih (2011) notes that the RCA method can provide early evidence of a problem, but goes on to say '*On no account should a signal be interpreted as indicating an association or causal relationship between vaccine and adverse event until confirmatory studies are conducted*'. Such confirmatory analyses include logistic regression to adjust for cofounders, analyses of clustering after vaccination to determine biological plausibility, and chart review. Examples of confirmed signals, and the evidence adduced to support this conclusion, may be found in Tse (2012), for influenza vaccine and convulsions, and Weintraub (2014), for rotavirus vaccine and intussusception.

[9] This section relies heavily on notes provided by Catherine Panozzo (Sanofi Pasteur), Caitlin Dodd (Erasmus University), Christel Saussier (ANSM) and Nick Andrews (PHE).

[10] For further information see http://www.cdc.gov/vaccinesafety/Activities/VSD.html.

7.2 The PRISM system and Mini-Sentinel[11]

The Sentinel Initiative is FDA's response to the 2007 US Congressional Mandate, the FDA Amendments Act (FDAAA), to create an active surveillance system utilizing electronic healthcare data. The goal of the Sentinel Initiative is to build and implement a new active surveillance system called the Sentinel System. The Sentinel System will be used to monitor the safety of all FDA-regulated products, including drugs, biologics, and medical devices (Mini-Sentinel 2011).

The Mini-Sentinel is a pilot programme designed to inform development of the Sentinel System by developing methods, resources, and procedures to facilitate active surveillance using routinely collected electronic healthcare data (Forrow 2012). Key features include rapid response time, transparency, and privacy (Mini-Sentinel 2013). Harvard Pilgrim Health Care is the coordinating center for Mini-Sentinel, and 31 sub-contracted public and private data and academic partners participate in the program through a data distributed approach which allows the partners to maintain physical and operational control over their electronic healthcare data (Platt 2012). The FDAAA set goals of accessing data from 25 million people by July 2010 and 100 million people by July 2012, and these goals have been surpassed (Curtis 2012).

Work focused specifically on vaccines comes under the Post-Licensure Rapid Immunization Safety Measurement (PRISM) program. This program was initially created to monitor the safety of 2009 pandemic H1N1 influenza vaccine, but was integrated into Mini-Sentinel in 2010 to ensure its sustainability (Nguyen 2012). PRISM consists of a distributed database of >30 million individuals enrolled in any of three national healthcare insurance plans (Aetna, HealthCore, and Humana) with additional data from eight state or city immunization registries (Florida, Michigan, Minnesota, New York State, New York City, Pennsylvania, Virginia, and Wisconsin).

Of note, in the US, approximately 52% of people aged 0-64 years have private health insurance that is generally provided at a reduced cost through employers. Almost all (97%) individuals 65 years of age and older receive health insurance through the publically funded program, Medicare. Individuals who are uninsured or insured through public programs would not be represented in the PRISM program.

Thus, compared with the other major US vaccine surveillance systems such as the Vaccine Adverse Event Reporting System (VAERS) and Vaccine Safety Datalink (VSD), PRISM data are more likely to be representative of the privately-insured US population. PRISM is also the largest general population cohort available for active vaccine surveillance in the US. However, limitations include the fact that the completeness of the vaccine exposure data in the claims and immunization registries is unknown; data updates are conducted only every quarter; and medical record validation can be a lengthy process (Nguyen 2012, Baker 2013).

Since monitoring the safety of the pandemic H1N1 influenza vaccine began in 2009, the PRISM population has been used to study the safety of other vaccine exposures and meet the various

[11] For further information see www.mini-sentinel.org.

objectives of the Mini-Sentinel. Examples of completed or ongoing projects include: intussusception and rotavirus vaccines; venous thromboembolic events and HPV vaccine; febrile seizures and influenza vaccines; pregnancy and birth outcomes and influenza vaccines. The research undertaken includes methodological elements, for example to develop new sequential methods incorporating confounders, and exploring data mining methods (Nguyen 2012).

7.3 VAESCO[12]

The VAESCO (Vaccine Adverse Event Surveillance and Communication) project, which began in 2008, was coordinated by ECDC and was implemented by a consortium composed of several European partners among which regulatory authorities, public health bodies and academia. Its objective was to explore the feasibility and demonstrate the benefits of collaborative post-licensure vaccine safety epidemiological studies. The long term aim of the work is to create an independent infrastructure and epidemiological resource in support of vaccine safety monitoring and investigation in Europe.

The VAESCO consortium has conducted collaborative vaccine safety studies through a common approach, a shared research infrastructure and standardized methodologies, facilitating subsequently data comparability and building collaborative networks.

The working model of a VAESCO study starts with a common protocol. Information collection is harmonized, and case definition of the outcome and local data management are standardized. However, the method for identifying patients and exposure and covariate information varies according to the event of interest and the available data sources in each country. Data may be captured both from electronic population-based health care databases (using Jerboa software) or from incompletely automated data sources using a standardised data entry system (CHAMELEON).

Several projects have already been conducted within the consortium, notably relating to pandemic influenza A(H1N1)2009 vaccines and their possible association with Guillain-Barré syndrome (Dieleman 2011) and narcolepsy, and MMR vaccines and idiopathic thrombocytopenic purpura (Andrews 2012). The methods used include case-control studies, self-controlled case series studies, and cohort studies.

The VAESCO initiative has suggested that there may be added benefit from applying common methods and using a shared infrastructure for data sharing across European countries. At the same time, heterogeneities between countries can be a problem and can cause delays while data issues are resolved, and can complicate the interpretation of results. To resolve these issues, it has been suggested that a flexible network approach be developed in which the countries participating on specific projects would vary depending on the purpose of the project and the capabilities of the individual countries. Thus, surveillance and rapid response activities could be restricted to countries and systems that can conduct rapid monitoring and signal assessment,

[12] For further information see www.brightoncollaboration.org/vaesco.html .

while in-depth studies could involve a variable membership depending on local interest and capabilities to contribute useful data.

7.4 OMOP and its successor OHDSI[13]

The Observational Medical Outcomes Partnership (OMOP) was a private-public partnership developed with the aim of '*identifying the most reliable methods for analyzing huge volumes of data drawn from heterogeneous sources*' (all quotations are taken from the OMOP website). The partnership was a collaboration among Pharmaceutical Research and Manufacturers of America , the US Food and Drug Administration, and the Foundation for the National Institutes of Health. This partnership was formed following the recognition that electronic health records were a huge and largely untapped source of data for monitoring the safety of drugs, and, potentially, the safety of devices, and procedures as well. Using these electronic health records databases, OMOP used interdisciplinary approaches including methods from epidemiology, computer science, and statistics to evaluate and compare methods for signal detection in pharmacoepidemiology.

To date, OMOP has assessed the performance of multiple methods for the analysis of observational data against a 'gold standard' of negative (biologically implausible and never reported drug/event combinations) and positive controls (known and verified drug/event combinations). These assessments were conducted to '*develop and evaluate standardized algorithms that can reliably discriminate the positive controls from the negative controls, and to understand how an estimated effect from an observational study relates to the true relationship between medical product exposure and adverse events*'. From these investigations, OMOP has determined that self-controlled designs perform well but that performance of various signal detection measures is dependent upon outcome, exposure, and database and that presently there does not exist one analytical approach which could be considered optimal across outcomes and databases.

Two defining characteristic of OMOP are transparency and open access. OMOP has made a large number of resources available on its website, including tools to interrogate data, a library of methods coded in SAS or R, and simulated data for testing methods.

OMOP has been featured in a special supplement of the journal Drug Safety (Drug Safety 2013; 36 Supplement 1). Evans (2013) contextualised the findings, arguing that the project demonstrates that both total pessimism and over-confidence in our ability to accurately detect new adverse events are misplaced. Of particular relevance to ADVANCE, the issue contains studies of case-control, cohort, and self-controlled methods. While OMOP focused on signal generation, and did not focus on vaccines, its overall finding that no single method is uniformly better than all others (Ryan 2013a), and that using different methods can lead to strikingly different results (Madigan 2013) are likely to be applicable to ADVANCE.

[13] For further information see www.omop.org and www.ohdsi.org.

OMOP has recently been disbanded and replaced with the Observational Health Data Sciences and Informatics (OHDSI) Program. Again, the OHDSI collaboration is interdisciplinary and open-source. However, OHDSI aims to expand upon the original goals of OMOP which were purely analytical and to explore application of the lessons from OMOP to real-world decision making.

7.5 GRiP[14]

Global Research in Paediatrics (GRiP) is European Union funded network of excellence which was formed to address the gaps in drug testing and surveillance in pediatrics. The goals of the GRiP network incorporate multiple aspects of pediatric drug research including training and education, development of methods for clinical trials in pediatrics, standardization of design, terminology, and reporting in pediatric research, among others. However, the goal of GRiP most applicable to the goals of ADVANCE is that which aims to improve epidemiological and post-marketing studies of drugs in children.

The GRiP Network plans to develop an integrated electronic infrastructure for epidemiological, pharmacovigilance and post marketing research. This infrastructure will exploit and link existing healthcare databases in Europe and the US to assess the occurrence of diseases in children, plus the use and effects of drugs (including vaccines) on a large scale. Methodologies for harmonization, data exchange across national boundaries (including ethical and governance issues), data mining and comparative safety and effectiveness studies will be developed.

Toward this goal of improving epidemiological and post-marketing studies of drugs in children, GRiP has listed five areas of focus:

1. Build an online platform for real time data sharing and scientific collaboration as a basis of an integrated research infrastructure.
2. Identify healthcare databases with population based information on drug use and vaccine outcomes in children and to describe their characteristics.
3. Describe the governance and ethical issues related to the use and linking of the healthcare databases.
4. Map disease and drug/vaccine coding terminologies across the healthcare and adverse events databases
5. Create a common methodology for drug and vaccine utilization studies, for disease incidence and prevalence studies, and for epidemiological ascertainment of drug and vaccine safety, all using a distributed data model.

The GRiP network has succeeded in identifying automated population-based healthcare databases globally which could be used for pharmacoepidemiological research and has approached those databases to request their participation and to request that they fill out a

[14] For further information see www.grip-network.org .

survey describing their database. The results of this work have been summarized in Deliverable number D2.2 which is available on the GRiP website. Also available is a description of spontaneous reporting databases including the FDA/CDC Vaccine Adverse Event Reporting System (VAERS).

Currently, Erasmus MC is working together with The Brighton Collaboration to map medical conditions plus drug/vaccine and dose terminologies through *Unified Medical Language System* (UMLS) concepts to the various terminologies (including the *Medical Dictionary for Regulatory Activities* - MEDDRA to link to ADE/AEFI databases). EMC has also developed a common data model for harmonization of spontaneous reporting databases. Data compiled under this common data model will be analyzed using traditional signal detection methods and will be used to modify existing methods or develop new methods specific to signal detection in pediatrics.

7.6 Other initiatives[15]

Several other projects related to risk assessment have been identified, some of which are still in progress. While these comprise a methodological element, this tends to be focused on the data gathering or informatics aspect. The projects are briefly listed here but not considered in further detail.

The EU-ADR project (www.euadr-project.org; completed) sought to design and validate a computerized system based on electronic healthcare records and biomedical databases for the early detection of adverse drug reactions.

The BIOVACSAFE project (www.biovacsafe.eu; ongoing) will develop new tools to speed up the testing and monitoring of vaccine safety, both before and after licensure.

The Global Collaborative Vaccine Safety Network is a global network to investigate vaccine safety (www.who.int/vaccine_safety/news/GVSI_P_Portofolio_2012-2020.pdf; ongoing). The network has undertaken meta-analyses (Dodd 2013); see Section 4.8.3.

The OpenPHACT project (www.openphacts.org; ongoing) will deliver an online platform with publicly available pharmacological data.

The SALUS project (www.salusproject.eu; ongoing) will seek to provide a standard-based interoperability framework to facilitate the execution of safety studies and the analysis of real-time data from disparate databases of electronic health records.

The EUROCAT project (www.eurocat-network.eu; ongoing) will provide data on congenital anomalies in Europe.

[15] This section is based on information gathered by Work Package 2.

The EUROmediCAT project (www.euromedicat.eu; ongoing) will develop and test a pharmacovigilance system for the safety of drugs taken during pregnancy; the drugs to be investigated do not (as yet) include vaccines.

8. METHODS FOR INDIRECT RISK ASSESSMENT

All the methods so far reviewed relate to evaluating the evidence for a direct causal link between vaccination and adverse events in individuals. However, as previously noted in Section 1.1, vaccination programmes are often large scale interventions that alter the ecology of host – pathogen interactions. The indirect benefits from herd immunity that result from such population-level effects are well known. There can also, however, be indirect disbenefits, which need to be considered in any comprehensive risk assessment.

Such effects can be of different types. If the infectious organism has several circulating strains, not all of which are included in the vaccine, mass vaccination can act as a selection mechanism to alter the strain distribution; see Miller (2011) for an example of such effects for pneumococcal conjugate vaccine. A second mechanism for indirect effects results from the increase in the average age at infection in the unvaccinated population, which under certain circumstances (notably inadequate vaccine coverage) can increase the number of susceptibles in older age groups. This has been observed in the case of mumps vaccine in Sweden (Sartorius 2005). This mechanism could, in theory, result in an increase in the numbers of adverse events, if these are more likely to result from infections at an older ages (Anderson and May 1991). This mechanism has been suggested as causing an increase in the number of cases of congenital rubella syndrome in Greece (Panagiotopoulos 1999). A third mechanism is the reduced exposure to circulating pathogens, and hence the reduction in boosting of immunity. For example, it has been suggested that this may in certain circumstances result in an increase in shingles following the introduction of varicella zoster vaccine (Brisson 2000). Some population-level effects, such as the emergence of adult pertussis as a major public health issue, may be related to some of these or other indirect effects, or more simply to waning with age of vaccine-induced protection (de Greeff 2010, van der Maas 2013).

As suggested in this brief overview, the indirect risks associated with vaccination stem from diverse sources, and it is difficult to deal in any generality with the methods available for studying them. Unlike the techniques used in observational pharmacoepidemiology, which are essentially empirical, the methods used to evaluate indirect effects sometimes involve mathematical modelling (Anderson & May 1991, Brisson 2000). Generally, however, indirect effects can be represented in terms of a quantity $\lambda(v)$, denoting the (possibly age-specific) population rate of a specified adverse event in a population with vaccine coverage v. A relevant absolute measure of total effect (direct plus indirect) is then $\lambda(v) - \lambda(0)$, while the indirect effect is $\lambda_0(v) - \lambda_0(0)$, where $\lambda_0(v)$ is the rate of the adverse event among the unvaccinated individuals within a population with vaccine coverage v (and $\lambda_0(0) = \lambda(0)$). The indirect effect sizes depend on the vaccine coverage achieved, but also on the characteristics of infection transmission in the population – notably its age structure and contact rates.

In some circumstances indirect effects, or total (direct plus indirect) effects can be quantified by comparing rates prior to and after the introduction of mass vaccination or by observing trends in the incidence of adverse events after the introduction of the vaccine (Panagiotopoulos 1999). Such ecological methods were discussed in Section 4.7.1. Serological survey data (de Greeff 2010) can help to obtain estimates which are unaffected by temporal variation in the completeness of reporting.

The only general approach so far discussed which in theory permits the evaluation of such measures of effect using contemporaneous comparisons (rather than historical before and after or trend-based comparisons), is the stepped wedge introduction of a new vaccine using cluster randomisation (see Section 4.1.2). However, even such a design would only permit the evaluation of those indirect effects that are rapidly manifested. In general, the overall indirect effects of vaccination are only likely to be estimable some time after the vaccine has been introduced.

9. FINDINGS AND RECOMMENDATIONS

In the next two sections, we describe the main findings. First, we consider the implications for integrating risk assessments in benefit-risk evaluations. Second, we outline areas where further research is required. In a final section, we end with a list of recommendations. These comprise four research components to be included in the Proof of Concept studies within Work Package 5.

9.1 Findings: integrating risk assessments in benefit-risk evaluation

In this section we consider more specifically the issues relating to integrating risk assessment in risk-benefit evaluations. We do so under three headings: study designs; measures of effect; and types of effect.

9.1.1 Study designs

The over-arching conclusion from this extensive review of risk assessment methods for vaccines is that no single methodological approach is suitable for evaluating vaccine-associated risks in every situation. Indeed, it is desirable to have available a range of different techniques, to cater with the wide diversity of problems and possibilities, and to be brought to bear on the same problem. This is borne out by the following considerations.

First, different situations call for different approaches. Monitoring risks for new vaccines (or new formulations of existing vaccines, or perhaps new vaccination schedules) is best undertaken using sequential methods, as these are best suited to providing the accumulating evidence required to develop a safety profile, and are likely to minimise the expected time to reach firm conclusions. In contrast, investigating signals associated with established vaccines are best studied using the full power of retrospectively available data, rather than sequential schemes.

Second, within each broad approach (sequential or non-sequential) a range of different designs are available. It is not possible, nor would it be desirable, to single out one such design as 'best'. Different methods are based on different assumptions, which may or may not be valid in different situations. Most of the methods reviewed have good properties in some circumstances, but may be inappropriate or unusable in others: for example, case-crossover methods cannot be used for indefinite risk periods. Often, however, it will not be clear a priori which method is 'best', and some flexibility in the choice of method, or methods, is desirable.

We resist developing a blueprint seeking to set out *a priori* what method is best for each situation. One reason is that these situations would have to be defined in unrealistic terms to avoid issuing misleading recommendations: for example, 'when all confounders are known and measured, then use a cohort design with adjustment for confounding' is correct, but wholly unhelpful, since confounders are seldom if ever all known and measured.

Third, there generally is value in applying contrasting methods to the same problem, and indeed to the same dataset, especially when these methods rely on different assumptions. For example, the case-control method is based solely on between-individual comparisons, and assumes that all fixed confounders are adjusted, whereas the self-controlled case series method uses only within-individual comparisons and assumes that events do not affect subsequent exposures. Applying both approaches to the same data provides added information: if the results are similar, evidence of robustness may be inferred; if the results are very different, information about biases may perhaps be deduced.

9.1.2 Measures of effect and heterogeneity

Integrating risk assessments into benefit-risk evaluations poses other, more fundamental problems than simply choice of design, namely, the transferability of measures of effect. Measures of effect being central to risk-benefit analyses, they were included as our first assessment criterion (Section 3.1). Often, the measures of effect which are integrated into benefit-risk analyses are obtained from different populations. Alternatively, it may be required to compare the risk-benefit profile in one population to that of another population. A key question is therefore the transferability from one population to another of risk estimates, owing for example to differences in the socio-economic environment, genetics, health care provision, levels of pre-existing immunity, circulating pathogens, and other factors.

As described in Section 3.1, and outlined formally in Section 2.5, benefit-risk evaluations require absolute, rather than relative, measures of risk. An absolute measure of effect often can be deduced from a relative measure, combined with a population-specific absolute baseline rate. However, such calculations are only valid if the relative measure used is relevant to the population under consideration. This raises the question of what effect measures, if any, are vaccine-specific, and to what extent they are population-dependent.

An understanding of the biological processes at work may perhaps help shed some light on the issue. However, we suspect that the primary source of evidence on transferability is likely to be

empirical. This would require comparing different estimates of different measures (both relative an absolute) obtained in different settings, to get some idea of which, if any, may be vaccine-specific and hence transferable between populations. Key to this is to gain a better understanding of the heterogeneity between studies (and between populations), using the meta-analytic methods reviewed briefly in Section 4.8.3.

A better understanding of such heterogeneity is also needed to quantify the uncertainty associated with benefit-risk evaluations. It may transpire that a Bayesian framework is better suited than the frequentist perspective, which so far in this account has been dominant, to the complex task of integrating data and prior understanding, along with the uncertainties associated with both.

9.1.3 Types of effect

The focus of this report has primarily been on the direct risks incurred by individuals as a result of vaccination. However, this discussion was framed by an overall context in which population effects are key, notably to public perceptions of risk (see Section 2.1), a perspective which ought not to be ignored in benefit-risk evaluations (see Section 2.5). (Similar considerations apply to direct and indirect benefits of vaccination.)

To this end we included a brief discussion of indirect risks associated with vaccination (Section 8). The difficulty is how to combine an evaluation of such indirect effects with more readily estimable direct effects. Only one design, the stepped wedge cluster-randomised vaccine introduction (Section 4.1.2) permits the estimation of both direct and indirect effects in the short term using contemporaneous data. Given sufficient time and data, before-and-after ecological comparisons can be made. Alternatively, some reliance may need to be placed on mathematical models, whose output is generally primarily of a qualitative nature.

The incorporation of quantitative information on indirect effects in benefit-risk evaluations is thus fraught with difficulties. Rather than ignoring them, indirect effects should at least be considered in qualitative terms, and preferably with some attempt at indicating their likely order of magnitude, in order to complement a more quantitative evaluation of direct risks.

9.2 Findings: some areas where new research is required

This review and assessment of methods for risk evaluation in relation to vaccines has highlighted several areas where some further methodological research may be warranted.

Heterogeneity

A key investigation so far largely lacking in the literature is a systematic study of heterogeneity of the associations found in different settings. In the context of ADVANCE, a key aspect of such a study would be to investigate the extent of heterogeneity associated with different measures of effect, but also to seek patterns which may explain some of the heterogeneity, for example using the techniques of meta-regression.

Designed vaccine introduction

A feasibility study ought to be undertaken for introducing design elements into vaccine introductions, so as to facilitate rapid benefit-risk evaluation, and facilitate future risk investigations. Such a study would focus on the practicalities of such schemes, but also on their possible benefits, perhaps via simulations. It is likely that numerous operational obstacles would be identified, but these need to be evaluated in the context of the possible benefits of such schemes.

Criteria for systematic reviews and meta-analyses

Systematic reviews tend to be based on assessment criteria which are appropriate for traditional designs, notably cohort and case-control studies, but may not be appropriate for other designs, notably self-controlled studies. Developing a set of assessment criteria suitable for vaccine studies, and notably meta-analyses of vaccine safety studies, is desirable.

Hypothesis generation, strengthening and confirmation

Higher prominence should be accorded to the issues and biases involved in undertaking hypothesis generation, strengthening and confirmation within the same database. This, together with strategies for bias reduction, for example by splitting the data into training and test sets, and how to do so, is worthy of further investigation.

Instrumental variables

No studies have been undertaken of vaccine safety using instrumental variables, though as reported in Section 4.8.1.4, some studies using instrumental variables have been done of vaccine efficacy. It may be useful to examine whether such studies could be done and which instruments could be used.

Healthy vaccinee effect

The healthy vaccinee effect is likely to impact upon all study designs, and when present will tend to bias relative risk estimates downwards. Research is required to study this effect in greater detail, ascertain its importance in different contexts, and investigate strategies for controlling it, for example using an adaptation of the intention-to-treat approach.

Propensity scores for vaccine-associated risks

Dynamic propensity scores have been suggested in vaccine safety studies to take account of the fact that individuals' vaccination histories evolve over time (Farrington 2009). These studies have not yet been done. Also, investigations of high-dimensional propensity scores, and validation of such methods, for vaccine-associated risks would be useful.

Bringing clarity to the nomenclature

Our review of methods (Section 4) has highlighted proliferation of new (and some not so new) designs which have appeared in pharmacoepidemiology. Most often, these designs have been described informally, without an explicit statistical model. It would be useful to undertake a more formal survey of these methods, in order better to understand the connections between them, if any, and their strengths and shortcomings.

Case-time-control methods for vaccine evaluation

The case-time-control method has been suggested when there are trends in exposure. It is not clear, however, to what extent the method would work when exposures are not monotone, but peaked or multimodal, as is typically the case with vaccination. The proposed analysis method of the case-case-time-control method also warrants investigation.

Sequential methods

In clinical trials, sequential methods are used to test key hypotheses. In vaccine safety surveillance, however, sequential methods have been used primarily for signal strengthening purposes. Developing sequential methods that are robust to confounding bias, and which can reliably be used to confirm hypotheses rather than strengthen (or weaken) them, is likely to be of benefit.

9.3 Recommendations: research components for inclusion in proof of concept studies

In the light of the findings described above, and in line with requirements of the ADVANCE project, we propose that the following research components are integrated in the proof of concept studies to be undertaken in appropriate databases. In proposing these research components, we fully acknowledge their potential limitations, notably that being conducted in particular populations and with specific vaccine-event pairs, we cannot guarantee that the conclusions derived from them are universally applicable. To mitigate this limitation, we recommend that each proof of concept study is designed to encompass a range of scenarios.

POC component 1: Study of heterogeneity of vaccine risk between databases

The objectives of POC component 1 are as follows:

- To investigate the extent of heterogeneity between databases (and hence between populations) of measures of effect for the same vaccine – event pair.
- To study how the heterogeneity varies for different effect measures, and to characterise the extent to which it is possible to transfer estimates between populations.

This study would involve several databases. A range of effect measures and baseline risks for a number of specified vaccine – event pairs, including some for which causality has been

established, some for which causality is suspected but not confirmed, and some for which no causal link is expected, would be estimated.

POC component 2: Evaluation of sequential methods for new vaccines

The objectives of POC component 2 are as follows:

- To investigate the performance of different sequential monitoring schemes for evaluating the safety of newly introduced vaccines.
- To compare the sequential evaluation with a definitive post-hoc safety evaluation involving all available data.

This study would be conducted in one or several databases in which information on a once new vaccine is available (for example, an influenza vaccine). For the first objective, the various sequential schemes that have been proposed in the literature (see Section 4.5) would be applied retrospectively from data of introduction of the vaccine, in order to mimic a real-time analysis. For the second objective, these results would be compared to those obtained from an analysis of the full database without using sequential methods. Several vaccine – event pairs would preferably be evaluated in this way. These would include some pairs with a known association, and some that should be unrelated. Simulated data could also be used.

POC component 3: Comparative evaluation of standard methods for established vaccines

The objectives of POC component 3 are as follows:

- To investigate the performance of different statistical methods of risk evaluation for established vaccines.
- To relate any differences to failures of assumptions and other characteristics of the methods.

This study would be conducted in one or several databases with information on established vaccines. The aim is to study in greater detail than currently available how the methods compare, and to better understand the reasons behind any differences. The methods could include standard cohort, case-control and self-controlled designs.

POC component 4: Signal detection, strengthening and confirmation within a single database

The objectives of POC component 4 are as follows:

- To quantify the biases involved in undertaking signal detection, strengthening and confirmation within the same database (rather than using different databases and methods for signal detection and signal evaluation as recommended).
- To investigate how to obtain the best division of the sample into a training and test set.

This study would be conducted in a single database, with real vaccination data but simulated events whose association with vaccination is controlled by the experimenter. The second objective is to help determine what proportion of the data should be kept aside for confirmation in order to achieve an optimal balance between overall sensitivity and specificity.

10. REFERENCES

Aaby P, Samb B, Simondon F, Knudsen K, Seck AM, Bennett J, Markowitz L, Rhodes P, Whittle H. Sex-specific differences in mortality after high-titre measles immunization in rural Senegal. Bulletin of the WHO 1994; 72: 761-770.

Abrahamocicz M, Beauchamp M-E, Sylvestre M-P. Comparison of alternative models for linking drug exposure with adverseeffects. Statistics in Medicine 2010; 31: 1014-1030.

Ad Hoc Group for the Study of Pertussis Vaccines. Placebo-controlled trial of two acellular pertussis vaccines in Sweden – protective efficacy and adverse events. Lancet 1988; i : 955-960.

Alderslade R, Bellman MH, Rawson NSB, Ross EM, Miller DL. The National Childhood Encephalopathy Study: A Report on 1000 Cases of Serious Neurological Disorders in Infants and Young Children From the NCES Research Team. Department of Health and Social Security. Whooping Cough: Reports from the Committee on the Safety of Medicines and the Joint Committee on Vaccination and Immunization. London: Her Majesty's Stationery Office; 1981:79 –169.

Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. International Journal of Epidemiology 2012; 41: 861-870.

Anderson RM, May RM. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford 1991.

Andrews N, Miler E, Grant A, Stowe J, Osborne V, Taylor B. Thimerosal exposure in infants and developmental disorders: a retrospective cohort study in the United Kingdom does not support a causal association. Pediatrics 2004; 114: 584-591.

Andrews N, Stowe J, Al-Shahi Salman R, Miller E. Guillain-Barré syndrome and H1N1 (2009) pandemic influenza vaccination using an AS03 adjuvanted vaccine in the United kingdom: self-controlled case series. Vaccine 2011; 29: 7878-7882.

Andrews N, Stowe J, Miller E et al. A collaborative approach to investigating the risk of thrombocytopenic purpura after measles-mumps-rubella vaccination in England and Denmark. Vaccine 2012; 30: 3042-3046.

Austin PC. A critical appraisal of propensity score matching in the medical literature between 1996 and 2003 (with Discussion). Statistics in Medicine 2008; 27: 2037-2069.

Autran B, Asturias EJ, Evans S et al. Global safety of vaccines: strengthening systems for monitoring, management and the role of GACVS. Expert Review of Vaccines 2009; 8: 705-716.

Baggs J, Gee J, Lewis E et al. The Vaccine Safety Datalink : a model for monitoring immunization safety. Pediatrics 2011; 127; S45-S53.

Baker MA, Nguyen M, Cole DV, et al. Post-licensure rapid immunization safety monitoring program (PRISM) characterization. Vaccine 2013; 31(S10): K98-K112.

Barlow WE, Davis RL, Glasser JW et al. The risk of seizures after receipt of whole-cell pertussis or measles, mumps and rubella vaccine. New England Journal of Medicine 2001; 345:656-661.

Bate A, Lindquist M, Edwards IR et al. A Bayesian neural network method for adverse drug reaction signal generation. European Journal of Clinical Pharmacology 1998; 54: 315-321.

Benichou J. A review of adjusted estimators of attributable risk. Statistical Methods in Medical Research 2001; 10: 195 – 216.

Black C, Kaye JA, jick H. MMR vaccine and idiopathic thrombocytopaenic purpura. British Journal of ClinicalPharmacology 2003; 55: 107-111.

Breslow NE, Day N. Statistical Methods in Cancer Research. Volume II: The Design and Analysis of Cohort Studies. IARC Scientific Publications 82, 1987.

Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. Journal of the American Statistical Association 1983; 78: 1-12.

Brisson M, Edmunds WJ, Gay NJ, Law B, De Serres G. Modelling the impact of immunization on the epidemiology of varicella zoster virus. Epidemiology and Infection 2000; 125: 651-669.

Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Medical Research Methodology 2006, 6: 54, doi: 10.1186/1471-2288-6-54.

Brown JS, Kulldorff M, Chan KA et al. Early detection of adversedrug events within population-based health networks: application of sequential testing methods. Pharmacoepidemiology and Drug Safety, 2007; 16: 1275-1284.

Bryan P, Seabroke S, Davies C. H1n1 vaccine safety: real-time surveillance in the UK. Lancet 2010; 376: 417-418.

Cepeda S. et al. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. American Journal of Epidemiology 2003; 158: 280 – 287.

Chen RT. Vaccine risks: real, perceived and unknown. Vaccine 1999; 17: S41-S46.

Chen Y, Briesacher BA. Useof instrumental variables in prescription drug research with observational data: A systematic review. Journal of clinical Epidemiology 2011; 64: 687-700.

Confavreux C, Suissa S, Saddier P, Bourdès V, Vukusic S. Vaccination and the risk of relapse in multiple sclerosis. New England journal of Medicine 2001; 344: 319-326.

Cox DR. Regression models and life tables (with discussion). Journal of the Royal Statistical Society, series B 1972; 34: 187-220.

Cox DR, Hinkley DV. Theoretical Statistics. Chapman & Hall/CRC Press, 1974.

Curtis LH, Weiner MG, Boudreau DM, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. Pharmacoepi demiology and Drug Safety 2012; 21(S1): 23-31.

da Silveira CM, Salisbury BM, de Quadros CA. Measles vaccination and Guillain–Barré syndrome. Lancet 1997; 349: 14–6.

Davis RL. Vaccine safety surveillance systems: critical elements and lessons learned in the development of the US Vaccine Safety Datalink's Rapid Cycle Analysis capabilities. Pharmaceutics 2013; 5: 168-178.

De Greeff S, de Melker HE, van Gageldonk PGM et al. Seroprevalence of pertussis in the Netherlands: Evidence for increased circulation of Bordetella Pertussis. Plos One 2010; 5: e14183.

De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Eurosurveillance 2013; 18: pii=20585.

Didelman J, Romio K, Weibel D et al. Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe. British Medical Journal 2011; 343: d3908.

Dodd C, Romio SA, Black S et al. International collaboration to assess the risk of Guillain Barré syndrome following influenza (H1N1) 2009 monovalent vaccines. Vaccine 2013; 31: 4448-4458.

Donegan K, Beau-Lejdstrom, King B, Seabroke S, Thomson A, Bryan P. Bivalent human papillomavirus vaccine and the risk of fatigue syndrome in girls in the UK. Vaccine 2013; 31: 4961-4967.

Dourado I, Cunha S, da Gloria Teixera M et al. Outbreak of aseptic meningitis associated with mass vaccination with a Urabe-containing measles-mumps-rubella vaccine. American Journal of Epidemiology 2000; 151: 524–30.

Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. American Statistician 1999; 53: 177-190.

Escolano S, Farrington CP, hill C, Tubert-Bitter P. Intussusception after rotavirus vaccination: spontaneous reports (Letter). New England Journal of Medicine 2011; 365: 2139.

Escolano S, Hill C, Tubert-Bitter P. A new self-controlled case series method for analysing spontaneous reports of adverse events after vaccination. American Journal of Epidemiology 2013; 178: 1496-1504.

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: 483-486.

Evans S. Moving along the yellow brick (card) road? Drug Safety 2013; 36 (Supplement 1): S3-S4.

Farrington CP. Estimation of vaccine efficacy using the screening method. International Journal of epidemiology 1993; 22: 742-746.

Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. Biometrics 1995; 51: 228-235.

Farrington P, Pugh S, Colville A et al. A new method for active surveillance of adverse events from diphtheria/tetanus/pertussis and measles/mumps/rubella vaccines. Lancet 1995; 345: 567 – 569.

Farrington CP, Miller E, Taylor B. MMR and autism: further evidence against a causal association. Vaccine 2001; 19: 3632-3635.

Farrington CP. Control without separate controls: evaluation of vaccine safety using case-only methods. Vaccine 2004; 22: 2064-2070.

Farrington CP, Whitaker HJ. Semiparametric analysis of case series data (with Discussion). Journal of the Royal Statistical Society Series C 2006; 55: 553-594.

Farrington CP, Firth MJ, Moulton LH, Ravn H, Andersen PK, Evans SJ. Epidemiological studies on the non-specific effects of vaccines: II – methodological issues in the design and analysis of cohort studies. Tropical Medicine and International Health 2009; 14: 977-985.

Farrington CP, Hocine MN. Within-individual dependence in self-controlled case series models for recurrent events. Journal of the Royal Statistical Society Series C 2010; 59: 457-475.

Farrington CP, Anaya-Izquierdo K, Whitaker HJ, Hocine MN, Douglas I and Smeeth L. Self-controlled case series analysis with event-dependent observation periods. Journal of the American Statistical Association 2011; 106: 417-426.

Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. Journal of the American Sattistical Association 1999; 94: 496-509.

Fine PEM, Chen RT. Confounding in studies of adverse reactions to vaccines. American Journal of Epidemiology 1992; 136: 121-135.

Fine PEM, Clarkson JA. Individual versus public priorities in the determination of optimal vaccination policies. American journal of Epidemiology (1986); 124: 1012-1020.

Fine PEM, Williams TN, Aaby P, Kallander K, Moutlon LH, Flanagan KL, Smith PG, Benn CS. Epidemiological studies of the 'non-specific effects' of vaccines: I – data collection in observational studies. Tropical Medicine and International Health 2009; 14: 969-976.

Fireman B, Lee J, Bembom O, van der Laan M, Baxter R. Influenza vaccination and mortality: differentiating vaccine effects from bias. American Journal of Epidemiology 2009; 170: 650-656.

Forrow S, Campion DM, Herrinton LJ, et al. The organizational structure and governing principles of the Food and Drug Administration's Mini-Sentinel pilot program. Pharmacoepi demiology and Drug Safety 2012; 21(S1): 12-17.

Frome EL. The analysis of rates using Poisson regression models. Biometrics 1983; 39: 665-674.

Galindo Sardinas MA, Zambrano Cardenas A, Coutin Marie G et al. Lack of association between intussusception and oral polio vaccine in Cuban children. European Journal of Epidemiology 2001; 17: 783-787.

Glavani AP, Reluga T, Chapman GB. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. PNAS 2007; 104: 5692-5697.

Gambia Hepatitis Study Group. The Gambia hepatitis intervention study. Cancer research 1987; 47: 5282-7.

Gangarosa EJ, Galazka AM, Wolfe CR et al. Impact of anti-vaccine movements on pertussis control : the untold story. Lancet 1999; 351: 356-361.

Garbe E, Kloss S, Suling M, Pigeot I, Schneeweiss S. High-dimensional versus conventional propensity scores in a comparative effectiveness study of coxibs and reduced gastrointestinal complications. Eur J Clin Pharmacol 2013; 69: 549-557.

Georgousakis M, Jayasinghe S, Brotherton J, Gilroy N, Chiu C, Macartney K. Population-wide vaccination against human papillomavirus in adolescent boys: Australia as a case study. Lancet Infectious Diseases 2012; 12: 627-634.

Greene SK, Rett M, Weintraub ES 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: 1100-1109.

Greenland S. An introduction to instrumental variables for epidemiologists. International Journal of epidemiology 2000; 29: 722-729.

Groenwold RHH, Hak E, Klungel OH, Hoes AW. Instrumental variables in influenza vaccination studies: Mission impossible? Value in Health 2010; 13: 132-137.

Garenne M, Leroy O, Beau JP, Sene I. Child mortality after high-titre measles vaccines : prospective study in Senegal. Lancet 1991; 338: 903-907.

Gustaffson L, Hallander HO, Olin P, Reizenstein E, Storsaeter J. A controlled trial of a two-component acellular, a five component acellular, and a whole-cell pertussis vaccine. New England Journal of Medicine 1996; 334: 349-355.

Gustaffson L, Hessel L, Storsaeter J, Olin P. Long-term follow-up of Swedich children vaccinated with acellular pertussis vaccines at 3, 5 and 12 months of age indicates the need for a booster dose at 5 to 7 years of age. Pediatrics 2006; 118: 978-984.

Hallas J. Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis. Epidemiology 1996; 7: 478-484.

Harpaz R, DuMouchel W, Shah NH et al. Novel data mining methodologies for adverse drug event discovery and analysis. Clinical Pharmacology and Therapeutics 2012; 91: 1010-1012.

Hernan MA, Jick SS, Olek MJ, Jick H. Recombinant hepatitis B vaccine and the risk of multiple sclerosis: A prospective study. Neurology 2004; 63: 838-842.

Hill, AB. The environment and disease: Association or causation? Proceedings of the Royal Society of Medicine 1965; 58: 295-300.

Hocine MN, Farrington CP, Touzé E, Whitaker HJ, Fourrier A, Moreau T, Tubert-Bitter P. Hepatitis B vaccination and first central nervous system demyelinating events: reanalysis of a case-control study using the self-controlled case series method. Vaccine 2007; 25: 5938-5943.

Hocine MN, Musonda P, Andrews NJ, Farrington CP. Sequential case series analysis for pharmacoepidemiology. Journal of the Royal Statistical Society Series A, 2009; 172: 213-236.

Huang W. Stepwise likelihood ratio statistics in sequential studies. Journal of th Royal Statistical Society Series B 2004; 66: 401-409.

Jefferson T, Rudin M, Di Pietrantoni. Systematic review of the effects of pertussis vaccines in children. Vaccine 2003a; 16: 2003-2014.

Jefferson T, Price D, Demicheli V, Bianco E. Unintended events following immunization with MMR : a systematic review. Vaccine 2003b; 21: 3954-3960.

Jefferson T, Rudin M, Di Pietrantoni C. Adverse events after immunisation with aluminium-containing DTP vaccines : systematic review of the evidence. Lancet Infectious Diseases 2004; 4: 84-90.

Jensen AK, Gerds TA, Weeke P, Torp-Pedersen C, Andersen PK. On the validity of the case-time-control design for autocorrelated exposure histories. Epidemiology 2014; 25: 110-113.

Jick H, Vasilakis-Scaramozza C, Jick SS. Live attenuated polio vaccine and te risk of intussusception. British Journal of Clinical Pharmacology 2008; 52: 451-453.

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. American journal of Epidemiology 2003; 157: 158-165.

Klein NP, Hansen J, Lewis E et al. Post-marketing safety evaluation of a tetanus toxoid, reduced diphtheria toxoid and 3-component acellular pertussis vaccine administered to a cohort of adolescents in a United States health maintenance organisation. Pediatric Infectious Diseases Journal 2010; 29: 613-617.

Kramarz P, DeStefano F, Garguillo PM et al. Does influenza vaccination exacerbate asthma ? Analysis of a large cohort of children with asthma. Arch Fam Med 2000; 9: 617-623.

Kristensen I, Aaby P, Jensen H. Routine vaccinations and child survival: follow-up study in Guinea-Bissau, West Africa. BMJ 2000; 321: 1435-1438.

Kuhnert R, Hecker H, Poethko-Muller C, Schlaud M, Venneman M, Whitaker HJ, Farrington CP. A modified self-controlled case series method to examine association between multidose vaccinations and deaths. Statistics in Medicine 2011; 30: 666-677.

Kulldorff M, Bavis RL, Koczak M, Lewis E, Lieu T, Platt R. A maximised sequential probability ratio test for drug and vaccine safety surveillance. Sequential Analysis, 2011; 30: 58-78.

Langholz B, Borgan O. Counter-matching: a stratified nested case-control sampling method. Biometrika 1995; 82: 69-79.

Langholz B, Borgan O. Estimation of absolute risk from nested case-control data.Biometrics 1997; 53: 767-774.

Li L, Kulldorff M. A conditional maximised sequential probability ratio test for pharmacovigilance. Statistics in Medicine, 2010; 29: 284-295.

Liddell FDK, McDonald JC, Thomas DC, Cunliffe SV. Methods of cohort analysis: Appraisal by application to asbestos mining. Journal of the Royal Statistical Society, Series A 1977; 140: 469-491.

Lieu TA, Kulldorff M, Davis RL et al. Real-time vaccine safety surveillance for the early detection of adverse events. Medical Care, 2007; 45 Supplement 1: S89-S95.

Lumley T, Levy D. Bias in the case-crossover design: implications for studies of air pollution. Environmetrics 2000; 11: 689-704.

Luo X, Sorok GS. Analysis of recurrent event data under the case-crossover design with applications to elderly falls. Statistics in medicine 2008; 27: 2890-2901.

Lupton D. Risk. Routledge, 2008 (2nd edition).

Madsen KM, Hviid A, Vestergaard M et al. A population-based study of measles, mumps and rubella vaccination and autism. New England Journal of Medicine 2002; 347: 1477-1482.

Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. American Journal of epidemiology 1991; 133: 144-153.

Madigan D, Ryan PB, Schuemie M. Does it matter? Systematic evaluation of the impact of analytical choices on effect estimates on observational studies. Therapeutic Advances in Drug Safety 2013; 53-60.

McCullagh P, Nelder JA. Generalized Linear Models. Second Edition. Chapman & Hall/CRC Press, Boca Raton 1989.

Meldrum M. 'A calculated risk': the Salk polio vaccine field trials of 1954. British Medical Journal 1998; 317: 1233-1235.

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. British Medical Journal 2013; 346: f794.

Miller E, Andrews NJ, Waight PA, Slack MPE, George RC. Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study. Lancet Infectious Diseases 2011; 11: 760-768.

Mini-Sentinel. About Us: Background. Mini-Sentinel Coordinating Center. 2011. Available at: http://www.mini-sentinel.org/about_us/default.aspx.

Mini-Sentinel. Principles and Policies. November 2013. Available at: http://www.mini-sentinel.org/work_products/About_Us/Mini-Sentinel-Principles-and-Policies.pdf.

Mohammed SM, Senturk D, Dalrymple LS, Nguyen DV. Measurement error case series models with application to infection-cardiovascular risk in older patients on dialysis. Journal of the American Statistical Association 2012; 107: 1310-1323.

Mohammed SM, Dalrymple LS, Senturk D , Nguyen DV. Design considerations for case series models with exposure onset measurement error. Statistics in Medicine 2013; 32: 772-786.

Moulton LH, Wolff MC, Brenneman G, Santosham M. Case-cohort analysis of case-coverage studies of vaccine effectiveness. American journal of Epidemiology 1995; 142: 1000-1006.

Murray CJL. Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bulletin of the WHO 1994; 72: 429 – 445.

Musonda P, Hocine MN, Andrews NJ, Tubert-Bitter P, Farrington CP. Monitoring vaccine safety using case series cumulative sum charts. Vaccine 2008; 26: 5358-5367.

Musonda P, Hocine MN, Whitaker HJ, Farrington CP. Self-controlled case series method: small sample performance. Computational Statistics and Data Analysis 2008a; 52: 1942-1957.

Nathanson N, Langmuir AD. The Cutter incident: poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United states during the spring of 1955, I. American Journal of Hygiene 1963, 78: 16-18.

Navidi W. Bidirectional case-crossover designs for exposures with time trends. Biometrics 1998; 54: 596-605.

Nazareth I, Tavares F, Rosillon D, Haguinet F, Bauchau V. Safety of AS03-adjuvanted split-virion H1N1 (2009) pandemic influenza vaccine: a prospective cohort study. BMJ Open 2013; 3: e001912.

Nelson JC, Yu O, Dominguez-Islas CP et al. Adapting group sequential methods to observational postlicensure vaccine safety surveillance: Results of a pentavalent combination DTaP-IPV-Hib vaccine safety study. American Journal of Epidemiology 2013; 177: 131-141.

Nguyen M, Ball R, Midthun K, et al. The Food and Drug Administration's post-licensure rapid immunization safety monitoring program: strengthening the federal vaccine safety enterprise. Pharmacoepi demiology and Drug Safety 2012; 21(S1): 291-297.

Offit P. The Cutter incident, 50 years later. New England Journal of Medicine 2005; 352: 14.

Olin P, Rasmussen F, Gustaffson L, Hallander HO, Heijbel H. Randomised controlled trial of two-component, three-component, and five-component acellular pertussis vaccines compared with whole-cell pertussis vaccine. Lancet 1997; 350: 1569-1577.

Olin P, Gustafsson L, Bareto L, hessel L, Mast TC, Rie AV, Bogaerts H, Storsaeter J. Declining pertussis incidence in Sweden following the introduction of acellular pertussis vaccine. Vaccine 2003; 21: 2015-21.

Orenstein EW, Se Serres G, Haber MJ et al. Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. International Journal of Epidemiology 2007; 36: 623-631.

Page ES. Continuous inspection schemes. Biometrika 1954; 41: 100-115.

Panagiotopoulos T, Antoniadou, Valassi-Adam E. Increasein congenital rubella occurrence after immunisation in Greece : retrospective survey and systematic review. British Medical Journal 1999; 319: 1462-1466.

Pearl J. Causal inference in statistics: an overview. Statistics Surveys 2009; 3: 96-146.

Platt R, Carnahan RM, Brown JS, et al. The US Food and Drug Administration's Mini-Sentinel: status and direction. Pharmacoepidemiology and Drug Safety 2012: 21(S1): 1-8.

Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika 1986; 73: 1-11.

Quantin C, Benzenine E., Velten M et al. Self-controlled case series and misclassification bias induced by case selection from administrative hospital databases: Application to febrile convulsions in pediatric vaccine pharmacoepidemiology. American Journal of Epidemiology 2013; 178: 1731-1739.

Ramsay ME, Moffatt D, O'Connor M. Measles vaccine: a 27-year follow-up. Epidemiology and Infection 1994; 112: 409-412.

Ray WA, Griffin MR. Use of MEDICAID data for pharmacoepidemiology. American Journal of Epidemiology 1989; 129: 837-848.

Rodrigues L, Smith PG. Use of the case-control approach in vaccine evaluation: efficacy and adverseeffects. Epidemiologucal Reviews 1999; 21: 56-72.

Rowhani-Rahbar A, Klein N, Lewis N, Fireman B, Ray P, Ragson B, Black S, Klein JO, Baxter R. Immunization and Bell's Palsy in children: a case-centered analysis. American Journal of Epidemiology 2012; 175: 878-885.

Ryan P, Schuemie M, Madigan D. Empirical performance of a self-controlled cohort method: Lessons for developing a risk identification and analysis system. Drug safety 2013; 36 Supplement: 95-106.

Ryan PB, Stang PE, Overhage JM et al. A comparison of the empirical performance of methods for a risk identification system. Drug Safety 2013a; 36 (Supplement 1): S143-S158.

Salmon DA, Proschan M, Forshee R et al. Association between Guillain-Barré syndrome and influenza A (H1N1) 2009 monovalent inactivated vaccines in the USA: a meta-analysis. Lancet 2013; 381: 1461-1468.

Sartorius B, Penttinen P, Nilsson J et al. An outbreak of mumps in Sweden, February-April 2004. Eurosurveillance 2005; 10: pii=559.

Savage LJ. The Foundations of Statistics. Wiley, New York 1954.

Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat – Which should we use? Value in Health 2002; 5: 431 – 435.

Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data.Epidemiology 2009; 20: 512-522.

Self SG, Prentice RL. Asymptotic distribution theory and efficiency results for case-cohort studies. Annals of Statistics 1986; 16: 64-81.

Senn S, Graf E, Caputo A. Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure. Statistics in Medicine 2007; 26: 5529-5544.

Shah B. R. et al. Propensity score methods gave similar results to traditional regression modeling in observationsl studies : a systematic review. J Clin Epidemiol 2005; 58: 550 – 559.

Shann F, Nohynek H, Scott JA, Hesseling A, Flanagan KL. Randomized trials to study the nonspecific effects of vaccines in children in low-income countries. Pediatric Infectious Diseases Journal 2010; 29: 457-61.

Shewhart WA . Economic Control of Quality of Manufactured Products. Van Nostrand Reinhold: Princeton, 1931.

Shields WD, Nielsen C, Buch D, Jacobsen V, Christenson P, Zachau-Christiansen B. Relationship of pertussis immunization to the onset of neurologic disorders: a retrospective epidemiologic study. J

Pediatrics 1988; 113: 801–5.

Shih M-C, Lai TL, Heyse JF, Chen J. Sequential generalized likelihood ratio test for vaccine safety evaluation. Statistics in Medicine, 2010; 29: 2698-2708.

Simpson SE. A positive event dependence model for self-controlled case series with applications in postmarketing surveillance. Biometrics 2013; 69: 128-136.

Slovic P. The Perception or Risk, Earthscan, 2000.

Slovic P. The feeling of risk: New perspectives on risk perception. Earthscan, 2010.

Smeeth L, Cook C,Fombonne E, Heavy L, Rodrigues LC, Smith PG, Hall AJ. MMR vaccination and pervasive developmental disorder: a case-control study. Lancet 2004; 364: 963-969.

Sturmer T, Brenner H. Degree of matching and gain in power and efficiency in case-control studies. Epidemiology 2001; 12: 101-108.

Sturmer T. et al. 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. J Clin Epidemiol 2006; 59: 437 – 447.

Suissa S. The case-time-control design. Epidemiology 1995; 6: 248-253.

Taylor B, Miller E, Farrington CP et al. Autism and measles, mumps and rubella vaccine: no epidemiologica; evidence for a causal association. Lancet 1999; 353: 2026-2029.

Thacker SB, Stroup DF, Carande-kulis V et al. Measuring the public's health. Public Health Reports 2006 ; 121: 14 – 22.

Trogdon JG, Nurmagambetov TA, Thompson HF. The economic implications of influenza vaccination for adults with asthma. American Journal of Preventive Medicine 2010; 13: 403-410.

Tse A, Tseng HF, Greene SK et al. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink Project, 2010-2011. Vaccine 2012; 30: 2024-2031.

Ury HK. Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data. Biometrics 1975; 31: 643-649.

Van Boven JFM, De Jong-van den Berg LTW, Vegter S. Inhaled corticosteroids and the occurrence of oral candidiasis: a prescription sequence symmetry analysis. Drug Safety 2013; 36: 231-236.

Van der Maas N, Mooi FR, de Greeff SC et al. Pertussis in the Netherlands : is the current vaccination strategy sufficient to reduce disease burden in young infants ? Vaccine 2013; 31: 4541-4547.

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: 603-610.

Van Holle L, Bauchau V. The upper bound to the relative reporting ratio: a measure of the impact of the violation of hidden assumptions underlying some disproportionality methods used in signal detection. Pharmacoepidemiology and Drug Safety 2014; DOI: 10.1002/pds.3556.

Van Puijenbroek EP, Bate A, Leufkens HGM et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiology and Drug safety 2002; 11: 3-10.

Verstraeten T, Davis RL, DeStefano F, Lieu TA, Rhodes PH, Black SB, Shinefield H, Chen RT. Safety of thimerosal-containing vaccines: a two-phased study of computerized health maintenance organization databases. Pediatrics 2003; 112: 1039-1048.

Vines SK, Farrington CP. Within-subject exposure dependency in case-crossover studies. Statistics in Medicine 2001; 20: 3039-3049.

Wald A. Sequential Analysis. Wiley: New York, 1947.

Wang S, Linkletter C, Maclure M, Dore D, Mor v, Buka S, Wellenius GA. Future-cases as present controls to adjust for exposure-trend bias in case-only studies. Epidemiology 2011; 22: 568-574.

Wei LJ, Gidden DV. An overview of statistical methods for multiple failure time data in clinical trials. Statistics in Medicine 1997; 16: 833-839.

Weintraub ES, Baggs J, Duffy J et al. Risk of intussusception after monovalent rotavirus vaccination. New England Journal of Medicine 2014; doi: 10.1056/NEJMoa1311738.

Weldeselassie YG, Whitaker HJ, Farrington CP. Use the self-controlled case-series method in vaccine safety studies: review and recommendations for best practice. Epidemiology and Infection 2011; 139: 1805-1817.

Whitaker HJ, Hocine MN, Farrington CP. On case-crossover models for environmental time series data. Environmetrics 2007; 18: 157-171.

Whitehead J. The Design and Analysis of Sequential Clinical Trials. Revised Second Edition. Wiley, Chichester 1997.

WHO. Causality assessment of adverse events following immunization. WHO Weekly Epidemiological Record 2001; 76: 85-88.

Wilson K, Hawken S. Drug safety studies amd measures of effect using the self-controlled case series design. Pharmacoepidemiology and Drug Safety 2013; 22: 108 – 110.

Wong K, Campitelli MA, Stukel TA, Kwong JC. Estimating influenza vaccine effectiveness in community-dwelling elderly patients using the instrumental variable analysis method. Archives of Internal Medicine 2012; 26: 484-491.

Yih WK, Kulldorff M, Fireman BH et al. Active surveillance for adverse events: the experience of the vaccine safety datalink project. Pediatrics, 2011; 127: 854-864.

Yoo BK, Frick Kd. The instrumental variable method to study self-selection mechanism: A case of influenza vaccination. Value in health 2006; 9: 114-122.

ANNEXES

Annex I. THE PERSPECTIVE OF PUBLIC HEALTH AGENCIES

Note prepared by Nick Andrews, PHE

Public Health Agencies (PHAs) are responsible for responding to health risks as well as improving the health of the nation. Vaccines have been one of the great success stories for public health and are a hugely important area of work for PHAs. Vaccine preventable diseases undergo enhanced surveillance at PHAs so that the impact and effectiveness of vaccines can be assessed alongside the uptake of vaccination, the costs and the assessment of any risks. Considering risks the following bullet points highlight keys priorities for PHAs.

1. **Passive post-licensing surveillance systems**
   A system for the reporting of suspected adverse events following immunisation (AEFI) is necessary to identify possible true risks. The system (vaccine specific pharmacovigilance) should be able to detect temporal /spatial clusters and novel/ serious adverse events. Also it should be able to assess whether the AEFIs are associated with a particular batch or manufacturer's vaccine and identify issues with vaccine delivery / cold-chain. The ability to do individual level causality assessment is also important.

   Statistical methods need to be employed to help identify signals from a potentially huge number of reports. Of particular benefit is to estimate the expected number of reports of a wide range of events to enable a rapid assessment of reports as they arise (i..e Observed v Expected analyses). It is recognised however that passive reporting will often lack sensitivity. Supplementing the passive reporting with active monitoring using larger routine dataset is therefore an important area for development.

2. **Active surveillance for identifying possible risks**
   In addition to passive surveillance having a system to actively and rapidly assess a range of adverse events of interest would be helpful since this addresses the sensitivity issue with passive surveillance. Rapid Cycle Analysis in the US within the vaccine safety datalink is an example of this.

3. **Epidemiological studies**
   There are many methods for conducting controlled epidemiological studies to assess the relation between the vaccine and the event of interest. It is important that the ascertainment of the events of interest should be unbiased with respect to vaccine history. For example the event of interest may be ascertained through a hospital database and the vaccine history ascertained from either writing to the General Practitioner or through linkage of datasets.

   Data sources: A key aspect is to have access and knowledge of suitable data sources for assessing risks of different types and being able to link data. For example having

primary care data, hospital data, vaccine registries, and disease registers. This also means having permissions in place to use the data rapidly for the adverse event under assessment- this may be on a patient level or a system/dataset level.
Pre-defined protocol: Before any data collection or analysis is carried out the purpose and the methods to be used need to be documented. This includes the hypothesis and objectives of the study and should include the risk periods to be used, how cases are to be selected, detailing what index dates are to be collected and what analysis is to be performed.

4. **Pregnancy Registers**
Developing registers of individuals given vaccines inadvertently – such as to pregnant women or other risk groups that are contraindicated.

5. **Communication**
Educating the public on how vaccine safety is monitored in their country before a suspected adverse event is published would give reassurance in the whole vaccine programme. Communicating the risk-benefit balance is complicated as it involves a risk assessment on the individual level but also the acceptance of the concept that each vaccinated individual is benefiting public health and each person has a responsibility to this. This is a difficult concept as the key question from a patient would be "what does it mean for me and my family". It is important to communicate what a "rare" adverse event is in relation to the disease risk. The risk/benefit assessment needs to be clear and backed by robust epidemiological studies.

To help attain this having good communication with other countries using the same vaccines to share safety concerns and potentially do combined studies if power is low within one country is useful. Also, being able to rapidly communicate with those in other countries using the same vaccines as well as EMA, CDC and WHO to see if they have any evidence of a risk or are doing studies is very helpful. This will enable the countries involved to provide a much stronger evidence based assessment of vaccine benefits and risk.

6. **Access to Experts in the Field of concern.**
As the risk to be assessed is nearly always new it is important to gain expert guidance on the condition under assessment. Experts can assist in the development of code lists and advise in the diagnostic pathway of the condition.

7. **Financial Capacity**
Having the capacity (financial) and experts to do the studies if essential. This means clinicians, statisticians, epidemiologists, researchers and access to dataset which can be costly.

8. **Methodological development**

Methodological development to enable available data to be best used to assess vaccine risks. The development of the self-controlled case-series method is a great example of this.

Implications of the needs of PHAs on methods for safety assessment

Considering the above priorities implications on methods are as follows…

a) How do we split precious data sources such as GP and hospital databases between signal detection and testing. Can a method be devised that could somehow do both or be seen as a 2-stage design?

b) When considering pooling of data or results from different data-sources or across countries methods need to be able to adjust for confounding effects that may be different between countries. In other words we can't have a situation where the most appropriate analysis of a countries own data gives substantially different results when pooled and analysed by a common method.

c) PHAs need a simple clear message - this may be difficult if methods are overly complex or if by pulling together data that are too different means it is very complicated to produce a sensible combined analysis.

d) PHAs need a result as quickly as possible, but not at the cost of a poor study. Having methods for both rapid evaluation (in days) and longer more detailed studies would be useful.

Annex II. THE PERSPECTIVE OF REGULATORY AUTHORITIES

Note prepared by Suzie Seabroke (MHRA)

Vaccine safety monitoring/rapid assessment

Observed vs. Expected analyses:

Over last 5 years the MHRA have used observed vs. expected (O/E) analyses to monitor events of special interest likely to be reported post-vaccination. These events are defined prior to a new vaccination campaign (e.g. HPV vaccine in 2008 and H1N1 influenza vaccine in 2009/10) as those likely to be reported as potential adverse events. These events are either potential ADRs identified from clinical trial data, potential ADRs identified from past experience with similar vaccines, or coincidental events that are prevalent in the target population and likely to be reported as potential ADRs.

We used spontaneous ADR data, collected principally via the UK Yellow Card scheme but also through reports in the media, as the observed count and derived an expected count using background incidence data from the Clinical Practice Research Datalink (CPRD) and vaccine exposure data.

We have used both the Maximised Sequential Probability Ratio Test (MaxSPRT) method for routine weekly analyses, starting immediately after the introduction of the vaccine into the national immunisation schedule, as this method adjusts for multiplicity and also a more simplistic snapshot method on an ad hoc basis. The MaxSPRT was felt to be more appropriate for the weekly analyses but less easy to communicate the results to the public in lay terms. The snapshot method was therefore used more ad hoc when findings were needed to be communicated to the public or when important ADRs other than those pre-defined prior to the introduction of the vaccine were reported.

Key points for this type of analysis:

- Helps put spontaneous data into context to better distinguish potential safety signals from coincidental event reporting
- Needs to be as near-real time as possible in order to rapidly detect any potential signals or to reassure that the observed is consistent with the expected
- Rapidity more important than definitiveness at this stage therefore methods can be crude
- Particularly for vaccines such as the seasonal flu vaccine that are administered over a short period of time or mass vaccination programmes such as the H1N1 flu vaccine, rapidity is critical as any serious potential safety concerns need to be identified in time to inform the vaccination campaign
- Rapidity also important to provide reassurance when observed events are consistent with expected in order to maintain public confidence in the vaccination campaign
- The results can be simply explained to the relevant stakeholders including healthcare professionals, marketing authorisation holders, the media, and the public

- Data on background incidence rates is required either to be calculated in advance or a data source to be available to calculate these rapidly as required
- Timely and frequent data on vaccine exposure is required

Other types of rapid assessment methods:

For vaccines with a high uptake rate (e.g. HPV vaccine in the UK has over 85% uptake in 12-13 year old girls), we have also used ecological analyses as a rapid assessment method. A potential advantage over the O/E analyses are that all data comes from same source (e.g. all from CPRD rather than a mixture of spontaneous data and CPRD data) but this method is not as rapid as the O/E methods given delays in recording. If both approaches are taken, and yield similar results, this can provide further confidence in the robustness of the results. Not all events are suitable for this type of analysis however.

For vaccination campaigns that are of particular pharmacovigilance importance, i.e. where there is little safety data already known, (for example in the recent campaign using the pertussis vaccine in pregnant women), in order to support the routine pharmacovigilance conducted using spontaneous reports we have again used additional rapid assessment methods. In the pertussis vaccine example a more robust analysis has been conducted identifying matched vaccinated and unvaccinated patients in the CPRD and comparing the risk of a range of pre-defined ADRs in the two growing cohorts on a regular basis. Again this type of analysis is not as rapid as the O/E methods due to delays in recording but is helpful for adverse events that may not be reported through spontaneous reporting schemes.

<u>Signal verification studies</u>

Choice of study method for a vaccine safety signal verification study has to be done on a case-by-case basis as different designs suit some signals/vaccines better than others. From experience however the self-controlled-case-series design has been particularly useful for vaccine signal verification.

Key points for choice of method

- Data sources available and how exposure and outcomes are recorded e.g. vaccines administered in a school-based setting are not well recorded in GP records and outcomes that require diagnosis by a hospital specialist also may not be well recorded.
- Availability of suitable controls – vaccination programmes with high uptake rates result in a lack of controls in both absolute number and controls likely to be different to vaccinated population.
- Power to detect rare outcomes – many of the recent vaccine safety concerns have involved rare conditions with national data sources struggling to identify sufficient patient numbers to calculate a precise estimate of risk. Differences in healthcare systems/medical practice across different countries however make pooling data from different countries in order to increase power problematic and may introduce bias. Under these circumstances our view would be that the benefits of different studies with

results that are replicated using different datasets, ideally in different countries and potentially using different designs (albeit individually still lacking in power) should not be underestimated. These benefits would need to be traded off against the power that a single larger study might provide (albeit with the possibility of methodological flaws. The ability to provide such results in a timely fashion might also influence the choice of study/studies to be performed.

- From a regulatory perspective it is critical to have some information on a potential signal very rapidly in order that action can be taken to protect subjects as quickly as possible if necessary. The rapid assessment methods above are generally used to provide such initial information. For signal verification although rapidity is still important definitiveness is also now a key priority. Careful consideration is therefore required in choice of study design to ensure that the method chosen can deliver robust results in the least amount of time possible. Methods for rapid signal verification are therefore of considerable interest to regulators but these must not be at the expense of an unacceptable loss of quality.

Annex III. THE PERSPECTIVE OF VACCINE MANUFACTURERS

Note prepared by Catherine Cohet and Dominique Rosillon (GSK)
Note: the views and opinions expressed in this document are of the authors and not an official GSK position.

Vaccine manufacturers continuously monitor the safety profile of their products, by collecting and analysing data from clinical trials, post-marketing studies and surveillance activities (pharmacovigilance). Manufacturers also follow up information on general safety discussions and concerns through reviewing the scientific literature. This aims at a rapid and effective identification of vaccine safety issues, the assessment of the associated risks - which is key to protect individuals and public health - and appropriate actions to maximize the benefit and minimize the risks.

This document briefly summarises the key features for vaccine safety risk assessment from a vaccine manufacturer's perspective, and provides insight on how pharmaco-epidemiological studies addressing potential risks related to vaccination are requested, designed, planned, implemented and reported. However, it does not intend to dive into the details of the methods, which are addressed in a separate WP4/Risk work stream.

Overall, the assessment of vaccine safety by the MAH (marketing authorisation holder) is driven or characterised by:

- Values (patient safety first, transparency, ethics…)
- Complex regulations
- Defined processes, periodically audited and inspected
- Competent staff with medical and scientific skills and experience

A. Structural, governance and process considerations

Vaccine manufacturers have established processes to collect adverse events (AEs), notify regulatory authorities, and systematically review ongoing safety data relating to their products, as defined by EU regulation. These processes include: maintenance of a worldwide clinical safety database of AEs from clinical trials, post-marketing studies, and pregnancy reports; literature reviews; management of individual case safety report (ICSRS); production of PSURs, DSURs, PBERs (Periodic Benefit Risk Evaluation Reports, Development Safety Updates, Periodic Safety Update Reports); safety reports to support license renewals; signal detection, evaluation and management, and labelling activities (pre- and post-marketing) in line with signal evaluation; and coordination of Risk Management and Pharmacovigilance Plans. In addition, the QPPV (Qualified Person for Pharmacovigilance) formally communicates both internal and external safety information.

Benefit-risk monitoring occurs as an integral part of this holistic safety management process, which includes review and evaluation, by multi-disciplinary product-specific safety review committees, of individual spontaneous AE reports and literature reports, batch reviews, enquiries from externals sources (including regulatory authorities and healthcare providers) and serious adverse events (SAEs) from clinical trials and post-marketing setting. Once a signal is

identified and evaluated, the appropriate action(s) is agreed upon within the vaccine-specific safety review committee and escalated to relevant governing bodies, with communication to competent authorities.

The potential impact of the finding on the benefit-risk balance of the vaccine, as well as its potential effect on public health, is a critical factor in the determination of planned actions, that can include: continuing routine surveillance; defining additional work to quantify the risk; changes to the vaccine prescribing information; changes to the Risk Management Plan (RMP); update of the clinical trials' investigator brochure; and/or referral to internal governing bodies and cascading to contributory disciplines and functions (such as regulatory affairs, clinical development, epidemiology/ pharmaco-epidemiology) for communication with regulators and implementation of pharmaco-epidemiological studies (e.g. PASS) if applicable. In addition to the assessment of changes to the risk profile of each vaccine during ongoing, routine surveillance and at safety review meetings, the risk-benefit balance is assessed at the time of periodic regulatory reporting (through DSURs, PSURs, PBRERs), file submissions to regulatory agencies, and license renewals.

The RMP is created and maintained by the manufacturer for all vaccine licence applications in the EU and follows the requirements set out in the Guideline on good pharmacovigilance practices (GVP). The RMP includes a description of the existing safety profile of products as well as potential and identified risks and outlines the risk minimisation activities. In order to assess the efficiency of the measures taken to minimize a specific risk, the RMP defines, in advance, the way this efficiency is measured. Potential risks can be identified from a number of sources: clinical trials, spontaneous reporting, health outcomes of special interest, "usual suspects" for a given vaccine class or history of risk with a previous vaccine (e.g. H1N1 pandemic influenza vaccines and Guillain Barré Syndrome, MMRV and febrile seizures, rotavirus vaccines and intussusception), and are updated based on the accumulation of incoming post-marketing data, in parallel with the emergence of newly identified unexpected signals.

Risk assessment and management, and further evaluation of risk through pharmaco-epidemiological studies are conducted following internal Standard Operating Procedures, strictly aligned with regulations and pharmacovigilance guidances. The ENCePP code of conduct is applied to the extent feasible when conducting PASS studies. Decision to conduct a formal risk assessment (e.g. via a pharmaco-epidemiological study), as opposed to routine pharmacovigilance surveillance, can be initiated as a company decision or following a requirement from a competent authority. With respect to labelling, the wider acceptance of quality safety methods for observational research in the labelling decisions may improve the timeliness of safety assessment as well as its relevance to real world practice. Finally, also part of the process are continuous interactions with regulators and (supra)national bodies to convey a common understanding of the challenges of designing and interpreting studies and identify studies that will adequately inform their assessment of the benefit-risk balance.

B. Feasibility considerations to inform the proper planning and design of pharmaco-epidemiological risk assessment studies

A substantial proportion of epidemiology/pharmaco-epidemiology studies conducted by the vaccine industry aim to fulfil a post-authorisation commitment, and this proportion varies along

the vaccine lifecycle, higher around the time vaccines are approaching licensure and launch, and somehow lower but important when undergoing labelling variations and changes in indications or formulation. Besides, the regulatory environment is evolving and becoming more stringent in terms of expectations from companies – recent examples are the 2013 updated RMP requirements (EMA/465932/2013 Rev1), the new PASS guidance implemented in July 2012 (EMA/813938/2011 Rev1), or the variation classification guideline published by the EC in May 2013 (EMA/427505/2013). This complex ever-changing regulatory framework is taken into consideration when designing, planning, implementing and reporting vaccine safety studies, together with scientific and expert determinants.

Moreover, it is important for the MAH to ensure that evidence generation is based on robust methods, best practices (e.g. applying ENCePP code of conduct), but also operational feasibility and adequate resourcing.

Methodological feasibility considerations in study planning and design include questions such as:

- What is the most appropriate data source? E.g. field study with primary prospective data collection vs. healthcare database; choice of country(ies), population, or health care setting (primary, hospital, public health programme);
- What is the most appropriate design for a given question / safety outcome, based on available data for exposure, outcome and covariates?
- Can pragmatic randomised clinical trials be envisaged?
- Can special populations who might be at increased risk be identified, and how?
- Is a quick assessment of the research question (e.g. evaluation of incidence vs. vaccine coverage in a relevant healthcare database) an appropriate option?
- What are the main sources of potential error and uncertainty (bias, confounding) and how can they be addressed in the design to ensure the results provide a valid answer?
- Of note, manufacturers have full access to clinical trial safety data, which can therefore be used for risk assessment at both pre- and post-licensure stages (e.g. using the pooled control harms, or pooling safety over an entire development program); this is potentially a major input in the expected benefit-risk assessment at licensure stage.

Operational feasibility considerations include:

- What are the options with regards to the type of collaboration and access to data– i.e. company-sponsored study, collaborative study, outsourced study;
- Governance and ethical considerations (e.g. on data access, study protocol review and approval);
- Availability of background incidence estimates of the outcome under study, in the population of interest, in particular for rare diseases;
- Are timelines for designing and delivering study results compatible with regulatory requirements and expectations?
- Are the level of resources and cost of the study acceptable?
- Is scientific advice available? Are the required inputs from technical and scientific experts available to the industry (e.g. blocked due to perceived conflict of interests)?

C. Methodological considerations: developing the most relevant design

Once the above feasibility considerations have been addressed, the following aspects are often taken into account in the design of a given risk assessment study:

- Methodology overall:
  - Active vs. passive surveillance or data collection?
  - Observed /expected analyses vs. analytical comparison vs. descriptive?
  - Hypothesis generating vs. hypothesis testing?
  - Retrospective vs. prospective data acquisition?
- Study design:
  - Confounding: what is the current knowledge of covariates and risk factors other than vaccination (e.g. age in the association between rotavirus vaccination and intussusception; H1N1 infection in the association between H1N1pandemic influenza vaccines and GBS); what is the capability for properly controlling for known confounding factors? Can methods be adapted to consider potentially unknown confounders?
  - Choice/availability of adequate controls;
  - Bias: selection of cases or controls; sampling individual vs. density sampling (time-at-risk periods).
- Exposure:
  - Is brand-specific data / batch number available for the vaccine under study?
  - Is exposure data validated?
  - Is date of exposure accurate? Prescription vs. dispensing vs. administration
  - In case of dose schedule, is dose sequence recorded?
  - Is the study setting covering all sources of vaccinations? Public health immunization programs vs. individual vaccination at specific clinics.

- Outcome:
  - Is the disease entity clearly clinically defined? (e.g. particularly an issue for some neurological outcomes e.g. neuritis);
  - Is an operational case definition available and suitable for the study? (E.g. is there a Brighton collaboration definition?);
  - For database studies, is there a validated set of codes and are coding aspects clear (incl. coding changes over time)? Is there a possibility to review case profiles or subjects' medical charts?
  - Is the case finding and ascertainment approach validated? (sensitivity, specificity, positive/negative predicted value);
  - What is the most appropriate design for acute vs. chronic outcomes; incident vs. recurrent outcomes ;
  - Post-vaccination risk period: what is the evidence? Are there reasonable/documented assumptions for the duration of the risk period?
- Sample size and power considerations (e.g. risk level that can be detected or excluded)
- Statistical methods in general

<u>Examples of GSK studies</u>

- HPV vaccine and spontaneous abortions; HPV vaccine and immune-mediated diseases: cohort studies in the UK CPRD GOLD; see ENCePP register for description:

  - An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom
  - Post-marketing safety study to assess the risk of spontaneous abortions in women exposed to Cervarix in the United Kingdom
- Pandemic H1N1 influenza vaccine and medically attended adverse events (MAEs), SAEs, AESIs and pregnancy outcomes: prospective cohort study (Nazareth et al, BMJ Open 2013; Tavares et al, Vaccine 2011)
- Pandemic H1N1 influenza vaccines and AESIs (adverse events of special interest: neuritis, convulsions, anaphylaxis, encephalitis, vasculitis, GBS, Bell's palsy, demyelinating disorders, vaccination failure): hypothesis-generating retrospective studies in Swedish and Canadian databases using historical cohort and matched cohort designs and SCCS (Arnheim-Dahlström et al, BMJ. 2012 for SCCS on epileptic seizures in Sweden; other components to be published)
- Pandemic H1N1 influenza vaccines and solid organ transplant rejection: SCCS in the UK CPRD GOLD (ClinicalTrials.gov # NCT01715792)
- Rotavirus vaccine and intussusception: active surveillance with a SCCS analysis (Velázquez et al, Pediatr Infect Dis J. 2012)
- Risk of febrile convulsions after MMRV vaccination in comparison to MMR or MMR+V vaccination (Schink et al, Vaccine 2014); modelling relative rates of hospitalisation for febrile convulsions and severe varicella under combined MMRV compared to separate MMR+V vaccination;
- Seasonal influenza vaccines and vaccine failure: prospective multi-centre, multi-year, hospital based, influenza surveillance study; test-negative case-control design to measure vaccine effectiveness in preventing flu-associated hospitalization in the elderly in Canada (ClinicalTrials.gov # NCT01517191)

Annex IV. DATA REQUIREMENTS FOR SOME KEY METHODS

Cohort Methods

For the Cox model: event times and, for each event time, the corresponding risk set, with covariate data including vaccine exposure on each case at its event time, and on the members of each risk set.

For the Nelson-Aalen estimator of the cumulative hazard: event and censoring times for all cohort members, stratified by covariates as required.

For the Poisson model: for each combination of vaccine exposure category, covariate level, and age/season group, the number of events observed and the total person-time at risk for the whole cohort within that combination.

Case referent methods

For matched case-control methods: for each 1:M matched set, a case-control indicator, vaccine exposure and covariates for all M+1 individuals, and the matched set identifier.

Nested case-control method: for each event time, a random selection of M controls from the corresponding risk set, with covariate information including vaccine exposure on the case and the M matched controls, and the matched set identifier.

Cumulative hazard estimator for nested case-control studies: in addition to the above, the event times and the sampling proportion for each risk set.

Self-controlled methods

For the case-crossover method: for each event time and preceding M control times, vaccine exposure information on the corresponding event and control windows, along with the case identifier. Fixed covariate information required only for interactions with vaccine effects.

For the standard self-controlled case series method: for each case and each vaccine exposure and age/season group, the person-time spent by the case within that group and the number of events observed, along with the case identifier. Fixed covariate information required only for interactions with vaccine effects.

For the semiparametric SCCS method: for each case, the list of distinct sample-wide event times within the case observation period, and the vaccine exposure at each of these, along with the case identifier. Fixed covariate information required only for interactions with vaccine effects.