COMMENTARY

Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines[†]

Olivia Mahaux*, Vincent Bauchau and Lionel Van Holle

Vaccine Clinical Safety and Pharmacovigilance, GSK Vaccines, Wavre, Belgium

INTRODUCTION

Observed-to-expected (OE) analyses, together with data mining algorithms[1–7] and pharmacoepidemiological studies,[8] are part of the quantitative pharmacovigilance toolkit for vaccines. While data mining algorithms generate hypotheses about potential safety concerns and pharmacoepidemiological studies test specific hypotheses or measure associations, OE analyses stand in between. The role of OE analyses is to refine previously detected signals when there is not enough information to determine whether further action is necessary.

In this paper, the focus is on the OE analyses of spontaneous reports, where the observed number of cases is obtained from a spontaneous reporting system and compared with the expected number of cases calculated based on background incidence rates from independent sources, such as epidemiological studies or national statistics. Note that disproportionality data mining algorithms estimate an "OE ratio" generated based on expected and observed numbers of cases from a single spontaneous reporting system.

The key requirements and statistical methods recommended for OE analyses are described in European guidelines.[9,10] Here, we discuss in more

detail how to perform the analysis and deal with uncertainties. Although described here in the context of vaccines, the methodology and recommendations are in principle also applicable for other medicinal products, but additional complexities would then have to be considered. We will not discuss the use of OE analyses for sequential monitoring, which has been described elsewhere.[11,12]

BASIC PRINCIPLE

Observe-to-expected analyses are generally used when a safety concern has been raised from such sources as literature reviews, medical reviews, disproportionate reporting,[1,5,7] or unexpected temporal relationship,[3,6] without clear knowledge about the causality or magnitude of the risk. OE analyses can help to monitor and provide insight into specified events by integrating medical evaluation and quantification of the unexpectedness of observing a given number of cases. Because of their first pass screening nature, routine signal detection methods developed for spontaneous report data usually use an arbitrary level of the medical dictionary[‡] and apply routine stratification independently of the specificities of the event or vaccine considered, leading to a risk of over-stratifying.[13] OE analyses are complementary to routine signal detection methods as they can combine spontaneously reported events to match medical conditions and apply ad hoc stratifications relevant to the event, vaccine, and/or population as considered.

*Correspondence to: O. Mahaux, Vaccine Clinical Safety and Pharmacovigilance, GSK Vaccines, Avenue Fleming 20, 1300 Wavre, Belgium. E-mail: Olivia.x.mahaux@gsk.com

[†]Prior presentation: These data were presented in part at the International Society for Pharmacoepidemiology—30th International Conference 2014 (Van Holle L, Mahaux O, Bauchau V. Beyond traditional "observed versus expected" analyses: a sensitivity analysis integrating uncertainties around reporting bias and background incidence rate. Pharmacoepidemiology and Drug Safety 2014; 23(S1):39).

[‡]Usually the Medical Dictionary for Regulatory Activities (MedDRA) and the preferred term level

In spontaneous reporting systems, adverse events are reported spontaneously, based on suspicion, not on actual causal association with the vaccine. Many adverse events occur after vaccination by pure coincidence, at a rate that would not be different if there had been no vaccination. The core principle of OE analyses is to estimate the expected number of these coincidental cases, under the null hypothesis of no association with the vaccine. Expected numbers are then compared with the number of cases actually reported.

CALCULATION OF THE EXPECTED NUMBER OF CASES

The number of cases of a particular event expected to occur by chance alone, within a particular risk period, is estimated based on background incidence rates for that event and total person-time at risk in the vaccinated population.

[**Expected number within the risk period**]
$$= [\text{Background incidence rate}]*[\text{Person-time at risk}] \tag{1}$$

Background incidence rate

The background incidence rate is the number of new cases occurring naturally in the population, expressed in person-time. Estimates of incidence rates for the event of interest are selected through literature reviews and/or database queries (e.g., observational or national health statistics databases). The background incidence rates should ideally be estimated from populations that have not been exposed to the vaccine of interest but that have similar demographic characteristics to the vaccinated population. The background (expected) incidence rates may need adjusting if the exposed population differs from the unexposed population from which the background incidence rate is calculated.

Formula (1) estimates the expected number of cases for an unstratified OE analysis using a reliable background incidence rate. When several incidence rates, relevant for the population under study, are available, the lowest estimated incidence rate can be used to improve the sensitivity of the OE analysis. However, if the lowest background incidence rate is considered an outlier, it should be discarded. Meta-analysis methods can be used to provide an overall weighted incidence rate.

Person-time at risk

The total person-time at risk reflects the cumulative time for all persons exposed to the vaccine during a risk period for which there is suspicion and/or medical plausibility that there is a vaccine-associated increased risk of experiencing the event.

In the simple case where the vaccine is administered with only one dose, the total person-time at risk is calculated by multiplying the number of persons vaccinated by the risk period. For vaccines scheduled with multiple doses, the calculations can be more complex, it is then important to assess whether there is a dose effect and whether the risk periods overlap.

Indeed, if the risk period is shorter than the average time window between two scheduled vaccine doses, the risk periods following each dose do not overlap. Assuming that the risk is identical after each dose (no dose effect), each dose contributes a fixed time at risk and can therefore be considered independently. The total person-time at risk can then be estimated by multiplying the risk period after a dose by the number of doses administered. In this case, no information about the compliance to each dose of the vaccination schedule is needed. In absence of other sources, the number of doses administered can be approximated by totality or a proportion of doses sold depending on the percentage of doses that has already been used (Box Example 1).

If the risk period is longer than the average time window between two vaccine doses, the risk periods overlap and the doses cannot be considered independently as it would overestimate the total time at risk by double counting. The risk period to consider for the first and the second dose should be limited to the average time between the two doses, while the risk period would be complete post dose 2 (this reasoning can be easily extended to three or more doses). To calculate the total person-time at risk, one will need the number of persons vaccinated and the average proportion of the individuals who received dose 1 and dose 2 (Figure 1). If these data are not available, some assumptions need to be chosen and documented; for example, the doses could be assumed to have been equally distributed, and this assumption may be very good or average depending on the time the vaccine is on the market and the compliance to the vaccination schedule (Box Example 2).

If a dose effect is suspected, a dose-specific analysis may be performed, where each dose is considered separately.
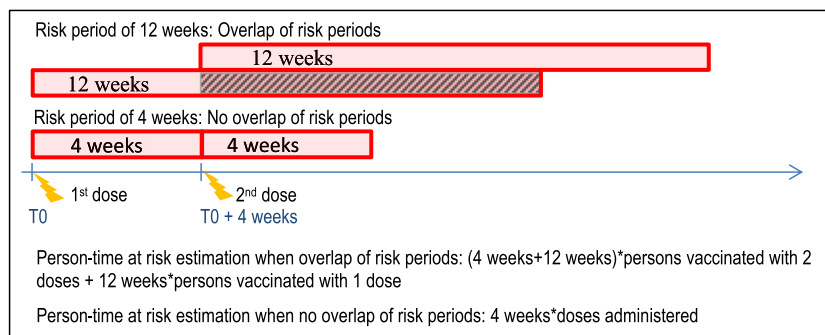
Figure 1. Examples to illustrate how person-time at risk is estimated

STRATIFICATION OF THE EXPECTED NUMBER OF CASES

The need for a stratified analysis may arise when background incidence rates differ between genders, age groups, geographical regions, or calendar time. As the distribution of the vaccinated population among the strata is rarely known, the demographic characteristics of the spontaneous cases may be used as a proxy of the demographic characteristics of the vaccinated population. However, this could lead to a biased demographic characterization of the vaccinated population, resulting in an overestimation of the expected number in strata characterized by a high reporting rate and an underestimation in strata characterized by a low reporting rate. As an alternative, stratified coverage data may be available from health authorities (as explained in Assumption 1) or may be estimated from observational databases when these include vaccination data up to the brand name.

The expected number of cases for each stratum is obtained by multiplying the incidence rate within the stratum by the corresponding person-time at risk. The overall expected number of cases is obtained by summing the expected numbers of cases over all strata (Formula (2)); however, it may be informative to look at the observed versus expected number within each stratum, as an excess risk might be specific to a particular stratum (Box Example 3).

[**Expected number within the risk period**]

$$= \sum_{s=1}^{N} [Incidence\ rate]_s * [\text{Person-time at risk within stratum } s]$$

$$(2)$$

ESTIMATION OF THE NUMBER OF OBSERVED CASES

The number of observed cases of a particular event following exposure to a given vaccine is available from spontaneous reporting systems. A thorough understanding of how the background incidence rates were estimated is crucial. Indeed, the method used to define cases for estimating the background incidence rate must be consistent with the one used during the medical assessment of spontaneous cases. This may be difficult because of differences between the dictionaries used in observational databases (e.g., International Classification of Diseases, Ninth Revision), from which epidemiological studies generating background incidence rates are increasingly performed, and MedDRA generally used in spontaneous reporting systems.

The medical evaluation of spontaneous cases should aim to classify the spontaneous cases into a confirmed diagnosis, an incomplete diagnosis due to lack of information or an alternative diagnosis. Only the "confirmed" cases are included in a best-case scenario.[§] For the uncertainty analysis around the observed number of cases, in addition to the "confirmed" cases, a proportion or the totality of the "incomplete" cases can be added for a mid-case or worst-case scenario,[§] respectively. On the other hand, cases should never be excluded based on causality assessment as it would bias downwards the observed count in contradiction with the null hypothesis.

Cases with a time-to-onset of the first symptoms of the event falling within the risk period that is used to calculate the expected number of cases are taken into account in a best-case scenario.[§] Cases with missing time-to-onset data may have occurred within the risk period considered and, thus, should be additionally used in an uncertainty analysis, in proportion to those in the risk period of interest for a mid-case scenario,[§] and in totality for a worst-case scenario.[§]

THE OBSERVED-TO-EXPECTED MEASURES

The OE analysis compares the observed and expected numbers of cases. This may be expressed as the ratio

---

[§]Best-case scenario refers to a best-case scenario analysis for the safety profile of the vaccine. The same logic applies for the mid-case and worst-case scenarios.

of the observed over the expected. An OE ratio of one means that the observed number of cases equals the expected number of cases, as stated by the null hypothesis. If the OE ratio is greater than one, then the observed is higher than the expected signaling an excess of risk.

The statistical uncertainty will often be driven by the observed number of cases, which is often small (rare events). To deal with this statistical uncertainty around the total number of cases observed over the risk period of interest, a 95% Poisson exact confidence interval (95%CI) can be calculated.[14] If the 95%CI lower limit is higher than the estimated expected number, the observed number is considered significantly higher than expected at a 95%CI. If the observed value is higher than the expected, with the 95%CI lower limit lower than the estimated expected value, the observed is considered higher than expected but not significantly at a 95% confidence level. The same logic applies when we consider the OE ratio: if the lower limit of the 95%CI of the OE ratio is greater than one, the observed value is considered significantly higher than expected.

ASSUMPTIONS

The OE analysis is based on a number of assumptions which, when violated, may generate biased estimates. We will first describe these assumptions and then describe a general method to address uncertainty generated by the deviations from assumptions.

Assumption 1: The number of doses administered to the population is known.

In some contexts, the authorities may make coverage data available (e.g., human papilloma virus mass vaccination data from Public Health England).[15] Additionally, coverage data may be reported by demographic characteristics or dose schedule (only one, two, or three doses), allowing more accurate calculations of the expected number of cases.

However, specific data on exposure are often lacking, and sales data are used as a proxy. Generally, not all doses sold are administered making sales data unreliable. Additionally, there is a lag between the sale and administration of a vaccine to the population. Depending on vaccine type (seasonal, pandemic, mass-vaccination, or routine) and the number of years on the market, the percentage of doses sold that are actually used may vary from less than 50% (as for the H1N1 pandemic)[16] to a theoretical value of 100%. The higher this percentage, the more sensitive

the OE analysis is when sales data are used as proxy of the doses administered.

Assumption 2: All cases presenting the event of interest after immunization are spontaneously reported.

Spontaneously reported events represent only a fraction of the events actually occurring after immunization. This so-called under-reporting[17] is dependent on the risk period considered, as discussed in Assumption 5. Under-reporting is also dependent on the plausibility of the event being causally associated with the vaccination. Other factors, such as the severity of the event, media coverage on the potential association between the vaccine and the event, public awareness, or the presence of the event in the label, also affect the extent of under-reporting. Under-reporting in vaccines spontaneous reporting systems varies and has been estimated for serious events at between 19% and 50%,[18,19] meaning that between 81% and 50% of the adverse events occurring after vaccination are being reported.

Observed-to-expected analyses generally focus on serious adverse events, often covered by the media (e.g., Guillain-Barre syndrome, Intussusception, Sudden infant death) and for which a potential causal association has been discussed in the literature. These events tend to be better reported particularly when they occur within a short time period after immunization. Nevertheless, the assumption that all cases are reported tends to lower the sensitivity of most OE analyses. Over-reporting (more cases reported than the number of cases actually occurring in the vaccinated population) may be observed following extensive media coverage and public awareness, such that an increased number of cases with similar symptoms are reported (over-diagnosing). Over-reporting may also occur because of multiple reports of the same case, where a lack of information makes it difficult to detect and delete duplicates.

Assumption 3: The background incidence rate in the vaccinated population is the same as the background incidence rate in the population used to calculate the expected.

The choice of the most relevant background incidence rate from multiple sources in the literature, characterized by different study designs and populations with different demographic characteristics, is often difficult but can dramatically impact the conclusions of the OE analysis. For some events, estimated background rates can differ by as much as a factor of 10 between literature sources, especially for rare events where the level of uncertainty is high.[20,21]

In addition, the "healthy vaccinee effect" occurs because of vaccinated individuals being on average healthier than unvaccinated individuals.[22] Using background incidence rates measured from the overall population could overestimate the background incidence rate of the vaccinated population. Consequently, it may lower the sensitivity of the OE analysis.

Assumption 4: The population on which the background incidence was measured is not exposed to the vaccine of interest.

Background rates on unexposed individuals are seldom available and often concern different geographical regions and/or time periods to those of the exposed population, resulting in geographical and/or secular trends. When the occurrence of vaccine exposure in the population used for measuring background incidence is non-negligible and when there is a non-negligible increased risk of experiencing the event because of exposure, the OE ratios may be biased towards one.[23]

Assumption 5: The risk period considered focuses on the time period for which an excess of risk occurs in case of causal association.

The risk period must correspond to the exact period of increased vaccine-associated risk. Overestimation of the risk period may dilute the excess of cases with the event by including periods beyond and/or before the true risk period, during which the vaccine did not generate extra risk for the event. When the risk period is underestimated, the sensitivity is also reduced because it is more difficult to reach statistical significance. Additionally, events occurring a long time after vaccination are less likely to be spontaneously reported than events occurring shortly after vaccination, especially if they are expected, common, or mild.[17] Consequently, a long risk period may include a period characterized by considerable under-reporting of the event, reducing the sensitivity of the analysis. Where no clear risk period for the event of interest is defined, the cumulative distribution of the OE ratio for each day over the whole time window can be used. This would allow potential sub-periods to be detected, where the number of observed cases is higher than expected.

UNCERTAINTY ANALYSES

Providing a single OE ratio estimate is not likely to be sufficient as the qualitative conclusion of the OE ratio could be reversed depending on how violated the above assumptions are. An uncertainty analysis should determine how much uncertainty would be needed to alter the qualitative conclusion.[24]

Most uncertainty analyses for OE consider only a limited number of values when accounting for the main sources of uncertainty (e.g., the lowest and highest published incidence rates, and 0–25% for under-reporting). These values remain arbitrary and may be subject to criticism.
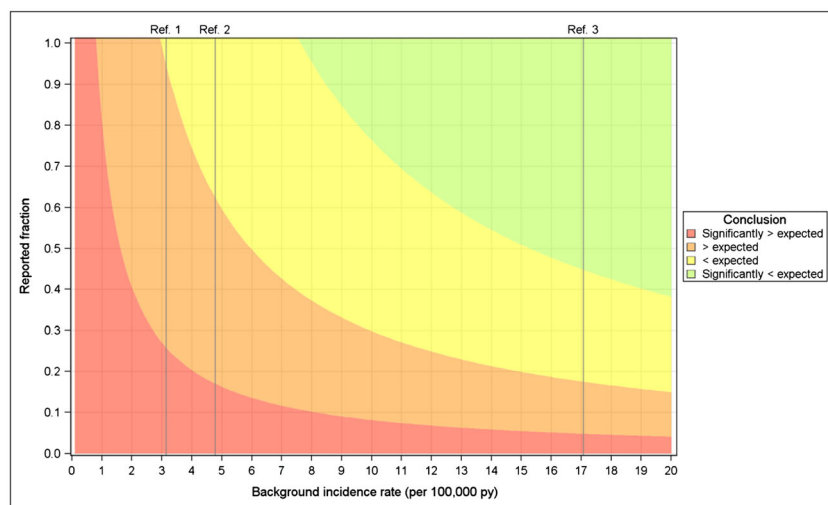

Figure 2. Heat map of the observed-to-expected analysis conclusion in the parameter plane defined by the background incidence rate and the reported fraction. Footnote. Figure 2, drawn from a theoretical example, shows that if Ref. 1 is the correct background incidence rate, the number of cases observed is lower than the number expected only if more than 95% of the cases occurring in the time window at risk were reported. If we take the background incidence rates Ref. 2 or Ref. 3, the number of cases observed is lower than expected only if, respectively, more than 62% or more than 18% of the cases occurring in the time window at risk are reported. Depending on how plausible these values are, an independent reviewer may draw his own conclusions. In most cases, there is no reason to consider that there is a protective effect of the vaccination, so having an observed reporting rate significantly lower than the expected may be an indicator of the range of reported fraction
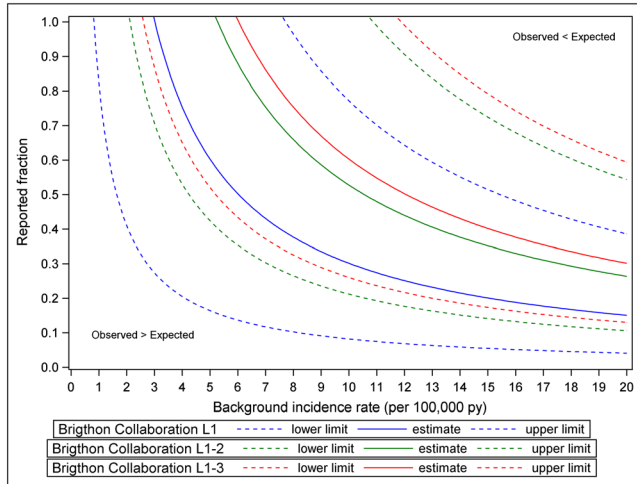
Figure 3. Observed-to-expected (OE) analysis conclusions depending on different scenarios for the reported fraction, the background incidence rate and the case confirmation level.
Footnote. Figure 3 shows the different OE conclusions for the complete range of reported fraction, a specific range of background incidence rate and three scenarios of case confirmation levels. However, this could apply to any other, or a combination of, uncertainty parameters considered as having a significant impact on the conclusion of the OE analysis

As an example of how to better account for uncertainties, we developed a visual framework that determines whether the observed number of events is (significantly) higher or lower than the expected number for simulated values of two sources of uncertainties around the expected. We illustrated this with an example (Figure 2) considering background incidence rates covering the range of estimates from the literature and under-reporting rates from 100% to 0% (equivalent to a reported fraction of zero to one). This visual framework enables independent reviewers such as regulatory authorities to draw their conclusions by making their own assumptions about two sources of uncertainty.

When additional sources of uncertainties are deemed to be important then the visualization can be adapted to include these additional uncertainties as illustrated in Figure 3 where the additional uncertainty around case confirmation (i.e., around the observed number of cases) was included in the visualization. This illustrates how additional sources of uncertainties could be incorporated.

Box. Examples of calculation of the expected number of cases for a theoretical event of interest

Example 1: 3,000,000 doses of vaccine X sold worldwide.
Increased *risk* of event Y within 30 days post immunization, whatever the dose.
Recommended *vaccination schedule* is three doses at 2, 4, and 6 months of age.
Assumptions: there is no dose effect and all 3,000,000 doses have been administered.
The *risk periods* following each dose *do not overlap*.
The *person-time at risk*: $3,000,000 * 30$ [person-days] or $3,000,000 * 30/365.2425 * 1/100,000 = 2.46$ [100,000 person-years].
Background incidence rate for event Y is 4.8 cases per 100,000 person-years (measured on unvaccinated population sharing similar demographic characteristics with the exposed population)
The *expected number of cases* of event Y: $2.46 * 4.8 = 11.8$.

Example 2: Same vaccine X as example 1 but increased *risk* of event Y within 90 days.
Assumptions: 100% compliance to the vaccination schedule, there is no dose effect and all 3,000,000 doses have been administered.
The *risk periods overlap* due to the vaccination schedule.
Risk periods to calculate the expected: 60 days post dose 1 and dose 2 and 90 days post dose 3.
The *person-time at risk* is $1,000,000 * 60 + 1,000,000 * 60 + 1,000,000 * 90$ [person-days] or $210,000,000/365.2425*1/100,000 = 5.75$ [100,000 person-years].
Same *background incidence rate* as for example 1: 4.8 cases per 100,000 person-years.
The *expected number of cases* would be $5.75 * 4.8 = 27.6$.

Example 3: Same as Example 2 but only females between 10 and 40 years are exposed to vaccine X.

Age group (years)	Stratified background incidence rate for females and event Y (per 100,000 person-years)	Coverage	Person-time at risk (100,000 person-years)	Expected number of cases of event Y
[10–25[	4.5	80%	$5.75 * 0.8 = 4.6$	$4.5 * 4.6 = 20.7$
[25–40]	2.3	20%	$5.75*0.2 = 1.15$	$2.3 * 1.15 = 2.6$

The total *expected number of cases* of event Y: $20.7 + 2.6 = 23.3$.

CONCLUSION

OE analyses are useful to strengthen safety signals, especially when rapid conclusions about the safety of a vaccine are needed and when the event of interest is short term and acute. However, these analyses rely on a number of assumptions, and these assumptions must be clearly described. Their impact on the qualitative conclusion of the analysis should be investigated through uncertainty analyses.

CONFLICTS OF INTEREST

O. M., L. V. H. and V. B. are all employees of the GSK group of companies. L. V. H. and V. B. also own restricted shares in the GSK group of companies as part of their employee remuneration.

KEY POINTS

- Observed-to-expected analyses are used to refine safety signals, especially for vaccines and when the event of interest is short-term and acute.
- In OE analyses, the observed is calculated from a spontaneous reporting system and compared with the expected calculated, based on background incidence rates from independent sources.
- The core principle is to estimate the expected number of coincidental cases, under the null hypothesis of no association with the vaccine, and compare with the number of cases actually reported.
- Conclusions of OE analyses are relying on a lot of assumptions.
- Simulations can be conducted to estimate how assumptions' violations could affect the conclusions of the OE analysis. These simulations can be represented in a visual framework that enables independent reviewers to draw their own conclusions independently from original assumptions.

CONTRIBUTORSHIP

All authors meet ICMJE criteria, they have contributed to the development of OE methodology, as well as writing and approval of the manuscript.

REFERENCES

1. van Puijenbroek E, Bate A, Leufkens H, *et al.* A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf* 2002; **11**: 3–10.
2. Van Holle L, Bauchau V. Use of logistic regression to combine two causality criteria for signal detection in vaccine spontaneous report data. *Drug Saf* 2014; **37**: 1047–1057.
3. Van Holle L. Using time-to-onset for detecting safety signals in spontaneous reports of adverse events following immunization: a proof of concept study. *Pharmacoepidemiol Drug Saf* 2012; **21**: 603–610.
4. Caster O, Juhlin K, Watson S, Noren N. Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank—retrospective evaluation against emerging safety signals. *Drug Saf* 2014; **37**: 617–628.
5. Bate A, Lindquist M, Edwards R, *et al.* A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998; **54**: 315–321.
6. Noren N, Hopstadius J, Bate A, *et al.* Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Disc* 2010; **20**: 361–387.
7. Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am Stat* 1999; **53**: 177–190.
8. Strom B, Kimmel S. *Textbook of pharmacoepidemiology.* John Wiley & Sons, Ltd: Chichester, England, 2006.
9. Guideline on good pharmacovigilance practices (GVP) – Product- or Population-specific considerations I: Vaccines for prophylaxis against infectious diseases, 9 December 2013. EMA/488220/2012.
10. The European Network for Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 3). EMA/95098/2010, http://www.encepp. eu/standards_and_guidances (Accessed 10 October 2014).
11. Bryan P, Seabroke S, Davies C. H1N1 vaccine safety: real-time surveillance in the UK. *Lancet* 2010; **376**: 417–418.
12. Donegan K, Beau-Legdstrom R, King B, *et al.* Bivalent human papillomavirus vaccine and the risk of fatigue syndromes in girls in the UK. *Vaccine* 2013; **31**: 4961–4967.
13. Hopstadius J, Noren N, Bate A, Edwards R. Impact of stratification on adverse drug reaction surveillance. *Drug Saf* 2008; **31**: 1035–1048.
14. Garwood F. Fiducial limits for the poisson distribution. *Biometrika* 1936; **28**: 437–442.
15. Provisional annual data for routine cohort as submitted by PCTs for first, second and third dose vaccine coverage by 31 August 2012, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/244479/HPV_AnnualVaccineUptake 2011-2012.pdf (Accessed 15 July 2014)
16. Folkenberg M, Callreus T, Svanstrom H, *et al.* Spontaneous reporting of adverse events following immunisation against pandemic influenza in Denmark November 2009-March 2010. *Vaccine* 2011; **29**: 1180–1184.
17. Hazell L, Shakir S. Under-reporting of adverse drug reactions – a systematic review. *Drug Saf* 2006; **29**: 385–396.
18. Lévy-Bruhl D, Desenclos JC, Rebière I, Drucker J. Central demyelinating disorders and hepatitis B vaccination: a risk–benefit approach for pre-adolescent vaccination in France. *Vaccine* 2002; **20**: 2065–2071.
19. Prevots R, Sutter R, Strebel P, *et al.* Completeness of reporting for paralytic poliomyelitis, United States, 1980 through 1991: implications for estimating the risk of vaccine-associated disease. *Arch Pediatr Adolesc Med* 1994; **148**: 479–485.

20. Sáez-Llorens X. A multi-country study of intussusception in children under 2 years of age in Latin America: analysis of prospective surveillance data. *BMC Gastroenterol* 2013; **13**: 95.
21. Bines JE, Liem NT, Justice FA, *et al.* Risk factors for intussusception in infants in Vietnam and Australia: adenovirus implicated, but not rotavirus. *J Pediatr* 2006; **149**: 452–460.
22. Fine P, Chen R. Confounding in studies of adverse reaction to vaccines. *Am J Epidemiol* 1992; **136**: 121–135.
23. Jones ME, Swerdlow AJ. Bias in the standardized mortality ratio when using general population rates to estimate expected number of deaths. *Am J Epidemiol* 1998; **148**: 1012–1017.
24. Rosenbaum PR. Safety in caution. *J Educ Stat* 1989; **14**: 169–173.