ORIGINAL RESEARCH ARTICLE

Drug Saf 2012; 35 (1): 61-78
0114-5916/12/0001-0061/$49.95/0

© 2012 Adis Data Information BV. All rights reserved.

Identifying Adverse Events of Vaccines Using a Bayesian Method of Medically Guided Information Sharing

Colin John Crooks,[1] *David Prieto-Merino*[2] and *Stephen J.W. Evans*[2]

1 Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
2 Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK

Abstract

Background: The detection of adverse events following immunization (AEFI) fundamentally depends on how these events are classified. Standard methods impose a choice between either grouping similar events together to gain power or splitting them into more specific definitions. We demonstrate a method of medically guided Bayesian information sharing that avoids grouping or splitting the data, and we further combine this with the standard epidemiological tools of stratification and multivariate regression.

Objective: The aim of this study was to assess the ability of a Bayesian hierarchical model to identify gastrointestinal AEFI in children, and then combine this with testing for effect modification and adjustments for confounding.

Study Design: Reporting odds ratios were calculated for each gastrointestinal AEFI and vaccine combination. After testing for effect modification, these were then re-estimated using multivariable logistic regression adjusting for age, sex, year and country of report. A medically guided hierarchy of AEFI terms was then derived to allow information sharing in a Bayesian model.

Setting: All spontaneous reports of AEFI in children under 18 years of age in the WHO VigiBase™ (Uppsala Monitoring Centre, Uppsala, Sweden) before June 2010. Reports with missing age were included in the main analysis in a separate category and excluded in a subsequent sensitivity analysis.

Exposures: The 15 most commonly prescribed childhood vaccinations, excluding influenza vaccines.

Main Outcome Measures: All gastrointestinal AEFI coded by WHO Adverse Reaction Terminology.

Results: A crude analysis identified 132 signals from 655 reported combinations of gastrointestinal AEFI. Adjusting for confounding by age, sex, year of report and country of report, where appropriate, reduced the number of signals identified to 88. The addition of a Bayesian hierarchical model identified four further signals and removed three. Effect modification by age and

sex was identified for six vaccines for the outcomes of vomiting, nausea, diarrhoea and salivary gland enlargement.

Conclusion: This study demonstrated a sequence of methods for routinely analysing spontaneous report databases that was easily understandable and reproducible. The combination of classical and Bayesian methods in this study help to focus the limited resources for hypothesis testing studies towards adverse events with the strongest support from the data.

Background

Serious adverse events following immunization (AEFI) can alter the important balance between the risks and benefits of performing routine vaccinations in healthy populations.[1,2] Their early detection is therefore vital in maintaining professional and public confidence in childhood vaccination programmes. Similarly, it is important to avoid inaccurate information that can result in health scares that reduce population coverage, lead to a loss of herd immunity and undermine vaccination programmes.[3] However, as many serious adverse events are rare and undetected in randomized controlled trials, the early signalling of potential associations, for further investigation as to whether the events are truly causal effects, requires a large worldwide spontaneous reporting system. This increases the power to detect rare AEFI, but does so by combining data from heterogeneous sources. Although such data sources might appear simple, data derived from many different countries are a consequence of different healthcare systems and vaccination policies that can also change over time. To minimize potential biases, confounders and interactions in the data, analyses need to be complex to model and adjust for this heterogeneity.

Furthermore, the methods chosen for categorizing events in these databases fundamentally alter the detection of potential signals of concern. If each specific adverse event code is assessed separately (i.e. splitting), this results in small numbers and low power for each comparison and also ignores useful information that might have been available from other related adverse events. If an effect is truly causal, it may be recorded using different but related terms and may cause several different related effects. Conversely, although grouping similar events together increases power, it can also lose more specific information about the precise types of events.

We propose that a Bayesian model of information sharing, as opposed to grouping, overcomes the difficulty of choosing between grouping or splitting data that is imposed by classical models. Berry and Berry[4] have shown that including a Bayesian hierarchy can be important in analysing adverse events, but the model they used was unadjusted, overly simplistic and only applied to clinical trial data. We therefore demonstrated a medically guided Bayesian hierarchy that is fully integrated with methods to assess for confounding and interactions, and that better optimizes the analysis of observational data, including spontaneous reporting, than previous methods. Each result is therefore based not only on the information in a specific vaccine adverse event, but also on the information from related associations, thereby maximizing the information extracted from the whole database.

Methods

Signal Detection

A signal is defined as an increase in adverse event reporting above that expected. Spontaneous reports are only a suspicion of an AEFI, not a proven association. Therefore, detecting an increase in spontaneous adverse event reporting (or signal detection), does not necessarily mean a true association has been found. Instead, it highlights a potential association for further investigation.[5]

© 2012 Adis Data Information BV. All rights reserved.

Database

The WHO global individual case safety reports database, VigiBase™, from the Uppsala Monitoring Centre (UMC), Uppsala, Sweden, is the largest spontaneous report database available worldwide.[6] A download from June 2010 was checked for duplicate reports, although before download the UMC uses its own software to detect and remove duplicates. Each variable was then assessed for its range, outliers and missing values. There were several possible combinations of vaccine codes and AEFI; therefore, this study was restricted to gastrointestinal AEFI. Apart from common AEFI, such as abdominal discomfort, diarrhoea and vomiting, few gastrointestinal AEFI have been previously reported from trials.[6]

Study Design and Population

The study design can be compared to a case-control study[7] and uses the Reporting Odds Ratio (ROR) as its effect measure ($ROR = \frac{a/c}{b/d}$ [see table I]). We analysed the data at the level of individual reports (by collapsing on the report identification for each vaccine and AEFI combination); therefore, each report was included only once for each vaccine and AEFI combination. The ROR measure adjusts for variations in background levels of reporting and can be easily modelled using logistic regression. Exclusions were applied to ensure comparability between reports in the numerator and denominator of the ROR. First, adverse events following medications (drugs as opposed to vaccines) were excluded from the background reporting group, as medications will have been given for different indications than routine vaccinations and are therefore associated with different adverse events. Second, reports were restricted to children under 18 years of age, as adults will have been vaccinated for different indications to children and will experience different adverse events related to age differences and indication. Where the report had no recorded age, it remained in the analysis as a separate category and the effect of this was examined later in a sensitivity analysis. Third, reports with non-sequential dates were excluded.

Exposures

The vaccination exposures were defined as a WHO Anatomical Therapeutic Chemical (ATC) classification level 4[8] code, consistent with one of the 15 most commonly prescribed vaccinations (excluding influenza, as this is usually given to an adult population). This defined categories for childhood vaccinations as diphtheria, *Haemophilus influenzae* type b, meningococcal, pertussis, pneumococcal, tetanus, tuberculosis, typhoid, hepatitis, measles, mumps, poliomyelitis, rubella, varicella zoster and rotavirus diarrhoea vaccination. Reports with vaccine codes not in these categories remained in the background information if they were from children under 18 years of age or had missing ages.

Outcomes and Hierarchy

Outcomes were defined as a WHO Adverse Reaction Terminology (WHO-ART)[9] preferred

Table I. Collapsing a spontaneous report database to calculate a reporting odds ratio (ROR)[a]

		ROR = (a∗d)/(c∗b)		
Selection of AEFI of interest	WHO-ART preferred terms	Drug of interest	All other drugs	
		drug 1	drug 2 →	drug K
AEFI of interest	AEFI-1	a	b	
All other AEFI	AEFI-2 ↓ AEFI-H	c	d	

a The analysis was performed by collapsing on individual reports; therefore, if a report recorded more than one vaccine, including the drug of interest, it was only counted in the numerator and not counted twice in the denominator.

AEFI = adverse events following immunization; **WHO-ART** = WHO Adverse Reactions Terminology.

© 2012 Adis Data Information BV. All rights reserved.

term with a system organ class (SOC) of gastrointestinal or hepatopancreatobiliary systems, respectively. Adverse events that occurred before the date of vaccination, or after the report's inclusion in the database, were excluded. Adverse events not coded to a gastrointestinal or hepatopancreatobiliary system were included in the analysis as background reports.

A hierarchy for the WHO-ART codes was constructed by a gastroenterologist (CC) [figures 1–3]. The highest hierarchical level grouped the luminal organs together as one system, and the liver, pancreas and biliary organs as a different system following the WHO-ART hierarchy. However, this makes assumptions of similarity between a wide range of codes. Therefore, a second hierarchical level was introduced, based on the specific organ involved in the AEFI. Lower levels of hierarchy were then added to group codes for similar physiological processes that differed only by severity, or to group codes with similar signs and symptoms. This multilevel hierarchy then ensures that the strongest effect of information sharing occurs at the lowest levels, where the codes are most similar, and the weakest information sharing occurs at the highest system group level, where the codes are only broadly similar.

Potential Confounders

To minimize strata with small numbers, potential confounders of age and year were categorized into quintiles of numbers of reports with missing value categories. Sex was categorized as male, female or missing. Age and sex were assessed as possible effect modifiers.[10,11] Countries that had <500 reports were combined regionally to avoid small numbers in strata. These regional groups were Africa, Europe, previous Soviet states, Middle East, South and East Asia, and South and Central America.

Year of event was a proxy for when changes in awareness about specific AEFI and vaccinations occurred. Two specific examples were assessed – one in 1995, when there was increased concern about Crohn's disease following measles vaccine being combined with mumps and rubella (further studies have found no such association[12]), and the second in 1999, when concerns about rotavirus vaccine and intussusception were corroborated, resulting in withdrawal of rotavirus vaccine.[13] Year was therefore examined as an effect modifier for these two examples to determine if the ROR estimates varied following these events.

Crude Analysis

The initial analysis was carried out using the Stata® 10 statistical package (StataCorp LP, College Station, TX, USA).[14] The database was collapsed to individual reports, for each vaccine and AEFI combination, and the ROR was calculated. A sensitivity analysis was performed by excluding non-study vaccine code reports from the background information and repeating the analysis. A second sensitivity analysis was conducted excluding reports from associations with an ROR >100.[5] This was because estimated measures of disproportionality will always depend on other associations in the database. For example, if a vaccine has a particularly strong association with an AEFI it will over-contribute to the denominator for that vaccine when calculating other AEFI and bias estimates towards the null.[7] The converse of a vaccine having a strong protective effect for an AEFI could theoretically cause overestimation of other AEFI for that drug, but this seems unlikely when most AEFI are rare following vaccines. To reduce these biases, AEFI that are strongly associated with a vaccine should be minimized or excluded from calculation of other AEFI for that vaccine.[15]

Stratified Analysis

A stratified analysis was performed separately for each potential confounder. A Chi-squared test for homogeneity was calculated across sexes and age groups for each combination of vaccine and AEFI, and adjusted with a Bonferroni correction by the number of vaccine-AEFI combinations tested.[16] Combinations with strong evidence against the null hypothesis, that variations of stratum RORs were due to chance, were assessed manually for interactions. The two associations with *a priori* effect modification by year were also assessed. If there were no important interactions,

© 2012 Adis Data Information BV. All rights reserved.

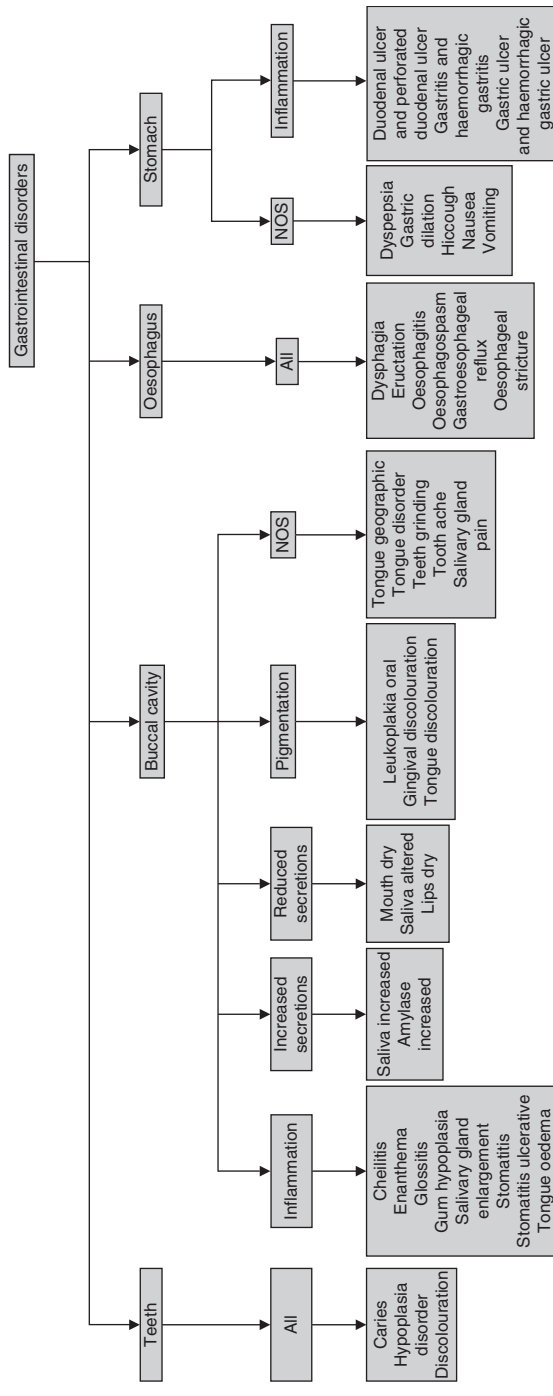

Fig. 1. Hierarchical grouping of gastrointestinal codes (1). **NOS** = not otherwise specified.

© 2012 Adis Data Information BV. All rights reserved.

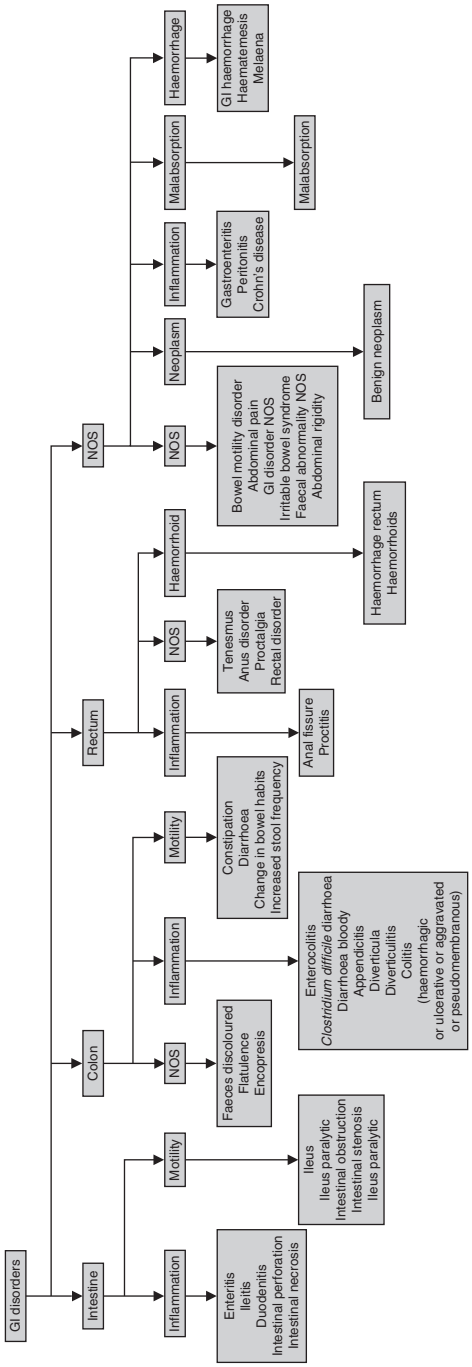

Fig. 2. Hierarchical grouping of gastrointestinal codes (2). **GI** = gastrointestinal; **NOS** = not otherwise specified.

the stratum estimates were combined using the Mantel-Haenszel method.[17,18] Crude and adjusted RORs were compared. If there was an important difference that altered the effect by around 10%, the co-variable was judged to be a confounder and used in the multivariable analysis. To assess the effect of including the missing values in their own strata, the reports with missing data for each analysis were excluded and the stratified analysis repeated.

Multivariable Analysis

Co-variables that were confounders in the stratified analysis were then used to construct separate logistic regression models for each vaccine and AEFI combination. Age group and sex were included as *a priori* confounders in the model. Then, country and year were included in order of magnitude of effect on the ROR in the stratified analysis. If there was a change in ROR by around 10% following the addition of a co-variable, it was judged to be a confounder and remained in the final model.

Bayesian Hierarchical Analysis

The adjusted log odds ratios and standard errors obtained from the multivariate analysis were included in a Bayesian hierarchical model, and re-estimated using Monte Carlo Markov Chain (MCMC) methods with the Gibbs sampler in the WinBUGS 1.4 package.[19] MCMC sampling methods involved sampling from a model's distribution, conditional on the observed data. The Gibbs sampler allows more complex models to be sampled, provided relationships between each adverse effect term are pre-defined by the hierarchical model. Higher levels in the hierarchy made more general assumptions of exchangeability than lower levels; therefore, the model was initially assessed with only the lowest level. Each additional level was subsequently added to the model and the Deviance Information Criterion (DIC) statistic was used to compare models. The DIC is a measure of the fit of the model to the data. The model with the lowest DIC was selected as the best fitting final model.

© 2012 Adis Data Information BV. All rights reserved.

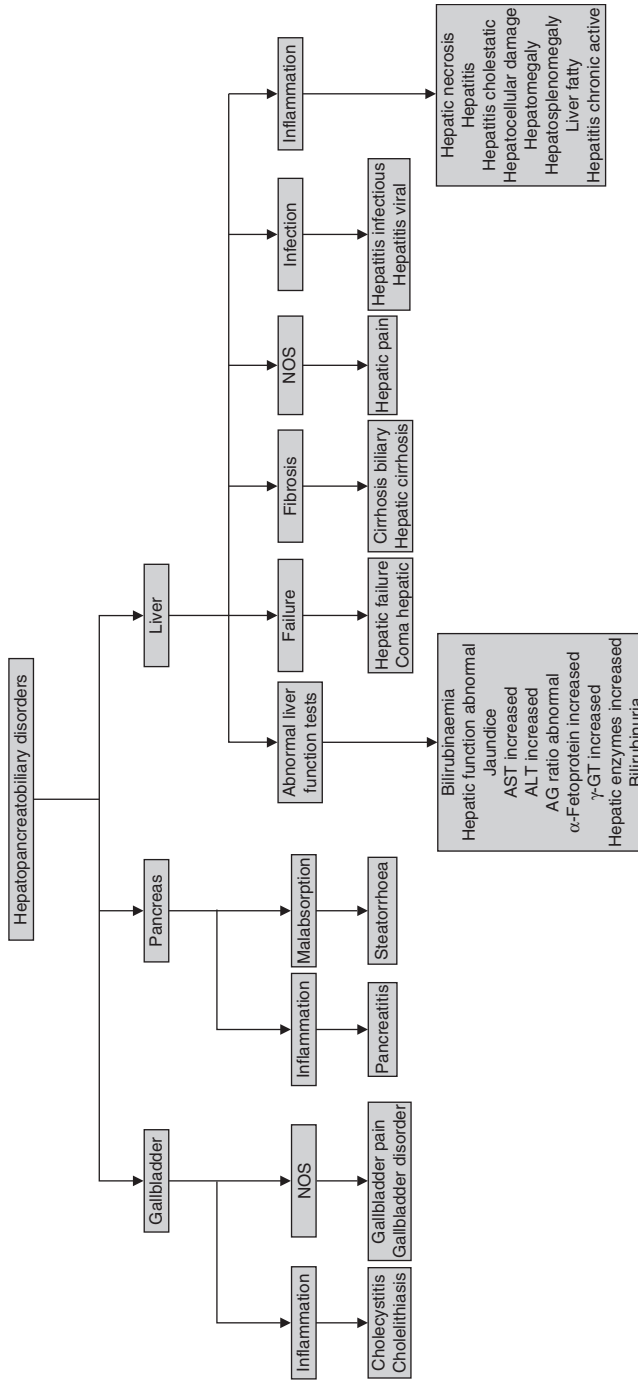

Fig. 3. Hierarchical grouping of hepatopancreatobiliary codes. **γ-GT** = γ-glutamate transferase; **AG** = albumin : globulin ratio; **ALT** = alanine aminotransferase; **AST** = aspartate aminotransferase; **NOS** = not otherwise specified.

© 2012 Adis Data Information BV. All rights reserved.

An important part of constructing a Bayesian model is deciding on the likely distribution for the estimated values based on prior knowledge. This information is called a 'prior'. Within the hierarchical model the priors are required for the highest (system) level means and for the variance of the group means at each hierarchical level. As this study was designed to detect potentially unknown signals, we chose priors that would allow a wide range of values with a slight weighting towards the null. The priors for the system group means on the log scale were normal distributions (mean = 0; variance = 1 000 000). The priors for the group variances on the log scale were the reciprocals of a gamma distribution (shape = 1; scale = 3). A threshold of the 99% credible or confidence interval (CI) excluding the null (Bayesian and frequentist models, respectively) was selected for signal detection. This was a compromise between type 1 and 2 errors and is similar to the method currently used by the WHO.[20]

Results

Study Population

A total of 372 206 reports were extracted with a WHO-ATC vaccine code (J07-). No duplicate reports were detected. Figure 4 shows the selected cases and exclusions. Reports with missing ages remained in the study sample.

Exposure

The numbers of reports for each vaccine are shown in table II. In total, 813 reports (0.3% of included reports) were not coded to a specific vaccine group, only to vaccines in general. These were included in the analysis as background reports.

Outcome

No reports had a missing WHO-ART code. A total of 492 reports (0.2% of included reports) were coded to a term that was not currently ac-

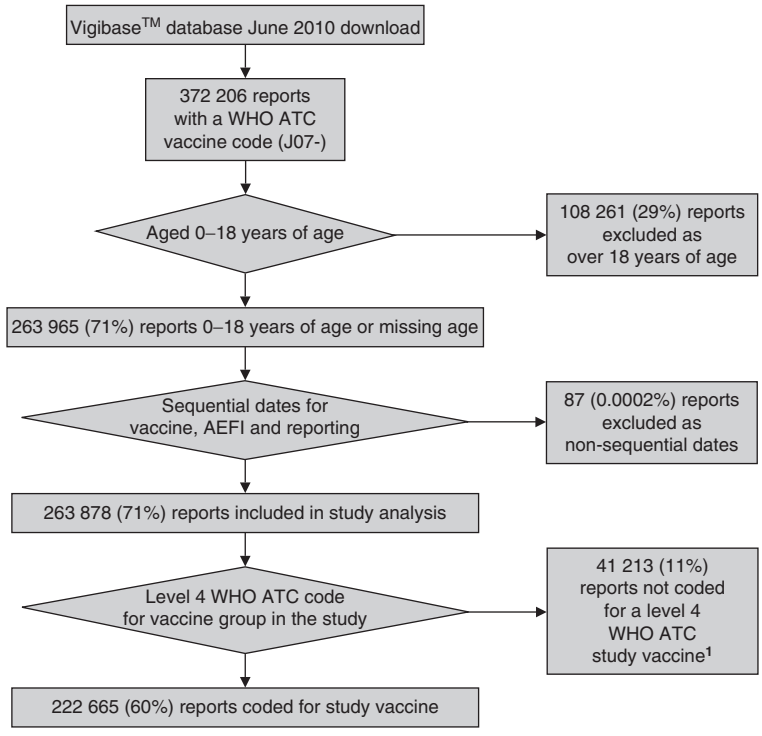

Fig. 4. Flowchart of selected and excluded case reports. **1** Included in study analysis as part of background reporting. **AEFI** = adverse effects after immunization; **ATC** = Anatomical Therapeutic Chemical classification.

© 2012 Adis Data Information BV. All rights reserved.

Table II. Number of childhood reports with each vaccine code[a]

Vaccine	Number of reports with each level 4 ATC vaccine code in order of magnitude
Mumps	196
Rubella	813
Typhoid	1 293
Rotavirus diarrhoea	1 451
Diphtheria	1 921
Tuberculosis	6 683
Pneumococcal	9 972
Varicella zoster	12 292
Tetanus	20 107
Meningococcal	22 992
Hepatitis vaccines	29 534
Non-study vaccine code	40 392
Measles	49 535
Haemophilus influenzae type b	50 737
Poliomyelitis	52 065
Pertussis	77 144

a A single report can be coded to more than one vaccine code; therefore, percentages are not shown. Combination codes will appear once according to the ATC coding system; for example, measles, mumps and rubella are coded under 'Measles'.

ATC = Anatomical Therapeutic Chemical classification.

cepted or was awaiting acceptance by the WHO. These were included as part of the background reporting, and 315 reports were coded to a study vaccine. Gastrointestinal AEFI were coded in 14 319 reports (5% of the included reports).

Potential Confounders

Overall, 16% of reports had no age recorded. The median age of reports with age recorded was 16 months and the distribution was skewed to the right (interquartile range 6–60 months). The quintile age groups were 0–3, 4–11, 12–35, 36–95 and 96–216 months. Forty-seven percent of reports were male, 47% female and 6% had no sex recorded. Countries with <500 reports were grouped as previously described. No reports had missing data on country and 11% of reports had missing information on year of report. The grouped year categories were <1990, 1990–4, 1995–9, 2000–4, 2005–present, and missing. Missing values were assumed to be missing at random, conditional on the vaccine type.

Crude Analysis

Estimation of crude RORs and 99% CIs generated 132 signals of 655 reported combinations. Restricting the analysis to the 222 665 reports coded to a study vaccine did not remove or identify any additional signals. Excluding reports with strong signals (13 signals mostly related to rotavirus vaccination) from the analysis also did not alter any signals.

Stratified Analysis

Effect Modification and Adjusted Reporting Odds Ratios

Five AEFI vaccine combinations were judged to have effect modification by age and one combination by sex (tables III and IV). For the combination of rotavirus and intestinal obstruction, age, sex and country adjusted $\log_e$ (ROR) before 1999 was 7.04 (99% CI 5.66, 8.42), during 1999 was 5.01 (99% CI 3.33, 6.68) and after 1999 was 5.72 (99% CI 4.31, 7.13). Interpretation of the signal did not change and year was not judged to be an effect modifier. It was only possible to calculate an ROR for measles and Crohn's disease after 1995 as there were no AEFI and Crohn's disease reports in the denominator before 1995.

To keep the model simple, the effect modifications for nausea, vomiting and diarrhoea identified above (tables III and IV) were not judged serious enough to be included in the multivariable analysis. However, it would be valuable to investigate these effect modifications in future studies. Summary RORs adjusting for each of the co-variables were then calculated, and all co-variables were judged to be confounders. A sensitivity analysis excluding missing data strata did not alter the signals detected and therefore the final analysis was conducted on the complete dataset that included strata for missing data.

Multivariable Analysis

A logistic regression model was constructed for each combination of vaccine and AEFI. For 158 of the combinations, the best model had sex and age included; for 56, the addition of year was necessary; for 91, the addition of country; and for

© 2012 Adis Data Information BV. All rights reserved.

Table III. Effect modification of adverse events following immunization (AEFI) by age group

AEFI	Vaccine	Age group	ROR (99% CI)
Vomiting	Diphtheria	0–3 mo	1.78 (1.37, 2.32)
		4–11 mo	1.31 (1.12, 1.52)
		12–35 mo	0.75 (0.55, 1.00)
		36–95 mo	0.16 (0.12, 0.23)
		8–18 y	0.61 (0.34, 1.11)
		Missing	0.60 (0.38, 0.95)
	Pneumococcal	0–3 mo	1.75 (1.60, 1.91)
		4–11 mo	1.42 (1.32, 1.52)
		12–35 mo	0.96 (0.87, 1.06)
		36–95 mo	0.65 (0.53, 0.79)
		8–18 y	0.56 (0.43, 0.74)
		Missing	0.51 (0.45, 0.58)
Salivation	Pertussis	0–3 mo	0.74 (0.60, 0.92)
		4–11 mo	0.58 (0.42, 0.80)
		12–35 mo	2.09 (1.74, 2.50)
		36–95 mo	3.88 (2.74, 5.49)
		8–18 y	0.56 (0.20, 1.55)
		Missing	0.31 (0.11, 0.85)
Diarrhoea	Tetanus	0–3 mo	1.82 (1.43, 2.30)
		4–11 mo	0.69 (0.56, 0.86)
		12–35 mo	0.63 (0.50, 0.80)
		36–95 mo	0.32 (0.26, 0.41)
		8–18 y	0.50 (0.42, 0.60)
		Missing	0.30 (0.24, 0.36)
Nausea	Hepatitis	0–3 mo	0.81 (0.50, 1.31)
		4–11 mo	2.19 (1.56, 3.05)
		12–35 mo	2.01 (1.53, 2.66)
		36–95 mo	1.27 (1.08, 1.49)
		8–18 y	0.67 (0.64, 0.70)
		Missing	2.92 (2.76, 3.10)

ROR = reporting odds ratio.

350 combinations, the best model required the addition of both year and country. Across all these models, 88 signals were identified in the dataset compared with the 132 signals from the unadjusted analysis.

Bayesian Hierarchical Modelling

Adjusted RORs and standard errors from the multivariable analysis were re-estimated in a Bayesian hierarchical model. The Markov chains were sampled for 200 000 iterations after a burn-in of 20 000 iterations, which, after thinning, provided a sample for analysis of 40 000 iterations. The full three-level hierarchical models were consistently the best fitting with the lowest DIC, although most improvement in model fit came from the two lowest hierarchy levels.

Supplementary figures S1–S6 (see Supplemental Digital Content, http://links.adisonline.com/DSZ/A57) show crude and adjusted RORs from the multivariable and Bayesian hierarchical models for detected signals only. No signals were identified for varicella or tuberculosis vaccines. The associations of three RORs from the multivariate analysis were reduced by the Bayesian analysis and lost their signals (table V). Conversely, four RORs from the multivariate analysis were increased and generated new signals (table VI).

A Worked Example Using Vaccination Against Hepatitis

Figure 5 shows the crude RORs for all reported adverse events following vaccination with a hepatitis vaccine. Compare this to figure 6, which shows the RORs adjusted for confounders by logistic regression. Overall, there is an adjustment of many of the signal estimates towards the null, showing that many of the generated crude signals can be explained by differences in the confounders. The aim of our study is to allow information sharing between similar adverse event codes. Within anatomical groups we want to allow some weak information sharing, so that if a vaccine has an adverse effect on the liver it will slightly increase the probability of observing other adverse events related to the liver. To help understand this, figure 7 shows only the adjusted RORs for adverse events from the liver anatomical grouping.

Table IV. Effect modication of adverse events following immunization (AEFI) by sex[a]

Sex	ROR (99% CI)
Male	0.76 (0.74, 0.80)
Female	1.42 (1.37, 1.50)
Missing	4.89 (3.59, 6.70)

a AEFI = vomiting; vaccine = *Haemophilus influenzae* type b.

CI = confidence interval; **ROR** = reporting odds ratio.

© 2012 Adis Data Information BV. All rights reserved.

Table V. Multivariate analysis signals removed by Bayesian analysis

Vaccine	AEFI	Multivariate analysis	Posterior distribution
		adjusted $\log_e$ [ROR] (99% confidence interval)	median $\log_e$ [ROR] (99% credible interval)
Tetanus	Gingivitis	1.58 (0.21, 2.95)	1.26 (–0.03, 2.57)
Measles	Tooth discolouration	2.36 (0.03, 4.69)	2.17 (–0.06, 4.44)
Poliomyelitis	Increased stool frequency	4.26 (0.54, 7.97)	2.76 (–0.33, 6.52)

AEFI = adverse events following immunization; **ROR** = reporting odds ratio.

To indicate the information that will be shared within the group, the black vertical block is added, representing the median ROR for the anatomical groups. Within the Bayesian model, this is the centre of a distribution that has a large range of uncertainty. To show the information that we want to be shared more strongly between events with similar symptoms or similar pathological processes, we have then added the grey vertical blocks to mark the median ROR values from the lower level groupings.

Now, when we use the Bayesian modelling to allow these individual RORs to share information within groups, the individual medians are pulled towards the group median (vertical grey blocks), with the RORs with the most uncertainty (the boxes with the widest CIs) pulled towards the group median the most. Figure 8a shows these Bayesian RORs for the liver grouping of adverse events following hepatitis vaccination. The CIs (technically now called credible intervals as they are from a Bayesian model) narrow as they are gaining information from their neighbours. The same process happens to a lesser extent when the information is shared within the larger grouping of adverse events from the liver.

Figure 8b shows the RORs from abnormal liver function tests at each stage of the process. Most of the RORs in the group of abnormal liver function tests are based on a large number of reports and therefore adjusting for confounders and information sharing has only a minimal effect on the estimates. However, for adverse events of a raised α-fetoprotein, there is a large amount of uncertainty in the crude estimate (shown by the wide white horizontal bar). After adjusting for confounders, this estimate shifts towards being a signal (light grey horizontal bar), and then when it shares the information from the other strong signals in the group, a signal is generated (dark-grey horizontal bar). Clearly, this is an extreme example presented to aid understanding of the methods, and the large effect is due to the crude ROR for a raised α-fetoprotein having a large degree of uncertainty whilst occurring in a group with strong signals present.

Discussion

This study demonstrated that almost one-third of gastrointestinal adverse events reported from children in VigiBase™ seem to be confounded by a combination of age, sex, year and country of report. Therefore, adjusting for these confounders had the largest impact, of all the applied methods, on the detection of signals for adverse

Table VI. Signals from the Bayesian analysis not detected by the multivariate analysis

Vaccine	AEFI	Multivariate analysis	Posterior distribution
		adjusted $\log_e$ [ROR] (99% confidence interval)	median $\log_e$ [ROR] (99% credible interval)
Hepatitis	α-Fetoprotein increased	3.84 (–0.30, 7.98)	3.11 (0.74, 5.68)
Hepatitis	Hepatocellular damage	0.95 (–0.01, 1.92)	1.05 (0.12, 1.98)
Hepatitis	Liver fatty	1.26 (–0.21, 2.72)	1.38 (0.03, 2.71)
Mumps	Stomatitis aphthous	2.83 (–0.16, 5.83)	2.68 (0.17, 5.26)

AEFI = adverse events following immunization; **ROR** = reporting odds ratio.

© 2012 Adis Data Information BV. All rights reserved.

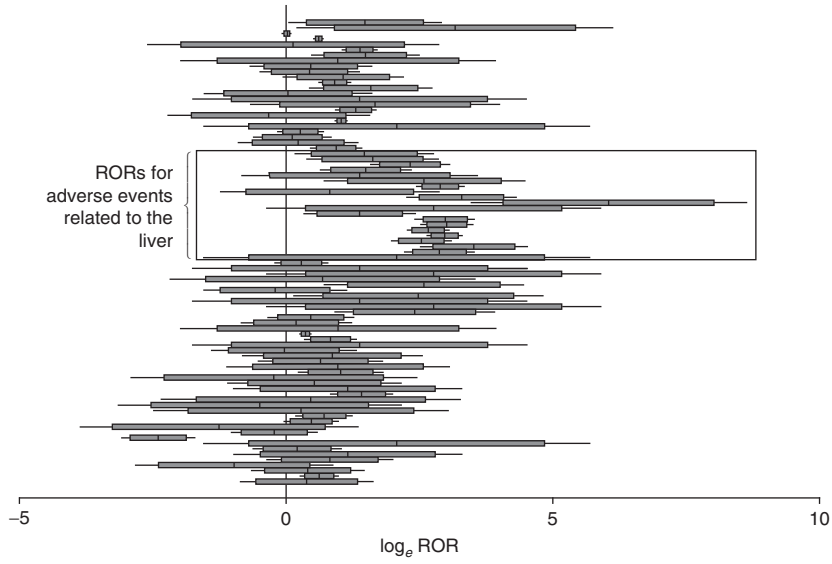

Fig. 5. Crude reporting odds ratios (RORs) for adverse events following hepatitis vaccination (boxes = 95% CIs; lines = 99% CIs).

events from a spontaneous report database. The addition of the Bayesian hierarchical model allowed information sharing and thus avoided the need to merge codes with small numbers into common groups, thereby preserving useful information. Consequently, combining classical and Bayesian methods will help to focus subsequent hypothesis testing studies towards those signals with the strongest support from the data, thus making the most efficient use of the limited resources available. In addition, testing for effect modification in this study identified potential interactions in

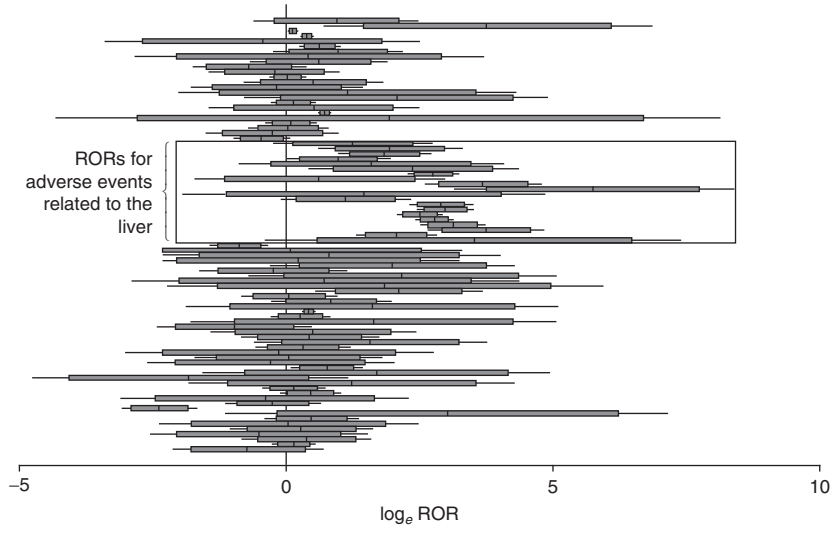

Fig. 6. Reporting odds ratios (RORs), adjusted by logistic regression, for adverse events following hepatitis vaccination (boxes = 95% CIs; lines = 99% CIs).

© 2012 Adis Data Information BV. All rights reserved.

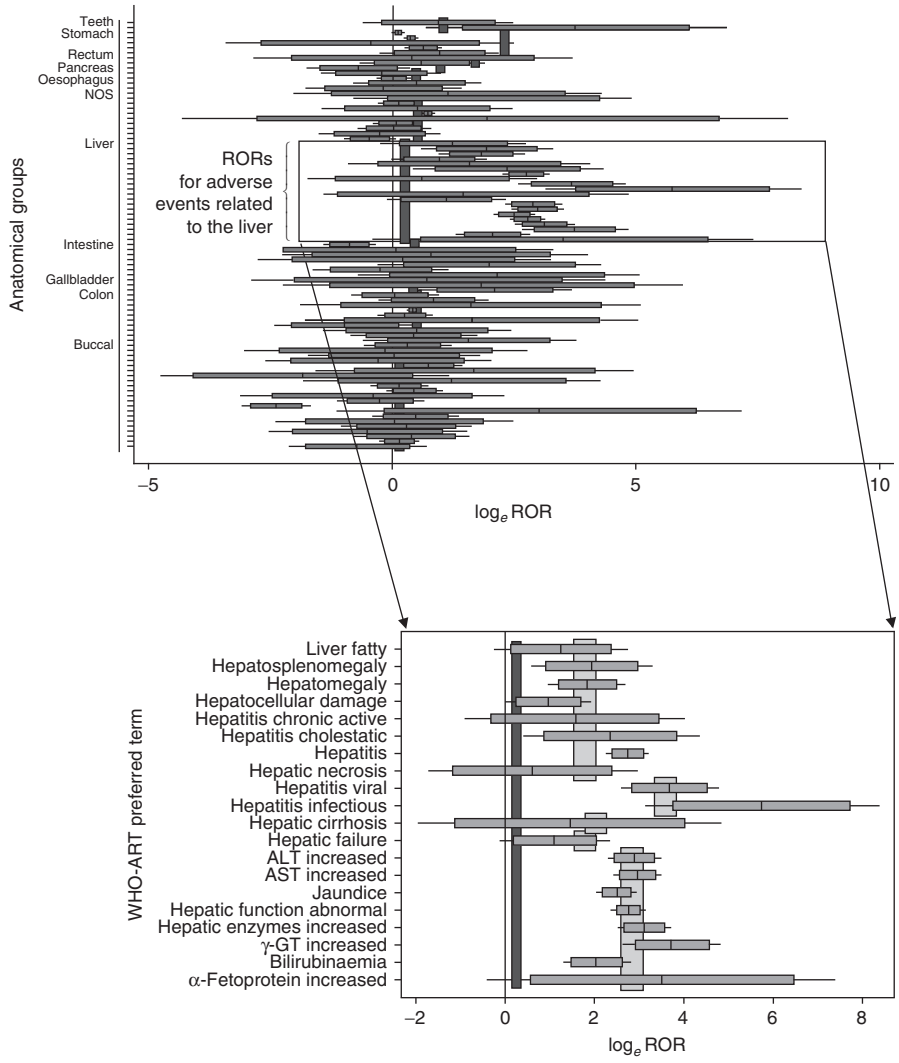

Fig. 7. Reporting odds ratios (RORs), adjusted by logistic regression, for liver adverse events following hepatitis vaccination (boxes = 95% CIs; lines = 99% CIs; vertical blocked lines = hierarchical group medians). Anatomical group median shown as black vertical line; lower-level group medians shown as grey vertical lines. **γ-GT** = γ-glutamate transferase; **ALT** = alanine aminotransferase; **AST** = aspartate aminotransferase; **NOS** = not otherwise specified; **WHO-ART** = WHO Adverse Reactions Terminology.

the outcomes of vomiting, nausea, diarrhoea and salivary gland enlargement, and these interactions need to be further investigated.

The strength of this study is that it demonstrated a novel sequence of methods to optimize analysis of the world's largest database of AEFI. It is therefore the first study to combine a Bayesian hierarchy with both multivariable regression to adjust for confounding and stratification to test for effect modification. Furthermore, the problems of multiple comparisons and small numbers were adjusted for by using a combined Bayesian model. The use of a Bayesian hierarchy for adverse event detection has only previously been demonstrated by Berry and Berry,[4] who found that including a single-level Bayesian hierarchy

© 2012 Adis Data Information BV. All rights reserved.

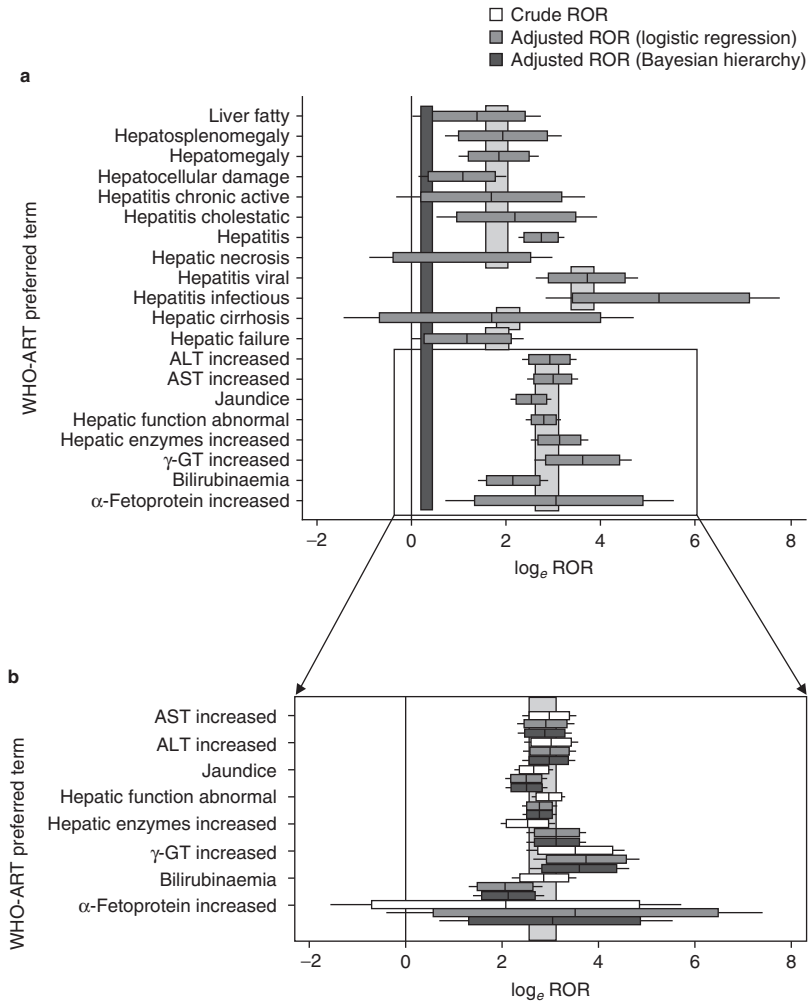

Fig. 8. (**a**) Reporting odds ratios (RORs), adjusted by a Bayesian hierarchy, for liver adverse events following hepatitis vaccination. Lower-level group medians shown as grey vertical lines; anatomical group median shown as black vertical line. (**b**) Crude and adjusted (with and without Bayesian hierarchy) RORs for abnormal liver function tests (boxes = 95% confidence/credible intervals; lines = 99% confidence/credible intervals; vertical blocked lines = hierarchical group medians). Lower-level group median shown as grey vertical line. **γ-GT** = γ-glutamate transferase; **ALT** = alanine aminotransferase; **AST** = aspartate aminotransferase; **WHO-ART** = WHO Adverse Reactions Terminology.

allowed sensible data sharing in a randomized controlled trial setting.

However, we have then further developed this method for use with spontaneous report data and have also used what we believe are more appropriate priors. The changes made by our more detailed hierarchy were consistent with clinical understanding and demonstrated that sensible assumptions were made of which codes should and should not share information and at which level. This was best demonstrated by the three new signals for hepatitis vaccines that occurred when analysed within the Bayesian model (see figure S5 and worked example). The two new signals of hepatocellular damage and fatty liver were only able to be identified because of the information

© 2012 Adis Data Information BV. All rights reserved.

provided by several other strong signals within the liver inflammation group. The same applies for the new signal for α-fetoprotein, which has large uncertainty when analysed alone, but when analysed with the other abnormal liver function tests the shared information reduces the variability in the model and it becomes a signal. One criticism of our Bayesian model is that only a few signals were altered by the Bayesian re-estimation; however, this was a consequence of our choice of non-informative priors and the three-level hierarchy. Although the study priors and hierarchy could be altered to increase their effect on the signals detected, this would be undesirable as we believe it would make unacceptably strong assumptions of exchangeability across the categories of adverse events. Another explanation may be that many of the signals were based on data collected over a long period of time and the data were therefore very extensive. We would expect utility of the method to be greater when data are more limited, for example in the early stages of postmarketing surveillance.

The largest criticisms of studies such as this one are that spontaneous reports are only an indication of a reporter's suspicion of an association rather than a real association. Therefore, signals generated from spontaneous reports are hypotheses that require appropriate testing in more detailed case-control or case-series analyses. However, in this study, the continuous data collection for VigiBase™ allowed us to investigate when changes in beliefs might have influenced reporting. This showed that the proven association of rotavirus and intussusception generated a signal before any increased concern had been reported, and this contrasted with the disproved association between measles, mumps and rubella (MMR) and Crohn's disease. Therefore, comparisons of spontaneous reporting databases before and after periods of increased awareness can help interpret associations where possible selection bias in reporting could have occurred.[21]

A second important criticism of our study is that reports in VigiBase™ come from many different countries with different criteria and systems, and these variations make it difficult to interpret the signals detected. However, as there were no missing data on the country of report, we were able to adjust for differences between countries by including the country of report in our models. Although there were missing data in other confounding co-variables, residual confounding should be minimal as a sensitivity analysis of complete records did not alter which signals were detected. More importantly, for our study there was enough detail recorded to identify a vaccine and adverse event in over 99.7% of the reports.

A potential limitation common to all signal detection studies is the need to choose cut-offs for confounding and the signal detection itself. As justified in the Methods section, sensible cut-offs were selected that were a compromise between type 1 and 2 errors, and these were similar to methods currently used by the WHO.[20] Finally, the use of ATC level 4 coding for signal detection does not identify the components of combination vaccines but codes for only one of the vaccine components. For instance, the MMR combination vaccine will only be coded under measles using ATC level 4 groupings. Further development of the hierarchical methods used in this study for the outcomes of AEFI could also be applied to vaccine exposure in future, but this extra complexity in the analysis was beyond the scope of this study.

It is difficult to assess performance of the methods used in this study without a gold standard for comparison. However, it is possible to compare the results to known associations. Although reports for combination vaccines will only be coded once under the level 4 ATC coding system, some combinations, such as mumps and measles, still had similar patterns of AEFI signals. This suggests that either separate vaccines are being used or that there is selective reporting of vaccines given at the same time. A more complex regression analysis could potentially disentangle the effects of separate antigens given at the same time. However, many of the final AEFI signals were similar to the signs and symptoms of the infectious diseases they prevent. For instance, measles, mumps and rubella vaccinations were associated with stomatitis, salivary gland problems and pancreatitis;[12] typhoid vaccinations were associated with gastrointestinal disturbances, as

© 2012 Adis Data Information BV. All rights reserved.

were oral vaccines such as poliomyelitis; and hepatitis vaccines were associated with abnormal liver function tests and hepatitis. This latter association has been observed following combined hepatitis A and B vaccination in adults,[22,23] but a study of children found no increased risk of gastrointestinal AEFI.[24] The association of diphtheria and pertussis with intestinal obstruction was probably related to the simultaneous administration of rotavirus vaccine, and this has been noted previously in the Vaccine Adverse Event Reporting System (VAERS).[25] The association of rotavirus and bloody diarrhoea has also been noted previously.[26]

Associations that were not so expected, and that could be of interest for further investigation, are poliomyelitis vaccine and dysphagia, diphtheria vaccine and hepatocellular damage, pneumococcal vaccines and haematemesis, typhoid and liver abnormalities, and hepatitis and tooth caries. Diphtheria and hepatocellular damage generated a signal based on only a single report. This is surprising as it would be expected that the Bayesian model would have shrunk the signal towards the null. However, on closer inspection of the results, there were three other associations within the liver group that shared information with the hepatocellular damage signal: hepatic necrosis, hepatitis and hepatitis viral. Although individually there was too much uncertainty for these associations to be signals, the Bayesian analysis allowed them to share their information and this strengthened the otherwise weak hepatocellular damage signal. This shrinkage towards a group median is a potential strength of our analysis over a solo Bayesian model that always shrinks towards the null. The sharing of group information allowed what might be considered an isolated signal to be flagged as potentially important when other information in the data supports it.

A previous study on adverse events found that up to 6% of adverse event signals were confounded.[27] However, this assessed only sex and age (and used only two categories for age) and was therefore not able to adjust for the extent of confounding we identified in our study. Literature on spontaneous reporting also describes the importance of testing for effect modification, but despite this it has rarely been performed.[5,28,29] However, we have been able to demonstrate that tests for effect modification can be easily integrated into a more complicated sequence of analysis and we believe our methods are understandable and potentially reproducible by other researchers. The current method of data analysis in regular use by the UMC includes calculation of the Bayesian information component as part of the screening process,[30] and this incorporates both adjustment for confounding and identification of subset-specific associations (effect modification).[31] Both are being currently applied, primarily in signal refinement, but await implementation in routine first-pass screening.

To further develop our methods, the effects of our hierarchical Bayesian model need to be tested in the whole VigiBase™ dataset for non-gastrointestinal AEFI, and this requires the development of more medically guided hierarchies to cover the other WHO-ART SOCs. In future, the more detailed Medical Dictionary for Regulatory Activities (MedDRA®) coding system could also be tested;[32] however, we chose not to use MedDRA®, as it is more changeable than WHO-ART coding. Furthermore, a few older reports entered into VigiBase™ before 2003 derived their MedDRA® coding from their existing WHO-ART coding. Further work is also needed to develop methods to allow the sharing of information between related codes in different organ classes: for example, codes for symptoms and signs of multisystem diseases or syndromes. It is possible that use of our methods involving information sharing may be more useful when using MedDRA® because the latter has very many more terms than WHO-ART, and therefore the amount of data per term is less.

Conclusions

The routine analysis of spontaneous report databases remains the only method of early identification of serious but rare AEFI for further investigation. The sequence of methods in this study maximizes information available in these datasets and can easily be applied to regular rou-

© 2012 Adis Data Information BV. All rights reserved.

tine data analysis. Importantly, the study avoids grouping similar adverse events together or splitting them into specific but small groups. The study provides logical, thorough and transparent methods that are interpretable and understandable. Therefore, our methods are appropriate for widespread use as they are easily reproducible by other researchers. We have demonstrated these methods in a limited childhood population for gastrointestinal adverse affects and the methods now need to be developed further to apply them to the whole VigiBase™ database and provide a comprehensive assessment tool to focus limited resources on the most appropriate signals for further investigation.

Acknowledgements

The information in this article does not represent the opinion of the WHO. We are extremely grateful to the UMC who have allowed us access to their data and who have helped with queries about the data files. A representative of the UMC read and commented on the manuscript to ensure factual accuracy in points relating to the data analysis process.

Colin Crooks is funded by a Medical Research Council population health scientist fellowship. The funding body was not involved in any aspect of the study. None of the authors has any conflicts of interest to declare.

References

1. Chen RT. Vaccine risks: real, perceived and unknown. Vaccine 1999; 17 Suppl. 3: S41-6
2. Fine PE, Clarkson JA. Individual versus public priorities in the determination of optimal vaccination policies. Am J Epidemiol 1986; 124 (6): 1012-20
3. Gangarosa EJ, Galazka AM, Wolfe CR, et al. Impact of anti-vaccine movements on pertussis control: the untold story. Lancet 1998; 351 (9099): 356-61
4. Berry SM, Berry DA. Accounting for multiplicities in assessing drug safety: a three-level hierarchical mixture model. Biometrics 2004; 60 (2): 418-26
5. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf 2009; 18 (6): 427-36
6. Expanded Programme on Immunization. Supplementary information on vaccine safety Part 2: Background rates of adverse events following immunization. Geneva: Department of Vaccines and Biologicals, World Health Organization, 2000
7. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Saf 2004; 13 (8): 519-23
8. WHO Collaborating Centre for Drug Statistics Methodology. International language for drug utilization research (ATC/DDD) [online]. Available from URL: http://www.whocc.no [Accessed 2011 Oct 14]
9. Uppsala Monitoring Centre. WHO-ART [online]. Available from URL: http://who-umc.org/DynPage.aspx?id=98107&mn1=7347&mn2=7252&mn3=7254&mn4=7339 [Accessed 2011 Oct 14]
10. Moulton LH, Rahmathullah L, Halsey NA, et al. Evaluation of non-specific effects of infant immunizations on early infant mortality in a southern Indian population. Trop Med Int Health 2005; 10 (10): 947-55
11. Aaby P, Jensen H, Gomes J, et al. The introduction of diphtheria-tetanus-pertussis vaccine and child mortality in rural Guinea-Bissau: an observational study. Int J Epidemiol 2004; 33 (2): 374-80
12. Demicheli V, Jefferson T, Rivetti A, et al. Vaccines for measles, mumps and rubella in children. Cochrane Database Syst Rev 2005; (4): CD004407
13. Centers for Disease Control and Prevention (CDC). Intussusception among recipients of rotavirus vaccine-United States, 1998-1999. MMWR Morb Mortal Wkly Rep 1999; 48 (27): 577-81
14. Stata Corporation. Statistical software: release 10.0. College Station (TX): Stata Corporation, 2007
15. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 2001; 10 (6): 483-6
16. Farrington CP, Firth MJ, Moulton LH, et al. Epidemiological studies of the non-specific effects of vaccines: II-methodological issues in the design and analysis of cohort studies. Trop Med Int Health 2009; 14 (9): 977-85
17. Bradburn MJ, Deeks JJ, Berlin JA, et al. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. Stat Med 2007; 26 (1): 53-77
18. Greenland S, Salvan A. Bias in the one-step method for pooling study results. Stat Med 1990; 9 (3): 247-52
19. Lunn D, Thomas A, Best N, et al. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000; 10: 325-37
20. Lindquist M, Stahl M, Bate A, et al. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf 2000; 23 (6): 533-42
21. Smeeth L, Rodrigues LC, Hall AJ, et al. Evaluation of adverse effects of vaccines: the case-control approach. Vaccine 2002; 20 (19-20): 2611-7
22. Geier DA, Geier MR. Hepatitis B vaccination and adult associated gastrointestinal reactions: a follow-up analysis. Hepatogastroenterology 2002; 49 (48): 1571-5
23. Woo EJ, Miller NB, Ball R. Adverse events after hepatitis A B combination vaccine. VAERS Working Group. Vaccine 2006; 24 (14): 2685-91
24. Niu MT, Rhodes P, Salive M, et al. Comparative safety of two recombinant hepatitis B vaccines in children: data from the Vaccine Adverse Event Reporting System (VAERS) and Vaccine Safety Datalink (VSD). J Clin Epidemiol 1998; 51 (6): 503-10
25. Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early

© 2012 Adis Data Information BV. All rights reserved.

detection of intussusception and other events after rotavirus vaccination. Vaccine 2001; 19 (32): 4627-34

26. Haber P, Chen RT, Zanardi LR, et al. An analysis of rotavirus vaccine reports to the vaccine adverse event reporting system: more than intussusception alone? Pediatrics 2004; 113 (4): e353-9
27. Woo EJ, Ball R, Burwen DR, et al. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug Saf 2008; 31 (8): 667-74
28. Evans SJW. Stratification for spontaneous report databases. Drug Saf 2008; 31 (11): 1049-52
29. Hopstadius J, Noren GN, Bate A, et al. Impact of stratification on adverse drug reaction surveillance. Drug Saf 2008; 31 (11): 1035-48
30. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998; 54 (4): 315-21
31. Norén GN, Bate A, Orre R, et al. Extending the methods used to screen the WHO drug safety database towards analysis of complex associations and improved accuracy for rare events. Stat Med 2006; 25 (21): 3740-57
32. Brown EG. Effects of coding dictionary on signal generation: a consideration of use of MedDRA compared with WHO-ART. Drug Saf 2002; 25 (6): 445-52

Correspondence: Dr *Colin Crooks*, Division of Epidemiology and Public Health, Clinical Sciences Building 2, Nottingham City Hospital, Nottingham NG5 1PB, UK.
E-mail: colincrooks@doctors.org.uk

© 2012 Adis Data Information BV. All rights reserved.