ORIGINAL RESEARCH ARTICLE

Drug Saf 2012; 35 (4): 335-346
0114-5916/12/0004-0335/$49.95/0

Vaccine-Based Subgroup Analysis in VigiBase

Effect on Sensitivity in Paediatric Signal Detection

Sandra de Bie,[1,2] *Katia M.C. Verhamme,*[1] *Sabine M.J.M. Straus,*[1,2] *Bruno H.Ch. Stricker*[1,3,4] and *Miriam C.J.M. Sturkenboom,*[1,3] *on behalf of the TEDDY European Network of Excellence*

1 Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands
2 Dutch Medicines Evaluation Board, Utrecht, the Netherlands
3 Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands
4 Inspectorate of Healthcare, The Hague, the Netherlands

Abstract

Background: Data mining of spontaneously reported adverse drug reactions (ADRs), using measures of disproportionality, is a valuable first evaluation step for drug safety signal detection. Of all ADRs reported for children and adolescents within VigiBase, vaccine-ADR pairs comprise more than half of the reports. ADRs concerning vaccines differ with respect to type and seriousness from other drugs, and therefore may influence signal detection for non-vaccine drugs if not accounted for appropriately.

The potential influence of vaccines on safety signal detection for drugs was recently raised by the CIOMS Working Group VIII, who proposed that it may be appropriate to undertake automatic signal detection using both medicines and vaccines, and some analysis using vaccines only. However, it has not described for which types of ADRs or drugs subgroup analysis is beneficial.

Objective: The aim of the study was to study the methodological aspects concerning the influence of a high prevalence of vaccine-related ADRs on signal detection within paediatric ADR data.

Methods: We analysed all paediatric Individual Case Safety Reports (ICSRs) received by VigiBase between 2000 and 2006, and calculated the reporting odds ratio (ROR) for all unique drug-ADR pairs with at least three reports. The ROR was additionally calculated in subgroups of vaccine-ADR pairs and non-vaccine-ADR pairs and further in different age groups. A proportional change in the ROR for the different subgroups was calculated and the change in the number of signals of disproportional reporting (SDRs) after subgroup analysis was assessed.

Results: Of all paediatric ICSRs (N = 218 840, of which 117 877 were vaccine-related), a total of 26 203 unique drug-ADR pairs were eligible for inclusion (5586 vaccine-related). A total of 1637 vaccine-related SDRs and 13 375 non-

vaccine-related SDRs were detected in the crude analysis. Subgroup analysis by restricting to either vaccines or non-vaccines revealed 494 additional SDRs for vaccines (+30.2%) and 821 additional SDRs for non-vaccines (+6.1%). Subgroup analyses were only beneficial for non-vaccines if the ADR of interest was reported uncommonly for non-vaccines and beneficial for vaccines if the ADR was reported uncommonly for vaccines. Subgroup analysis for ADRs that were reported commonly for either vaccines or non-vaccines led to the disappearance of 272 SDRs for vaccines and 2721 SDRs for non-vaccines. We could empirically derive a model that predicts the change in ROR in the subgroups based on the proportion of vaccines within the total dataset.

Conclusion: The high proportion of vaccine-related reports within paediatric ADR data has a large and mathematically predictable impact on signal detection in paediatric ADR data. Subgroup analysis reveals new SDRs that potentially represent genuine safety signals. The most inclusive and sensitive signal detection method would be the combination of a crude and subgroup-based data mining approach, based on the ratio between the proportion of vaccines within the ADR of interest and within all other ADRs.

Background

Spontaneously reported adverse drug reactions (ADRs) are an important source for identifying drug safety signals.[1] For efficient signal detection, data mining methods have been developed that mostly are based on measures of disproportionality. Well known examples are the Reporting Odds Ratio (ROR), the Proportional Reporting Ratio (PRR), the Information Component (IC), and the Empirical Bayes Geometric Mean (EBGM).[2-6] These data mining methods are used as a first signal identifying method. Subsequently, further case evaluation is necessary to determine whether the signal of disproportionality is a real safety signal.[7,8]

Although useful, data mining methods are subject to bias and confounding. Effort has been made to allow for dealing with possible confounding factors such as age, sex and time, but with varying results.[6,9-11] Other factors that might influence the disproportionality estimates include the number of serious versus non-serious individual case safety reports (ICSRs), consumer versus healthcare professional reported ICSRs, company owned databases versus databases of national competent authorities and differences in the distribution of the population characteristics within the different databases or of the different outcomes.[2,12] All these factors might lead to a relative increase of specific groups of reports by type of ADR, drug class or age group. Such clusters of reports may influence the distribution of drugs and ADRs within the data and thereby jeopardize the assumption that reporting should be non-differential in order to guarantee unbiased estimates of measures of disproportionality.

The phenomenon of clusters of reports of a specific group of drugs is very clearly observed in paediatric safety signal detection. Within national compilations of paediatric ICSRs, vaccines make up 45–69% of the suspected drugs within the ADR reports.[13,14] ADRs reported for vaccines differ from non-vaccines with respect to seriousness and type.[13-15] Many studies have already addressed issues around data mining within vaccine-related ADRs, mostly boosted by the work on the US' specific Vaccine Adverse Event Reporting System (VAERS) database.[9,11,16,17] Little is, however, known about the influence of vaccines on safety signal detection for non-vaccines and *vice versa* in a mixed ADR database, containing both vaccines

and non-vaccines. Due to the high exposure to vaccines in the paediatric age group, this influence will be most pronounced when data mining is applied to paediatric ADR reports.

The potential influence of vaccines on safety signal detection for drugs was recently raised in the report of the CIOMS Working Group VIII.[8] The working group proposed that it may be appropriate to undertake automatic signal detection using both medicines and vaccines, and some analysis using vaccines only. However, it has not described for which types of ADRs or drugs subgroup analysis is beneficial.

Therefore, we studied the methodological aspects of signal detection within paediatric ADR data where the prevalence of vaccine-related ADRs is high. We studied how restriction to either vaccine- or non-vaccine-related ADRs influences disproportionality analyses, whether this affects the number of detected signals of disproportional reporting (SDRs) and for which ADRs subgroup analysis using restriction is beneficial.

Methods

Setting

We used data from the VigiBase database of suspected ADRs. This WHO global ICSR database system was established in 1968, and in March 2007 it held more than 3.8 million ICSRs.[18,19] VigiBase is maintained on behalf of the WHO Programme by the Uppsala Monitoring Centre (UMC). In March 2007, more than 80 countries participated in the WHO International Drug Monitoring Programme and another 17 countries were associate members who did not yet actively contribute data. ICSRs are submitted through the national pharmacovigilance centres. The WHO Programme member countries submit ICSRs to the UMC on a regular basis; preferably once a month, but at least every quarter.[19]

At the time of data extraction in 2006, all ADRs within VigiBase were coded using preferred terms of the WHO Adverse Reaction Terminology (WHO-ART) coding dictionary. Reported drugs were recoded using the WHO drug dictionary, and were also coded according to the Anatomical Therapeutic Chemical (ATC) classification system of the WHO Collaborating Centre for Drug Statistics Methodology.[20]

Selection of Individual Case Safety Reports (ICSRs) and Drug-ADR Pairs

From VigiBase we extracted all ICSRs on children, aged 0 to ≤18 years that were received or occurred between January 2000 and December 2006. Only ICSRs received through spontaneous reporting and where the drug was reported as suspected were included. The information in these reports included country of origin, type of reporter, age at onset, year of onset, sex, suspected drugs, ADRs, starting and stopping date of the suspected drugs, starting and stopping date of the ADRs, dosing regimen of the drugs, administration route and causality assessment of the event. We excluded ICSRs where the reported drug or ADR could not be coded in the WHO drug dictionary or WHO-ART.

An ICSR can contain more than one suspected drug, or more than one ADR. We defined a drug-ADR pair as a unique combination of a single drug and a single ADR. Hence, an ICSR containing two ADRs with one suspected drug for both ADRs counted as two pairs.

Vaccine-ADR pairs were defined as drug-ADR pairs in which a vaccine, coded using the ATC code J07, was reported as the suspected drug. All other drug-ADR pairs were considered as non-vaccine-related drug-ADR pairs.

Statistical Analyses

Disproportionality Calculations and Subgroup Analysis

For all possible, unique, drug-ADR pairs with at least three records within the database we calculated the RORs with 95% CIs.[4,5] All disproportionality analyses were conducted on a drug-ADR-pair level. Calculation of the ROR is based on a two-by-two contingency table, containing all drug-ADR pairs within the dataset, in which cell A represents the number of pairs for the combination of interest, cell B all other ADRs for the drug of interest, cell C the number of pairs for the ADR of interest for other drugs and cell D all other pairs without the ADR and without the

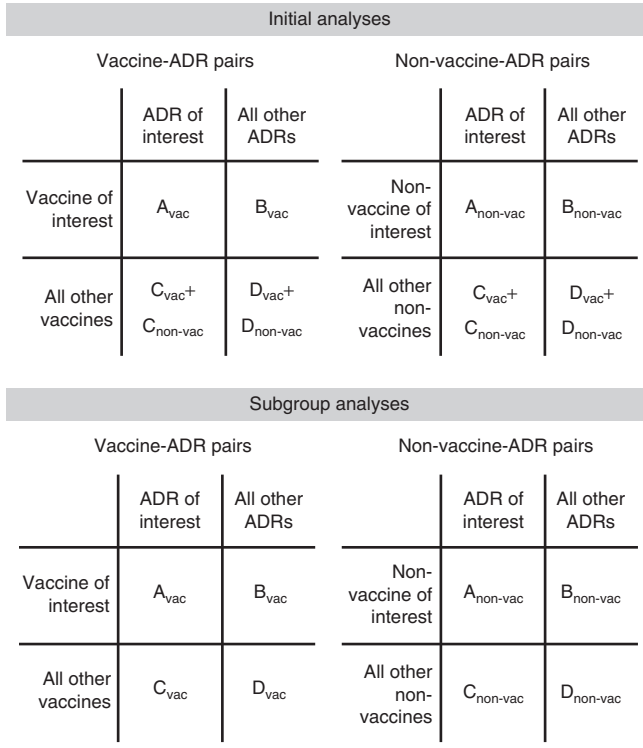

Fig. 1. Distribution of vaccine-ADR and non-vaccine-ADR pairs within the two-by-two contingency table. **ADR** = adverse drug reaction; $\mathbf{A_{non-vac}}$ = non-vaccine-ADR pair of interest; $\mathbf{A_{vac}}$ = vaccine-ADR pair of interest; $\mathbf{B_{non-vac}}$ = all other ADRs for non-vaccine of interest; $\mathbf{B_{vac}}$ = all other ADRs for vaccine of interest; $\mathbf{C_{non-vac}}$ = ADR of interest for all other non-vaccines; $\mathbf{C_{vac}}$ = ADR of interest for all other vaccines; $\mathbf{D_{non-vac}}$ = all other non-vaccine-related ADRs; $\mathbf{D_{vac}}$ = all other vaccine-related ADRs.

drug of interest (figure 1). When the drug-ADR pair of interest concerns a non-vaccine, cells A and B of the two-by-two table in a mixed ADR database as VigiBase will only contain non-vaccine-ADR pairs, while cells C and D will contain both vaccine-ADR and non-vaccine-ADR pairs. For vaccine-related combinations, cells A and B consist solely of vaccine-ADR pairs, and cells C and D will be mixed with non-vaccines.

To explore the effect of vaccine-ADR pairs on the number of detected SDRs we first performed a crude analysis (independent of the type of drug) and subsequently subgroup analyses, where we split the total ADR dataset in vaccine-ADR pairs or non-vaccine-ADR pairs (figure 1). In this subgroup analyses we calculated the ROR for all vaccine-ADR pairs after restriction to vaccine-ADR pairs, and for non-vaccine ADRs the calculation was restricted to non-vaccine-ADR pairs.

RORs, both in the initial crude analyses and within the subgroup analyses, were additionally calculated within predefined subgroups of age in order to observe whether the effect differed across age categories. Age at the time of the event was categorized into three categories according to the guidelines of the International Conference of Harmonization (ICH): 0 to <2 years, 2 to ≤11 years and 12 to ≤18 years.[21] The ROR was only calculated within those age groups that contained at least three reports of the unique drug-ADR pair of interest.

Vaccine- and Non-Vaccine Proportion Ratio

Subgroup analysis of vaccine-ADR pairs or non-vaccine-ADR pairs only influences cells C and D of the two-by-two contingency table (figure 1). For each drug-ADR pair in this study the proportion of vaccine-ADR pairs and the proportion

of non-vaccine-ADR pairs within cells C and D was calculated. The ratio between the proportion of vaccine-related pairs in cell C and the proportion of vaccine-related pairs in cell D was defined as the 'vaccine proportion ratio'. The 'non-vaccine proportion ratio' was defined as the ratio between the proportion of non-vaccine-related pairs in cell C and the proportion of non-vaccine-related pairs in cell D.

Effect of Subgroup Analysis on Reporting Odds Ratio (ROR)

In order to compare the subgroup-specific ROR estimates with the crude overall ROR, a proportional change in the ROR was calculated. This change was calculated as follows: $[(ROR_{subgroup}/ROR_{crude})-1]\times 100\%$. A mathematical model was constructed to describe the relationship between the change in ROR and the proportion ratio of vaccines or non-vaccines. To compare the change in ROR based on the mathematical model with the observed change in ROR, an R-square statistic was calculated.

Signals of Disproportional Reporting (SDRs)

For all unique drug-ADR pairs with at least three records it was determined whether the combination was an SDR, defined as an ROR with a lower limit of the 95% CI >1.[5] Evaluation of potential SDRs was performed both during the crude analysis and following restriction to vaccines or non-vaccines.

For SDRs that were newly detected after the subgroup analysis, a random sample of 10% was taken to evaluate whether this was a false positive or a true positive association. This was obtained by evaluating whether the ADR was listed or covered in the summary of product characteristics (SPC) of the drug or vaccine of interest. For SDRs that disappeared after subgroup analysis, a random sample of 5% was taken for comparison with the SPC to evaluate whether these disappearing SDRs were false negative or true negative SDRs.

Comparisons

Characteristics of the vaccine-ADR pairs and non-vaccine-ADR pairs (and within the age categories) were compared using chi-square (to compare proportions) and Mann-Whitney tests (to compare means). A p-value <0.05 was considered to be statistical significant.

Results

ICSRs

In the period between January 2000 and December 2006, 221 508 ICSRs involving children and adolescents aged ≤18 years were received by the WHO-UMC. ICSRs that lacked information on the reported drug or reported ADR were excluded (n = 2668; 1.2%). The remaining ICSRs (n = 218 840) contained 812 415 drug-ADR pairs, with a median of two pairs per ICSR (table I).

The median age in the reported ICSRs was 5.0 years. The distribution of the ICSRs within the three predefined age categories was 34.2% in the 0 to <2 years group, 38.7% for children aged 2 to ≤11 years and 27.1% in children aged 12 to ≤18 years. More drug-ADR pairs per ICSR were reported for the youngest children (0 to <2 years) than for the two other age categories (mean 4.9 vs 3.1 record per ICSR) [p = 0.000]. Consequently, the highest proportion of drug-ADR pairs is present in the youngest category (45.1%) [p = 0.000].

Vaccine-related reports made up 53.9% (n = 117 877) of all ICSRs. These vaccine-related reports had a higher number of drug-ADR pairs per ICSR than non-vaccine-related reports (4.4 vs 2.9) [p = 0.000]. None of the ICSRs had both a vaccine and a non-vaccine reported as the suspected drug.

Vaccines accounted for 63.7% of the total number of drug-ADR pairs. In the youngest children (0 to <2 years), the proportion of vaccine-ADR pairs was highest at 87.6% (p = 0.000). This proportion decreased with age to 56.1% in children aged 2 to ≤11 years, and 27.1% in children aged 12 to ≤18 years (table I).

Unique Drug-ADR Pairs and SDRs

The 218 840 ICSRs included in the analyses contained 90 441 unique drug-ADR pairs, of which 13% (n = 11 478) concerned vaccine-ADR pairs. Of the unique pairs, 26 203 (29.0%) were reported at least three times and were included in the subgroup analyses. The number of unique

Table I. Characteristics of the individual case safety reports

Characteristic	ICSRs [n (%)]	Drug-ADR pairs [n (%)]	No. of pairs/ICSR[a]
N	218 840 (100)	812 415 (100)	3.7
Sex [male]	110 271 (50.4)	417 463 (51.4)	3.8
Age [y]			
Median [range]	5.0 [0.0–18.0]	3.0 (0.0–18.0)	
0 to <2	74 735 (34.2)	366 023 (45.1)	4.9
2 to ≤11	84 744 (38.7)	261 488 (32.2)	3.1
12 to ≤18	59 361 (27.1)	184 904 (22.8)	3.1
Vaccine-related [y][b]	117 877 (53.9)	517 642 (63.7)	4.4
0 to <2	58 835 (78.7)	320 814 (87.6)	5.5
2 to ≤11	42 505 (50.2)	146 784 (56.1)	3.5
12 to ≤18	16 537 (27.9)	50 044 (27.1)	3.0
Non-vaccine-related [y][b]	100 963 (46.1)	294 773 (36.3)	2.9
0 to <2	15 900 (21.3)	45 209 (12.4)	2.8
2 to ≤11	42 239 (49.8)	114 704 (43.9)	2.7
12 to ≤18	42 824 (72.1)	134 860 (72.9)	3.1

a Mean number of drug-ADR pairs within an ICSR.

b Percentage of vaccines or non-vaccines in this group relative to all ICSRs or drug-ADR pairs in this age category.

ADR = adverse drug reaction; **ICSR(s)** = individual case safety report(s).

pairs with at least three records per age stratum was 6818 (25.1%) for the age category 0 to <2 years, 11 074 (25.3%) for the age category 2 to ≤11 years and 11 659 (24.1%) for the age category 12 to ≤18 years (table II).

Within the unique drug-ADR pairs with at least three records, 21.3% (n = 5586) concerned vaccines. The proportion of unique vaccine-ADR pairs within the combinations with at least three records decreased with age; 55.1% (n = 3757) within children aged 0 to <2 years, 24.4% (n = 2699) within children aged 2 to ≤11 years and 13.9% (n = 1626) within children aged 12 to ≤18 years.

In the crude analysis, 15 012 unique drug-ADR pairs with at least three records (57.3%) were considered an SDR. These concerned 1637 vaccine-ADR pairs (10.9% of all SDRs) and 13 375 non-vaccine-ADR pairs (89.1% of all SDRs) [table II].

Change in ROR and SDRs after Subgroup Analyses

The mean proportion of vaccines in cell D of the two-by-two contingency table for the unique vaccine-ADR pairs was 63.8% (95% CI 63.8, 63.8). For the unique non-vaccine-ADR pairs, the mean proportion of non-vaccines in cell D was 36.2% (95% CI 36.2, 36.2). The median vaccine proportion ratio was 0.98 and ranged from 0.01 to 1.68. The median non-vaccine proportion ratio was 1.89 and ranged from 0.01 to 2.92. The proportional change in ROR after restriction to either vaccine-related pairs or non-vaccine-related pairs depended on the vaccine-proportion and non-vaccine proportion ratio, respectively (figures 2a and b).

When the vaccine- or non-vaccine proportion ratio was <1, the ROR increased after subgroup analysis; when the ratio was >1 the ROR after subgroup analysis decreased in comparison to the crude ROR. The shape of the relationship between the proportional change in ROR and the proportion of vaccines within the data was modelled mathematically (figure 3). The step-by-step explanation on the construction of this model can be found in the Supplemental Digital Content (http://links.adisonline.com/DSZ/A66). The observed RORs after subgroup analyses were perfectly predicted by the model (R-square = 1.0).

The shape of the relationship between the change in ROR after subgroup analysis and the

Table II. Number of signals of disproportional reporting

	Vaccines				Non-vaccines				Overall			
	Total database	0 to <2 y	2 to ≤11 y	12 to ≤18 y	Total database	0 to <2 y	2 to ≤11 y	12 to ≤18 y	Total database	0 to <2 y	2 to ≤11 y	12 to ≤18 y
Crude analyses												
Unique drug-ADR pairs [n (%)][a]	11 478 (12.7)	7865 (28.9)	6166 (14.1)	3999 (8.3)	78 963 (87.3)	19 320 (71.1)	37 561 (85.9)	44 371 (91.7)	90 441 (100.0)	27 185 (100.0)	43 727 (100.0)	48 370 (100.0)
Unique drug-ADR pairs with at least 3 records [n (%)][a]	5586 (21.3)	3757 (55.1)	2699 (24.4)	1626 (13.9)	20 617 (78.7)	3061 (44.9)	8375 (75.6)	10 033 (86.1)	26 203 (100.0)	6818 (100.0)	11 074 (100.0)	11 659 (100.0)
Number of SDRs [n (%)][a,b]	1637 (10.9)	1053 (29.9)	811 (12.8)	716 (10.6)	13 375 (89.1)	2468 (70.1)	5518 (87.2)	6016 (89.4)	15 012 (100.0)	3521 (100.0)	6329 (100.0)	6732 (100.0)
Subgroup analyses												
Unique drug-ADR pairs with at least 3 records per subgroup[a]												
Vaccine- or non-vaccine proportion ratio <1 [n (%)][b,c,d]	2879 (51.5)	1749 (46.6)	1302 (48.2)	591 (36.3)	15 291 (74.2)	2285 (74.6)	6458 (77.1)	7060 (70.4)	NA	NA	NA	NA
Vaccine- or non-vaccine proportion ratio >1 [n (%)][b,c,d]	2707 (48.5)	2008 (53.4)	1397 (51.8)	1035 (63.7)	5326 (25.8)	776 (25.4)	1917 (22.9)	2973 (29.6)	NA	NA	NA	NA
Became SDR (n)[b]	494	153	202	115	821	159	306	327	NA	NA	NA	NA
Vaccine- or non-vaccine proportion ratio <1 [n (%)][b,c,d]	494 (100.0)	153 (100.0)	202 (100.0)	115 (100.0)	821 (100.0)	159 (100.0)	306 (100.0)	327 (100.0)	NA	NA	NA	NA
Vaccine- or non-vaccine proportion ratio >1 [n (%)][b,c,d]	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Became non-SDR (n)[b]	272	89	205	306	2721	370	833	439	NA	NA	NA	NA
Vaccine- or non-vaccine proportion ratio <1 [n (%)][b,c,d]	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	NA	NA	NA	NA
Vaccine- or non-vaccine proportion ratio >1 [n (%)][b,c,d]	272 (100.0)	89 (100.0)	205 (100.0)	306 (100.0)	2721 (100.0)	370 (100.0)	833 (100.0)	439 (100.0)	NA	NA	NA	NA
Increase in number of SDRs [%][e]	30.2	14.5	24.9	16.1	6.1	6.4	5.5	5.4	NA	NA	NA	NA

a Relative frequency within vaccines or non-vaccines compared with total.

b All pairs with at least three records are included in the analyses. Since a drug-ADR pair can occur in either of the age-categories, age-categories total above 100% of the database.

c Vaccine proportion ratio: proportion of vaccines in cell C of the two-by-two table divided by the proportion of vaccines in cell D of the two-by-two table. Non-vaccine proportion ratio: proportion of non-vaccines in cell C of the two-by-two table divided by the proportion of non-vaccines in cell D of the two-by-two table.

d Percentage based on the total number of SDRs within this category.

e Increase in number of SDRs (%) = (newly detected SDRs)/initial SDRs*100%.

ADR = adverse drug reaction; **NA** = not applicable; **non-SDR** = not a signal of disproportional reporting; **SDR(s)** = signal(s) of disproportional reporting.

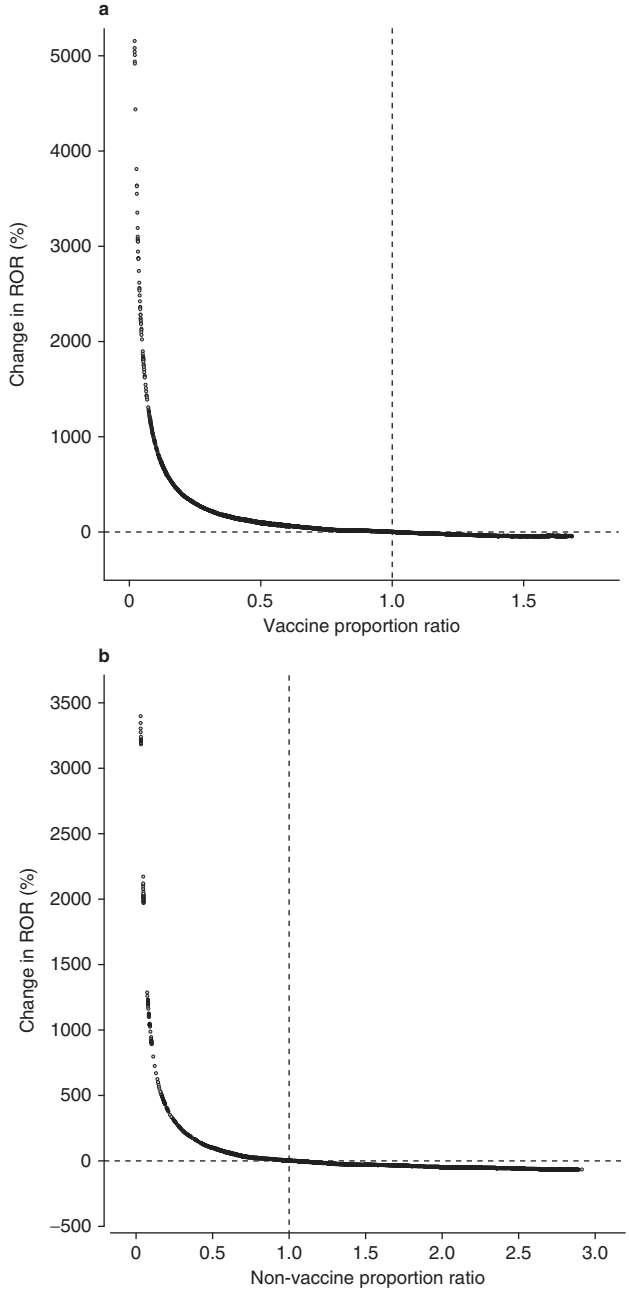

Fig. 2. (**a**) Change in ROR (%) after subgroup analysis within vaccine-ADR pairs. All unique vaccine-ADR pairs with at least three records are included. On the x-axis, the vaccine proportion ratio is given: the proportion of vaccines in cell C of the two-by-two table divided by the proportion of vaccines in cell D of the two-by-two table. The y-axis has been cut off at 5000%; the actual data on the y-axis ranges from −41% to 19 055% for vaccines. (**b**) Change in ROR (%) after subgroup analysis within non-vaccine-ADR pairs. All unique non-vaccine-ADR pairs with at least three records are included. On the x-axis, the non-vaccine proportion ratio is given: the proportion of non-vaccines in cell C of the two-by-two table divided by the proportion of non-vaccines in cell D of the two-by-two table. The y-axis has been cut off at 3500%; the actual data on the y-axis ranges from −66% to 6990%. **ADR** = adverse drug reaction; **ROR** = reporting odds ratio.

vaccine- and non-vaccine proportion ratios is similar for all age categories, although the mean proportion of vaccines in cell D of the two-by-two table and, consequently, the point where the proportion ratio equals 1, differed. The point of equality was 87.2% (95% CI 87.2, 87.2) for children aged 0 to <2 years, 55.8% (95% CI 55.8, 55.9) for children aged 2 to ≤11 years and for children aged 12 to ≤18 years the point of equality was 26.9% (95% CI 26.9, 26.9). The mean proportions of non-vaccines were 12.8% (95% CI 12.6, 12.8), 44.2% (95% CI 44.1, 44.2) and 73.1% (95% CI 73.1, 73.1) for the increasing age categories (data not shown).

Consequently to the change in ROR upon subgroup analyses, the number of detected SDRs changed. Of the unique vaccine-related pairs that were not an SDR in the crude analysis, 494 (12.5%) became an SDR in the vaccine-specific analysis, while 272 (16.6%) of the initial vaccine-related SDRs were no longer an SDR. Within the non-vaccine-ADR pairs, 821 (11.3%) became an SDR and 2721 (20.3%) were no longer an SDR after subgroup analysis. The effect of subgroup analysis on the number of detected or disappearing SDRs depended on vaccine- and non-vaccine proportion ratios (table II). When the vaccine- or non-vaccine proportion ratio was <1, additional SDRs were detected after subgroup analysis, while no SDRs disappeared. When the ratio was >1, SDRs disappeared after subgroup analysis, while no additional SDRs were detected. When subgroup analysis is restricted to ADRs with a proportion ratio <1, 30.2% more SDRs (n = 272) were detected for vaccines and 6.1% more SDRs (n = 2721) were detected for non-vaccines.

Vaccine-ADR pairs

Change in ROR (%) after restriction to vaccine-ADR pairs =

$$100\% \times \left(\frac{D_{vac}/D}{C_{vac}/C} - 1\right)$$

ROR after restriction to vaccine-ADR pairs =

$$\frac{D_{vac}/D}{C_{vac}/C} \times \mathrm{ROR}_{crude}$$

Non-vaccine-ADR pairs

Change in ROR (%) after restriction to non-vaccine-ADR pairs only =

$$100\% \times \left(\frac{D_{non\text{-}vac}/D}{C_{non\text{-}vac}/C} - 1\right)$$

ROR after restriction to non-vaccine-related pairs only =

$$\frac{D_{non\text{-}vac}/D}{C_{non\text{-}vac}/C} \times \mathrm{ROR}_{crude}$$

Fig. 3. Estimation of proportional change in ROR and ROR after subgroup analyses. **ADR** = adverse drug reaction; $C_{non\text{-}vac}/C$ = proportion of non-vaccines in C; C_{vac}/C = proportion of vaccines in C; $D_{non\text{-}vac}/D$ = proportion of non-vaccines in D; D_{vac}/D = proportion of vaccines in D; **ROR** = reporting odds ratio; ROR_{crude} = reporting odds ratio based on crude analysis using both vaccines and non-vaccines.

Characteristics of Newly Detected and Disappearing SDRs

A random sample of 10% of the newly detected SDRs (n = 134) concerned 50 vaccine-related and 84 non-vaccine-related SDRs. For the vaccines, 40.0% (n = 20) of the newly detected SDRs after subgroup analysis were true positives, associations that were listed in the SPC of the vaccine. For the non-vaccines 63.1% (n = 53) of the newly detected SDRs after subgroup analysis concerned associations that were true positives.

A random sample of 5% of the SDRs (n = 139) that disappeared after subgroup analysis was reviewed. These concerned ADRs of 14 vaccines, of which 8 were listed in the SPC (57.1%) [false negatives], and ADRs of 125 non-vaccines, of which 67 were listed in the SPC (53.6%) [false negatives].

Within all newly detected vaccine-related SDRs, the following ADRs were most frequently reported: oedema periorbital (n = 10; 2.0%), rash erythematous (n = 10; 2.0%) and abdominal pain (n = 8; 1.6%). Of the non-vaccine-related SDRs, fever (n = 94; 11.4%), agitation (n = 52; 6.3%) and rash (n = 37; 4.5%) were most frequently reported.

For the disappearing vaccine-related SDRs after subgroup analysis, the following ADRs were most frequently reported: fever (n = 22; 8.1%), injection site reaction (n = 18; 6.6%) and injection site mass (n = 13; 4.8%). Of the non-vaccine-related SDRs that disappeared, medication error (n = 92; 3.4%), death (n = 65; 2.4%) and coma (n = 56; 2.1%) were the most frequently reported ADRs.

Discussion

In this study we explored methodological aspects concerning signal detection for non-vaccine

drugs and vaccines within a mixed dataset of paediatric ADR data and how this is influenced by the proportion of vaccines within the data. For vaccine-related ADRs that were less frequently reported for vaccines compared with all other ADRs in the dataset, subgroup analysis by restriction to vaccine-related pairs led to an increase in the ROR, resulting in new vaccine-related SDRs. Equally, for non-vaccine-related ADRs, additional SDRs were detected for ADRs that were less frequently reported for non-vaccines compared with all other ADRs in the dataset.

In 2008, both Hopstadius et al.[10] and Woo et al.[11] studied the impact of stratification as a method to deal with confounding within data mining for signal detection. Hopstadius et al.[10] used VigiBase to study the effect of stratification as a method to adjust for possible confounding by age, sex, time of reporting and country of origin. Based on their results, they concluded that the possible improvement of the data mining methods by stratification is smaller than previously assumed. Woo et al.[11] used the VAERS database for their study. Stratification by age and sex did reveal some signals that were previously undetected.

In a comment on the studies of Hopstadius et al.[10] and Woo et al.[11], Evans[22] touched upon the differences between conventional drugs (non-vaccines) and vaccines, and highlighted the differences in health, population and spectrum of adverse events between the two groups of drugs. Evans[22] concluded that data mining within vaccine data only has gains and losses since reactions that are commonly reported for other vaccines might be missed as a potential safety signal if they are mined in a vaccine stratum solely. This issue has also been touched upon in the recently published report on ''Practical aspects of signal detection in pharmacovigilance'' of the CIOMS Working Group VIII in which the impact of vaccines on signal detection is considered.[8] In their report, the CIOMS Working Group states that restricting to vaccines probably does not solve all problems that are related to signal detection within vaccines, and proposed that it may be appropriate to undertake automatic signal detection using both medicines and vaccines, and some analysis using vaccines only.

The concern of Evans[22] on the risk of missing vaccine-related signals when restricting to vaccine-related pairs, especially for ADRs commonly reported for vaccines, is confirmed in our study. Using measures we defined as the 'vaccine proportion ratio' (the proportion of vaccine-related pairs in cell C divided by the proportion of vaccine-related pairs in cell D) and the 'non-vaccine proportion ratio' (the proportion of non-vaccine-related pairs in cell C divided by the proportion of non-vaccine-related pairs in cell D), we were able to identify subsets of reports for which subgroup analysis by restriction to either vaccine-related or non-vaccine-related pairs is either harmful or beneficial.

In the current study, ADRs that were commonly reported for either vaccines or non-vaccines were reflected by a vaccine- or non-vaccine proportion ratio >1. For these ADRs, subgroup analysis led to a decrease in the ROR, resulting in the disappearance of SDRs that were detected in the crude analysis. This can be regarded as harmful since, based on a random sample of 5%, a majority of the vaccine-related pairs and non-vaccine-related pairs are included in the SPC and might represent real associations. ADRs that were reported less frequently for either vaccines or non-vaccines, were reflected by a vaccine- or non-vaccine proportion ratio <1. For these ADRs, the ROR increased after subgroup analysis, resulting in SDRs that were not detected in the crude analysis. This can potentially be regarded as beneficial since, based on a random sample of 10%, 40.0% of the vaccine-related pairs and 63.1% of the non-vaccine-related pairs are included in the SPC and might represent real associations.

The potential gain or loss of SDRs should be kept in mind when applying safety signal detection in databases containing both vaccines and conventional drugs. Ignoring subgroup analysis will decrease the sensitivity of the data mining algorithm due to false negative SDRs. Restricting to subgroup analysis only will also increase the number of false negative SDRs, resulting in a lower sensitivity.

We explored the conditions under which this approach, both crude analysis and subgroup analysis, is most efficient (no loss of SDRs) and will increase sensitivity (figure 4). In this method,

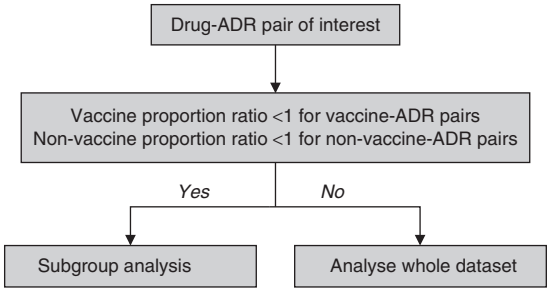

Fig. 4. Flowchart showing a method of determining when to perform subgroup analysis. For non-vaccine-ADR pairs, subgroup analysis by restriction to non-vaccine-ADR pairs is beneficial if the non-vaccine proportion ratio is <1. For vaccine-ADR pairs, subgroup analysis by restriction to vaccine-ADR pairs is beneficial if the vaccine proportion ratio is <1. **ADR** = adverse drug reaction.

subgroup analyses are applied only to ADRs with a vaccine- or non-vaccine proportion ratio <1, while for ADRs with a ratio >1 the ROR is calculated using the whole dataset. Following this approach will lead to additional SDRs being detected, while no SDRs will be lost, and will increase the sensitivity of the data mining algorithm.

Subgroup analysis for ADRs with a vaccine- or non-vaccine proportion ratio >1 will lead to a decrease in ROR and SDRs, but can, however, be informative to compare the magnitude of class effects for the individual drugs in the class. Studying fever convulsions using vaccines only will give an estimate of the disproportionality compared with the other vaccines. In other words, for which vaccines are fever convulsions most frequently reported compared with all other vaccines.

Strengths and Limitations of this Study

To our knowledge, this study is the first to present a method that identified groups of drug-ADR pairs for which subgroup analysis using restriction is beneficial and can increase the sensitivity of the data mining algorithm. We used a large sample size and were able to precisely predict the change in ROR due to restriction to either vaccines or non-vaccines. Paediatric pharmacovigilance is still in its infancy and more studies and research are needed to further develop pharmacovigilance tools that can be applied within this special population.

Our study also has limitations. First, we only investigated the influence of the proportion of vaccines within the database on the estimates of the ROR. We did not study any of the other data mining algorithms for signal detection. However, given the relatedness of analyses and dependence on the same underlying data, it is likely that the results would be comparable. Second, in this study we stratified by age, but we did not take other important confounding factors such as sex, time of reporting, country of origin and seriousness of the reports into account. These factors are also to be considered when studying individual safety signals. Third, it should be emphasized that we used paediatric ADR data only. When applying similar methods to adult data, clustering factors might be less important, resulting in fewer newly detected SDRs after subgroup analysis. Studying the influence of other factors, such as seriousness of the report, might also influence the number of detected SDRs within adult data. This might have a larger impact and implication on the daily practice of signal detection. Fourth, subgroup analysis as performed in our study was only possible since we used ADR data on a drug-ADR pair level, which allowed us to make two mutually exclusive datasets. When performing the analysis on an ICSR level, division of the data into two strict groups might not be possible. Fifth, the data used for this study dates back to 2006. More up-to-date data is currently available; however, using more recent data would not have influenced the methodological approach and conclusions of this study apart from the comparisons of SDRs with information from the SPCs. Finally, a subgroup analysis by restriction to either vaccines or non-vaccines has an influence on the number of drug-ADR pairs that can be used to calculate an ROR. Since restriction decreases the number of pairs included, this also has its effect on the 95% CI of the estimates. This can, in part, explain why certain genuine drug-ADR pairs were no longer SDRs after subgroup analysis (false negatives).

Conclusions

The high proportion of vaccine-related reports within paediatric ADR data has a large and mathematically predictable impact on signal detection in paediatric ADR data. Subgroup anal-

ysis reveals new SDRs that potentially represent genuine safety signals. However, depending on the distribution of vaccine/non-vaccine reports, SDRs can also disappear after subgroup analysis. The most inclusive and sensitive signal detection method would be the combination of a crude and a subgroup-based data mining approach based on the vaccine- or non-vaccine proportion ratio.

Acknowledgements

TEDDY (Task force in Europe for Drug Development for the Young) was funded by the European Community's 6th Framework Programme, project No LSHB-CT-2005-005216. The funding agency had no role in the collection of data, analysis or interpretation of the data, or the decision to submit.

Data from the WHO Collaborating Centre for International Drug Monitoring was used. The information is not homogeneous at least with respect to origin or likelihood that the pharmaceutical product caused the adverse reaction.

The opinions and conclusions expressed in this article are not necessarily those of the UMC, the various national centres or the WHO.

Katia Verhamme has been involved as project coordinator in analyses conducted by various pharmaceutical companies, and received unconditional research grants from Pfizer, Yamanouchi and Boehringer-Ingelheim, none of which are related to the subject of this study. Miriam Sturkenboom leads a research group that is conducting research for pharmaceutical companies through non-conditional grants, none of which is related to this research. All other authors have no conflicts of interest to declare that are directly relevant to the content of this study.

We would like to thank the anonymous peer reviewers whose comments greatly improved this manuscript.

References

1. Hartmann K, Doser AK, Kuhn M. Postmarketing safety information: how useful are spontaneous reports? Pharmacoepidemiol Drug Saf 1999 Apr; 8 Suppl. 1: S65-71
2. Hauben M, Madigan D, Gerrits CM, et al. The role of data mining in pharmacovigilance. Expert Opinion Drug Saf 2005 Sep; 4 (5): 929-48
3. Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf 2009 Jun; 18 (6): 427-36
4. 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 Oct-Nov; 10 (6): 483-6
5. van Puijenbroek EP, Bate A, Leufkens HG, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf 2002 Jan-Feb; 11 (1): 3-10
6. Almenoff JS, LaCroix KK, Yuen NA, et al. Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. Drug Saf 2006; 29 (10): 875-87
7. Meyboom RH, Lindquist M, Egberts AC, et al. Signal selection and follow-up in pharmacovigilance. Drug Saf 2002; 25 (6): 459-65
8. Practical aspects of signal detection in pharmacovigilance. Report of CIOMS Working Group VIII. Geneva: CIOMS, 2010
9. Zeinoun Z, Seifert H, Verstraeten T. Quantitative signal detection for vaccines: effects of stratification, background and masking on GlaxoSmithKline's spontaneous reports database. Hum Vaccin 2009 Sep; 5 (9): 599-607
10. Hopstadius J, Noren GN, Bate A, et al. Impact of stratification on adverse drug reaction surveillance. Drug Saf 2008; 31 (11): 1035-48
11. 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
12. Hauben M, Patadia V, Gerrits C, et al. Data mining in pharmacovigilance: the need for a balanced perspective. Drug Saf 2005; 28 (10): 835-42
13. Aagaard L, Weber CB, Hansen EH. Adverse drug reactions in the paediatric population in Denmark: a retrospective analysis of reports made to the Danish Medicines Agency from 1998 to 2007. Drug Saf 2010; 33 (4): 327-39
14. Kimland E, Rane A, Ufer M, et al. Paediatric adverse drug reactions reported in Sweden from 1987 to 2001. Pharmacoepidemiol Drug Saf 2005 Jul; 14 (7): 493-9
15. Aagaard L, Hansen EW, Hansen EH. Adverse events following immunization in children: retrospective analysis of spontaneous reports over a decade. Eur J Clin Pharmacol 2011; 67 (3): 283-8
16. Niu MT, Erwin DE, Braun MM. Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. Vaccine 2001 Sep 14; 19 (32): 4627-34
17. Braun MM. Toward better vaccine safety data and safer vaccination. Pediatrics 2008 Mar; 121 (3): 625-6
18. Report from the WHO Collaborating Centre for International Drug Monitoring. Activities: July 2005-June 2006 [online]. Available from URL: http://www.who-umc.org/graphics/9594.pdf [Accessed 2012 Feb 1]
19. Lindquist M. VigiBase, the WHO Global ICSR database system: basic facts. Drug Inf J 2008; 42: 409-19
20. WHO Collaborating Centre for Drug Statistics Methodology. Guidelines for ATC classification and DDD assignment [online]. Available from URL: http://www.whocc.no/ [Accessed 2012 Feb 1]
21. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Clinical investigation of medicinal products in the pediatric population E11. 2000 (updated 2000) [online]. Available from URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E11/Step4/E11_Guideline.pdf [Accessed 2012 Feb 1]
22. Evans SJ. Stratification for spontaneous report databases. Drug Saf 2008; 31 (11): 1049-52

Correspondence: *Sandra de Bie*, MD, MSc, Erasmus University Medical Center, Department of Medical Informatics, Room Ee-2159, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands. E-mail: s.debie@erasmusmc.nl