PHARMACOEPIDEMIOLOGY AND DRUG SAFETY 2014; **23**: 787–794
Published online 3 January 2014 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/pds.3556

ORIGINAL REPORT

The upper bound to the Relative Reporting Ratio—a measure of the impact of the violation of hidden assumptions underlying some disproportionality methods used in signal detection

Lionel Van Holle*,† and Vincent Bauchau†

Vaccine Safety Research Group, Vaccine Clinical Safety and Pharmacovigilance, GlaxoSmithKline Vaccines, Wavre, Belgium

ABSTRACT

Purpose For disproportionality measures based on the Relative Reporting Ratio (RRR) such as the Information Component (IC) and the Empirical Bayesian Geometrical Mean (EBGM), each product and event is assumed to represent a negligible fraction of the spontaneous report database (SRD). Here, we provide the tools for allowing signal detection experts to assess the consequence of the violation of this assumption on their specific SRD.

Methods For each product–event pair (P–E), a worst-case scenario associated all the reported events-of-interest with the product of interest. The values of the RRR under this scenario were measured for different sets of stratification factors using the GlaxoSmithKline vaccines SRD. These values represent the RRR upper bound that RRR cannot exceed whatever the true strength of association.

Results Depending on the choice of stratification factors, the RRR could not exceed an upper bound of 2 for up to 2.4% of the P–Es. For Engerix™, 23.4% of all reports in the SDR, the RRR could not exceed an upper bound of 2 for up to 13.8% of pairs. For the P–E Rotarix™-Intussusception, the choice of stratification factors impacted the upper bound to RRR: from 52.5 for an unstratified RRR to 2.0 for a fully stratified RRR.

Conclusions The quantification of the upper bound can indicate whether measures such as EBGM, IC, or RRR can be used for SRD for which products or events represent a non-negligible fraction of the entire SRD. In addition, at the level of the product or P–E, it can also highlight detrimental impact of overstratification. © 2014 The Authors. *Pharmacoepidemiology and Drug Safety* published by John Wiley & Sons, Ltd.

KEY WORDS—upper bound; conservative bias; overstratification; Relative Reporting Ratio; EBGM; IC; pharmacoepidemiology

Received 24 September 2013; Revised 31 October 2013; Accepted 5 November 2013

INTRODUCTION

The most frequently used methods for detecting signals in spontaneous reports are numerator-based, the most widely used being disproportionality analyses (DPAs).[1–6] These methods partially overcome the main limitation of passive surveillance data: the lack of reliable estimates of the exposed population.

Some assumptions must be respected for generating unbiased estimates of association between products and events from spontaneous report databases (SRDs). These assumptions are often not explicit in the literature and are often violated in practice. However, their full understanding is needed for performing efficient quantitative signal detection based on DPA.

When using disproportionality measures (DPMs) based on the Relative Reporting Ratio (RRR)[5], such as the Information Component (IC)[1,6] or Empirical Bayesian Geometric Mean (EBGM)[2,5], an observed number of reports is compared with the expected number derived from all events and all products (including the product and event of interest). The assumption being that the observed and the expected numbers can be treated as independent. However, the larger the fraction of the safety database represented by the product (or event) of interest, the less independent is the observed from the expected number evaluated using the RRR, IC, and EBGM. Here, we show that this non-independence between observed and expected numbers results in an artefactual upper

*Correspondence to: L. Van Holle, Vaccine Safety Research Group, Vaccine Clinical Safety and Pharmacovigilance, GlaxoSmithKline Vaccines, Parc de la Noire Epine, Avenue Fleming, 20, 1300—Wavre, Belgium. E-mail: lionel.f.van-holle@gsk.com

†Both authors state that the manuscript has not been published elsewhere.

bound to the observed–expected (OE) ratios and a conservatively biased measure of the association. Similar findings have been previously demonstrated for standardized mortality ratios, which are also OE ratios.[7]

Some measures of disproportionality, such as the Proportional Reporting Ratio (PRR)[3] and the Reporting Odds Ratio (ROR)[4], use the other products, at the exclusion of the product of interest, as a comparator group for testing if the event of interest is associated with the product of interest. Gogolak[8] discussed the fact that the differences between RRR, PRR, and ROR calculations are small when large data sets are involved and no product dominates the population. However, if the selected product is large compared with other products, dramatically different ratios can result.[8]

There is currently no established criterion to determine whether the hidden assumption that each product and event is a negligible fraction of the SRD is respected for ensuring independence between the observed and expected. Important violation of this assumption should prevent the use of DPMs based on the RRR (DPM_{RRR}) for some product–event pairs (P–Es), to avoid potential misinterpretation caused by the artefactual upper bound and conservative bias. There are a number of general recommendations, guidelines, and warnings in the literature, but they are hard to apply in practice to SRDs as their size and diversity evolve continuously. For example, Almenoff stated that the safety database "should be of *sufficient size* and *diversity* to serve as a suitable *background* for evaluating disproportionate reporting."[9] Norén's thesis discussing the IC[10] warned about the limitation of the OE ratio, which cannot exceed the inverse of the marginal relative reporting for each event. He concluded that "this limits the usefulness of OE ratios as a measure of disproportionality to events that are *reasonably rare*."[10]

The fact that DPM_{RRR} measures are "always applicable" is sometimes presented as an advantage[4] as in practice, safety physicians and reviewers consider the disproportionality scores not only for rare events but for all events. Indeed, real-life use of the products can potentially involve subpopulations excluded from clinical trials, justifying that a safety surveillance system is not restricted to rare events.[11] Without a quantifiable criterion to ascertain the violation of the underlying DPM_{RRR} assumptions, their use could result in a misleading confidence in the absence of association between products and reported events characterized by low DPM_{RRR} values.

The first objective of this article was to define mathematically the artefactual upper bound to the RRR. The second objective was to illustrate mathematically the conservative bias of the RRR compared with the PRR. The third objective, based on quantification of the artefactual upper bound, was to determine the impact of violating this assumption through assessing the proportion of vaccine–event pairs for which the DPM_{RRR} cannot detect a signal of disproportionate reporting. The fourth and final objective was to provide the tools for signal detection experts to determine *a priori* whether DPM_{RRR} are suboptimal for their application to a specific SRD or to a product or P–E pair.

METHODS

The measures of disproportionality

The RRR is defined as in Equation (1), with terms defined in Table 1.

$$\mathrm{RRR} = \frac{\text{Observed}}{\text{Expected}} = \frac{A}{\frac{(A+B)*(A+C)}{(A+B+C+D)}} = \frac{A*(A+B+C+D)}{(A+B)*(A+C)} \quad (1)$$

Other DPM are based on the RRR: the Bayesian EBGM and IC. Their implementation differs, but both provide shrinkage towards the null.[6]

The PRR[3] is defined as

$$\mathrm{PRR} = \frac{A*(C+D)}{(A+B)*C} \quad (2)$$

For signal detection, these measures are calculated for all observed P–E pairs in the SRD. Signals of disproportionate reporting are generally defined when the DPM, or its lower confidence interval limit, is above a given threshold.[12]

The RRR upper bound

For each P–E, the RRR upper bound was defined as the value that RRR cannot exceed should all the reported cases with the event of interest be associated with the product of interest. It corresponds to a

Table 1. Two-by-two contingency table

Entire safety database	Event of interest	Other events
Product of interest	*A*	*B*
Other products	*C*	*D*

DOI: 10.1002/pds

worst-case safety scenario defined in a framework where the constraints inherent to the SRD of interest are applied: typically, the fraction each reported product (or event) represents compared with all reported products (or events).

When simulating this worst-case safety scenario, we ended up with the following contingency table (Table 2):

Table 2. Two-by-two contigency table under the worst-case safety scenario

Entire safety database	Event of interest	Other events
Product of interest	A'	B'
Other products	C'	D'

where $C' = 0$ and $A' = A + C$.

Consequently, the RRR upper bound (RRR_{ub}) is

$$RRR_{ub} = \frac{A'*\left(A' + B' + C' + D'\right)}{\left(A' + B'\right)*\left(A' + C'\right)} = \frac{\left(A' + B' + D'\right)}{\left(A' + B'\right)}$$

$$= \frac{(A + B + C + D)}{(A + B + C)} \leq \frac{(A + B + C + D)}{(A + B)} \quad (3)$$

For example, inequality (3) indicates that for a product representing one quarter of the SRD, the unstratified RRR, and then the unstratified EBGM, could not exceed 4 for any events reported with the product, even in the worst-case safety scenario where all reported cases with the event of interest would be associated to the product of interest.

The majority of the published criteria for defining signals of disproportionate reporting using the EBGM involve a threshold value of 2 for the EBGM or for the EB05 (the lower limit of the 90% credibility interval around the EBGM).[12] In Table 3, we defined an intuitive categorization of the upper bound to help quantify and interpret the RRR upper bound. This categorization should be considered to apply in general terms to a number of situations but may not strictly apply in specific cases.

The conservative bias of RRR relative to PRR as measures of the strength of association

Let us define P as the prevalence of the product of interest among the reported cases presenting the event of interest ($P = \frac{A}{(A+C)}$) and P_V the prevalence of the product of interest in the SRD ($P_V = \frac{(A+B)}{(A+B+C+D)}$).

Let us define the Conservative Bias Factor (CBF) as the ratio between RRR and PRR:

$$CBF = \frac{(A + B + C + D)*C}{(C + D)*(A + C)} = \frac{(A + B + C + D)*[(A + C) - A]}{[(A + B + C + D) - (A + B)]*(A + C)} = \frac{1 - P}{1 - P_V} \quad (4)$$

Under the null assumption of no association between the product of interest and the event of interest, $P = P_V$, and thus, there is no difference between RRR and PRR. When the product and event of interest are assumed to represent a negligible fraction, then P and P_V would tend towards 0 and CBF would tend towards 1.

When the product or event of interest does not represent a negligible fraction, RRR and PRR values may diverge in cases where there is an association between the product and the event of interest. When the association is positive ($P > P_V$), $CBF < 1$ and RRR values are smaller than PRR values. When the association is negative ($P < P_V$), $CBF > 1$ and RRR

Table 3. Categorization of the upper bound

Range for the RRR upper bound	Interpretation
[0, 2]	Product–event pairs with an upper bound in this range are ineligible to be a signal of disproportionate reporting by DPM_{RRR} considering the wide use of threshold values based on EBGM larger than 2 or EB05 larger than 2.
[2, 5]	Product–event pairs with an upper bound in this range are theoretically eligible to be signals of disproportionate reporting by DPM_{RRR}. However, the product (or event) of interest represents more than 20% of the database. For these product–event pairs, the conservative bias is considerable and only very frequent events characterized by a very strong association with the product can overcome the conservative bias to be eligible as signals of disproportionate reporting (especially when the threshold is based on the EB05).
[5-10]	Product–event pairs with an upper bound in this range are theoretically eligible to be signals of disproportionate reporting by DPM_{RRR}. The range for the upper bound means that the product (or event) of interest is between 10% and 20% of the database. For these product–event pairs, the conservative bias is moderate and only frequent events characterized by a strong association with the product can overcome the conservative bias to be eligible as signals of disproportionate reporting (especially when the threshold is based on the EB05).
>10	With the product (or event) of interest representing less than 10% of the database, we can assume the conservative bias to have only a moderate impact in masking some signals of disproportionate reporting.

DOI: 10.1002/pds

values are larger than PRR values. This demonstrates that the RRR provides a more conservative measure of the association than the PRR. The higher the association is, the more divergent the RRR and PRR values are.

Assessment of the impact of stratification on the upper bound

The stratification factors investigated are shown in Table 4.

Table 4. Definition of the stratification factors

Stratification factor	Categories
Sex (*S*)	Male, Female, Unknown
Age (*A*)	[0, 0.5 year], [0.5, 1 year], [1, 10 years], [10, 19 year], [19, 65 years], [65+ years], [Unknown]
Year (*Y*)	[1960, 1990], [1991, 1993], [1994, 1996], [1997, 1999], [2000, 2002], [2003, 2005], [2006, 2008], [2009, 2011]
Region (*R*)	USA, Canada + New Zealand + Australia, Europe, Latin America, Rest of the World, Unknown

For a stratified RRR, the upper bound can still be derived by applying the same principle ($A' = A + C$) across the different strata (e.g., N different strata, $i = 1, \ldots, N$) before calculating the common RRR (using the Mantel-Haenszel method)[13].

$$\mathrm{RRR}_{\mathrm{ub}}^{\mathrm{strat}} \leq \frac{(A+C)}{\sum_{i=1}^{N} \frac{(A_i+B_i+C_i)*(A_i+C_i)}{(A_i+B_i+C_i+D_i)}} \qquad (5)$$

with terms defined in Table 5. The impact of different stratifications (None, *S*, *A*, *R*, *Y*, *SA*, *SR*, *SY*, *AR*, *AY*, *RY*, *SAR*, *SAY*, *SRY*, *ARY*, *SARY*) on the $\mathrm{RRR}_{\mathrm{ub}}$ was evaluated.

Table 5. Two-by-two contingency table for stratum i ($i = 1, \ldots, N$)

Stratum i ($i = 1, \ldots, N$)	Event E	Other events
Product P	A_i	B_i
Other products	C_i	D_i

The company safety database

The GlaxoSmithKline (GSK) vaccines SRD is the Operating Company Event Accession and Notification System (OCEANS). As of 1 February 2010, OCEANS contained 147 015 vaccine-related spontaneous reports distributed across 28 425 distinct vaccine–event pairs, involving 45 distinct GSK suspected vaccines and 4331 distinct MedDRA[‡] preferred terms.

The quantification of the $\mathrm{RRR}_{\mathrm{ub}}$ was performed at different levels:

- The GSK SRD restricted to vaccines.
- One product for which the hidden assumptions underlying $\mathrm{DPM}_{\mathrm{RRR}}$ was suspected to be violated: Engerix™.
- A specific P–E pair: Rotarix™-Intussusception.

RESULTS

In this section, we present the quantification of the RRR upper bound at the following: (1) the GSK vaccines SRD; (2) the single product Engerix™; and (3) the P–E pair Rotarix™-intussusception.

At the GSK vaccines SRD level

Under the worst-case safety scenario, up to 2.4% of all P–E pairs had an upper bound below 2 for the full stratification (SARY) by Sex, Age, Region, and Year of the RRR (Figure 1). Depending on the stratification factor used, between 5.8% and 12.5% of the P–E pairs had an $\mathrm{RRR}_{\mathrm{ub}}$ below 5. And at least 13.3% of the P–E pairs had an $\mathrm{RRR}_{\mathrm{ub}}$ below 10 whatever the choice of stratification factors. Should the SARY stratification be used, then this percentage may increase up to 24%.

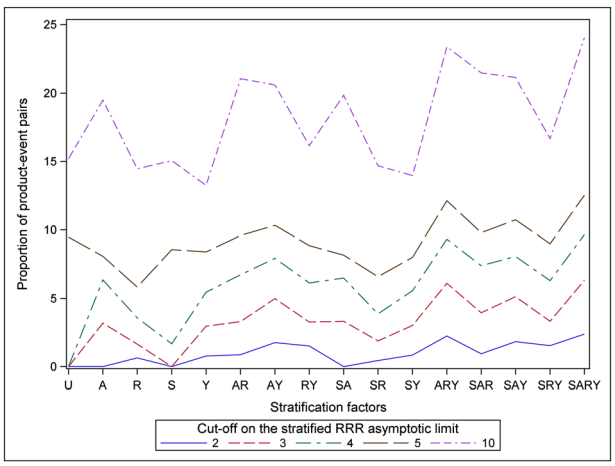

Figure 1. Proportion of product–event pairs having a stratified Relative Reporting Ratio upper bound below the cutoff value for different choices of stratification

[‡]Medical Dictionary for Regulatory Activities (MedDRA) is a clinically validated international medical terminology used by regulatory authorities and the regulated biopharmaceutical industry throughout the entire regulatory process, from pre-marketing to post-marketing activities, and for data entry, retrieval, evaluation, and presentation.

DOI: 10.1002/pds

At the level of the product Engerix™

Engerix™ represents 34 347 reports in OCEANS, 23.4% of all reports in the GSK vaccines SRD. As such, we can mathematically conclude that no P–E pair can have an unstratified RRR exceeding the upper bound of $1/0.234 = 4.28$ when using DPM_{RRR} in OCEANS (inequality in equation (3)).

Figure 2 highlights the asymptotic behavior of the RRR when tending towards the worst-case safety scenario where the percentage of cases with the event of interest tends to be 100% reported with Engerix™. The simulations were performed for all events reported with Engerix™. Under that worst-case scenario, the unstratified RRR had an upper bound, whereas the PRR tended towards infinity. Figure 2 also highlights the conservative bias of the RRR relatively to the PRR. Indeed, for $PRR < 1$, the RRR values tend to be higher than the PRR values, whereas the opposite is true for $PRR > 1$ with extreme differences observed for events that are mainly reported to Engerix™ (i.e., for events with the strongest association).

In practice, when attributing all cases with the event of interest to the product of interest, we slightly increased, in that worst-case safety scenario, the total number of reports for the product of interest (Equation (3)). That slight increase is of different magnitude depending of the prevalence of event of interest in the safety database. That explains the width of both red and blue lines in Figure 2: the lower part of the two lines corresponds to the disproportionality values derived for the pair Engerix™-Pyrexia. Indeed, Pyrexia is the most frequently reported event in the SRD and attributing all Pyrexia events to Engerix™ resulted in an upper bound lower than for the other events characterized by a much smaller prevalence in the SRD.

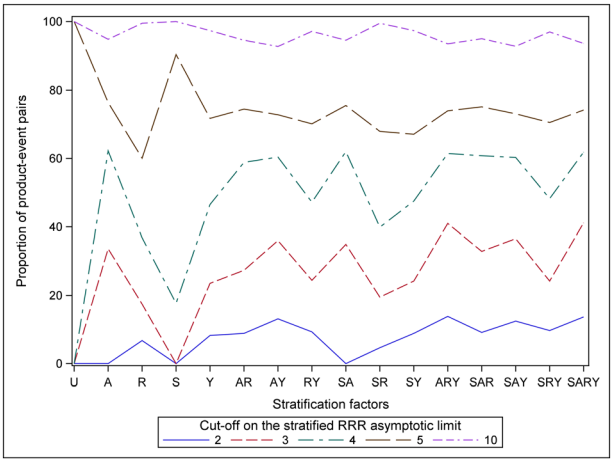

Figure 3. Proportion of product–event pairs having a stratified Relative Reporting Ratio upper bound below the cutoff value for different choices of stratification for Engerix™

The computation of the RRR_{ub} allowed us to determine that up to 13.8% of the Engerix™-event pairs had an upper bound below 2 for the RRR stratified by Age, Region, and Year (Figure 3). Under the worst-case scenario, at least 60% of the P–E pairs had an upper bound below 5 whatever the choice of stratification of the RRR. Similarly, at least 92.7% of the P–E pairs had an upper bound below 10.

The product–event pair Rotarix™-Intussception

The choice of stratification factors dramatically impacted the RRR_{ub} from 52.5 for the unstratified RRR to 2.0 for a fully stratified RRR (Figure 4, Table 6). The calculated RRR was always very close to its upper bound whatever the choice of stratification (Table 6). Under the standard

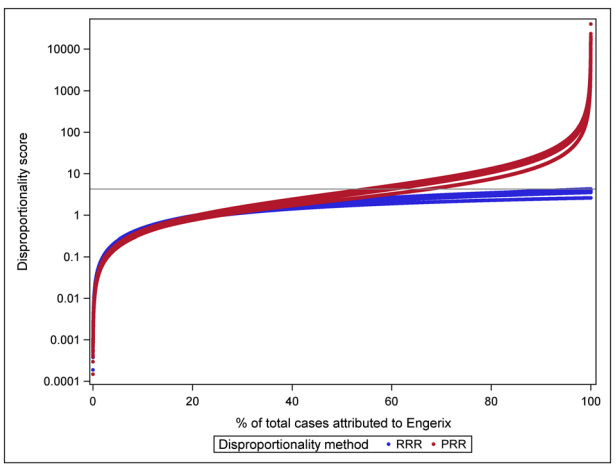

Figure 2. Illustration of the upper bound associated with the disproportionality score Relative Reporting Ratio for Engerix™

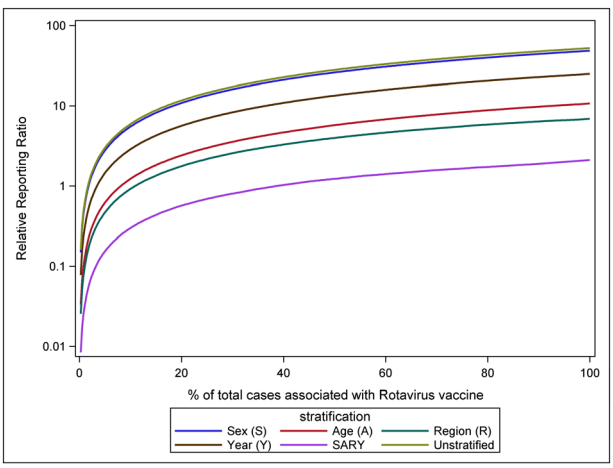

Figure 4. Illustration of the upper bounds by choice of stratification factors for the Rotarix™-Intussusception pair

DOI: 10.1002/pds

Table 6. Evolution of the RRR upper bound of the Rotarix™-Intussusception pair for different stratification factors

Stratification	RRR	RRR_{ub}
None	49.8427	52.5029
A	10.2208	10.7663
R	6.5118	6.8594
S	46.0812	48.5406
Y	23.7123	24.9778
SARY	1.9352	2.0384

SARY stratification used in routine signal detection, the P–E pair Rotarix™-Intussusception could not exceed the fixed, predefined threshold of 2 on the EB05[15] whatever the strength of association. Consequently, the fully stratified EBGM was underpowered for detecting the Rotarix™-Intussusception pair as a signal of disproportionate reporting in OCEANS (Table 6).

DISCUSSION

The quantification of the RRR upper bound illustrated the ability of this measure, or *tool*, to detect and quantify potential issues related to the use of DPM_{RRR}. It gave an estimation of the percentage of P–E pairs in an SRD for which the hidden assumptions behind DPM_{RRR} use were violated to an extent that it led to an estimation of the strength of association by DPM_{RRR} so conservatively biased that signals of disproportionate reporting were impossible or unlikely to be detected for these pairs even in the worst-case safety scenario in which all reported cases with the event of interest are associated with the product of interest.

For the GSK vaccines SRD, the high proportion of P–E pairs having an RRR upper bound below 10 suggests that the use of DPM_{RRR} on this SRD is not recommended. The use of DPM_{RRR} for products representing a large proportion of reports in the SRD can result in very low sensitivity (as illustrated with Engerix™). The quantification of the upper bound sheds new light on a previous study that concluded that the use of a threshold of 2 on the EB05 resulted in an undersensitive signal detection algorithm[15] for screening GSK vaccines SRD.

The cutoff value of 2 used for defining categories on the RRR_{ub} was chosen for highlighting the inability to detect signals of disproportionate reporting when they are defined based on a cutoff value of 2 on the EB05 or EBGM as it is a commonly performed.[12] The cutoff values of 5 and 10 were arbitrarily chosen.

The key literature about DPM_{RRR} to date has been based on analyses performed on large regulatory databases (FDA, WHO).[1,2,14] These SRDs are characterized by a very large number and diversity of products and events allowing the assumptions underlying DPM_{RRR} to be better respected than in small or medium company SRD. Some comparisons of DPM_{RRR} (IC) versus other DPMs (PRR, ROR, and so forth) have been performed previously but were based on the Netherlands Pharmacovigilance Foundation Lareb SRD,[4] which, although smaller than FDA and WHO SRDs, contains a substantial number of distinct drugs and events. They concluded that "the different measures used are broadly comparable when four or more cases per combination have been compared." However, for a vaccine or small/medium-size drug manufacturer's SRD, the hidden assumptions behind DPM_{RRR} are less likely to hold. It cannot be assumed up-front that each product represents only a negligible fraction of the SRD and dramatic differences between RRR and PRR can arise, as shown by Gogolak[8] and demonstrated here mathematically and empirically. The danger is to extrapolate some conclusions from a context where the underlying assumptions are likely to be respected to a context where they are more likely not to be. The calculation of the RRR upper bound can help the signal detection expert to determine how relevant the DPM_{RRR} and stratification choices are for a specific SRD (especially for small or medium companies).

Even in very large SRDs characterized by a very large number of products or events, some P–E pairs could have a very low RRR upper bound when the event of interest (or product of interest) occurs in a population having very different demographic characteristics compared with the other events (products) reported in the SRD. Indeed, a DPM_{RRR} stratified by these demographic characteristics will result in a significant number of strata not contributing to any observed cases and consequently discarded from DPM_{RRR} estimation. In such situations, the product of interest represents a non-negligible fraction of the total number of cases observed among the contributing strata, resulting in a violation of the underlying assumption and in a conservatively biased estimate of the strength of the association.

The quantification of the RRR upper bound also provides the signal detection expert with the means to interpret the differences in DPM_{RRR} values under different choices of stratification factors. Without these upper bounds, a signal detection expert would have tended to give more credibility to the fully stratified DPM_{RRR} compared with the unstratified for Rotarix™-Intussusception P–E pair, for example. Replicating the epidemiological interpretation of stratification, he would have attributed the difference between the two estimates to better comparability between the product of interest and all products after

DOI: 10.1002/pds

stratification by sex, age, region, and year of reporting. By keeping in mind some general warnings about the danger of overstratification,[16] he might have oscillated between the two explanations. The upper bounds allow him to see that the fully stratified RRR is strongly conservatively biased as the corresponding upper bound is below 2. He can consequently dismiss this DPM for signal detection purpose and decide to look at other DPMs, like ROR or PRR, or look at DPM_{RRR} measures with other sets of stratification factors with high values for the upper bound, or rely on qualitative clinical criteria.

CONCLUSION

Disproportionality measures based on the RRR (such as IC and EBGM) are characterized by an asymptotic upper bound and a conservative bias in the estimation of the strength of association compared with the PRR. The higher the association between a product and an event, the greater the conservative bias is. The larger the fraction represented by a product in a given SRD, the lower the RRR upper bound.

Under the assumption of each product and event representing a negligible fraction of the entire SRD, the upper bound tends towards infinity and the conservative bias is null. Large regulatory SRDs, characterized by a large number of products and events, are likely to meet the criteria underlying this assumption. However, small-sized and medium-size SRDs may violate these assumptions with consequences that have not been previously assessed.

The calculation of the RRR upper bound can help to determine how relevant the DPM_{RRR} and stratification choices are for a specific SRD, product or, P–E pair under surveillance. It can help guide the decision to use other DPMs that do not rely on the same assumptions, like PRR or ROR, or avoid overstratification for a given P–E pair, for example. It can also help in interpreting low values of DPMs based on the RRR.

We advise systematic computation of the RRR upper bound along with the measure of disproportionality, whenever using one based on the RRR.

CONFLICT OF INTEREST

Engerix and Rotarix are trademarks of the GlaxoSmithKline Group of Companies.

Both authors are employees of GlaxoSmithKline Vaccines.

KEY POINTS

- The assumption that each product and event represents a negligible fraction of the Spontaneous Report Database (SRD) underlies disproportionality measures based on the Relative Reporting Ratio (DPM_{RRR}).
- There is currently no established measure to determine the impact of the violation of this assumption.
- The DPM_{RRR} are characterized by an artefactual upper bound to the observed–expected ratio and a conservatively biased measure of the association compared to the Proportional Reporting Ratio.
- The RRR upper bound can be derived under a worst-case safety scenario in which every reported event of interest is considered as reported with the product of interest.
- Based on the RRR upper bound, conclusions can be drawn on the capability of DPM_{RRR} to detect signals depending on the SRD, product, product–event pair, and choice of stratification.

ETHICS STATEMENT

A poster was presented at the International Conference of Pharmacoepidemiology, Montreal (2013).

ACKNOWLEDGEMENTS

Editing and publication co-ordinating services were provided by Juliette Gray (XPE Pharma & Science, Wavre, Belgium), Veronique Delpire (Words and Science, Brussels, Belgium), and Mandy Payne (Words and Science, Brussels, Belgium). GlaxoSmithKline Biologicals SA funded all costs associated with the development and the publishing of the present manuscript.

REFERENCES

1. Bate A. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998; **54**: 315–321.
2. Dumouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am Stat* 1999; **53**: 177–190.
3. Evans S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001; **10**: 483–486.
4. van Puijenbroek EP. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf* 2002; **11**: 3–10.
5. Almenoff JS. Novel statistical tools for monitoring the safety of marketed drugs. *Clin Pharmacol Ther* 2007; **82**(2): 157–166.
6. Bate A, Evans SJW. Quantitative signal detection using spontaneous ADR reporting. *Pharmacoepidemiol Drug Saf* 2009; **18**: 427–436.
7. Jones ME. Bias in the standardized mortality ratio when using general population rates to estimate expected number of deaths, *Am J Epidemiol* 1998; **148**(10): 1012–1017.
8. Gogolak VV. The effect of backgrounds in safety analysis: the impact of comparison cases on what you see, *Pharmacoepidemiol Drug Saf* 2003; **12**: 249–252.

9. Almenoff J. Perspectives on the use of data mining in pharmacovigilance. *Drug Saf* 2005; **28**(11): 981–1007.
10. Norén N. Statistical methods for knowledge discovery in adverse drug reaction surveillance, Matematiska Institutionen 2007, http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-6764 (last accessed on 28 September 2012).
11. Kulldorf M. A maximized sequential probability ratio test for drug and vaccine safety surveillance. *Sequential analysis* 2011; **30**: 58–78,
12. Deshpande G. Data mining in drug safety – review of published threshold criteria for defining signals of disproportionate reporting, *Pharmaceut Med* 2010; **24** (1): 37–43.
13. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease, *J Natl Cancer Inst* 1959; **22** (4): 718–749.
14. Szarfman A. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous report database. *Drug Saf* 2002; **25**(6): 381–392.
15. Van Holle L. Optimization of a quantitative signal detection algorithm for spontaneous reports of adverse events post immunization. *Pharmacoepidemiol Drug Saf* 2012. DOI: 10.1002/pds.3392.
16. Hopstadius J. Impact of stratification on adverse drug reaction surveillance. *Drug Saf* 2008; **31** (11): 1035–1048.