LEADING ARTICLE

Drug Safety 2005; 28 (10): 843-850
0114-5916/05/0010-0843/$34.95/0

Impact Analysis of Signals Detected from Spontaneous Adverse Drug Reaction Reporting Data

Patrick Waller,[1] *Emma Heeley*[2] and *Jane Moseley*[3]

1 Patrick Waller Limited, Consultancy in Pharmacovigilance and Pharmacoepidemiology, Southampton, UK
2 Post-Licensing Division, Medicines and Healthcare products Regulatory Agency, London, UK
3 Pharmacoepidemiology Research Team, Medicines and Healthcare products Regulatory Agency, London, UK

Abstract

This paper describes a new method of prioritising signals of potential adverse drug reactions (ADRs) detected from spontaneous reports that is called impact analysis. This is an interim step between signal detection and detailed signal evaluation. Using mathematical screening tools, large numbers of signals may now be detected from spontaneous ADR databases. Regulatory authorities need to rapidly prioritise them and focus on those that are most likely to require significant action. Using two scores ranging from one to 100, each with three input variables, signals may be categorised in terms of the strength of evidence (E) and the potential public health impact (P). In a two-by-two figure with empirically derived cut-off points of ten (the logarithmic mean) for each score, signals are placed in one of four categories (A–D) that are ranked according to their priority (A being the highest and D the lowest). A sensitivity analysis is then performed that tests the robustness of the categorisation in relation to each of the six input variables. A computer program has been written to facilitate the process and reduce error. Further work is required to test the feasibility and value of impact analysis in practice.

The main purpose of spontaneous adverse drug reaction (ADR) reporting schemes is to provide 'signals' of previously unrecognised hazards related to the use of marketed medicines. A signal may be defined as an alert from any available data source[1] or, as in the following WHO definition,[2] based on reporting of individual cases:

"Reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously. Usually more than a single report is required to generate a signal, depending upon the seriousness of the event and the quality of the information".

By its nature, a signal cannot be regarded as definitive but indicates the need for further enquiry or action.[2] Detailed signal evaluation using all the relevant data that are available is a complex[3] and resource intensive process. Hence, there is a need for regulatory authorities to screen and prioritise signals with a clear focus on their potential impact on public health. This has increasingly become the case with the routine use of mathematical approaches (e.g. proportional reporting ratios [PRRs]) as aids to signal detection from spontaneous ADR databases,[4] as it has been shown that large numbers of signals may be detected.[5] At present only principles[6-8] and qualitative criteria are available – for

example 'SNIP', which takes into account the strength of signal, whether it is really new, the clinical importance of the ADR and the potential for prevention,[9] and a system of triage recently proposed by the WHO that is based on eight algorithms.[10]

This paper describes 'impact analysis', a new, quantitative tool for prioritising signals as an interim step between signal detection and detailed signal evaluation. The purpose of the tool is to focus detailed signal evaluation on those signals that are the strongest and are most likely to have an impact on public health. The method is only applicable to signals derived from spontaneous ADR reporting. Although it utilises a disproportionality score, the method could be used regardless of whether or not a signal has been detected by a mathematical approach. Impact analysis has been developed for use at the UK Medicines and Healthcare products Regulatory Agency (MHRA) but could be applied to other settings with appropriate modification and testing.

1. Development of Impact Analysis

The concept of signal impact analysis was put forward during the development of a future model for pharmacovigilance[11] and is an attempt to improve on the SNIP criteria[9] using a quantitative approach. It was suggested that an empirical mathematical approach incorporating sensitivity analysis could be developed to support this function.[11] The suggested key factors were causality, frequency, health consequences and risk predictors. The original idea was that these might be synthesised into an output that indicated both the public health importance of the suspected ADR and potential for prevention, based on the attributable risk amongst exposed patients. It should be recognised that the purpose of impact analysis is not to decide whether or not a signal represents a true drug effect but what priority should be given to a detailed evaluation (which, when performed, would include an assessment of causality and might lead to detailed benefit-risk analysis).

The approach described in section 2 was subsequently developed from first principles by considering all the potential factors that might be taken into account in assessing the importance of a signal from the perspective of a regulatory authority. It was decided to focus entirely on signals detected from spontaneous ADR reporting data for two reasons. First, there were practical difficulties in the application of the same scales to data arising from multiple sources and, second, because signals that arise from sources other than spontaneous ADR reporting almost invariably require detailed evaluation.

The scoring system and figures were developed empirically by using examples and calculating all the possible combinations of the input variables. A multiplicative approach was used whereby each score was determined by multiplying the three input variables. The underlying concept was to consider the first input to each scale as a key indicator that is then potentially downgraded by multiplication of inputs that are invariably less than one. Although it was recognised that cut-off points on the scales could potentially be defined at any point, the initial aim was to develop a system based on the logarithmic mid-point for both component scales (see section 2). Once the basic method was firmly established a purpose-designed computer program was developed to automate the process.

2. Description of the Method Used to Perform Impact Analysis

Impact analysis is a mathematical tool for prioritising signals arising from spontaneous ADR data. It summarises the impact of a signal through two scores on scales that both range from a minimum of one to a maximum of 100. The two scales are as follows.

- Evidence score – this summarises the strength of evidence for a causal association.
- Public health score – this summarises the potential public health implications.

Using empirical cut-off points of 10 (the logarithmic mean) for both scales, each signal can be placed into one of four categories, each of which defines the level of priority and implies a particular course of action (figure 1).

2.1 Inputs Required

Each of the two scores takes into account three inputs that are either intrinsically numerical or use an ordinal scale to translate a qualitative assessment

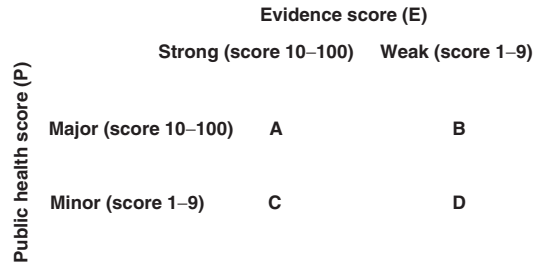

Fig. 1. Impact analysis categories. **A** = high priority – detailed evaluation needed; **B** = there is a need to gather more information; **C** = low priority but still needs to be addressed; **D** = no action warranted at the present time.

into a numerical variable (see figure 2 for scoring figures).

The components of the evidence score are:
(a) PRR and its lower 95% confidence limit (this is a simple means of taking into account both the magnitude of the signal and the degree of precision of the estimate);
(b) strengths/weaknesses of the case series being considered;
(c) biological plausibility of the putative reaction based on the number of factors supporting plausibility within the range of 0–3.

The components of the public health score are:
(d) the number of cases of the ADR in the population per year since the first ADR was reported for the drug;
(e) the potential health consequences of the ADR (fatal and non-fatal);
(f) the order of magnitude of the reporting rate for the drug/reaction combination during the previous year.

2.2 Calculation of Scores and Categories

Details of the scoring systems for each component (a) to (f) are given in figure 2. When information is not available for a particular input, a default may be entered (as indicated in each figure). To allow for the unforeseen, defaults have been provided for all the inputs although their use should be unnecessary for most of them. The (f) input is most likely to require a default because it requires some exposure data, which may not be immediately available. Use of a default generally tends to move the score towards the logarithmic mean (except for biological plausibility), therefore use of multiple defaults should be avoided.

The evidence score (E) is calculated as:

$E = a \times b \times c$ (range 1–100; 1–9 is categorised as weak and 10–100 as strong).

The public health score (P) is calculated as:

$P = d \times e \times f$ (range 1–100; 1–9 is categorised as minor and 10–100 as major).

Using the values of E and P each signal is categorised into one of four categories: A, B, C or D (as shown in figure 1). The following consequential actions continue logically from the categorisation.

- A is a high priority signal and, therefore, detailed evaluation is required rapidly.
- B is a signal that, were it real would be important, but currently the evidence is weak and, therefore, there is a need to gather more information (e.g. request new studies) in a timely manner.
- C is a signal that is likely to be real but of limited public health importance and it should, therefore, be given low priority and dealt with in due course.
- D is a signal that is unlikely to be important and for which current evidence is weak; therefore, no action is warranted at the present time.

2.3 Comments Relating to Use of the Scoring Figures

Part (a) scoring PRR: this element is based on the mean of the PRR for the drug/reaction combination and its lower confidence limit, thereby reflecting both the strength and precision of the estimate. No judgement is required.

Part (b) scoring for strengths and weaknesses of the evidence: this element is based on a rapid judgement of the overall quality of the relevant series of case reports. The case series should be briefly reviewed and a scientific judgement made as to which of the following five categories should be applied. It is not intended that detailed assessment of causality in individual cases should be undertaken at this stage (as would be necessary in a detailed signal evaluation). When assessing this aspect of the evidence producing the signal, the key issues are the completeness of the data and whether at least some of the cases support a causal effect in terms of temporal association (including the presence or ab-

Part (a): Scoring for proportional reporting ratio (PRR)

Score	Mean of PRR and its lower 95% confidence limit
100	>20
80	>15–20
60 (default)	>10–15
40	>5–10
20	≤5

Part (b): Scoring for strengths and weaknesses of the evidence

Score	Strength of evidence
1	Strong
0.8	Fairly strong
0.6 (default)	Average
0.4	Fairly weak
0.2	Weak

Part (c): Scoring for biological plausibility

Score	No. of factors supporting plausibility
1	3
0.75	2
0.5	1
0.25 (default)	0

Part (d): Scoring for number of cases of the ADR of interest with the suspect drug in the population per year

Score	Cases/year
100	>10
80	>6–10
60	>3–6
40 (default)	>1–3
20	≤1

Limited population adjustment: add 20 to score (to be used when drug use is entirely confined to a particular subpopulation).

Part (e) : Scoring for health consequences

Score	Case fatality	Score	Non-fatal outcome
0.5	>20%	0.5	Permanent major disability
0.4	>10–20%	0.4	Major transient sequelae
0.3	>5–10%	0.3 (default)	Permanent minor disability
0.2 (default)	>1–5%	0.2	Minor transient sequelae
0.1	≤1%	0.1	Inconvenience

Scores for case fatality and non-fatal outcome are added to produce the overall score that has a range of 0.2–1; non-fatal outcome is based on the worst known outcome.

Part (f) : Scoring for order of magnitude of the reporting rate

Score	Reporting rate during the previous year
1	>1 in 1000
0.75	>1 in 10 000 to 1 in 1000
0.5 (default)	>1 in 100 000 to 1 in 10 000
0.25	≤1 in 100 000

Fig. 2. Scoring tables for impact analysis. **ADR** = adverse drug reaction.

sence of positive rechallenges or dechallenges) and the presence or absence of alternative causes. The following guidelines should be used to determine the most appropriate category:

- strong – a series of well documented cases with no alternative causes and normally with at least one positive rechallenge;
- fairly strong – a series of generally well documented cases with few alternative causes and at least one positive dechallenge;
- average – a series of cases of variable quality;
- fairly weak – a series of cases that have significant limitations regarding plausible temporal associations and/or for which there are likely alternative explanations;
- weak – a series of cases that are generally incompletely documented, lack plausible temporal associations and/or are generally explainable by alternative causes.

Part (c) scoring for biological plausibility: this element is based on the number of known plausible factors within a range of 0–3. The possible factors to consider that might support plausibility are:

- similar drugs are known to produce such an effect;
- a mechanism can be postulated;
- any other supportive data are known.

It is not intended that a detailed literature search should be conducted in order to determine the plausibility category; the user should apply their own pharmacological knowledge and judgement, supplemented by standard reference sources (e.g. the summary of product characteristics, the British National Formulary, etc.).

If no relevant information is found then a score of 0.25 should be applied. In cases where there is uncertainty, a possibility should generally be counted as a positive finding. In cases where one of the previous factors applies but there is also evidence of implausibility (e.g. the reaction may be related to the treatment indication for the suspect drug), it may be appropriate to reduce the score by one level. In most cases a judgement is required and, where there is doubt, particular attention should be paid to this variable in the sensitivity analysis. The default is set at the lowest level for this variable, recognising that many type B ADRs initially seem implausible because there is no obvious biological basis. Thus, providing the signal is quite strong and based on cases of at least average quality, it is still possible for a reaction with no plausible factors to meet the threshold for the evidence score.

Part (d) scoring for number of cases of the ADR of interest with the suspect drug in the population per year: this element is intended to reflect the burden of the ADR in the whole population based on the total number of reported cases of the reaction of interest with the suspect drug divided by the time since the first reaction to the drug was reported (as a proxy for the time on the market).

In some situations the public health score will tend to underestimate the potential importance of the signal because use of the drug is generally limited to a particular sub-population (e.g. children). This problem is dealt with by making an adjustment to the scoring of this element. If this situation is considered to apply, 20 should be added to the score (assuming it is not already at the maximum level). In general, the limited population adjustment is appropriate if ≥50% of the population is ineligible to receive the drug.

Part (e) scoring for health consequences: this element summarises the potential health consequences of the ADR and requires two inputs that relate to case fatality and non-fatal outcomes. The inputs are added to produce a total score that ranges from 0.2 to 1.0 with increments of 0.1. The default value for the case fatality is 0.2 and for non-fatal outcome it is 0.3, giving a default total of 0.5.

Case fatality is the proportion of all people with the ADR whose outcome is fatal. Case fatality should not be based on data specific to the drug/reaction combination but on data derived from all similar reactions to all drugs on the database.

The non-fatal outcome should be decided on the basis of the worst known outcome in the list. If this worst outcome is unlikely to occur and less serious outcomes are much more likely, this should be downgraded one category (where possible). Determination of the non-fatal outcome score is necessarily subjective and medical judgement is required. Once the non-fatal outcome score for a specific reaction has been determined in an impact analysis it should be stored for future use in relation to that reaction.

Drug	a	b	c	d	e	f	Evidence	Public health	Category
Sildenafil	100	0.4	0.5	40	0.5	0.5	20	10	A

Reaction	a	b	c	d	e	f	Evidence	Public health	Category
Retinal vein thrombosis	100	0.4	0.5	40	0.5	0.5	20	10	A
	80	0.4	0.5	40	0.5	0.5	16	10	A
	100	0.6	0.5	40	0.5	0.5	30	10	A
	100	0.2	0.5	40	0.5	0.5	10	10	A
	100	0.4	0.75	40	0.5	0.5	30	10	A
UK impact analysis	100	0.4	0.25	40	0.5	0.5	10	10	A
	100	0.4	0.5	60	0.5	0.5	20	15	A
	100	0.4	0.5	**20**	0.5	0.5	20	5	**C**
	100	0.4	0.5	40	0.7	0.5	20	14	A
	100	0.4	0.5	40	**0.3**	0.5	20	6	**C**
	100	0.4	0.5	40	0.5	0.75	20	15	A
Name of assessor: Waller	100	0.4	0.5	40	0.5	**0.25**	20	5	**C**

Date and time: 01/12/2003 15:12

Inputs

PRR	32.6
Lower 95% CL	14.4
Number of cases	6
Date of first reaction	01/06/97
Limited population adjustment	Y
Case fatality score	0.1
Non-fatal outcome score	0.4
Number of cases in past year	2
Exposure in past year	41 000

a	PRR
b	Strength
c	Plausibility
d	Cases per year
e	Health consequences
f	Reporting rate

Fig. 3. Impact analysis program output: sildenafil and retinal vein thrombosis. **CL** = confidence limit; **PRR** = proportional reporting ratio.

Part (f) scoring for order of magnitude of the reporting rate: this element uses the reporting rate in the past year as a very crude indicator of absolute risk, bearing in mind that substantial under-reporting is possible. It is calculated from the number of cases of the reaction of interest with the suspect drug reported in the past year divided by the most recent and best estimate of patient exposure available for a 1-year period. Since it is only necessary to estimate the reporting rate within an order of magnitude, it will rarely matter whether the exposure estimate reflects prescriptions, numbers of users or person-time of use. Patient years will normally be used as the estimate of exposure for impact analysis, as it is possible to calculate this from both primary and secondary care data.

2.4 Sensitivity Analysis

The sensitivity analysis is an integral part of the tool. It is used to test how sensitive the categorisation is to changes in each of the variables. This is done by moving the score for each input variable up and down one level (where this is feasible, given that only values within the specified ranges can be generated) and recalculating the category (12 times in total). If the score is already at the minimum or maximum point on the scale, no change can be applied in that direction. In the case of the health consequences input, if the score is 0.9 or 0.3, an increment of 0.1 is applied in the upper or lower direction, respectively, so as to reach the maximum or minimum score for this input.

2.5 Examples

Examples of outputs from the impact analysis program for two signals, one in category A and the other in category D, are given in figures 3 and 4.

3. Discussion

Impact analysis is a tool that has been developed to guide decision making on the initial prioritisation and handling of signals arising from spontaneous ADR data. Impact analysis should focus attention on the most important signals and enable those that are weak and/or have limited public health significance to be discarded, at least until further data become available. The method has largely been developed

empirically (i.e. the scoring figures and cut-off points have been chosen and modified on the basis of experience) using the Adverse Drug Reaction Online Information Tracking (ADROIT) database, which is held at the MHRA. The method may yet require some modifications on the basis of systematic evaluation. Signal impact analysis has been developed specifically for regulatory use at the MHRA, but it could potentially be adapted for use in other settings (e.g. by other regulators or by pharmaceutical companies) where suitable input data are available, subject to testing. Such adaptations might, for example, include use of other measures of disproportionality rather than the PRR.

To be of value in saving resources, impact analysis needs to be performed rapidly and, with the aid of a purpose-designed computer program, this has proved possible – about half an hour is the usual time taken for one signal. The computer program is based on a spreadsheet and all the calculations, including the sensitivity analysis, are performed automatically once the specified data have been input, thereby minimising the scope for mathematical error and providing an audit trail. There is scope for possible future enhancements, for example some of the parameters used to assess the quality of cases (e.g. the presence of rechallenge, completeness of information) could be inputted automatically from the database and other sources of data might be used to assess health consequences.

During the development process some modification of the original thinking[11] has been necessary, the most important change being that the tool has been limited to signals arising from spontaneous ADR data. Also, we have not yet incorporated risk factors or preventability into the tool. However, it may be argued that these are rarely a major consideration in the immediate prioritisation of signals arising from spontaneous ADR data. An important aim was to keep the method as simple as possible. In practice, the most difficult issues involved in conducting impact analysis relate to scoring non-fatal outcomes (which requires medical judgement) and the availability of suitable exposure data to calculate reporting rates.

It is important not to place too much weight on the precise value of the scores generated by impact analysis and to realise that estimation and judgement

Drug	a	b	c	d	e	f	Evidence	Public health	Category
Rabeprazole	20	0.4	0.5	20	0.7	0.5	4	7	D
Reaction	40	0.4	0.5	20	0.7	0.5	8	7	D
Interstitial nephritis	20	0.4	0.5	20	0.7	0.5	4	7	D
	20	0.6	0.5	20	0.7	0.5	6	7	D
	20	0.2	0.5	20	0.7	0.5	2	7	D
	20	0.4	0.75	20	0.7	0.5	6	7	D
UK impact analysis	20	0.4	0.25	20	0.7	0.5	2	7	D
	20	0.4	0.5	**40**	0.7	0.5	4	14	**B**
	20	0.4	0.5	20	0.7	0.5	4	7	D
	20	0.4	0.5	40	0.9	0.5	4	9	D
	20	0.4	0.5	40	0.5	0.5	4	5	D
	20	0.4	0.5	40	0.7	**0.75**	4	11	**B**
Name of assessor Waller	20	0.4	0.5	40	0.7	0.25	4	4	D

Date and time 01/12/2003 15:14

Inputs

PRR	6.2
Lower 95% CL	2
Number of cases	3
Date of first reaction	01/09/98
Limited population adjustment	N
Case fatality score	0.2
Non-fatal outcome score	0.5
Number of cases in past year	2
Exposure in past year	139 000

a	PRR
b	Strength
c	Plausibility
d	Cases per year
e	Health consequences
f	Reporting rate

Fig. 4. Impact analysis program output: rabeprazole and interstitial nephritis. **CL** = confidence limit; **PRR** = proportional reporting ratio.

are required in its application. It should also be emphasised that the choice of cut-off points was arbitrary and requires further testing in practice. The sensitivity analysis is an integral and vital part of the method that is used to consider the robustness of the findings. When a change of score by one level for any of the variables results in a different overall categorisation, the scoring for that variable needs to be revisited and consideration given as to whether or not a different score might have been input. Ultimately a judgement has to be applied using the findings of impact analysis and the expectation is that the tool should add value by improving the consistency and transparency of such judgements.

The findings of an impact analysis represent the situation at a given point in time and, potentially, any of the variables may change with time (the health consequences of the ADR input are least likely to do so). When signals that have been categorised as 'B', 'C' or 'D' become significantly stronger (e.g. because the PRR or reporting rate increases markedly), the scores can be recalculated. Once a signal has been categorised as 'A' the signal should be thoroughly investigated and the tool is unlikely to be of further use for that drug/reaction combination. It is important to realise that an 'A' categorisation merely implies that there is a good basis to spend time fully evaluating the signal. Some such signals will, inevitably, turn out to be false positive or not to require regulatory action.

There are several limitations to our method. In particular, spontaneous ADR reporting data are well known to be subject to various biases that impact analysis does not eliminate. The input variables are not completely independent of each other (for example inputs a, d and f are inter-related), but each input has a specific conceptual value. We considered whether or not to include other variables that could be relevant to the need for signal evaluation (for example, recency of introduction of the drug to the market) but chose not do so because such factors are less important and increased complexity could be disadvantageous.

4. Conclusion

In conclusion, a mathematical approach to analysing the impact of signals arising from spontaneous ADR reporting data has been developed along with a supporting computer program. The tool now needs to be evaluated in comparison to a currently used alternative approach (such as the collective judgement of experienced pharmacovigilance personnel). The results of such a study are described by Heeley et al. (see article in this issue of *Drug Safety).*

Acknowledgements

We thank Stephen Evans and Rosalind Coulson for valuable discussions and Peter Waller for his help in developing a computer program to facilitate the method described.

Funding was provided internally by the MHRA. The authors have no conflicts of interest that are directly relevant to the content of this study.

References

1. Waller PC, Wood SM. Regulatory aspects of adverse drug reactions. In: Davies DM, Ferner RE, de Glanville H, editors. Davies's textbook of adverse drug reactions. 5th ed. London: Chapman & Hall Medical, 1998: 20-28
2. World Health Organisation definition of a signal [online]. Available from URL: http://www.who-umc.org/defs.html [Accessed 2004 Aug 31]
3. Meyboom RHB, Egberts ACG, Edwards IR, et al. Principles of signal detection in pharmacovigilance. Drug Saf 1997; 16: 355-65
4. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiol Drug Saf 2003; 12: 559-74
5. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 2001; 10: 483-6
6. Meyboom RH, Lindquist M, Egberts AC, et al. Signal selection and follow-up in pharmacovigilance. Drug Saf 2002; 25: 459-65
7. Bright RA, Nelson RC. Automated support for pharmacovigilance: a proposed system. Pharmacoepidemiol Drug Saf 2002; 11: 121-5
8. van Puijenbroek EP, van Grootheest K, Diemont WL, et al. Determinants of signal selection in a spontaneous reporting system for adverse drug reactions. Br J Clin Pharmacol 2001; 52: 579-86
9. Waller PC, Lee EH. Responding to drug safety issues. Pharmacoepidemiol Drug Saf 1999; 8: 535-52
10. Stahl M, Lindquist M, Edwards IR, et al. Introducing triage logic as a new strategy for the detection of signals in the WHO drug monitoring database. Pharmacoepidemiol Drug Saf 2004; 13: 355-63
11. Waller PC, Evans SJW. A model for the future conduct of pharmacovigilance. Pharmacoepidemiol Drug Saf 2003; 12: 17-29

Correspondence and offprints: Dr *Patrick Waller*, Patrick Waller Limited, Consultancy in Pharmacovigilance and Pharmacoepidemiology, 15 Tamella Road, Botley, SO30 2NY, UK.
E-mail: patrick.waller@btinternet.com