Drug Safety 2006; 29 (3): 273-275
0114-5916/06/0003-0273/$39.95/0

Extension of Points on Clarifying Terminology in Drug Safety

Congratulations to Aronson and Ferner[1] on their comprehensive and thoughtful exploration of fuzzy ideas and terminology in drug safety. We accept the authors' invitation to use their article as a starting point for further discussion by extending points that might be of specific practical interest to the 'drug safety expert' and could be of potential assistance in identifying/prioritising those reports of adverse drug effects (ADEs) that might require more in-depth analysis. We make a point about rethinking the original alphabetic (A/B) classification of Rawlins and Thomas.[2] We believe it should be rethought, as do Aronson and Ferner[1] with their discerning reformulation, and will show how its extension to signal detection makes this a valuable approach. We also will provide two descriptive categories of ADEs ('surprise' and 'paradoxical' reactions), which we believe represent a useful conceptual typology for the purposes of pharmacovigilance. Finally, we would like to comment on causality assessments and the use of the term 'signal generation'.

Pharmacovigilance includes signal detection, prioritisation, evaluation, and finally resolution (e.g. label change) followed by further monitoring for signal detection. The cornerstone of signal detection is the 'prepared mind' that can recognise the hitherto undetected potentially clinically significant drug-related adverse effects.[3] Recently, statistical data mining algorithms (DMAs) are being promoted as adjuncts to this activity. The manual case and data review has often been termed the 'traditional' method. There is considerable debate about the optimum deployment of DMAs and traditional methods for signal detection; we have found that considering new descriptive terminology has helped us formulate rationales for deploying both these tools.

Initially, when DMAs were introduced in pharmacovigilance, they were believed to have the potential to detect drug-related adverse effects with significant public health impact that would otherwise have gone undetected by traditional methods. There is conflicting evidence in this regard and many feel that their experience with DMAs has not so far fulfilled this potential in that most or all credible associations highlighted by these tools are already known. A corollary to this is our sense of a shift in the discourse to include whether case-by-case clinical assessment by medically qualified reviewers (a significant component of the traditional method) is inefficient, impractical and a poor use of resources, and that physician involvement should only begin with associations highlighted by DMAs. This concern is particularly relevant to large health authorities with significant disparity between the amount of data to be reviewed and existing resources. This viewpoint merits serious consideration but we do not subscribe to the 'either/or' approach. For example, it is generally agreed that medical reviewers should focus more on intake assessment of unlabeled and serious ADEs. In addition, and more controversially, the case intake process needs to have sufficient medical supervision to ensure an accurate level of ADE coding in order that DMAs and other pharmacovigilance tools can be utilised to maximum efficiency.

We have found that new ADE descriptive categories can provide a useful conceptual framework for formulating a pharmacovigilance logic. A currently used example is the designated medical event (DME) that may be considered as a sentinel effect ('worst first doctrine'), which may be rare, serious, have a high drug attributable risk, and have an asymmetric error cost of misclassification.[4] We do not rely on contemporary DMAs to detect these events. An example of an additional descriptive category of ADE that has helped us evaluate signal detection strategies is that of 'paradoxical reactions'. These have been reported in many forms including paradoxical seizures with antiepileptic medications, anaphylaxis with corticosteroids, urticaria with antihistamines, and worsening of tuberculosis with antituberculosis therapy. These reports are often not coded as 'paradoxical drug reaction' but may be coded as the treatment indication or its worsening. These reports together with legitimate reports reflecting the natural history/complications

of the treatment indication can mask the paradoxical effect itself or its true significance. One cogent example is a study conducted by the US FDA, which involved a systematic clinical review of all reports of paradoxical bronchospasm in their database that identified 126 reports.[5] This number is notable in light of the absence of reports of paradoxical bronchospasm coded as such in the database at the time and the current paucity of such reports. Because it may go undetected in statistical analysis of aggregate data, this is another example of why intake medical review of spontaneous reports may be critical in some cases.

We would also like to comment on what we call 'surprise' reactions that suggest that DMAs may highlight certain medically relevant ADEs in advance of traditional approaches. They are infrequently or rarely reported, are often of intermediate seriousness, not typical of adverse drug reactions in general, do not have the classical hallmarks of reactions traditionally classified as 'idiosyncratic' (type B ['bizarre']) adverse drug reactions and are not obviously explainable based on the primary pharmacological/therapeutic activity of the drug. They may only become recognised when *post hoc* explanations are sought based on more detailed pharmacological knowledge of the formulation. Anecdotally, we have found that they can be associated with early statistical disproportionalities based on small numbers of reports and while not qualifying as DMEs, they are objective medical events that can have significant medical implications (e.g. cost of medical evaluation and sometimes medically serious as well). We find these associations 'surprising' since the individual reports are often well documented yet could be easily overlooked or discounted based on manual review of ADE listings as opposed to full review of individual cases. An example of this would be topiramate-induced sweat gland dysfunction (oligohidrosis). Retrospective data mining of this drug event combination in a public database by two of the authors identified a statistical disproportionality in 1997 based on four reports. The first literature report appeared in 2001 where a potential mechanism of action (i.e. carbonic anhydrase inhibition) was suggested.[6] A 2003 pilot study using iontophoretic evaluation of sweat gland function on and off drug strongly supported the association.[7]

Such effects are not widely known to be based on known primary pharmacological/therapeutic activities at the time of recognition. Thus, in this instance, DMAs may have their greatest value by usefully focusing attention of safety reviewers to reports of 'surprise reactions' that might otherwise be overlooked or discounted based on manual review. However, a limitation to confirming our hypothesis is the inability to pinpoint when these associations were first detected by traditional methods. Their utility in naturalistic pharmacovigilance settings would depend on the opportunity cost associated with false positive findings.

These examples, two possibly amenable to traditional methods (DMEs and paradoxical) and one (surprise) favoring the quantitative approach, emphasize the value of broadening the view of ADE descriptors. They also reinforce the importance of a holistic approach to signal detection using a comprehensive suite of signal detection strategies.

Aronson and Ferner[1] make several cogent points about the term 'signal' and related terms, especially the problematic nature of the term 'signal generation'. We would first note that there is considerable semantic ambiguity about the meaning of the term 'signal' itself and the potential for confusion may be amplified with increasing use of DMAs, each of which has its own model assumptions, metrics, *ad hoc* thresholds and available configurations.[8] In this context, we agree that 'signal generation' is a problematic term. We slightly diverge on the point that signals are detected and reported, and not generated. The reality for better or worse, is that with the large space of available metrics, threshold and data mining procedures associated with specific DMAs, there may be numerous opportunities to 'generate signals' by sequentially trying different data mining exercises and choosing an analysis based on the fit of the results to pre-existing expectations and, in effect, retrofitting an analysis to pre-existing beliefs.

Finally, Aronson and Ferner[1] make the distinction between causality assessment of an individual

case versus a case series. The authors cite the Bayesian confidence propagation neural network's positive predictive value of 50% and negative predictive value of 85% in the context of causal probability assessment of case series[9] (similar to values reported for another Bayesian algorithm, the multi-item gamma Poisson shrinker when using a comparable reference standard[10]). However, it is important to note, as the authors reporting the original findings discussed, that the reference standard used (amendments to the product label in a standard drug compendia) may not always reflect causality.[9] Therefore, it would be incorrect to conclude that the figures cited represent the positive and negative predictive value for causality.

Manfred Hauben,[1,2,3] *Lester Reich*[1] and *Flic Gabbay*[4]

1 Risk Management Strategy, Pfizer Inc., New York, New York, USA

2 Department of Medicine, New York University School of Medicine, New York, New York, USA

3 Departments of Pharmacology and Community and Preventive Medicine, New York Medical College, Valhalla, New York, USA

4 Transcrip, Southampton, UK

Acknowledgements

No sources of funding were used to assist in the preparation of this letter. The authors have not conflicts of interest that are directly relevant to the content of this letter.

References

1. Aronson JK, Ferner RE. Clarification of terminology in drug safety. Drug Saf 2005; 28 (10): 851-70
2. Rawlins MD, Thomson JW. Pathogenesis of adverse drug reactions. In: Davies DM, editor. Textbook of adverse drug reactions. Oxford. Oxford University Press, 1977: 44
3. Trontell A. Expecting the unexpected: drug safety, pharmacovigilance, and the prepared mind. N Engl J Med 2004; 351 (14): 1385-7
4. Hauben M, Madigan D, Gerrits CM, et al. The role of data mining in pharmacovigilance. Expert Opin Drug Saf 2005; 4 (5): 929-48
5. Nicklas RA. Paradoxical bronchospasm associated with the use of inhaled beta agonists. J Allergy Clinical Immunol 1990; 85 (5): 959-64
6. Arcas J, Ferrer T, Roche MC, et al. Hypohidrosis related to the administration of topiramate to children; Epilepsia 2001. 4 2(10): 1363-5
7. Ben-Zeev B, Watemberg N, Augarten A, et al. Oligohydrosis and hyperthermia: pilot study of a novel topiramate adverse effect. J Child Neurol 2003; 18 (4): 254-7
8. Hauben M, Reich L. Communication of findings in pharmacovigilance: use of the term "signal" and the need for precision in its use. Eur J Clin Pharmacol 2005; 61 (5-6): 479-80
9. Lindquist M, Stahl M, Bate A, et al. A retrospective valuation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf 2000; 23: 533-42
10. Hauben M, Walsh L, Reich L. Predictive value of a computerized signal detection algorithm (MGPS) when applied to FDA AERS database. In: 21st International Conference on Pharmacoepidemiology and Therapeutic Risk Management; 2005 Aug 21-24; Nashville (TN). Pharmacoepidemiol Drug Saf 2005 August; 14 (2 Suppl): S17-8