Drug Safety 2003; 26 (5): 364-366
0114-5916/03/0005-0364/$30.00/0

The authors' reply

We thank Lilienfeld et al. for commenting on an article discussing the Bayesian confidence propagation neural network (BCPNN) as presented at the Drug Safety Research Unit (DSRU) signal detection conference[1] and proposing stratification in the BCPNN analysis of the WHO Adverse Drug Reaction database.

They suggest that our procedures as described in the paper[1] produce 'noise with little consequent protection or improvement in public health'. Although we wish to minimise noise in signal detection and hypothesis generation, we accept that an effective signalling tool must produce some false positive signals. Previously, because we wished to know the merits of our method, we performed a critical evaluation of our initial process.[2] This evaluation gave a positive predictive value of close to 50% and a negative predictive value of about 85%, for signals which reach the standard literature over a 7-year period. That such a high percentage of quantitatively highlighted drug-adverse reaction combinations could be considered signals, even before further filtering (as is done routinely) by clinical review was highly promising. Few, if any, other signalling procedures have been investigated for their performance at all, which can in part be explained by the impossibility of a perfect evaluation of a method for which no gold standard exists. We consider that the performance we demonstrated was satisfactory for a general signal finding tool, and clearly does not produce too much noise.

The use of stratification in highlighting associations in signal detection will prevent some false positive signals, as associations arising from confounding variables will be dampened and may no longer erroneously stand out (this is, of course, only true, when one has chosen to stratify by the variable that confounds the association). It is too simplistic to suggest that the routine use of stratification in analysis will prevent all false positive associations or account for all problems and heterogeneity in any very large, continuously changing data set. Indeed, stratification may in certain situations reduce the potential for the early detection of signals. The possibility of confounding factors and biases is myriad, stretching to all variables in the data set and beyond. Stratification analyses by all possible confounders on a routine basis is clearly impossible. Stratification by all variables in the data set would lead to an insurmountably large amount of data to store, sort or assess, and thus to combinatorial explosion. To stratify by one or two might suggest erroneously that all possible confounding factors have been considered.

We do, however, recognise that stratification by some variables can be useful in signal analysis, that is, after combinations have been highlighted using the BCPNN. Indeed, stratified BCPNN analyses on the WHO database have been illustrated in an earlier publication.[3] Comparison to drug group has also been done for classifying signals as drug group or drug specific effects.[4] Quantitative analysis of why or why not a drug-adverse reaction combination stands out may often provide added information. However, we emphasise that clinical evaluation of the information remains essential.

In our use of the BCPNN in routine signal detection, we analyse how the measure of disproportionality changes over time, for each drug-adverse reaction combination in our database. Combinations reaching a predetermined threshold as new data is accumulated are then reviewed by clinical experts.[5] Using this approach, the numbers of associations for consideration by the expert panel are substantially reduced. We believe that this emphasis on change over time focuses clinical interest on the likely signals more effectively than stratification can at this stage. As stated above, it is only after BCPNN *and* human analysis that we consider that a signal has been found.[2]

Since the work Lilienfeld et al. refer to,[1] we have further improved the effectiveness of our routine use of the BCPNN as a signalling tool using a triage process.[6] This approach allows for the detection of nonquantitative signals in specially defined circumstances where other information may be sufficiently compelling to suggest the need for signalling.[6] Moreover, we have developed an unsupervised pattern recognition procedure using the

neural network, for automated knowledge discovery.[7] This methodology captures complex relationships in the data, amongst other things making dependencies due to confounding variables much more apparent, and therefore easy to reject when necessary, but at the same time also finding the totally unexpected. We are currently using this tool experimentally, but it will be used in our routine work in the near future.

Knowledge discovery must be the aim of signal detection. This process can even be aided by heterogenous data, since, for example, patients treated in different countries may have different metabolic phenotypes, or, more important, different background disease burdens (e.g. malnutrition) and ways of medical practice that provide unique challenges for drugs resulting in adverse drug reactions in special subgroups. Such interesting associations might be missed if an adjusted estimate is used to control for the influence of these extra variables. Stratum-specific values might be quoted and analysed, but would mean that much more data would need to be clinically reviewed. Also, the impact of the increased sparsity of the data through the strata should not be ignored: the detection of true signals might be delayed, or more false positives generated.

In our view, signal detection is likely to be hindered by an obsessive attempt to remove what some may consider noise at too early a stage. The logic of signal detection is totally different from that of signal analysis. There are a number of *general premises* that must be applied to avoid missing signals, and to generate hypotheses, but there are *strict parameters* which need be applied to see whether or not an hypothesis is correct. Some of these premises and parameters are given in table I, to illustrate the differences in logical approach.

Data mining and pattern recognition seek only to determine if an undefined aggregate of data differs in some way from a general background, and to a level of significance which can be pre-set. Human review and understanding is then applied to determine whether or not the uncovered information is useful. Lilienfeld et al. are suggesting that one decides first what one would like to find, and then looks for it. That is fine if one has an hypothesis, but not if one is looking for the unusual: it is only too easy to throw out the needles with the hay.

Table I. Basic premises and parameters in pharmacovigilance and epidemiology

General premises in pharmacovigilance
All drugs cause adverse drug reactions
Adverse drug reactions are rare
There are geographic differences in occurrence and reporting
Similarly there is phenotypic variation
There is bias and confounding
Detection depends on clinical and decision making skills
'Noise' must be considered during the consideration of each case
Depends on voluntary reporting of concerns
The degree of all the above is unknown
All that is known is that the above exist and that they vary!
One tries to make as much use of the data as one can while accepting its limitations
Better data often cannot be collected for ethical reasons (e.g. poisoning or re-challenge)
Gives hypotheses
Working parameters of epidemiology
Must have working hypothesis
Consider defined populations
Carefully considered sampling technique
Must have a comparison group
One must try to avoid bias and confounding factors
Aim to quantify hypothesis
Result is a probability
With a certain confidence associated?

Andrew Bate

WHO Collaborating Centre for International Drug Monitoring, Uppsala, Sweden

Department of Clinical Pharmacology, Umeå University, Umeå , Sweden

Ralph I. Edwards

WHO Collaborating Centre for International Drug Monitoring, Uppsala, Sweden

Marie Lindquist

WHO Collaborating Centre for International Drug Monitoring, Uppsala, Sweden

Roland Orre

Department of Mathematical Statistics, Stockholm University, Stockholm, Sweden

References

1. Bate A, Lindquist M, Edwards IR, et al. A data mining approach for signal detection and analysis. Drug Saf 2002; 25 (6): 393-7
2. Lindquist M, Stahl M, Bate A, et al. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf 2000; 23 (6): 533-42
3. Orre R, Lansner A, Bate A, et al. Bayesian neural networks with confidence estimations applied to data mining. Computational Statistics Data Anal 2000; 34 (8): 473-93
4. Bate A, Lindquist M, Orre R, et al. Data-mining analyses of pharmacovigilance signals in relation to relevant comparison drugs. Eur J Clin Pharmacol 2002; 58 (7): 483-90
5. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998; 54 (4): 315-21
6. Lindquist M, Edwards IR, Ståhl M, et al. The effect of different strategies of signal finding in WHO International ADR data [abstract L14]. Proceedings of the International Society of Pharmacovigilance (IsoP) Conference; 2001 Oct 17-21: Tunis
7. Bate A, Orre R, Lindquist M, et al. Pattern recognition using a recurrent neural network and its application to the WHO database [abstract]. Pharmacoepidemiol Drug Saf 2001; 10 (S1): S163