

Signal Detection in EudraVigilance

Jim Slattery, European Medicines Agency, London, England

ABSTRACT

This paper does not present finished work but discusses a current problem faced by the EMEA in detecting when a new safety problem with a marketed pharmaceutical product is emerging. This detection is based on analysis of a very large database of spontaneous reports by healthcare professionals. The mixture of complexity but also known structure within the data appear to call for a combination of formal statistical modeling and data-mining.

INTRODUCTION

EudraVigilance, the European drug safety data processing network and management system, was launched in December 2001. It is now possible to exchange post-authorisation Pharmacovigilance data efficiently between the EMEA, Competent Authorities and Pharmaceutical Companies that hold Marketing Authorisations in the European Economic Area (EEA).

EudraVigilance contains two reporting modules: the EudraVigilance Post Marketing module (EVPM) collects reports of suspected adverse drug reactions associated with medicinal products authorized for use in the EEA while the EudraVigilance Clinical Trials module collects such reports from pre-authorisation drug trials. The work reported here is the planning phase of a project to decide which of the many thousands of reports received by EVPM each year represent true adverse drug reactions (ADRs) and which are prompted by chance temporal associations of clinical adverse events (AEs) and use of a medicinal product. A process referred to as 'signal detection'.

The main challenges associated with the work are the very large size of the dataset, the arbitrarily many different codes which may be used for medicinal products and for adverse events, and the complex social and clinical mechanisms underlying the reporting process.

As part of a tender to redesign the EudraVigilance Data Warehouse system in 2005 the EMEA have adopted SAS as the analysis package within which to build the signal detection algorithms.

SPONTANEOUS REPORTING SYSTEMS FOR PHARMACOVIGILANCE

The EudraVigilance dataset is formed by collection of voluntary reports of adverse events (AEs) which occur during or following the administration of medicines. Which types of adverse events should be reported and under which circumstances is governed by a set of well defined rules and hence the dataset is not simply an unsystematic collection of case reports. However, the rules are such that their implementation may vary between individual reporters and with factors associated with the medicine – for example how long it has been on the market or whether the adverse event is known to be related to it or currently debated in the media – and with other factors – for example the likelihood that the patient would have suffered the adverse event in the absence of the medicine or whether the event happened within the EU. Thus drawing sound conclusions concerning the likelihood of a causal relationship between a medicine and an adverse event requires modelling not only of what we know about the patient, the medicine and the adverse event, but also of what we know about the characteristics of the reporting mechanisms.

A number of similar pharmacovigilance systems exist: the AERS system operated by the US Food and Drug Administration; the WHO Collaborating Centre for Drug Monitoring database in Uppsala Sweden; and several European National systems. However, in contrast to the complexity of the data structure, the methods of analysis currently used by pharmacovigilance departments are generally quite simple. They are based on measures of the extent to which a chosen AE occurs in association with a particular medicine compared with other medicines – disproportionality measures. Development of a system which models the pharmacological, clinical and sociological aspects of the data is a considerable task and this document will address only some, primarily clinical, issues.

ESTABLISHED METHODS OF ANALYSIS

Some discussion has been devoted to the merits of different ways of expressing disproportionality and to different inferential procedures for signal detection. Practical comparison has suggested that these methods provide very similar performance (Van Puijenbroek EP 2002) and this is a credible finding given the very close relationships between them.

THE PRR MODEL

The automated detection method currently used with the pilot EudraVigilance system involves the calculation of a very simple disproportionality measure, the Proportional Reporting Ratio. For an adverse event A, defined by a code or group of codes within MedDRA (Ω_A), and a treatment T, defined by a code or group of codes within the EudraVigilance Medicinal Products Dictionary or some other dictionary (Ω_T), we say:

For and ICSR i let A_i be the adverse events listed on i and T_i be the treatments listed on i then define:

$$X(A,T) = \{i : A \cap A_i \neq \emptyset \text{ and } T \cap T_i \neq \emptyset\}$$

In other words, the number of ICSRs containing an element of A and of T.

$$PRR(A,T) = [X(A,T)/X(\Omega_A,T)]/[X(A,\Omega_T)/X(\Omega_A,\Omega_T)]^*$$

The ordering of pairs (A,T) based on the PRR values gives a baseline against which to compare other methods of signal detection (i.e. other orderings of (A,T)).

In this simple implementation the PRR compares the drug of interest and the adverse event of interest with all drugs and events in the database. If no data other than the marginal counts of drugs given and the adverse events suffered are available, such a comparison appears the only rational course and hence the PRR is essentially identical in terms of information to other disproportionality measures based on the same data.

There are, of course, other measurements collected in the ICSR databases and the multivariate distribution of the adverse events and drugs are also informative. Hence the PRR is unlikely to be the most sensitive instrument for signal detection that can be constructed.

REPORTING ODDS RATIO

The disproportionality can be expressed as a 'reporting odds ratio' as follows:

$$ROR(A,T) = [X(A,T)/(X(\Omega_A,T) - X(A,T))]/[X(A,\Omega_T)/(X(\Omega_A,\Omega_T) - X(A,\Omega_T))]$$

This gives similar results to the PRR and has been argued to have a clearer clinical interpretation. We would argue that it has some advantage over the PRR but this lies in the same fact that favours odds ratios over risk ratios in general – that it consists of ratios of counts of disjoint sets of observations and hence the mathematics of models based on it is somewhat more elegant.

BCPNN

The WHO use a method of analysis called a Bayesian Confidence Propagation Neural Network to look for associations in their database of spontaneous reports of suspected adverse drug reactions. The measure of disproportionality used is called the information component which is essentially the logarithm of the observed count divided by the expected count. When applied to a simple 2x2 cross classification of product and AE this is equivalent to log(PRR), however the properties of the IC are calculated using a Bayesian approach under the assumption that the marginal probabilities have beta distributions. The methodology was developed in large part by R.Orre (Orre 2005) who now works through his company, Neurologic.

MIGPS

* Note that there is some confusion about the definition of PRR. This definition conforms to that originally used by Evans et al but other have quoted $PRR(A,T) = [X(A,T)/X(\Omega_A,T)]/[(X(A,\Omega_T) - X(A,T))/(X(\Omega_A,\Omega_T) - X(\Omega_A,T))]$

The FDA use a method developed by Lincoln Technology (DuMouchel W 1999) called the multi-item gamma Poisson shrinker. An introduction to the ideas involved is given in Szarfman et al (Szarfman A 2002).

The expected values are calculated within marginal tables of degree chosen at a level depending on the order of interaction to be investigated. For instance, if three way interactions are sought the expected values are calculated within three-way marginal tables. Some more details are included below under Data Model. The empirical Bayes method provides a way of smoothing the ratios of observed to expected values.

MODELING CONSIDERATIONS

It appears to be true in many areas of statistics that the competing standard processes of inference tend to result in very similar deductions. This has been demonstrated in practice as regards signal detection in pharmacovigilance (van Puijenbroek 2002). Thus the route to more efficient signal detection is unlikely to lie in increased mathematical sophistication but in more complete and realistic modelling of the known facts concerning the nature of the data and the way they are collected.

Without exception the signal detection processes discussed above rely on a comparison of the observed counts of specific reports and the number of such reports which is predicted by a mathematical model based on our understanding of the way these reports would arise if no causal association existed between a specific product and a specific adverse event. Viewed in this way the importance of selecting the most appropriate model is immediately clear.

The categorical nature of the outcome variable (AE) and of the primary explanatory variable (product) make log-linear modelling procedures an obvious candidate for analysis of databases of spontaneous reports. Under such a model the simplest disproportionality measure is the standardised residual and this may be viewed for any cell in the full cross-classification or for a cell within a marginal table when we are interested in low order interactions.

Different forms of standardised residuals have been proposed of which the most commonly used are probably $(o-e)/\sqrt{e}$ and $2o \ln(o/e)$ [†]. o and e being the observed and expected probabilities in the cell.

In deciding how to model complex data it is worth spending some time deciding which features of the data are likely to be important in answering the questions posed and which may be ignored. The complexity of this dataset arises from the very large number of different product and AE codes, any of which may occur on any report. Thus the most basic representation of the data is as a C-dimensional table, where C is the total number of products and AEs, and two cells in each dimension, corresponding to presence or absence of the code. This full 2^C cross-tabulation is very sparse – for instance, if no safety report includes more than, say, 10 products and AEs combined, we can immediately say that the proportion of non-empty cells will be of the order of $C^{10}/[10! \times 2^C]$ – ie about 1 cell in 10^{16} when C is only 100 and fewer for larger values of C. DuMouchel presents a notation to concentrate on specific (low) levels of complexity in the dataset. He simply refers to any m-way marginal table by listing the codes which label the dimensions of the marginal table – which he refers to as an ‘item set’.

This appears a helpful approach to restricting the size of the problem though the implications of this model and, in particular, the effect of ignoring structure in the complete database which is not reflected in the marginal tables are not fully apparent. It is clear that some sensible approach to reducing the complexity of the problem is essential to any realistic model fitting process.

A conceptually simpler model for estimating the main effects (one drug, one AE) of products would result if we could concentrate on essentially one principal reaction and one principal product per report. We could then see the data as a two way table with many levels for AE and many levels for product. This appears to have advantages in terms of model simplicity but clearly loses some of the structure which is essential for higher interactions and may well be important even for main effects. Whether any information is lost by attempting to reduce the relationship between products and AEs to a two dimensional table is a question which will be addressed carefully in the project as, if it is possible, it may provide an efficient way to examine the main effects. The PRR approach to modelling is one example of reducing the dimensionality of the problem where the entire 2^{P+A} table is analysed using only 2×2 marginal tables for one product and one AE.

[†] Note that when these formula are applied to the 2×2 marginal tables for a single product and AE both are closely similar to a PRR. If o and e are the observed and expected counts for the cell associated with positive observations of the product and AE $(o-e)/\sqrt{e} = (PRR-1) \times \sqrt{e}$ and $2o \ln(o/e) = 2o \ln(PRR)$. Hence what is proposed may be viewed as a generalisation of the PRR.

Reductions to this form of flat table could be achieved by two mechanisms either prioritising events and products in such a way as to be able to select one of each per report or, at least for AEs, grouping them into clinically sensible hierarchies.

The nature of the model which will supply the expected cell values against which disproportionality can be assessed will be determined by the information available. Most signal detection processes have started with a basic model of independence between products and adverse events. This makes sense as a basic model. However, known associations between products and particular AEs could affect the ability to detect associations between other products and the same AEs. See Annex 2. In a standard log-linear model these effects would be fitted as terms specific to each known association. This would be simple if reduction to a 2 dimensional table proves possible but, if not, may require careful selection of marginal tables of fairly large dimension for investigation of some associations.

Associations between drugs will occur because drugs have an indication in common. Such associations are not unexpected and are not indicative of safety problems. However, within a spontaneous reporting system drugs could also be associated because they represent a greater risk in combination than would be expected on the basis of individual effects. These latter effects can only be seen within the dataset (i.e. without using additional data) by contrasting the profile of AEs for the combination with other drugs. However, the expected associations caused by prescribing practice may produce artefactual associations between drugs and AEs and hence an exploration of the pairwise associations of drugs would be worth undertaking to understand the problem in more depth. From a modelling point of view, we may wish to fit the marginal distributions of some combinations of products exactly to discriminate the effects of each product.

Structural zeroes may occur in the dataset, and in marginal tables, for a number of reasons. Example are when a particular AE cannot be associated with a product (e.g. injection site reactions for oral products) or two products cannot be co-prescribed. The effect of these zeroes on signal detection needs to be carefully considered.

Products and AEs are not the only data of interest and adjusting for covariates also needs to be considered.

CONCLUSION

The problem described in this paper presents a challenge. Standard statistical approaches are available for modeling many features of the data but the size and complexity of the dataset make these models difficult to fit. Data-mining techniques can handle this form of data but would need modification to allow formal modeling of the known structures present in the data. To achieve such modifications without incurring unreasonable penalties in calculation time or simple failure to converge seems likely to require innovative methods. Without such modifications such techniques appear identical to the current simplistic methods used in the Eudravigilance pilot dataset.

REFERENCES

DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *The American Statistician* 1999; 53: 177-90

Orre R, Bate A, Noren G, Swahn E, Arnborg S, Edwards IR. A Bayesian recurrent neuralnetwork for unsupervised pattern recognition in large incomplete data sets. *International Journal of Neural Systems* 15(3):207-222 June 2005

Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reporting database. *Drug Safety* 2002; 25(6): 381-92

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

Van Puijenbroek EP, Dietmont WL, van Grootheest K. Application of quantitative signal detection in the Dutch spontaneous reporting systems for adverse drug reactions. *Drug Safety* 2003; 26 (5): 293-301

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Jim Slattery
European Medicines Agency
7 Westferry Circus
Canary Wharf
London / Postcode
Work Phone: 0044 (0)207
Email: jim.slattery@emea.eu.int
Web:

DISCLAIMER

The views expressed are those of the author and not necessarily the official views of the EMEA

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.