

Drug Safety, the Food and Drug Administration and Statistical Data Mining

Enabling data mining, statistical analysis and graphical reporting to run smoothly

BY MICHAEL O'CONNELL, P.H.D.

Drug safety is an urgent public issue. Popular brand-name drugs, like Vioxx, have been withdrawn from the market; and multi-million-dollar legal actions are in the courts. Meanwhile, big pharmaceutical companies, faced with huge potential payouts, have suffered shocks to profitability and dramatic share price erosion.

In this volatile climate, an ounce of prevention, in the form of early detection of adverse event signals and accurate drug risk assessment in pre- and post-marketing data analysis and graphical data review, can help to avert hazards for both consumers and businesses. With this in mind, the U.S. Food and Drug Administration (FDA) issued three safety guidances in 2005:

- Pre-marketing Risk Assessment;
- Development and Use of Risk Management Plans
- Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment

These guidances provide the framework and the impetus for a new standard of data analysis and reporting throughout the industry.

In short, the FDA has raised the bar on safety reporting and, together, these guidances form a prequel to a proposed rule still under discussion. Regardless of when the new rule comes into effect, intensified scrutiny of pre-marketing and post-marketing data is already reality. In practical terms, this means ever greater quantities of data need to be analyzed, more accurate inferences need to be drawn, and results need to be reported unambiguously to wider audiences with differing levels of technical expertise — all in shorter time windows. The fate of a drug, developed through years of diligent

clinical research, may ultimately be decided in a matter of hours. With so much at stake, published safety reports, whether paper or digital, must be based on rigorous statistics and presented with visual punch. Investigators, collaborators, institutional review boards and data safety management boards all want to see accurate and compelling presentations of safety data during pre-marketing clinical trials, as well as from post-marketing adverse events (AE's) databases.

This increased pressure to perform both deeper analysis and to create sharper presentations falls on data managers, statisticians, epidemiologists, clinicians and business managers at a time when operating budgets are often shrinking. Pharmaceutical companies and regulatory agencies need to streamline the workflows among programmers, statisticians, epidemiologists and other team members without requiring them to be conversant in every discipline. They need intuitive, rigorous statistical analytics with highly customizable tables and scientific graphics to enable a full range of analysis and visual expression across all stakeholder audiences. To accomplish this, they need tools designed to make the processes of data mining, statistical analysis and graphical reporting run smoothly.

Comparing adverse event rates

Statistics and data mining techniques are currently applied to the analysis of adverse event counts in pre-marketing clinical trials and post-marketing observational data. The goal in both areas is to detect elevated adverse events in particular drug treatments of interest. In pre-marketing clinical studies, adverse event rates between the treatment and placebo-control arms of the study are the focus. Results are summarized with statistics

including relative risks and odds ratios.

In post-marketing data, the number of subjects who are using the drugs of interest is typically unknown, and the above summary statistics cannot be calculated. As such, comparisons of adverse event rates between drugs typically use a score comparing the fraction of all reports for a particular event for a specific drug with the fraction of reports for the same particular event for all drugs. This analysis can be refined by adjusting for aspects of reporting or characteristics of the patient that might influence the amount of reporting. In addition, it may be possible to limit the data mining to an analysis for drugs of a specific class, or for drugs that are used to treat a particular disease.

The statistic generated by the data mining quantifies the disparity between the observed and expected values for a given product-event combination. In the usual frequentist statistical analysis, this statistic is typically compared to a threshold that is chosen by the analyst.

Acronyms and Terms

- **AE**
Adverse Event
- **AERS**
Adverse Event Reporting System, U.S. FDA
- **CDER**
Center for Drug Evaluation and Research, U.S. FDA
- **MedDRA**
Medical Dictionary for Regulatory Activities
- **Medwatch**
Safety Information and Adverse Event Reporting Program, U.S. FDA
- **Gamma-Poisson Model**
A Statistical Modeling Method
- **Hierarchical Bayes model**
A Statistical Modeling Method

A potential excess of adverse events is often defined as any product-event combination with a score exceeding the specified threshold (FDA Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment, 2005).

Testing multiplicity

Testing multiplicity arises when there are thousands of adverse events that need to be compared between treatment and control populations. In the usual frequentist statistical hypothesis testing framework, if a statistical test has an alpha level of five percent and 20 comparisons are made, one test result will be wrong just due to the frequentist inference framework. As such, the alpha level and testing framework must be carefully controlled in order to draw valid scientific inference.

In double-blind pre-marketing clinical trials, statistical analysis of efficacy has strict controls regarding statistical power, sample size requirements and alpha spending. The FDA requires that efficacy trials are suitably powered for the primary comparison, and that secondary comparisons and interim analyses in sequential trials are planned and pre-specified. Drug safety testing grows out of an entirely different premise. It may be necessary to conduct hundreds, possibly thousands of tests to discover elevated occurrence of any particular adverse event (AE) from the multitudes of AE's as listed in the medical dictionary for regulatory activities, MedDRA. And, since the FDA is concerned about identifying potential safety issues, multiple tests at a five percent alpha level are conservative and accepted. From a statistical and scientific inference point of view, however, the multiplicity testing issue must be addressed in AE analysis as in any other setting.

Hierarchical Bayes models follow a more direct path to describe probabilities of the events of interest, for example, elevation of the AE rate in the treatment versus placebo arm of the trial. The Bayes posterior distribution can be simply interrogated *a posteriori* and the

testing multiplicity issue can be readily addressed in this way. As such, Bayesian modeling provides a natural approach to handling the multiple testing issues involved in the comparison of many AEs between treatment and placebo.

Sparse data

While there are thousands of potential adverse events, very few are typically observed for many of the MedDRA preferred terms. With sparse data, statistical comparisons are difficult and potentially unreliable. The Bayesian modeling approach addresses this situation by shrinking very high and/or low relative risks toward the mean. This mechanism handles the often-erratic nature of AE count data. The “shrinkage” process can be usefully employed to elicit meaningful inferences, even when counts are sparse and potentially unreliable. In light of the issues described above, the FDA has encouraged use of Bayesian approaches in the analysis of safety data. In the 2002 Pharmaceutical Report, George Chi, James Hung and Robert O'Neill (Center for Drug Evaluation and Research) agree, “Safety assessment is one area where frequentist strategies have been less applicable. Perhaps Bayesian approaches in this area have more promise.”

Data quality

The FDA adverse event reporting system (AERS) database includes data that is reported on a voluntary ad hoc basis from doctors and drug companies via MedWatch 3500 forms. AERS data are tables of frequencies of AE reports classified by drug and reported by AE. There are many issues with this database, such as missing information, duplicate records, coding and classification errors and inconsistent spelling of drug names. In order to use this data in any statistical analysis significant clean up is necessary. Ideally the analytic solution will provide for data clean up, pre-processing and dataset preparation.

Summary

The drug safety landscape is rapidly transforming due to market forces and

regulatory pressures. Drug companies and regulatory agencies, such as the FDA, need better solutions to meet the increasing demands placed on pre- and post-marketing drug safety analysis and reporting as outlined in the recent FDA guidances. They need a solution that is fully equipped to manage data, apply rigorous statistical analysis and data mining techniques and produce flexible graphical/tabular reports of drug safety information.

With many different study designs, datasets and statistical inference needs, the solution and tools need to be highly flexible and customizable. In particular, a rich collection of rigorous statistical analysis methods and a highly customizable statistical graphics palette are crucial for accurate and powerful analysis and reporting of pre and post-marketing drug safety data.

Bayesian modeling methods can provide a natural, accurate and powerful approach to the comparative analysis of adverse event rates. The gamma-Poisson model is one of many hierarchical Bayes models that are suitable for analysis of adverse event analysis for example, the beta-binomial in the case where the denominators are known as in pre-marketing clinical studies.

There are several other approaches to the comparative analysis of adverse events including machine-learning methods such as neural nets, trees and forests; and specialized biostatistics modeling methods and their Bayesian analogues. Solutions such as these enable rigorous statistical analysis and data mining of pre-marketing clinical studies and post-marketing safety databases such as AERS. These advanced statistical and graphical techniques enable signal detection from both pre-marketing and post-marketing data sources, allowing end-users in statistics, epidemiology and medical affairs groups to extract more information from their data and to anticipate downstream safety issues. **SC**

Michael O'Connell is director of Life Sciences at Insightful. He may be reached at editor@ScientificComputing.com.