14th International Conference on Information Fusion
Chicago, Illinois, USA, July 5-8, 2011

Pharmacovigilance, Signal Detection and Signal Intelligence Overview

Atsuko Shibata
Global Safety
Amgen Inc.
Thousand Oaks, CA, U.S.A.
ashibata@amgen.com

Manfred Hauben
Risk Management Strategy
Pfizer Inc.
New York, NY, U.S.A.
Manfred.Hauben@Pfizer.com

Abstract - The field of pharmacovigilance (or drug safety surveillance) has evolved significantly over the last few years. Most notably, various statistical methods have been developed and applied to the scan of large-scale adverse event (AE) databases to identify disproportionality in the reporting of specific product-AE combinations against marginal distributions as a background. This paper provides an overview of the safety signal detection framework and how statistical data mining methods complement traditional safety monitoring approaches in the framework. In addition, commonly used statistical methods and AE databases are described with some analytical examples. Finally, emerging trends and future considerations in safety signal detection are discussed. This paper provides a context for the topics related to cross-functional signal intelligence.

Keywords: Drug safety; pharmacovigilance; safety surveillance; signal detection; signal intelligence.

1 Pharmacovigilance

1.1 What is pharmacovigilance?

Pharmacovigilance is defined as "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any drug-related problem" [1]. Medicinal products (drugs) are required to follow a rigorous and highly regulated development process before they are allowed to be brought to market. Controlled clinical trials conducted prior to the granting of market authorization involve systematic and organized collection and analysis of adverse event data, as well as other data relevant to drug safety (e.g., electrocardiograms, laboratory measurements for liver function). Although the controlled clinical trials are considered a hallmark of demonstrating the efficacy of a drug, safety data available from those trials have well-recognized limitations, such as the limited number of study subjects included in the trials (compared with the size of patient populations that may be exposed to the drug once on the market), the limited duration of exposure to the drug per study subject (particularly in the case of a drug intended for long-term use), limited or no data for potentially higher-risk patient sub-populations that are often excluded from controlled clinical trials (e.g., patients with organ impairment, pediatric and geriatric patients, and women of childbearing age who may be treated during pregnancy and lactation). These limitations make it necessary that the marketing authorization holder of a drug and regulatory authority continue to collect, analyze, and interpret data relevant to patient safety that become available after the drug is introduced to market.

1.2 Recent trends in pharmacovigilance

Over the past few years, relative emphasis of pharmacovigilance activities has shifted (Figure 1). In the past, most pharmacovigilance departments at biopharmaceutical companies focused on the handling of individual adverse event reports (called individual case safety reports or ICSRs) with less attention paid to systematic analysis and review of aggregate adverse event data and risk management planning. Individual adverse event case handling remains highly regulated and timely reporting of ICSRs to regulatory authorities and other stakeholders continues to be an important compliance matter. At the same time, expectations for marketing authorization holders and clinical trial sponsors have increased in the areas of signal detection and risk management. Some of these expectations have been delineated specifically in regulatory guidelines and requirements; other parts are left for the sponsor company to interpret and appropriate processes and methods must be developed. Because of the volume and complexity of adverse event data subjected to detection of safety signals, various statistical methods have been developed to aid routine monitoring of data, which are used in conjunction with more traditional pharmacovigilance approaches.

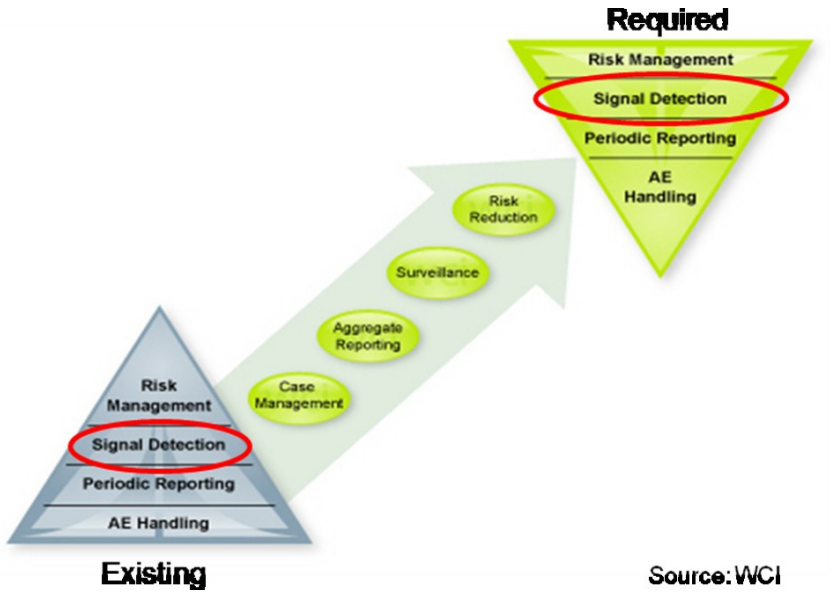

Figure 1. Recent trends in pharmacovigilance (reproduced with permission from WCI)

2 Signal Detection in Pharmacovigilance

2.1 What is a safety signal?

While the term "signal" has been used commonly and widely in the area of pharmacovigilance for years, its definition has evolved over the past few years. Terminology can have profound impacts on clarity and consistency in communication on drug safety to patients, prescribers, manufacturers, and regulators [2] and therefore establishing a common and clear definition of a safety signal is essential.

Most recently, Working Group VIII of the Council for International Organizations of Medical Sciences (CIOMS VIII) [3] defined a drug safety signal as follows, adapting the definition proposed by Hauben and Aronson [2]:

> "Information that arises from one or multiple sources (including observations and experiments), which suggests a new potentially causal association, or a new aspect of a known association, between an intervention and an event or set of related events, either adverse or beneficial, that is judged to be of sufficient likelihood to justify verificatory action."

Some of the keywords and concepts in the CIOMS VIII definition of a safety signal are worth a discussion:

- "Information that arises from one or multiple sources": An earlier definition of a safety signal [4] referred to "report(s) of an event," implying that adverse event reports are the primary, if not the only, source of safety signals. The CIOMS VIII definition acknowledged that new information relevant to drug safety may arise from other sources, such as clinical and nonclinical experiments and published articles on clinical study results.
- "suggests": A safety signal is not synonymous with a confirmed safety issue. The information must be suggestive of something new that would be worth further investigation, after which the suggested association may or may not be confirmed.
- "new": The concept of newness has always been an important part of safety signal detection. It should be noted that newness may be on emerging trends and changes in the specificity, severity, and/or rate of occurrence (frequency) of a previously known (thus may not be totally new) adverse drug reaction.
- "judged to be of sufficient likelihood to justify verificatory action": As discussed above (regarding the word "suggests"), a safety signal by the CIOMS VIII definition precedes further investigation ("verificatory action"), not at the conclusion of such investigation. Importantly, this definitional element emphasizes the crucial role of clinical and scientific judgment in determining whether or not a possible association rises to the level warranting further action.

2.2 Natural history of signals

A safety signal has its lifecycle, as pharmacovigilance and signal detection are ongoing activities spanning the lifecycle of a medicinal product, both before and after marketing authorization. Figure 2 provides a hypothetical depiction of the natural history that safety signals may follow.

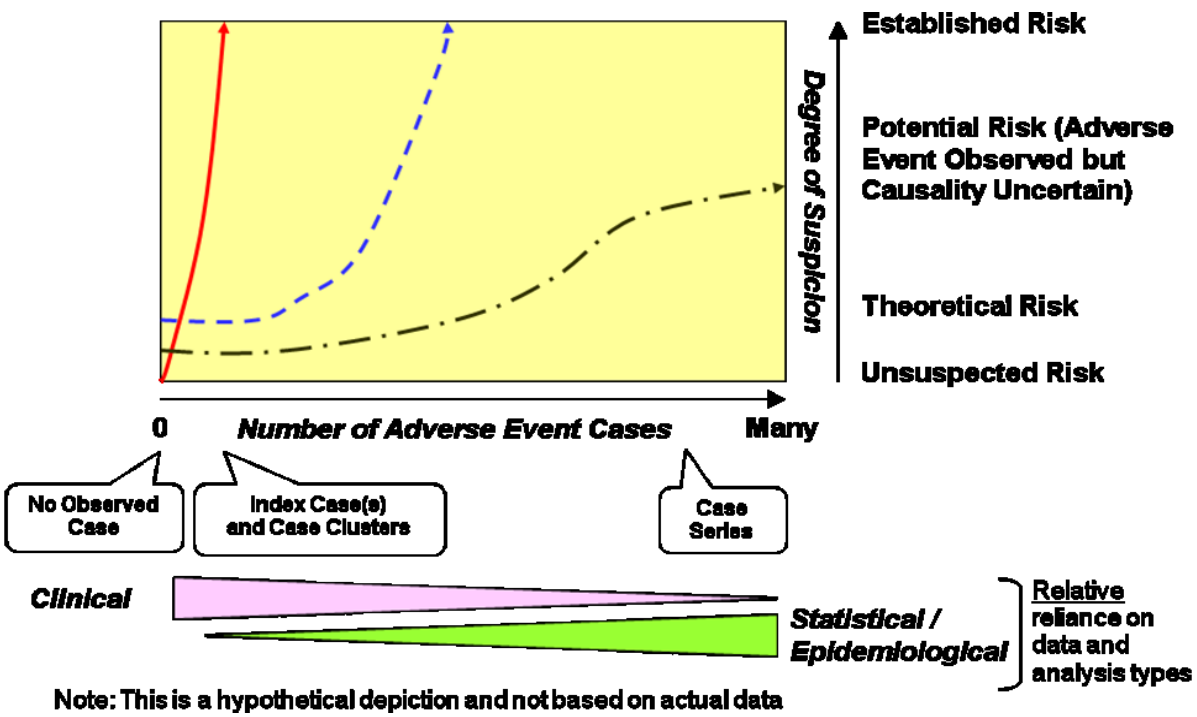

Figure 2. The natural history of safety signals

When clinical trials involving human subjects are initiated, a drug starts with zero adverse event case in humans. As the drug proceeds through a clinical development program and the post-marketing phase, adverse event reports start to come in and, for certain types of events, may form some patterns. For the types of events that rarely occur in the patient populations being treated, a relatively small number of cases may require special attention and serve as index cases. In some situations, the same or similar kinds of adverse events may be reported from the same clinical site or in the same geographic location, forming a cluster.

As the number of adverse event cases increases, relative reliance on different types of data and analytic techniques would change. When there are a relatively small number of cases, pharmacovigilance professionals can review the data, focusing on clinical characteristics of individual cases. Statistical and epidemiological methods would be useful to summarize a large amount of data. The large amount of data, however, does not preclude the importance of clinical judgment and assessment.

Three hypothetical scenarios for how safety signals may evolve are illustrated in Figure 2.

Scenario 1 (dash line) :

Some suspicion about a safety risk exists at the time of starting a safety monitoring program. The suspicion could be theoretical based on the biological mechanism action of the drug or the risk associated with other drugs in the same therapeutic class, or it may have been observed in animal studies. As the sample size of the clinical data (from clinical studies and post-market experience) grows, adverse event cases are reported. After a careful assessment of a case series, it is concluded that that this drug is very likely causally associated with a particular adverse event. The conclusion may result in risk communication actions (e.g., updating the prescribing information), and/or additional risk assessment activities (e.g., investigations in independent data sources).

Scenario 2 (solid line)

At the beginning of clinical development, no anticipation exists for a particular type of safety risk. However, after just a small number of adverse event cases start to be reported, particularly some on rare conditions that have high drug attributability in general, a causal relationship between the drug and the adverse event is suspected. After further investigation, the level of suspicion reaches the point where risk communication actions, such as label change, are deemed necessary.

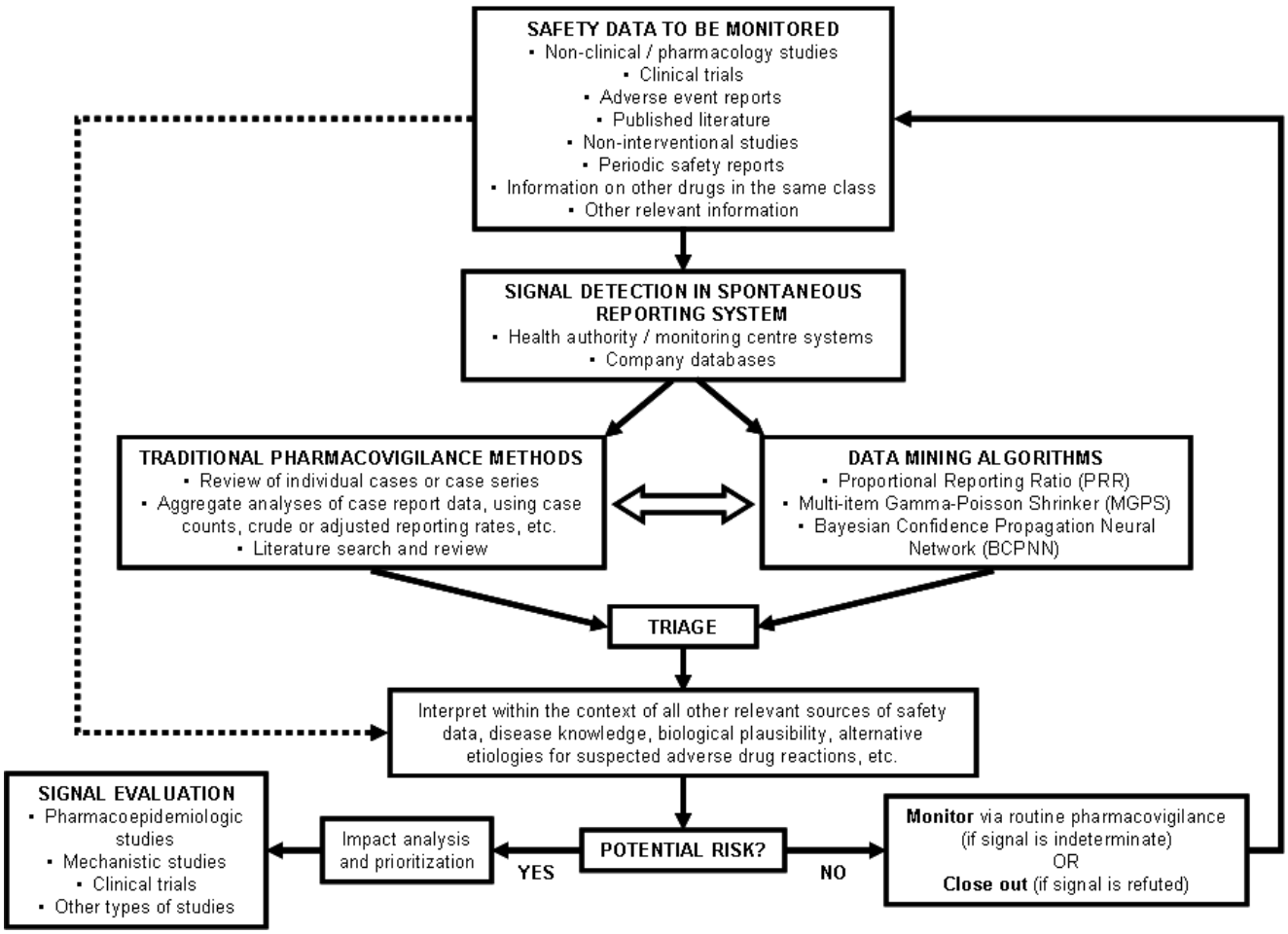

Figure 3. Signal detection and evaluation steps (reproduced with permission from the Report of CIOMS Working Group VIII [3]; © CIOMS)

Scenario 3 (long-dash-and-dot line)

In this situation, a suspicion is raised but does not reach the level for initiating risk communication actions. Safety signals are put under active review, but more data do not necessarily help make a definitive conclusion because of inherent limitations of data or existence of plausible alternative explanations (e.g., confounding). In such situations, investigators may turn to independent datasets for further insights.

2.3 A Framework of Safety Signal Detection and Evaluation Process

A typical signal detection program (here the term "program" refers to a set of planned activities and business procedures to be followed, not computer or software program) consists of a flow of sequential steps of signal detection, prioritization, and evaluation (Figure 3) as well as its linkage to risk management activities, [3].

Pharmacovigilance and drug safety departments at biopharmaceutical companies may be organized differently, but many follow their adaptations of this framework explicitly or implicitly. Information relevant and important to the safety monitoring of medicinal products, coming from multiple sources, not just adverse event reports, is reviewed and observations prioritized by the safety team responsible for a given product. Prioritized safety observations are then subjected to further investigation or assessment and the team makes recommendations based on assessment results. After the team assessment (signal evaluation), recommendations are brought to a cross-functional safety committee, which makes decisions on risk communication and other actions to be taken. The level of urgency for risk communications and other actions is driven by clinical and public health impact of the new information on patient safety. As safety signal detection is an iterative process throughout the product's lifecycle, observations and decisions at each of

these workflow steps may lead to the adjustments to how routine monitoring is conducted.

The term "signal detection" has been used by some authors and pharmacovigilance professionals synonymously with statistical data mining or disproportionality analysis of large adverse event database. This is a misnomer, however, as signal detection is not limited to a narrow range of data sources or analytical approaches.

There is no single right approach that would be optimal for all medicinal products in all situations. Just as diversification is important in financial investment, balancing the portfolio of various analytical methods aligned to a strategy and objectives is important in safety signal detection. Furthermore, no statistical methods or algorithms would replace the importance of medical and scientific judgment of trained pharmacovigilance professionals. At the same time, as the amount of information increases, codifying human experts' tacit reasoning and consistent application of sound pharmacovigilance logic become more and more important in supporting proactive and scalable safety surveillance.

2.4 Data Sources and Statistical Data Mining Methods Used in Safety Signal Detection

Statistical data mining methods for application in pharmacovigilance emerged in the late 1990s, originally as a means of performing systematic signal detection in large databases from the spontaneous reporting systems (SRSs) of adverse event information maintained by health authorities and drug monitoring centers [3]. Some of the databases that can be used for signal detection in the post-authorization drug safety surveillance are listed in Table 1.

Table 1. Databases that can be used for signal detection in post-authorization drug safety surveillance

Type of databases	Examples
Spontaneous reporting system (SRS) databases	Vigibase (WHO) EudraVigilance (EEA) AERS (USA) Sentinel (UK)
Prescription event monitoring databases	Drug Safety Research Unit (UK) Intensive Medicines Monitoring Program (New Zealand)
Large linked administrative databases, or electronic health records (EHR) databases	Healthcare insurance claims databases (managed by for-profit Managed Care Organizations or government agencies)
Electronic medical records (EMR) databases	General Practice Research Database (UK)

Adapted with permission from the Report of CIOMS Working Group VIII [3]; © CIOMS

Data mining methods aim to examine a large dataset of adverse event report records by using statistical or mathematical tools. These methods are based on the concept of disproportionality. It is assumed that, in the absence of disproportionality, the distributions of reported adverse events are the same across drugs. If a specific adverse event were associated with a given drug, however, this event would have higher reporting frequency and create reporting disproportionality. Statistics of disproportionate reporting are screened based on the ranking of drug-event combinations by the level of disproportionality, or statistical "unexpectedness."

Several statistical algorithms have been developed for disproportionality analysis. The data construct underlying all the disproportionality analysis methods are shown in Table 2. Commonly used statistical measures of association are listed in Table 3. Some of these statistics (e.g., RR and IC) have been integrated into Bayesian approaches, such as Multi-item Gamma-Poisson Shrinker (MGPS) and Bayesian Confidence Propagation Neural Network (BCPNN), which have been developed in part to account for the variability associated with small numbers of reports [3]. Figure 4 shows an example of disproportionality analysis using the proportional reporting ratio (PRR).

While 2x2 table-based disproportionality analysis is the most commonly used in the contemporary statistical methods in pharmacovigilance, a variety of multivariate methods and sequential methods are being increasingly tested and applied for purposes of signal detection and preliminary evaluation.

Table 2. 2x2 table of adverse event report data for disproportionality analysis

	Reports for event of interest	Reports for all other events	Total
Reports for drug of interest	A	B	A+B
Reports for all other drugs	C	D	C+D
Total	A+C	B+D	A+B+C+D

Reproduced with permission from the Report of CIOMS Working Group VIII [3]; © CIOMS

Table 3. Statistics for disproportional reporting

Statistic	Formula
Proportional reporting ratio (PRR)	A(C+D) / C(A+B)
Reporting odds ratio (ROR)	AD/CB
Relative reporting (RR)	A(A+B+C+D)/(A+C)(A+B)
Information component (IC)	Log_2[A(A+B+C+D)/(A+C)(A+D)]

Adapted with permission from the Report of CIOMS Working Group VIII [3]; © CIOMS

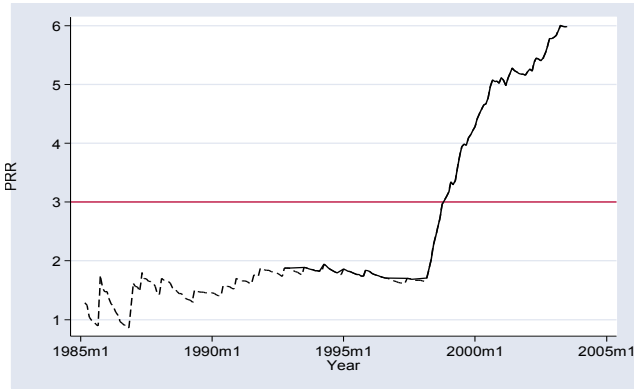

Figure 4. Time trend of PRRs for Isotretinoin and reports of depression (reproduced with permission from the Report of CIOMS Working Group VIII [3]; © CIOMS)

Disproportionality analysis is typically used to identify product-event combinations of unusually high frequency compared with the reference distributions based on data for other medicinal products represented in the database (inter-product analysis). The analytical methods can also be applied to examine time trends for a given drug, where reference distributions are established with historical data for the same drug (intra-product analysis).

3 Cross-functional Signal Intelligence

An interest and opportunities seem to exist for monitoring the data sources that have not commonly been used to date for safety signal detection, which would augment the pharmacovigilance toolkit.

For example, product complaints filed to the drug manufacturers are routinely reviewed and actions taken, when needed, according to local and regional regulations and guidelines for good manufacturing practices (GMP). It makes sense that some of the product complaints may be indicative of product defects or suboptimal presentation of information (e.g., instructions in package inserts regarding how to administer drugs), which may manifest as adverse event reports in the same or other patients. In addition, biopharmaceutical companies receive inquiries about their products from healthcare professionals, patients, and other stakeholders. These inquiries are driven by a desire for more information and may not directly involve any specific instances of product complaints or observed adverse events. It is possible, however, the patterns of inquiries (e.g., a large number of inquiries on the same topic, particularly clustered in time and place) may indicate some underlying issues about a medicinal product and/or associated information available to the stakeholders, for which negative consequences could be corrected or prevented if recognized timely. These types of logic underlie discussions and exploration of cross-functional signal detection and assessment capabilities with intent to generate novel insights on which we can act to promote the safe and appropriate use of medicinal products. Other records related to the production and use of medicinal products may also be useful. To our best knowledge, except for the statistical data mining methods developed for large-scale adverse event databases (see Tables 1-3 and Figure 4), no analytical methods for safety signal intelligence, including non-adverse event data sources, have been published. Systematic scanning of data from cross-functional sources would add value by identifying otherwise unnoticed patterns.

While considering the expansion of safety signal detection to the monitoring of data from cross-functional sources beyond adverse event report data, some of the principles for data analysis and visualization, including those summarized below, should be considered. Some of these points may not be unique to cross-functional situations. Nonetheless they may merit special attention when specialists from various functions, including those from outside the pharmacovigilance and drug safety department, jointly develop a surveillance strategy and analytical approaches.

- The objective and method of analysis must be informed by what the decision maker needs to know. This may be guided by the Key Intelligence Topic and Key Intelligence Question (KIT/KIQ) methodology.
- Data must be available in the form that supports the objective of the analysis. Data quality, completeness, and granularity, as well as consistent application of coding conventions, are among the important factors.
- Logical connection of concepts across functional data through vocabulary linkage (e.g., anatomic reference layer) would facilitate cross-functional data analysis.
- Signals may manifest in the form of an outlier (deviation from the "norm"), a spike (sudden increase in number), or a gradual change in the report count, reporting rate, and other data points and statistics.
- Underlying causes for observed changes and variability may be limited to one or more of the following: geography, time, product (drug, medical device, etc.), patient or reporter demographics, various reporting artifacts (e.g., media influence), and other data elements.
- The same underlying cause may impact data from different sources on different time scales. Analysis with built-in time lag may be informative.

4 Conclusion

Statistical data mining approaches have been developed and applied in the field of drug safety surveillance, adding to the toolkit of pharmacovigilance professionals. Use of data sources other than adverse event spontaneous reporting systems has gained some interest in recent years. Further development of statistical methods and technological solutions to analyze large amounts of data to detect signals for potential safety issues, while minimizing noise, would enhance the efficiency and effectiveness of pharmacovigilance activities.

References

[1] World Health Organization. *The importance of pharmacovigilance – safety monitoring of medicinal products.* World Health Organization, Geneva, 2002.

[2] Hauben M and Aronson JK, *Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions.* Drug Safety, 2009, 32(2):99-110.

[3] CIOMS Working Group VIII. *Practical aspects of signal detection in pharmacovigilance.* Council for International Organizations of Medical Sciences (CIOMS), Geneva, 2010.

[4] CIOMS Working Group IV. *Benefit-risk balance for marketed products: evaluating safety signals.* Council for International Organizations of Medical Sciences (CIOMS), Geneva, 1999.