Signal detection and management for pharmacovigilance – Traditional methods and their inefficiencies (Part 01)
Abstract

The signal management in pharmacovigilance is a process of multiple activities, which aim to determine whether there are new risks associated with a particular medicinal product, or whether known risks associated with a particular drug have changed in frequency or severity. The signal management process follows a systematic approach: Signal detection, signal validation, signal prioritization, signal assessment, recommendation for action, and finally exchange of information. Signal detection is the most vital part of signal management. Traditional methods include usage of solitary algorithms which are not accompanied by additional features of qualitative data mining resulting in mediocre performance. Intense efforts are invested to develop the quantitative and qualitative detection algorithms. To prevail over this limitation, we came up with a hybrid approach to effectively solve this drawback in which the algorithm is fueled to assess a signal, and also manage it for effective grading based on a qualitative approach. In our two-part article series, we would be providing a comprehensive and detailed description of our hybrid approach.

Introduction

Every pharmaceutical product approved to be used in a marked setting has proven benefits but is also associated with adverse effects. Timely detection of such unknown risks is pivotal to ensure the patient’s safety. The detection process applies to all medicinal products, covering their entire life cycle, specifically including clinical development and post-market phases, for any type of adverse event, serious or non-serious. Signal detection and management in pharmacovigilance involves the ongoing monitoring of individual case safety reports (ICSR) to identify case reports or case report series of adverse events (AE) that are worthy for further exploration, and potentially requires safety actions such as a safety signal investigation. Traditionally, signals are detected either qualitatively or quantitatively. The former involves qualitative analysis through the manual assessment of individual case safety reports (ICSR) in an individual or cumulative manner. The latter, on the other hand, involves the more common quantitative approach that makes use of statistical techniques - the most often used being disproportionality analysis.

To identify the disproportionate reporting ratios, the most common technique used is data mining. Data mining techniques, commonly known as signal disproportionality analysis (SDAs), are used to explore a wide variety of databases of spontaneous reports. This analysis helps to find the previously unknown associations between medicinal products and reported adverse events (AEs) that might have escaped monitoring during the manual case assessment. Quantitative signal detection, or more specifically disproportionality analysis, is done through disproportionality statistics i.e., by considering the ratio of the proportion of spontaneous ICSRs to the proportion that would be expected if no association had existed between the medicinal product and the event. Numerous ways are available
to calculate disproportionality. The most common methods are: The proportional reporting ratio (PRR) or the reporting odds ratio (ROR) as well as using the Bayesian methods such as the multi-item gamma Poisson shrinker (MGPS) and the information component (IC).

A detailed analysis on spontaneous reports is performed to identify adverse drug reactions (ADRs). SDAs and the nature of the reporting databases vary between operators. Given the wide range of environments, it is unclear whether any algorithm can be expected to forecast reliable prediction or estimation of an ADR. Although the current clinical assessment mostly relies on disproportionality analysis, however, it is solely based on aggregate numbers of reports and hence, overlooks the quality and nature of the content of the report.

As mentioned earlier, the signal management process includes the following steps:

1) Signal detection
2) Signal validation
3) Signal prioritization
4) Signal assessment
5) Recommendation for action, and
6) Exchange of information

Typically, traditional software solutions lack a functional user interface for overall signal management. An exponential rise in data volume has been seen in the last decade with the proliferation of solicited and unsolicited safety reports. This will result in increased and multiple changes in surveillance data which will be reported in coming years. However, not only the volume of safety cases is increasing but also the source and type of records, including reports from electronic health records and claims, personal health records, standards for health data, data from Federal and private sector mobile devices for tracking health, and data from social websites (blogs, patient advocacy group sites, and search term logs). Hence a comprehensive signal management system would be required to handle such challenges.

vigiRank - a data-driven screening algorithm for identifying potential, causally associated safety signals - can be a good hybrid approach. It accounts for report quality and content along with disproportionate reporting. Signal detection and prioritization is done by using the predictive vigiRank algorithm on the computed vigiRank variables. The computed variables are disproportionality reporting ratio (specifically IC), recent reporting, geographic spread, informative reports, time-to-onset, dechallenge, rechallenge, solely reported, and multiple reporting elements. The algorithm has been implemented using LASSO logistic regression on the data made available from the FDA Adverse Event Reporting System (FAERS) dataset.

After identifying safety signals in an ad hoc manner, the next phase is signal prioritization. vigiRank provides an output score. This score can be an additional factor to consider in this phase. One can use reaction outcomes or the seriousness features for prioritizing signals with serious medical conditions. The European Medicines Agency (EMA) has provided a set of serious events called designated medical events (DMEs). This list acts as an important confounding factor for signal prioritization. EMA suggests that the signals with these events should be assessed on high priority irrespective of the disproportionality ratios and one should not entirely depend on this list alone. Important medical event (IME) is also a list of this kind which facilitates prioritization. Similarly, every pharma company will have its own list of events called targeted medical events (TMEs) for its products.
Signal validation includes validating a signal both analytically and through existing literature. The literature in the form of research articles can be found on sites like PubMed, drug-labelling information sites etc. Analytical charts give clarity on data like gender distribution, percentage of reports with particular drug characteristics etc. This step is like an initial assessment for each detected signal. After validation, one may classify each signal into one of the three categories i.e., (1) Valid signal, (2) Not a signal, and (3) Worthy of further analysis. According to the EMA GVP Module IX, the various considerations that are useful in this phase include:

- Previous awareness on the reaction, and the available data on the summary of product characteristics (SmPC) of medicinal products
- Strength of the evidence like disproportionality, quality of data, dose-response relationship etc
- Clinical relevance and context, which include an understanding of drug-drug reactions, severity, and consideration of medication errors

The solution presented thus provides various functionalities including, but not limited to, PubMed and drug label literature review along with summarization functionality, annotations as well as graphs to illustrate various statistical figures. Sections like ‘Open signals’, ‘Closed Signals’, ‘Further Evaluation’, ‘Keep Under Monitoring’, and ‘Archive’ are created to manage the flow of a signal, and all these sections have tables representing all the signals present in that stage of signal management. The quality and content of individual reports is of fundamental importance. Along with the use of vigiRank algorithm and an efficient user interface, we aim to integrate the value of automation with the breadth of aspects used in clinical assessment.

**Datasets**

While building the solution, volumes of medical data was exhaustively analyzed which was of paramount importance to prove the scalability, robustness, and superiority of the proposed system. The findings have been compiled below, subject to certain limitations:

a) FAERS is a public spontaneous adverse event reporting system by the FDA. Its data is open.

b) EudraVigilance is also an open safety database from the EMA.

c) Drug-labelling information of each medicinal product is used in the signal validation phase

d) PubMed and the Cochrane Library are used to verify existing research literature and reviews for a particular drug-event combination.

**Signal detection and management**

A response to a drug which is noxious and intended, and which occurs at doses, normally used in humans for the prophylaxis, diagnosis, or therapy of disease, or for the modification of physiological function. An active surveillance system is the collection of case safety information as a continuous preorganized process. Spontaneous reporting is done by the system whereby case reports of adverse events are voluntarily submitted. These cases are submitted to the national regulatory monitoring authorities from health professionals and pharmaceutical manufacturers.

Signal is a reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously. An adverse event can be classified into four parts:
1) Severity: Mild, moderate, or severe
2) Seriousness: Non-serious or serious
3) Expectedness: Expected or unexpected
4) Causality: Related or unrelated

Multiple factors contribute to ADRs, such as:

- Poor knowledge of pharmacology, adverse effects of drugs
- Irrational use of drugs, poor prescribing patterns
- Promotional activities by pharmaceutical company detailers
- Lack of authentic sources of information
- Liberal drug outlets and unhealthy pharmaceutical practices
- Liberal over the counter (OTC) and self-medication practices
- Ignorant, illiterate public

A signal is a possible association relationship between AE and drugs. The relationship between the adverse event and a drug is either unknown or incompletely documented previously. Usually more than a single report is required to generate a signal. Depending on the attributes of a signal such as seriousness, reaction outcome of event and the quality of the information a signal is considered to be evaluated further. It cannot be regarded as definitive as an evaluation and validation of the signal is necessary. Traditional approaches use data mining algorithms to generate a signal.

There exist certain challenges related to databases which need to be taken into consideration. Solution to mitigate problems related to data mining for databases are as follows:

- Missing, incorrect, or vague information
- Separate reports about the same incident
- Events may be due to the treated condition, another condition, or another product.
- Over-reporting
- Timeliness of reporting and processing

**Conclusion**

In this paper, we discussed signal detection, signal prioritization, signal validation, how a signal is processed in native methods, its flaws, and how to overcome those challenges using the hybrid methodology. Though signal detection is our priority, signal management is also a major topic we need to focus on. We also discussed vigiRank and vigiGrade, the predictive algorithm-based tools that can help predict a score as an output for each signal. We also proposed the hybrid methods in Part 2 that incorporates major signal management techniques to produce unambiguous results for users as well as coping up with the ever-increasing size of the report database without compromising on the speed of identification of potential safety issues and their prioritization.
References


About the authors

Shubham Chaudhari
Domain Consultant, TCS ADD

Achyuth Tanay Kota
Domain Consultant, TCS ADD

Ashutosh Pachisia
Data Scientist, TCS ADD

Dr. Neelesh Joshi
Lead Life Sciences NA

Dr. Alejandra Guerchicoff
Industry Leader – TCS ADD Platform

Aakash Anand
Domain Consultant, TCS ADD

Aditya Chouhan
Domain Consultant, TCS ADD

Kamal Khumar
Domain Consultant, TCS ADD

Surbhi Nair
Domain Consultant, TCS ADD

Pratik Gada
Team Lead and Safety surveillance scientist Life Sciences

Dr. Varun Vijay Vaidya
Lead Life Sciences UK Project Management

Tata Consultancy Services
Awards and accolades

About TCS ADD Platforms
TCS ADD is a modern and open drug development platform for life sciences that enables digital ecosystems, simplifies data complexity and provides faster access to new and effective drugs for patients in need. The platform is powered by our proprietary cognitive intelligence engine data driven smart analytics and Internet of Things (IoT) that makes clinical trials more agile and safe. TCS ADD leverages the best of cloud architecture and personalized user experience design in compliance with quality guidelines and privacy regulations.

To know more
Email: add.platform@tcs.com

About Tata Consultancy Services Ltd (TCS)
Tata Consultancy Services is a purpose-led transformation partner to many of the world’s largest businesses. For more than 50 years, it has been collaborating with clients and communities to build a greater future through innovation and collective knowledge. TCS offers an integrated portfolio of cognitive powered business, technology, and engineering services and solutions. The company’s 500,000 consultants in 46 countries help empower individuals, enterprises, and societies to build on belief.

Visit [www.tcs.com](http://www.tcs.com) and follow TCS news @TCS.