Smarter Signal Management:

AI, big data, and predictive analytics

eBook

Contents

Executive summary 2

Introduction 3

Artificial intelligence 4

New signal detection methods 5

Neural signal detection 6

Multimodal signal detection 8

Predictive signal detection 10

Conclusion 12

Executive summary

Safety signal detection and management is facing four main challenges today: reducing false positives to minimize the workload for signal evaluators, reducing false negatives to lessen the exposure of patients to adverse reactions, making sense of signal scores across multiple data sets, and providing better safety information before clinical trials even begin, especially in this time of expedited product development with the global pandemic.

New advances in methodologies and technologies such as artificial intelligence are needed if we are to make significant progress in addressing these challenges.

This eBook outlines three such advances: neural signal detection, multimodal signal detection, and predictive signal detection. These advances have the potential to revolutionize signaling to improve our ability to protect patients and make the safety department even more of a critical asset to its organization.

"With the current urgency to bring products to market very quickly, the need for better and more predictive signal detection is becoming paramount to keeping patients safe. The right methodologies and technologies can turn the vast amounts of real-world evidence available today from overwhelming obstacles into critical assets."

BRUCE PALSULICH
VICE PRESIDENT
PRODUCT STRATEGY
ORACLE HEALTH SCIENCES

Introduction

Pharmaceutical, biotech, and medical device companies are increasingly concerned with safety. Advanced therapies have become more complex to develop as well as more costly to move through the approval process to market authorization. When those products do reach the marketplace, safety surveillance remains a major focus for companies, because of the potential human and financial consequences of unforeseen complications. Safety is always paramount.

Unfortunately, some safety issues—including interactions—are statistically unlikely to be observed, even in very large clinical trials with tens of thousands of participants. Adverse reactions that are rare or very rare will only readily appear once the product is marketed and available to hundreds of thousands or even millions of patients.

In the midst of a public health crisis such as the COVID-19 pandemic, there's also enormous pressure on companies developing vaccines or virus treatments to move as quickly as possible. Shortened development time frames make the choice of a molecule to develop an even more significant decision for companies. Additionally, the potential for early approvals increases the chances that smaller or shortened clinical trials won't uncover a critical safety issue before the product is in widespread use.

In response, innovative new signal detection and management approaches have been developed to help clinical trial sponsors, manufacturers, and CROs combat these safety-related challenges. Such advances provide insights that companies can use to predict potential safety issues even before clinical development begins and apply those learnings to their choice of initial research candidates. Ultimately that will mean safer medicines for patients.

What if you could predict potential safety issues before clinical development begins?

Artificial intelligence

AI is rapidly becoming a game-changer for safety. It is providing a crucial component for complying with changing global regulatory frameworks, lowering costs, and shifting human interaction into higher-value tasks — ultimately improving patient safety.

Because AI is such a frequently discussed term, it may be useful to examine what we mean by AI and some related concepts:

Artificial intelligence (AI) is the science of making intelligent computer systems that mimic cognitive human functions, such as learning and problem solving, using techniques such as statistical analysis of data, expert systems that rely on if-then statements, and machine learning

Natural language processing (NLP) is a subfield of AI concerned with the processing and analysis of human language data, typically in the form of unstructured free-text

Natural language generation (NLG) is a subfield of AI concerned with the production of human language, typically in the form of unstructured free-text

Image processing is a subfield of AI concerned with the processing and analysis of digital images, such as optical character recognition (OCR)

Machine learning (ML) is an AI technique used to train software algorithms to learn from data without using explicit instructions, relying on patterns and inference instead

Deep learning is the most advanced ML method, using artificial neural networks with multiple layers to progressively extract higher-level features from the raw input

Neural networks are computer systems composed of artificial neurons which are connected and organized into multiple layers, inspired by the human brain

New signal detection methods

Three new signal detection methods can assist safety teams in finding hidden signals, assessing information from multiple big data sets, and making predictions about the safety of products in development.

Neural signal detection

Using AI, specifically neural networks, neural signal detection enables greater modeling power and flexibility than the non-AI-based methods in common use today, greatly improving detection accuracy. Multimodal signal detection represents another advance that can be employed when searching through large volumes of data located in disparate data sets—it generates a composite signal score across those data sources that is superior to any individual set's score.

Multimodal signal detection

Predictive signal detection

Finally, predictive signal detection utilizes AI, specifically machine learning, to enable researchers to more accurately predict potential safety issues for a candidate compound prior to beginning development. Together, these new technologies and methodologies will enable companies to make better decisions, make them faster, and improve the safety profiles of their products.

Neural signal detection

Enhancing the capability to find signals accurately has been a problem for many years. Now, modern AI techniques are providing the path forward to make neural signal detection a reality.

Neural networks are AI systems composed of artificial neurons which are connected and organized into multiple layers. This approach, inspired by the similar organization of neurons in the human brain, enables the system to work with highly advanced models as well as learn and adapt over time. Neural networks are the foundation for a deep-learning methodology known as neural signal detection (NSD). In NSD, deep neural networks are used in lieu of other statistical/machine-learning methodologies, such as logistic regression, to model relationships between observed/reported adverse events, products, and other relevant information.

The first advantage of NSD over other approaches is that it enables the modeling of more complex dependencies between adverse events, products, and other factors, including confounding and masking effects, with greater accuracy. Secondly, NSD can model multiple events simultaneously in the context of event or product hierarchies. Finally, it easily supports the inclusion of external information, such as biochemical properties of drugs.

For these reasons, NSD translates to greater signaling accuracy, i.e. a reduction of false alerts and missed signals, and an improvement in the timeliness of detection. An evaluation of NSD based on the Observational Medical Outcomes Partnership (OMOP) gold standard demonstrated a 19%-29% improvement in signal detection accuracy compared to other signaling methodologies that do not use AI, as shown in Figure 1.

Neural signal detection *continued*

Many signal analysts today rely on algorithms developed by statistical safety expert Dr. William DuMouchel, now Chief Statistician at Oracle Health Sciences, together with the FDA. His original data mining algorithm, known as Gamma Poisson Shrinker (GPS), and later developed into an improved method known as Multi-item Gamma Poisson Shrinker (MGPS), has been the standard for several years.

It uses sophisticated regression analysis to examine data, and has been further refined, together with Dr. Rave Harpaz, Oracle Health Sciences Senior Director Research and Data Science, in a method known as Regression-adjusted Gamma Poisson Shrinker (RGPS). Neural Signal Detection (NSD) is the next evolutionary step, and a focus for Drs. DuMouchel and Harpaz.

Area Under the ROC Curve

	MGPS	RGPS	NSD
GI bleed	0.83	0.82	0.88
Liver injury	0.71	0.80	0.76
Myocardial infarction	0.68	0.69	0.80
Renal failure	0.79	0.86	0.89
Average	0.76	0.79	0.83

Error Reduction

	NSD
MGPS	29%
RGPS	19%

Figure 1. Performance of NSD compared to MGPS and RGPS based on the OMOP gold standard.

An evaluation of NSD demonstrated a 19%-29% improvement in signal detection accuracy

Multimodal signal detection

One of the complex issues companies face in signal management today is dealing with the sheer proliferation of available data sources that may contain potential safety signals. Instead of just examining data from clinical trials or post-marketing spontaneous reports, companies are looking at new "secondary use" data sources such as electronic healthcare records, administrative claims data, social media, web search logs, and pharmacokinetics/-dynamics/-genomics data.

Each data source may produce a different signal score; the problem arises when trying to understand the strength of the signal across these multiple data sets and make sense of the varying results. The solution to this problem is called multimodal signal detection, in which signal statistics are aggregated across the available data sets in order to generate a more accurate composite signal score, as illustrated below (Figure 2).

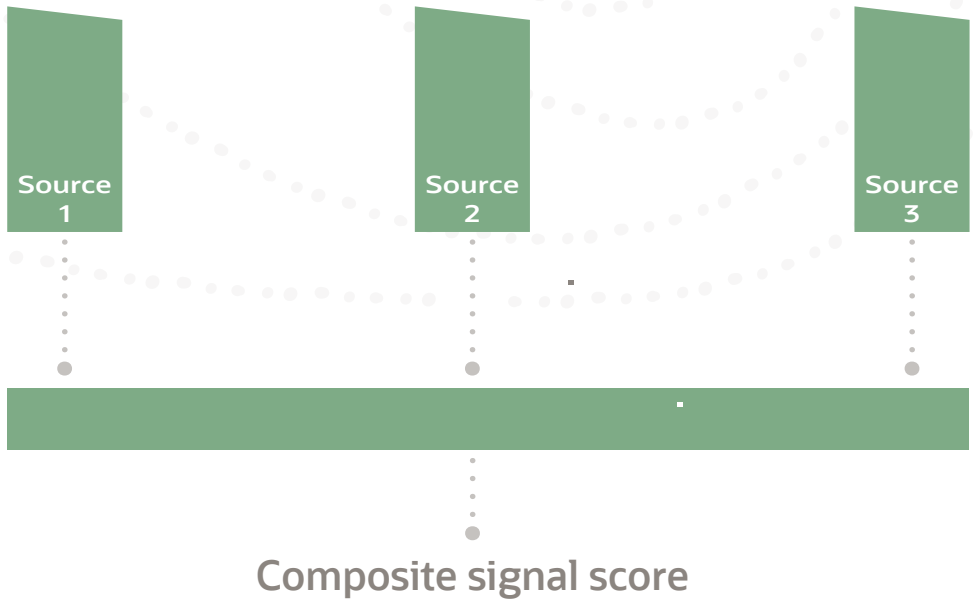

Figure 2. Architecture of a multimodal signal detection system.

Multimodal signal detection *continued*

Oracle co-developed the multimodal approach to help companies understand safety signals in multiple big data sets. The process of aggregation in multimodal signaling takes advantage of the built-in pros of each individual data set while at the same time negating the cons, resulting in more accurate signal statistics.

In a recent study[1], it was demonstrated that a multimodal system built on signal statistics from the FDA Adverse Event Reporting System (FAERS), MEDLINE, and claims data not only improves the accuracy of signal detection, but also increases the lead time to detection. This is illustrated in Figure 3, where, regardless of the rate of false alerts, the combined signals provided several months of additional lead time over the uncombined signals.

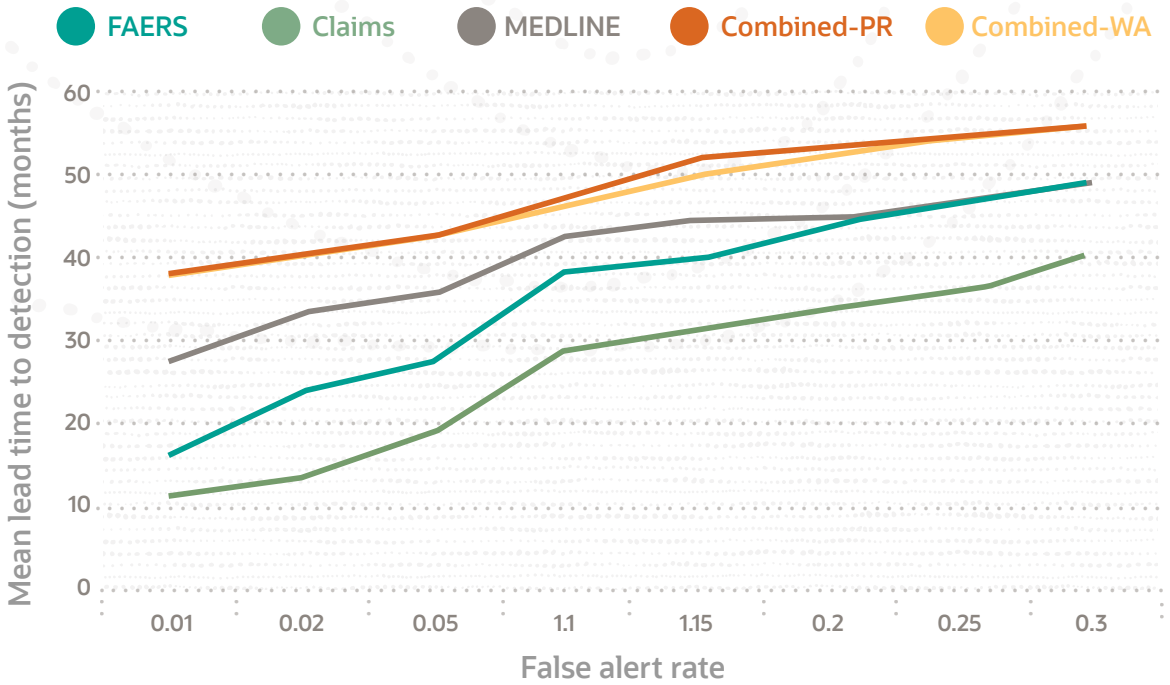

Figure 3. Increased lead time with multimodal signal detection.

Predictive signal detection

One of the most crucial decisions a company makes is selecting the best candidate as a starting point for R&D. Companies may have access to a large library of molecules or the opportunity to license them from others. There may be several choices whose structures or known mechanisms of action align with the needs of researchers.

What's just as important to the eventual outcome, however, is the likely safety profile of the resulting product during clinical trials and beyond. The current practice for addressing this problem is to examine known class effects—the safety profiles of medicines with molecular structures similar to the R&D candidate—and assume a high probability that a new medicine will share the general profile of its class. Yet those assumptions are still guesses—there could be a different safety profile that would make even a highly-effective new product unacceptable in the marketplace or face regulatory rejection.

What if it were possible to predict the safety profile with a higher degree of confidence? That would reduce uncertainty about the choice of a starting compound, in turn reducing the risk of harming subjects/patients or investing heavily in unsafe products. Predictive AI and machine learning promise to make this process possible.

Making more informed research decisions at the outset could save patient lives, millions of dollars, and years of work when developing a new product.

Predictive signal detection *continued*

A recent paper in The Lancet journal EBioMedicine[2] reported on research by several companies and academic institutions—Novartis Institutes for BioMedical Research, Harvard Medical School, MIT, Oracle Health Sciences, and others—that examined how this could be done.

The researchers looked at in vitro data on the pharmacology of common targets for more than 2,100 drugs on the market, searching for proteins associated with the treatments. They then referenced the safety issues in the FAERS database involving those drugs. Finally, they created machine learning models that were designed to predict future adverse events directly from the in vitro data in databases such as PubMed abstracts.

Not only were the researchers able to identify 221 instances of such links, the models also suggested a number of previously unknown and unexpected safety issues associated with one specific protein that warranted further investigation. While the results are preliminary and involve a limited set of indications, the evidence points to the ability for machine learning to predict these safety issues on a wider scale, which can have massive implications for safety.

As this technology develops, it promises to be a huge game-changer for the industry. Making more informed research decisions at the outset could save patient lives, millions of dollars, and years of work when developing a new product.

Conclusion

Safety is incredibly important to medicinal product development, from clinical trials to post-marketing surveillance, but can sometimes still be seen as simply a cost center. More sophisticated and automated methods of signal detection and management, with the application of AI and machine learning, can greatly streamline the process as well as open up new avenues for finding hidden signals and shifting into more predictive data science, confirming safety to be a strategic part of the business. With the rapid development of this new technology, companies are well served by adopting software platforms like Oracle Empirica that have the necessary built-in capabilities as well as the flexibility to add newer AI-based functionality as it becomes available. Such platforms will be key to the elevation of the safety organization to becoming a critical component of the company's decision-making process.

References

1. Harpaz R, DuMouchel W, Schuemie M, Bodenreider O, Friedman C, Horvitz E, Ripple A, Sorbello A, White RW, Winnenburg R, Shah NH. Toward multimodal signal detection of adverse drug reactions. Journal of Biomedical Informatics 2017, 76, pp 41-49.
2. Ietswaarta R, Arat S, Chen AX, Farahmand S, Kim B, DuMouchel W, Armstrong D, Fekete A, Sutherland JJ, Urban L. Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology. EBioMedicine 2020.

About Oracle Health Sciences

As a leader in Life Sciences cloud technology, Oracle Health Sciences' Clinical One and Safety One are trusted globally by professionals in both large and emerging companies engaged in clinical research and pharmacovigilance. With over 20 years' experience, Oracle Health Sciences is committed to supporting clinical development, delivering innovation to accelerate advancements, and empowering the Life Sciences industry to improve patient outcomes. *Oracle Health Sciences. For life.*

Learn more about Empirica at
oracle.com/empirica

CONTACT

+1 800 633 0643

healthsciences_ww_grp@oracle.com

oracle.com/healthsciences

CONNECT

blogs.oracle.com/health-sciences

facebook.com/oraclehealthsciences

twitter.com/oraclehealthsci

linkedin.com/showcase/oracle-health-sciences