Linguamatics

The power of natural language processing to transform Real World Data into insights

Enhancing drug development with actionable intelligence for better decision-making

In pharma and healthcare, understanding the real world (i.e. outside clinical trials) impact of therapies on patients is critical. Real World Evidence (RWE) can inform all phases of drug development and commercialization. RWE can shed light on real world clinical effectiveness or safety profiles of products across a broad patient community. It can be used to assess patient-reported outcomes, provide input for product reputation management, help with key opinion leader engagement, and more.

However, many Real World Data (RWD) sources contain unstructured text, which hinders easy analysis. **I2E text analytics can unlock the value from real world sources such as electronic health records (EHRs), adverse event reports, social media, and customer call transcripts.**

The real world challenge

RWE can inform all phases of drug development, commercialization, and drug use in healthcare settings. However, many sources of RWD contain large amounts of valuable information hidden in unstructured text (e.g. in EHRs, patient-reported outcomes such as forums and social media, and insurance claims data). There are many challenges in creating value from RWD. These include:

◆ data access (which may involve patient privacy issues);

◆ data quality (e.g. missing data, coding errors);

◆ data structure (e.g. complex grammar in tweets or customer calls);

◆ data extraction and integration of structured and unstructured fields; and

◆ mapping to standards (e.g. medical codes, vocabularies, formats).

The NLP-based text mining solution

Linguamatics' natural language processing (NLP) text mining platform, I2E, can provide a solution to these challenges—extracting the key facts from these unstructured documents, using relevant ontologies and focused queries, and transforming RWD into actionable intelligence for decision-making.

Queries can be written to extract information on: Treatment patterns, e.g. drug switching or discontinuation; numerics such as lab values or dosage information; or patient data such as history of disease, problem list, demographics, social factors, and lifestyle. The agile, iterative nature of query development in I2E means that business rules can be encoded to suit the particular data set, whether sentiments from tweets, or treatment pattern choices and resulting outcomes from EHRs.

Benefits to pharma organizations

◆ For pharma companies, effectively mining RWD provides value from bench to bedside, and feeds enhanced drug discovery, development, and post-market delivery.

◆ RWE supports product development and commercial decision-making, based on a better understanding of disease states and treatment patterns across a broad population.

◆ RWE can be used for health economics and outcomes research, comparative effectiveness research, and post-market product lifecycle management (disease forum engagement, reputation management, key opinion leader engagement, safety profiles, treatment regime effectiveness).

◆ RWD can be used to understand treatment effectiveness, as well as to provide insight into patterns of care, long-term drug safety, healthcare resource utilization, and disease epidemiology.

Robust RWE that weaves together different sources of both structured and unstructured data—such as clinical data, genomic data, and socioeconomic data—yields a better picture of individual patient characteristics and responses to therapeutics. This improves our understanding of a drug's ability to treat individual patient needs, and supports precision medicine initiatives.

Figure 1: Data relevant to RWE comes in multiple types and forms, for example: Clinical setting data derived from patient medical records, nurses' notes, and pathology reports; claims data derived from insurance reimbursements; patient registries; patient-reported outcomes derived directly from the patient experience, such as patient surveys, forum sites, and social media; customer-support emails and call centre feeds; and adverse events reports.

Real world use cases

RWD in pharma for medical affairs: Johnson & Johnson

Patient- and customer-call transcripts are a rich seam of potential patient-reported outcomes, side effects, drug interactions, and more, providing access to insights that can have a huge impact on commercial business decisions, and assist post-launch product marketing and planning.

Johnson & Johnson uses NLP to annotate and categorize "voice of the customer" (VoC) call feeds, to gain insights into the real world use of their drugs. Researchers in the Predictive Analytics group have built an end-to-end workflow to process the call transcripts, using agile text mining to make sense of the unstructured feeds. The calls are categorized and tagged for key metadata such as caller demographics and reason for calling (e.g. complaint, formulation information, side effect, drug–drug interactions). Using I2E in this workflow has more than doubled the efficiency in analysis; the accuracy of I2E mining is at 95%, allowing the Medical Affairs teams to do longitudinal analysis of real world patient outcomes.

Comparing adverse event profiles from clinical trials to patient forum data: AstraZeneca

Nausea is a common, often debilitating, adverse reaction (AR) associated with many medicines. Yet under-reporting of a subjective adverse event such as nausea can give rise to disparities in clinical trial results and real world occurrences. AstraZeneca wanted to examine the differences between nausea AR frequencies from clinical trials and those from a source of RWD, patient-reported outcomes. Through collaboration with PatientsLikeMe (PLM), researchers were able to access patient online self-reported data, and used I2E to extract nausea AR frequencies reported in clinical trials from FDA Drug Product Labels. The resulting data were used to compare and contrast nausea from the self-reported PLM community and FDA Drug Labels, and demonstrated different reporting rates, some of which may be due to dosing and usage differences.

Why wait?

I2E is a world-leading, agile, scalable, real-time NLP-based text mining solution. I2E already helps top pharmaceutical companies and healthcare providers speed effective drug discovery, development, and delivery of healthcare therapeutics.

To understand the power of NLP text analytics to transform your RWD into actionable RWE, contact us at: **enquiries@linguamatics.com**