DATABASE
The Journal of Biological Databases and Curation

Challenges and opportunities for mining adverse drug reactions: perspectives from pharma, regulatory agencies, healthcare providers and consumers

Graciela Gonzalez-Hernandez[1], Martin Krallinger[2], Monica Muñoz[3], Raul Rodriguez-Esteban[4], Özlem Uzuner[5] and Lynette Hirschman [6,*]

[1]Department of Computational Biomedicine, Cedars-Sinai Medical Center, 700 N. San Vicente Blvd., West Hollywood, CA 90069, USA
[2]Life Sciences—Text Mining, Barcelona Supercomputing Center, Plaça Eusebi Güell, 1-3, Barcelona 08034, Spain
[3]Division of Pharmacovigilance, Office of Surveillance and Epidemiology, Center of Drug Evaluation and Research, FDA, 10903 New Hampshire Ave, Silver Spring, MD 20993, USA
[4]Roche Innovation Center Basel, Roche Pharmaceuticals, Grenzacherstrasse 124, Basel 4070, Switzerland
[5]Information Sciences and Technology, George Mason University, 4400 University Dr, Fairfax, VA 22030, USA
[6]MITRE Labs, The MITRE Corporation, 202 Burlington Rd., Bedford, MA 01730, USA

*Corresponding author: Tel: +1 617 893 9608; Fax: +1 781 271 2780; Email: lynette@mitre.org

Abstract

Monitoring drug safety is a central concern throughout the drug life cycle. Information about toxicity and adverse events is generated at every stage of this life cycle, and stakeholders have a strong interest in applying text mining and artificial intelligence (AI) methods to manage the ever-increasing volume of this information. Recognizing the importance of these applications and the role of challenge evaluations to drive progress in text mining, the organizers of BioCreative VII (Critical Assessment of Information Extraction in Biology) convened a panel of experts to explore 'Challenges in Mining Drug Adverse Reactions'. This article is an outgrowth of the panel; each panelist has highlighted specific text mining application(s), based on their research and their experiences in organizing text mining challenge evaluations. While these highlighted applications only sample the complexity of this problem space, they reveal both opportunities and challenges for text mining to aid in the complex process of drug discovery, testing, marketing and post-market surveillance. Stakeholders are eager to embrace natural language processing and AI tools to help in this process, provided that these tools can be demonstrated to add value to stakeholder workflows. This creates an opportunity for the BioCreative community to work in partnership with regulatory agencies, pharma and the text mining community to identify next steps for future challenge evaluations.

Introduction

The biomedical text mining community has recognized that there is a strong need for text mining in the drug development pipeline and for pharmacovigilance (PV) in particular. Monitoring drug safety is a central concern throughout the drug life cycle, from drug discovery, through pre-clinical and clinical research, regulatory review and, finally, in post-market surveillance (see Figure 1). Information about toxicity and adverse drug reactions (ADRs) is generated at every stage in this life cycle, leading to ever-increasing volumes of information.

Stakeholders have an urgent need for text mining and artificial intelligence (AI) systems to manage their information requirements, to detect signals in noisy domains and to support human decision-makers in complex data management and decision-making tasks.

These stakeholders are a diverse group (see Figure 2). Key players are the consumers and patients who consume the drugs, the pharmaceutical companies who develop and manufacture the drugs, the healthcare providers and researchers who administer and monitor the drugs and, central to the process, the regulatory agencies who are responsible for overseeing the drug approval process from clinical trials through post-market surveillance. Each of these stakeholders provides input to the PV process but in different forms. For example, consumers may submit spontaneous reports of adverse events (AEs), or they may discuss problems and side effects informally on social media (SM). During drug development, pharmaceutical companies file toxicology reports, information about clinical trials and filings for a new product for regulatory approval. For post-market surveillance, pharmaceutical companies must provide and maintain appropriate
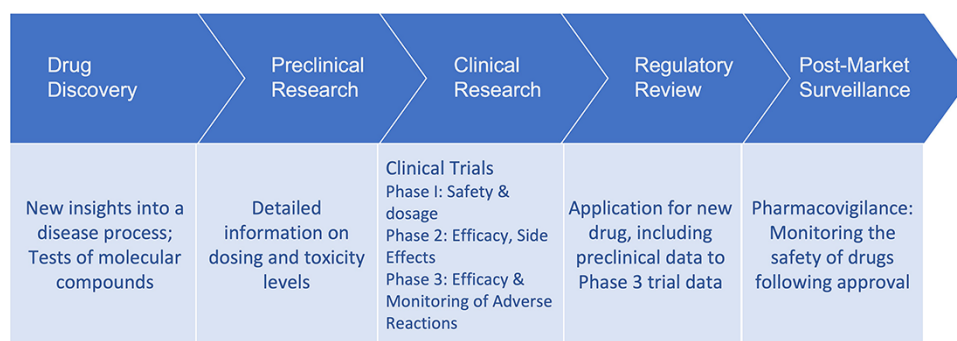
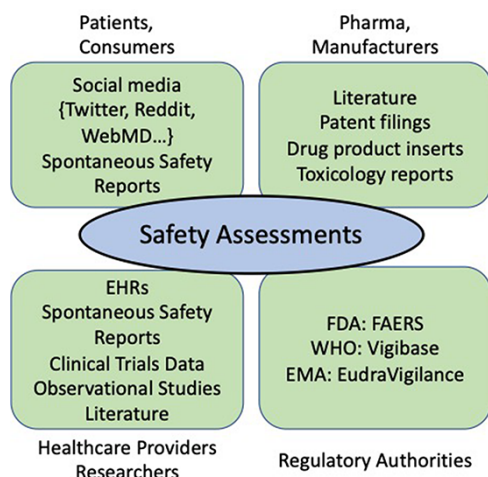**Figure 1.** Schematic of the drug discovery and development process.

Figure 2. Stakeholders and information sources in the drug development process.

drug product labels, and they are required to report adverse drug events (ADEs) to the appropriate regulatory agency, including reports from the literature. Clinicians and healthcare providers also file reports of AEs through the spontaneous reporting system (SRS) and may document observations in case reports in the literature. Regulatory agencies oversee this process from drug development through post-market surveillance.

The complex process of drug development, approval and post-market surveillance present many opportunities for text mining of toxicity and ADEs. Possible applications span multiple disciplines, including pharmacology, clinical medicine, toxicology, genomics, proteomics and computational biology. The applications also span multiple languages and different data sources [electronic health records (EHRs), journal publications, toxicology reports, patient safety reports, drug labels and SM] and may have specific requirements regarding data access, privacy, output encoding or detecting rare events.

Recognizing the importance of these applications, the organizers of BioCreative VII (Critical Assessment of Information Extraction in Biology) convened a panel of experts to discuss opportunities for text mining challenge evaluations in this area. BioCreative has a long history of organizing challenge evaluations (1–4) to promote the development of text mining and text processing tools useful to communities of researchers, publishers and database curators in the biomedical sciences. Since its inception in 2004, each BioCreative

workshop has provided challenge evaluation tasks to facilitate putting text mining tools into the workflow of end users. Previous BioCreative workshops included related evaluations: the first chemical entity and drug mention detection track took place at BioCreative IV (the CHEMDNER track) (5), motivated by the need for more efficient and quality-evaluated chemical entity recognition tools for toxicology text mining applications in the context of the eTOX project (6). Subsequent BioCreative evaluations have continued and extended their focus on chemical/drug interactions with the genome and proteome (7).

For the panel, the BioCreative VII organizers invited five experts with experience in applications related to drug discovery and PV to discuss 'Challenges in Mining Drug Adverse Reactions'. Each of the five panelists has contributed a section highlighting an aspect of text mining related to toxicity or ADRs, based on their research and their experience with text mining challenge evaluations. The contributions below do not attempt to cover the whole of this complex drug life cycle. Rather, each section provides a snapshot of an opportunity where text mining for ADRs could contribute to a specific aspect, often in combination with other AI and text mining technologies.

Each contribution discusses a specific extraction task in the drug life cycle and provides a sampling of the multiple dimensions of the problem space, including:

- Type of input: biomedical literature, toxicology reports, EHRs, drug product labels and SM reports;
- Application for ADR extraction for different stages of the drug discovery, approval and post-market surveillance pipeline, from pre-clinical studies to post-market surveillance;
- Target output formats and terminologies/ontologies, including Medical Subject Headings (MeSH), Medical Dictionary for Regulatory Activities (MedDRA) and RxNorm;
- Technical challenges, including sparse mentions of drug events (e.g. in SM or medical records); use of formal vs informal names or distinguishing ADEs from other medical information, such as patient history or indications for a medical condition;
- Access barriers, such as paywalls for literature or privacy issues in accessing patient safety reports.

In the final section, we review the landscape for future challenge evaluations that would help to drive the field forward, including potential sources of data, gold standard annotations

needed for evaluation and the need to create resources to handle additional languages.

Opportunities for applying text mining to drug safety surveillance

PV systems are designed to detect ADRs that were not identified during a medicine's pre-approval clinical development program and to further characterize a product's known risks. For Food and Drug Administration's (FDA) Center for Drug Evaluation and Research, routine surveillance activities include screening of multiple sources, including AE reports, the published literature and periodic safety reports submitted by the pharmaceutical industry [8]. Safety concerns may also arise from other data sources such as post-marketing studies, exchanges between regulators, inspectional data and SM. After the identification of a signal, preliminary knowledge on the causal relationship between the product and the safety concern and potential public health impact are used for signal triage and prioritization. The signals that represent new potential risks are then formally evaluated by multidisciplinary teams that consider all available data. Based on the evaluation, the agency will take an action, which may include further assessment, communication or other risk minimization actions in accordance with the importance of any identified risks.

Text mining applications can help improve the efficiency of PV activities throughout signal management. This includes the extraction and organization of AEs from various clinical texts, but also identifying and classifying important attributes connected to those AEs. Examples of text mining applications with various data sources related to PV are described below.

Individual case safety reports

For decades, individual case safety reports (ICSRs) housed in various drug manufacturer safety databases and in surveillance systems such as the FDA's Adverse Event Reporting System (FAERS), European Medicines Agency's (EMA) EudraVigilance and the World Health Organization's VigiBase have been the main sources of post-marketing ADR information [9, 10]. These systems rely on healthcare professionals, patients and others to report AEs voluntarily either to the product's manufacturer, which will subsequently report them to health authorities or directly to authorities. Over 23 million ICSRs are currently in FAERS, with >2 million reports received annually since 2018. FAERS adheres to the international safety reporting guidance issued by the International Council on Harmonisation and utilizes the MedDRA for coding AEs [11]. ICSRs contain both structured and unstructured data fields, including an AE narrative. The narratives often contain information that may not be well represented by coded fields; thus, an assessment of case narratives is necessary to further characterize the AE and assess causal relationships. Consequently, natural language processing (NLP) tools coupled with machine learning (ML) techniques are increasingly explored to extract key information from the ICSR's free text narrative. The accurate extraction and organization into structured data fields can support deduplication of cases, data mining algorithms, identification of relevant cases and efficient narrative evaluations [12–17].

Product labeling

Product labeling plays an important role in signal identification. A product's approved labeling is used to determine the expectedness of an AE during signal identification (i.e. has the AE been previously observed—that is, mentioned explicitly in the product's labeling?) [8]. However, drug labels are complex free text, frequently updated; AEs are not required to be described in MedDRA terminology in product labeling by FDA. Manually extracting AE terms from labeling for thousands of FDA-approved products would be very resource intensive and require frequent revisions, as hundreds of safety-related labeling changes occur annually. Integrating the labeling status of each drug product's AEs into safety databases can increase the efficiency of case review and aggregate analyses. Consequently, the use of a semi- or fully automated process for extracting labeled AEs from prescribing information and mapping them to MedDRA has been evaluated [18, 19]. Most recently, the MITRE Corporation and the FDA organized a shared task to evaluate the performance of existing automatable techniques for identifying AEs in labeling [20]. Although the evaluated tools did not perform at levels that would allow use without human intervention, they performed well enough to consider implementing in decision-support workflows.

Sentinel system

Within FDA's Sentinel Initiative, the Active Risk Identification and Analysis (ARIA) system uses standardized claims and claims linked with EHR data to monitor the safety of medicines. A barrier to the evaluation of safety issues in the ARIA system is the ability to identify the health outcomes of interest (HOI) with an acceptable positive predictive value [21, 22]. To address this limitation, proof of concept work has demonstrated that the use of NLP on unstructured EHR data may improve the classification accuracy of HOIs like anaphylaxis [23]. The Sentinel Innovation Center has recently initiated several projects to build on this work that aims to develop and validate algorithms for the identification of HOIs using NLP tools combined with ML approaches [24].

Mining ADRs in the pharmaceutical industry

PV of marketed drugs is a basic function in the pharmaceutical industry and part of its legal obligations [25]. While regulatory authorities, such as the EMA and the FDA, can directly collect PV reports, it is pharmaceutical companies that gather the largest share or about 95% in the USA [26]. PV data are collected and reviewed following a schedule based on company-internal and regulatory guidance [27]. Additionally, pharmaceutical companies address PV-related questions and inspections from regulatory agencies worldwide (which can be requested on short notice) and from business partners.

Source heterogeneity

Pharmaceutical companies are confronted with a broad range of content source types that are relevant for PV [26] and include several that are amenable to NLP, such as reports from patients, consumers and healthcare professionals as well as clinical trial and scientific literature reports. Additionally, EHRs are starting to be considered for the purpose of PV [28, 29].

Pharmaceutical companies apply different strategies for each type of source. Clinical trial reports and the scientific literature can be mined with the help of various databases. Spontaneous reports can be received through phone calls, emails, face-to-face communications, SM messages, etc. Each of these media requires its own monitoring procedures and infrastructure, leading to the creation of individual reports that can be tractable with NLP. The content on which such reports are based is naturally multilingual and can contain both lay language from patients and medical language from healthcare professionals.

Access to licensed or restricted content and regulatory expectations are important factors in restricting the scope of sources used in PV activities. Ideally, any existing document describing a drug's effects should be in scope of PV activities. However, unequal and changing access to different content types, for instance for licensed content not available to companies with smaller budgets, limits the scope of sources that are commonly used. For example, the complex and costly licensing landscape involved in the text mining of the full-text scientific literature prevents its establishment as a reference source. On the other hand, abstract databases such as MEDLINE (https://www.nlm.nih.gov/medline/medline_overview.html) and EMBASE (https://www.embase.com/) are well established and mentioned by regulatory authorities. The fact that these, and similar, databases are mainstays in PV practice highlights the importance of regulatory expectations, which are determined not only by the regulatory framework but also by past experience. Any changes in PV practice bring the increased scrutiny that can only be addressed with sound validation practices (30).

Opportunities for text mining

Because of the costly nature of PV monitoring, which currently involves extensive manual work, and benefits from economies of scale, many of the required activities are outsourced to specialized companies that provide extensive manual support. This approach may not scale up on short notice, as seen with the rollout of the Corona virus disease 2019 (COVID-19) vaccines, which was associated with a high number of spontaneous reports (31). To address this and other challenges, computational approaches to PV (32–35) have the potential to reduce costs, address demand surges and even lead to insourcing of some activities.

Some important computational challenges for PV from the point of view of pharmaceutical companies are:

- Extracting ADRs from the scientific literature (36–38);
- Linking existing PV databases to incoming PV reports;
- Extracting, highlighting and/or summarizing key facts from PV-related documents to lighten the work of the safety experts that review them (33, 34, 39);
- Collecting relevant information or documents to address ad hoc regulatory requests on specific known or suspected ADRs;
- Establishing whether incoming ADRs merit further study as safety signals.

Pharmaceutical companies have already used internally gathered spontaneous reports for computational PV research

(33, 34, 39), but their use in an open computational challenge would require careful anonymization and legal clearance. Otherwise, open data, such as literature reports from FAERS or drug labeling from SIDER (Side Effect Resource) database (http://sideeffects.embl.de/), can still be used in such challenges.

Overall, technological, validation and content-access limitations represent an obstacle to more efficient and effective PV processes in the pharmaceutical environment.

ADRs from a toxicology perspective

The importance of text mining and information extraction applications is not limited to data directly associated with observations found in patients. In fact, before carrying out drug safety studies in humans during clinical trials, extensive in vitro testing as well as toxicology in vivo studies using animal models is carried out (6). In the case of in vitro studies, they generate valuable results to better understand and characterize candidate drug molecular targets as well as potential off-target interactions that could lead to undesired adverse reactions. Text mining strategies to extract automatically from the literature drug–target interactions as well as drug-metabolism and drug-induced gene expression alterations have been addressed by the recent BioCreative DrugProt track (40). Moreover, pre-clinical in vivo animal studies represent a key step toward further safety testing in humans, required for subsequent regulatory approval. Pre-clinical safety studies provide a wealth of data and information sources for the design and analysis of human clinical trial studies. They are key to determine if a candidate drug has a low incidence of adverse side effects (41) or whether there might be some safety concern for clinical trials.

During pre-clinical studies, animal models are used to characterize experimentally in detail potential toxic effects of the administered chemicals under different conditions, typically including different dosages and treatment durations among other aspects. Pre-clinical animal studies serve to characterize duration and dosage-dependent toxicological outcomes at the organ level, providing, for instance, insights into pathological and molecular alterations leading to drug-induced liver injuries or cardio- or nephrotoxicity. Attempts have been made to use text mining strategies to be more systematic in the extraction of chemically induced adverse reactions from scientific literature, patents, pharma company legacy reports as well as other data sources (6, 41). The systematic extraction of associations between chemicals and organ-level toxicities is critical to generate large-scale data repositories that can be exploited by cheminformatics modeling systems. These characterize the associations between certain chemical structures or substructures and toxic AEs, which in turn can be used to guide the molecular drug design by modifying drugs in a way that the unwanted potential toxicity can be reduced. Under such application scenarios, the implemented text mining solutions require not only the detection of chemical entity mentions from a variety of free text but also their normalization to their corresponding chemical structure representations.

Medical chemistry and toxicology experts require text mining solutions that not only support chemical named-entity recognition and text query searches but also enable chemical compound structure search capabilities that require sophisticated name-to-structure conversion components, optical

Source	Score	Sentence
FullText	6.59	Liver function monitoring After a change in the prescribing information for *Troglitazone* in November 1997 liver function was assessed before starting *Troglitazone* treatment and repeated every treatment cycle to detect any elevation of *liver enzymes* nbsp nbsp
PubMed	4.16	Identification of *glutathione* conjugates of *Troglitazone* in *human* hepatocytes
FullText	3.26	Therefore the dual effect of *Troglitazone* in inducing a marked cellular acidosis and decreasing assayable *ALT* activity may account for the extraordinary potency of this compound to inhibit a major metabolic pathway sustaining growth in normal (18) and specifically in malignant cells (3 12) as well as to contribute to its *potential hepatotoxicity*
FullText	2.84	No effect was observed with 1 nbsp 181 M *Troglitazone* and 100 nbsp 181 M *Troglitazone* was highly toxic for the cells as shown by a strong increase in the extracellular *lactate dehydrogenase* activity (data not shown)
PubMed	2.41	Furthermore a *GSH* adduct of troglitazone M2 limited to *GSTM1* wild-type hepatocytes was detected
PubMed	2.14	Careful monitoring of serum *liver enzymes* during *Troglitazone* therapy is therefore essential
PubMed	2.04	The decline followed increasingly stringent requirements for *liver function test* monitoring and may have been due to improved *patient* selection and management as a result of the widely publicized association between *Troglitazone* and *hepatotoxicity*

Figure 3. Examples of entities and relations from toxicology text mining.

chemical structure recognition tools as well as chemical data integration and harmonization strategies (42).

Moreover, chemically induced AEs are described in the scientific literature in terms of traditional drug safety and toxicology studies; these are also a valuable information source for experimentally induced toxicity studies trying to examine the protective effect of certain substances, for instance with respect to hepatotoxicity (43).

AEs caused by chemical compounds and drugs observed in pre-clinical toxicology studies have been harmonized through controlled terminologies such as MedDRA, although some examinations from animal necropsies and pathology observations cannot easily be harmonized to current standard vocabularies due to lack of granularity or coverage. Organ-level AEs observed in pre-clinical, but also clinical studies, may also be reported in the form of changes (e.g. increase or decrease) of certain measurable observable entities like biochemical markers, lab test results or cell counts. For instance, liver toxicity is often studied by measuring changes in biochemical markers like bilirubin or enzymes (ASAT or ALAT). Figure 3 provides examples of different entity types and relations of relevance for toxicology text mining applications.

Challenges of evaluating ADE extraction from clinical narratives

Clinical narratives document the details of patient condition, providing a valuable resource that can complement the structured data for clinical care, outcomes prediction, clinical research (44) and other applications. These narratives detail the observations of the caregivers on the patient state, discuss the pertinent medical history of the patient, explain the rationale for the actions taken in response to the findings and diagnoses and even provide details on the patient's response to treatments (45). ADEs (46) are among the information documented in the narratives.

NLP methods can extract the salient information discussed in clinical narratives. The extracted information can then be put into a structured format for use by downstream computerized applications (47, 48), including PV applications, as noted in earlier sections (23, 28, 29). Information from clinical narrative is drawn from a much wider range of patients than clinical trial data, so these data are particularly valuable for pharmacogenomics studies (49) or for detecting hypersensitivity reactions (50).

Discharge summaries are particularly rich in narrative descriptions of the patient state. However, they contain few instances of ADEs. The relative infrequency of ADEs in these notes limits the sample sizes for NLP system development for their automatic extraction. Additionally, their infrequency presents challenges for data set creation, gold standard generation and evaluation of NLP methods. In this section, we review the challenges of ADE extraction, from the perspective of evaluating NLP methods in clinical records. We also present a discussion of some possible ways of overcoming these challenges.

Since 2017, there have been multiple efforts on evaluating ADE extraction from clinical narratives (51, 52), drug labels (18, 20) and even SM (53, 54). The observations here are based on the 2018 n2c2 shared task in ADE extraction (51), which focused on discharge summaries from the MIMIC III data set (55). This shared task aimed to evaluate NLP systems for their ability to extract spans and relations of drugs and their strengths, dosages, forms, routes, frequencies, durations, reasons for administration and ADEs.

Challenges of data set creation for ADE extraction

Data set creation for infrequent events, such as ADEs, requires identification of samples that sufficiently represent the event, so that annotation resources and system development efforts can focus on those samples. Possible methods for data set creation include utilizing existing information such as International Classification of Diseases (ICD) codes to filter out the abundant irrelevant samples and to focus on the relevant samples, pre-screening the data or pre-annotating the data with existing NLP methods. Often, structured information such as ICD codes provides a good starting point for identifying narratives with ADE information. However, not all notes with ADE-related ICD codes will have mentions of ADEs, and conversely, some notes that do not have such ICD codes will have mentions of ADEs. The more promising approach of manually pre-screening the notes to identify samples with ADEs remains intractable in terms of the labor it requires. The third possible approach of utilizing existing automated methods to pre-annotate ADEs introduces biases of the utilized methods into the data set. As a result, data set creation efforts may combine these approaches, so that they can balance out the need for identifying a good, representative data set with the biases and labor that come with each approach.

Challenges of gold standard generation

Generating gold standard annotations for tasks such as ADE extraction can be challenging. The ADEs and the information related to them, such as drugs, their strengths, dosages, forms, routes, frequencies, durations and reasons, present a complex web of relations that must be carefully considered and interpreted. When the sought-after information is so complex, annotation errors are common and can lead to low inter-annotator agreement. Pre-annotation with existing systems can help with managing the cognitive overload that comes with the task complexity; however, it comes with its own downsides: the bias introduced by the system and complexity of correcting any pre-annotations to fit into the desired annotation schema.

In addition to task complexity, the sheer semantics of ADEs make them challenging to annotate. In particular, ADEs and reasons for medication administration are ambiguous and can be confused with each other, since both are medical problems and differ from each other only in the way they relate to the drug. Differentiating and disambiguating them correctly requires careful consideration of the context in which they are discussed and often requires domain knowledge.

Challenges of automatic ADE extraction

Systems for automatic ADE extraction face two challenges: the complexity of the task especially with regards to ambiguity of ADEs vs. reasons for medication administration, and the infrequency of ADEs. The NLP systems in the 2018 n2c2 evaluation had an average F-measure of 0.84 across all tasks (extracting drugs, their strengths, dosages, forms, routes, frequencies, durations, reasons and ADEs); however, the best performance on ADEs was under 0.60, and the best performance on reasons was under 0.8. Unsurprisingly, most of the system errors on ADE extraction result from incorrect disambiguation from reasons for medication administration.

Mining ADRs from SM

Application and end uses/end users—why do it?

SRSs like the FAERS are at the core of current post-marketing medication surveillance (PV). Although mining SRS data has proven useful for identifying safety signals of serious effects (56), the weaknesses of SRSs are well-documented, including, among others, significant reporting bias (57) and incompleteness of the reports (58).

Reporting bias refers to the fact that healthcare providers report to the SRSs what they deem important, such that serious events are over-represented, while bothersome side effects that may be of great importance to patients and lead to non-adherence and non-persistence are under-represented (59). When patients are asked to list and score AEs on how bothersome they were, experiencing at least one AE perceived as 'extremely' bothersome more than doubled the odds of complete non-adherence (60, 61). Perceived weight gain and difficulty thinking or concentrating attributed to medication cause high distress to patients and are also leading causes of non-adherence to antipsychotic medications (62, 63). High levels of medication adherence have been shown to decrease hospitalization risk and lower overall medical costs for patients (64, 65). However, about 50% of medications are not taken

or not taken as prescribed, leading to increased morbidity and mortality and at an estimated cost of $100–$239 billion per annum to the US healthcare system (66).

Incompleteness refers to the fact that SRSs often lack crucial information such as pre-existing behaviors and medical conditions (e.g. allergies, pregnancy, smoking, alcohol use and hepatic/kidney renal insufficiency) and concomitant treatments. These problems greatly reduce the value of SRS data for many PV goals, such as identifying events in patients with characteristics not encoded in discrete fields, e.g. pregnancy, or evaluating rates and sources of intolerability that lead to non-adherence and non-persistence, which are major clinical and public health problems that increase healthcare costs (67).

The limitations of SRSs have prompted researchers and regulatory agencies in charge of drug safety to explore additional data sources for more effective ADE monitoring, such as EHRs, claims data and, relatively recently, SM. SM, which is the focus of this section, may be particularly useful for identifying adverse effects that are sources of intolerability and lead to non-adherence/persistence but may not be reported by physicians because they are not serious or unexpected. Indeed, a recent systematic review by Golder *et al.* (68) concluded that mild and symptom-related effects are over-represented in SM as compared to other data sources, making it an ideal complement to balance reporting bias in SRSs. SM may also include other important health information and health behaviors not often available through SRSs (and thus, addressing incompleteness) as it is a rich source of self-reports on health behaviors (69).

Existing approaches (what is being done now?)

In general, for the direct application of SM as a source of signal for the detection of AEs, after collecting patient-generated data from SM that mention a medication, it is necessary to solve three problems: (i) determining whether such mention includes the mention of an AE, (ii) the span of the mention and (iii) determining which AE is mentioned (specific to a given standard vocabulary, such as MedDRA). These three problems are usually referred to as classification, extraction and normalization of AEs. A recent publication (70) discusses the state of the art for these three problems; a detailed discussion is beyond the scope of this overview.

Data sources and annotated gold standards

Data and comments about medications directly from the patient are abundant. There are multiple potential data sources that can bring the patient's voice directly, each with different advantages and limitations. In general, usable data specific to medication intake and adverse effects can be found in specialized health forums for a specific condition, as well as generic health forums (such as WebMD, DailyStrength or Drugs.com), or on Reddit and Twitter. The use of health forum data is more limited, with terms of use that generally prohibit their use by commercial entities and might even restrict any use of the data. The use of Reddit and Twitter data is more open, with no restrictions placed on its use, as long as the specific terms of use for sharing are followed.

Table 1. Table of sample applications and insertion points for text mining in drug discovery and PV

Application	Input source	Stakeholder	Drug discovery cycle	Challenges for NLP	Access barriers	Evaluations or resource
Drug mechanism	Literature	Academic and pharma	Discovery and repurposing	Drug interaction w gene/protein	Literature paywalls and competitive intelligence	DrugProt (BioCreative VII)
Toxicity studies	Literature and toxicology reports	Pharma	Pre-clinical	Chemical-organ interactions; vocabulary for animal studies	Literature paywalls and competitive intelligence	eTOX
Clinical trials	Clinical trial data and EHR	Pharma and regulatory agencies	Clinical trial	Finding relevant occurrences in EHR; distinguishing AEs from other medical conditions	Privacy and local language access	n2c2, MADE 1.0
Marketing intelligence	Patent filings	Pharma	Marketing	Drug interaction w gene/protein	Multilingual	CHEMD-NER Patents BioCreative V (2015)
PV	Spontaneous reports	Pharma, regulatory agencies, clinicians and consumers	Clinical trials and post-market surveillance	Completeness of reports, duplicate reports and timeline of drug administration and adverse reaction	Privacy and local language access	
PV	Drug product inserts	Regulatory agencies and consumers	Post-market surveillance	Distinguishing AEs from other medical conditions	XML labels available @ DailyMed	NIST TAC (2017) ADR; ADE Eval
PV	Literature (case reports)	Pharma, regulatory agencies and clinicians	Clinical trials and post-market surveillance	Handling article full text; finding relevant case studies; temporal and causal relation between drug administration and AE	Literature paywalls	
PV	SM	Regulatory agencies, clinicians and consumers	Post-market surveillance	Finding relevant tweets; use of informal language	Identifying useful samples; access to feeds	SMM4H (2017, 2018, 2019, 2021)

Conclusion

The preceding sections provide examples of opportunities for text mining to aid in the complex process of drug discovery, testing, marketing and post-market surveillance. These are summarized in Table 1. While the specific applications described above have focused on English language sources and resources, many of the relevant types of input are written in languages other than English. Therefore, there is growing activity to develop capabilities in multiple languages to capture information in medical records, clinical trials, SRSs and SM (71). The requirements include not only language-specific text mining tools but also appropriate language-specific resources and nomenclatures, e.g. MeSH or MedDRA.

From an information extraction perspective, it is worth noting that different applications need to capture different concepts and relations, depending on the types of input and the application. When extracting information from EHRs or case reports for purposes of detecting possible ADRs, it is important to capture specifics about all the patient's medications (and dosing information), as well as signs, symptoms, lab values and medical conditions—and the timeline that connects these, since a patient may be on multiple medications for different medical conditions. Therefore, the extraction of ADEs may be best framed as capture of a relation between drug and AE or, alternatively, as an attribute on the drug or the medical condition. In other contexts such as drug product labels, where the label concerns a specific drug and certain sections of the drug product label are devoted to listing adverse reactions, extraction of ADRs can be treated more simply as a concept

extraction problem—although there is still potential ambiguity between other medical conditions such as indications or pre-existing conditions. In still other sources, such as SM or reports from SRSs, it is often challenging even to identify what drug or medication is being discussed, given issues of informal language and telegraphic text.

The stakeholders are eager to embrace NLP and AI tools to help in this process, provided that these tools can be demonstrated to add value to stakeholder workflows. In the pharmaceutical industry, the push toward digitalization of healthcare is putting the focus on more computationally driven data analysis, of which text mining approaches to PV are part. Advances in NLP/AI keep the momentum behind the exploration of applications that had not been considered before due to uncertainty of their regulatory acceptance. Validation, both retrospective and prospective, within a changing scientific environment remains the ultimate hurdle for the establishment of productive pipelines.

Regulatory agencies have also shown interest in experimenting with the application of these technologies across a range of input types (72). This creates an opportunity for the BioCreative community to work in partnership with regulatory agencies, pharma and the text mining community to identify next steps for future challenge evaluations.

Disclaimer

This report includes an account of work sponsored by an agency of the US Government. Neither the US Government

nor any agency thereof, nor any of their employees, makes any warranty, express or implied or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the US Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the US Government or any agency thereof.

Supplementary data

Supplementary data are available at Database Online.

Acknowledgements

Funding

List of Abbreviations

AE Adverse event
ADE Adverse drug event
ADR Adverse drug reaction
AI Artificial intelligence
EHR Electronic health record
EMA European Medicines Agency
FAERS FDA Adverse Event Reporting System
FDA Food and Drug Agency
HOI Health outcomes of interest
ICD International Classification of Diseases
ICSR Individual case safety report
ML Machine learning
MedDRA Medical Dictionary for Regulatory Activities
NLP Natural language processing
PV Pharmacovigilance
SM Social media
SRS Spontaneous reporting system
WHO World Health Organization

Conflict of interest

R.R.-E. is an employee of F. Hoffmann-La Roche Ltd.

References

1. Arighi,C.N., Wu,C.H., Cohen,K.B. *et al.* (2014) BioCreative-IV virtual issue. *Database (Oxford)*, **2014**, 1–6. 10.1093/database/bau039.
2. Arighi,C.N., Lu,Z., Krallinger,M. *et al.* (2011) Overview of the BioCreative III Workshop. *BMC Bioinform.*, **12**, S1. 10.1186/1471-2105-12-S8-S1.
3. Krallinger,M., Morgan,A., Smith,L. *et al.* (2008) Evaluation of text-mining systems for biology: overview of the Second BioCreative community challenge. *Genome Biol.*, **9 Suppl 2**, S1. 10.1186/gb-2008-9-s2-s1.
4. Hirschman,L., Yeh,A., Blaschke,C. *et al.* (2005) Overview of BioCreAtIvE: critical assessment of information extraction for biology. *BMC Bioinform.*, **6 Suppl 1**, S1. 10.1186/1471-2105-6-S1-S1.
5. Krallinger,M., Leitner,F., Rabal,O. *et al.* (2015) CHEMDNER: the drugs and chemical names extraction challenge. *J Cheminform*, **7**, S1. 10.1186/1758-2946-7-S1-S1.
6. Canada,A., Capella-Gutierrez,S., Rabal,O. *et al.* (2017) LimTox: a web tool for applied text mining of adverse event and toxicity associations of compounds, drugs and genes. *Nucleic Acids Res.*, **45**, W484–W489. 10.1093/nar/gkx462.
7. Database (Oxford). *Virtual BioCreative Issue*. https://academic.oup.com/database/pages/BioCreative_virtual_issue (21 August 2022, date last accessed).
8. US Food and Drug Administration. *Best Practices in Drug and Biological Product Postmarket Safety Surveillance for FDA Staff*. November 2019. https://www.fda.gov/media/130216/download (21 August 2022, date last accessed).
9. Lester,J., Neyarapally,G.A., Lipowski,E. *et al.* (2013) Evaluation of FDA safety-related drug label changes in 2010. *Pharmacoepidemiol. Drug Saf.*, **22**, 302–305. 10.1002/pds.3395.
10. Croteau,D., Pinnow,E., Wu,E. *et al.* (2022) Sources of evidence triggering and supporting safety-related labeling changes: a 10-year longitudinal assessment of 22 new molecular entities approved in 2008 by the US Food and Drug Administration. *Drug Saf.*, **45**, 169–180. 10.1007/s40264-021-01142-3.
11. Efficacy Guidelines. *International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH)*. https://www.ich.org/page/efficacy-guidelines (21 February 2022, date last accessed).
12. Kreimeyer,K., Foster,M., Pandey,A. *et al.* (2017) Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J. Biomed. Inform.*, **73**, 14–29. 10.1016/j.jbi.2017.07.012.
13. Botsis,T., Scott,J., Goud,R. *et al.* (2014) Novel algorithms for improved pattern recognition using the US FDA Adverse Event Network Analyzer. *Stud. Health Technol. Inform.*, **205**, 1178–1182. 10.3233/978-1-61499-432-9-1178.
14. Botsis,T., Woo,E.J. and Ball,R. (2013) Application of information retrieval approaches to case classification in the vaccine adverse event reporting system. *Drug Saf.*, **36**, 573–582. 10.1007/s40264-013-0064-4.
15. Wunnava,S., Qin,X., Kakar,T. *et al.* Towards transforming FDA adverse event narratives into actionable structured data for improved pharmacovigilance. In: *2017 Proceedings of the Symposium on Applied Computing*, Marrakech, Morocco. pp. 777–82. 10.1145/3019612.3022875.
16. Pham,P., Cheng,C., Wu,E. *et al.* (2021) Leveraging case narratives to enhance patient age ascertainment from adverse event reports. *Pharmaceut. Med.*, **35**, 307–316. 10.1007/s40290-021-00398-5.
17. Kreimeyer,K., Dang,O., Spiker,J. *et al.* (2021) Feature engineering and machine learning for causality assessment in

pharmacovigilance: lessons learned from application to the FDA Adverse Event Reporting System. *Comput. Biol. Med.*, **135**, 104517. 10.1016/j.compbiomed.2021.104517.

18. Roberts,K., Demner-Fushman,D. and Tonning,J.M. (2017) Overview of the TAC 2017 adverse reaction extraction from drug labels track. In: *Proceedings of the 2017 Text Analysis Conference*, Gaithersburg, MD. NIST. https://tac.nist.gov/publications/2017/additional.papers/TAC2017.ADR_overview.proceedings.pdf (21 February 2022, date last accessed).

19. Ly,T., Pamer,C., Dang,O. *et al.* (2018) Evaluation of natural language processing (NLP) systems to annotate drug product labeling with MedDRA terminology. *J. Biomed. Inform.*, **83**, 73–86. 10.1016/j.jbi.2018.05.019.

20. Bayer,S., Clark,C., Dang,O. *et al.* (2021) ADE eval: an evaluation of text processing systems for adverse event extraction from drug labels for pharmacovigilance. *Drug Saf.*, **44**, 83–94. 10.1007/s40264-020-00996-3.

21. Brown,J.S., Maro,J.C., Nguyen,M. *et al.* (2020) Using and improving distributed data networks to generate actionable evidence: the case of real-world outcomes in the Food and Drug Administration's Sentinel system. *J. Am. Med. Inform. Assoc.*, **27**, 793–797. 10.1093/jamia/ocaa028.

22. Anklam,E., Bahl,M.I., Ball,R. *et al.* (2022) Emerging technologies and their impact on regulatory science. *Exp. Biol. Med. (Maywood)*, **247**, 1–75. 10.1177/15353702211052280.

23. Ball,R., Toh,S., Nolan,J. *et al.* (2018) Evaluating automated approaches to anaphylaxis case classification using unstructured data from the FDA Sentinel System. *Pharmacoepidemiol. Drug Saf.*, **27**, 1077–1084. 10.1002/pds.4645.

24. Desai,R.J., Matheny,M.E., Johnson,K. *et al.* (2021) Broadening the reach of the FDA Sentinel system: a roadmap for integrating electronic health record data in a causal analysis framework. *NPJ Digit. Med.*, **4**, 170. 10.1038/s41746-021-00542-0.

25. Pharmacovigilance Risk Assessment Committee (PRAC). (2015) Good practice guide on recording, coding, reporting and assessment of medication errors. EMA/762563/2014.

26. Center for Drug Evaluation and Research. (2016) U.S. Food and Drug Administration. Drug Safety Priorities 2016: Initiatives and Innovation. https://www.fda.gov/drugs/drug-safety-and-availability/drug-safety-priorities-2016-initiatives-and-innovation (21 August 2022, date last accessed).

27. European Medicines Agency. (2022) Guidelines on good pharmacovigilance practices (GVP). Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev. 2) [online]. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices (18 March 2022, date last accessed).

28. Gini,R., Sturkenboom,M.C.J., Sultana,J. *et al.* (2020) Different strategies to execute multi-database studies for medicines surveillance in real-world setting: a reflection on the European model. *Clin. Pharmacol. Ther.*, **108**, 228–235. 10.1002/cpt.1833.

29. Star,K., Watson,S., Sandberg,L. *et al.* (2015) Longitudinal medical records as a complement to routine drug safety signal analysis. *Pharmacoepidemiol. Drug Saf.*, **24**, 486–494. 10.1002/pds.3739.

30. Huysentruyt,K., Kjoersvik,O., Dobracki,P. *et al.* (2021) Validating intelligent automation systems in pharmacovigilance: insights from good manufacturing practices. *Drug Saf.*, **44**, 261–272. 10.1007/s40264-020-01030-2.

31. Kant,A., van Hunsel,F. and van Puijenbroek,E. (2022) Numbers of spontaneous reports: how to use and interpret? *Br J Clin Pharmacol*, **88**, 1365–1368. 10.1111/bcp.15024.

32. Lewis,D.J. and McCallum,J.F. (2020) Utilizing advanced technologies to augment pharmacovigilance systems: challenges and opportunities. *Ther. Innov. Regul. Sci.*, **54**, 888–899. 10.1007/s43441-019-00023-3.

33. Abatemarco,D., Perera,S., Bao,S.H. *et al.* (2018) Training augmented intelligent capabilities for pharmacovigilance: applying deep-learning approaches to individual case safety report processing. *Pharmaceut. Med.*, **32**, 391–401. 10.1007/s40290-018-0251-9.

34. Mockute,R., Desai,S., Perera,S. *et al.* (2019) Artificial intelligence within pharmacovigilance: a means to identify cognitive services and the framework for their validation. *Pharmaceut. Med.*, **33**, 109–120. 10.1007/s40290-019-00269-0.

35. Wadhwa,D., Kumar,K., Batra,S. *et al.* (2021) Automation in signal management in pharmacovigilance-an insight. *Brief. Bioinformatics*, **22**, bbaa363. 10.1093/bib/bbaa363.

36. Gurulingappa,H., Mateen-Rajput,A. and Toldo,L. (2012) Extraction of potential adverse drug events from medical case reports. *J. Biomed. Semantics*, **3**, 15. 10.1186/2041-1480-3-15.

37. Yeleswarapu,S., Rao,A., Joseph,T. *et al.* (2014) A pipeline to extract drug-adverse event pairs from multiple data sources. *BMC Med. Inform. Decis. Mak.*, **14**, 13. 10.1186/1472-6947-14-13.

38. Negi,K., Pavuri,A., Patel,L. *et al.* (2019) A novel method for drug-adverse event extraction using machine learning. *Inform. Med. Unlocked*, **17**, 100190. 10.1016/j.imu.2019.100190.

39. Eskildsen,N.K., Eriksson,R., Christensen,S.B. *et al.* (2020) Implementation and comparison of two text mining methods with a standard pharmacovigilance method for signal detection of medication errors. *BMC Med. Inform. Decis. Mak.*, **20**, 94. 10.1186/s12911-020-1097-0.

40. Miranda-Escalada,A., Mehryary,F., Luoma,J. *et al.* (2021) *Overview of DrugProt BioCreative VII track: quality evaluation and large scale text mining of drug-gene/protein relations Proceedings of the BioCreative VII Challenge Evaluation Workshop*, 11. https://biocreative.bioinformatics.udel.edu/media/store/files/2021/Track1_pos_1_BC7_overview.pdf (21 August 2022, date last accessed).

41. Natsiavas,P., Malousi,A., Bousquet,C. *et al.* (2019) Computational advances in drug safety: systematic and mapping review of knowledge engineering based approaches. *Front. Pharmacol.*, **10**, 415. 10.3389/fphar.2019.00415.

42. Zheng,S., Dharssi,S., Wu,M. *et al.* (2019) Text mining for drug discovery. *Methods Mol. Biol.*, **1939**, 231–252. 10.1007/978-1-4939-9089-4_13.

43. Jornod,F., Jaylet,T., Blaha,L. *et al.* (2021) AOP-helpFinder webserver: a tool for comprehensive analysis of the literature to support adverse outcome pathways development. *Bioinformatics*, **38**, 1173–1175. 10.1093/bioinformatics/btab750.

44. Stephens,K.A., Au,M.A., Yetisgen,M. *et al.* Leveraging UMLS-driven NLP to Enhance Identification Of Influenza Predictors Derived from Electronic Medical Record Data. https://www.biorxiv.org/content/10.1101/2020.04.24.058982v1?rss=1 (21 August 2022, date last accessed).

45. Gu,Q., Dillon,C.F. and Burt,V.L. (2010) Prescription drug use continues to increase: U.S. prescription drug data for 2007-2008. NCHS Data Brief, 1–8.

46. Flockhart,D.A.,Yasuda,S.U., Honig,P. *et al.* Center for Education and Research on Therapeutics (CERT), Georgetown University. *Preventable Adverse Drug Reactions: A Focus on Drug Interactions*. https://www.fda.gov/drugs/drug-interactions-labeling/preventable-adverse-drug-reactions-focus-drug-interactions (21 August 2022, date last accessed).

47. Demner-Fushman,D., Chapman,W.W. and McDonald,C.J. (2009) What can natural language processing do for clinical decision support? *J. Biomed. Inform.*, **42**, 760–772. 10.1016/j.jbi.2009.08.007.

48. Uzuner,O. and Stubbs,A. (2015) Practical applications for natural language processing in clinical research: the 2014 i2b2/UTHealth shared tasks. *J. Biomed. Inform.*, **58**, S1–S5. 10.1016/j.jbi.2015.10.007.

49. Wilke,R.A., Xu,H., Denny,J.C. *et al.* (2011) The emerging role of electronic medical records in pharmacogenomics. *Clin. Pharmacol. Ther.*, **89**, 379–386. 10.1038/clpt.2010.260.

50. Wong,A., Seger,D.L., Lai,K.H. *et al.* (2019) Drug hypersensitivity reactions documented in electronic health records within a large health system. *J Allergy Clin Immunol Pract*, **7**, 1253–1260 e1253. 10.1016/j.jaip.2018.11.023.

51. Henry,S., Buchan,K., Filannino,M. *et al.* (2020) 2018 n2c2 shared task on adverse drug events and medication extraction in electronic health records. *J. Am. Med. Inform. Assoc.*, **27**, 3–12. 10.1093/jamia/ocz166.

52. Jagannatha,A., Liu,F., Liu,W. *et al.* (2019) Overview of the first natural language processing challenge for extracting medication, indication, and adverse drug events from electronic health record notes (MADE 1.0). *Drug Saf.*, **42**, 99–111. 10.1007/s40264-018-0762-z.

53. Sarker A, Belousov M, Friedrichs J, *et al.* (2018) Data and systems for medication-related text classification and concept normalization from Twitter: insights from the Social Media Mining for Health (SMM4H)-2017 shared task. *J Am Med Inform Assoc*; **25**, 1274–1283. 10.1093/jamia/ocy114.

54. Sarker,A. and Gonzalez-Hernandez,G. Overview of the second social media mining for health (SMM4H) shared tasks at AMIA 2017. Training 1(10,822): 1239.

55. Johnson,A.E., Pollard,T.J., Shen,L. *et al.* (2016) MIMIC-III, a freely accessible critical care database. *Sci. Data*, **3**, 160035. 10.1038/sdata.2016.35.

56. Michel,C., Scosyrev,E., Petrin,M. *et al.* (2017) Can disproportionality analysis of post-marketing case reports be used for comparison of drug safety profiles? *Clin. Drug Investig.* **37**, 415–422. 10.1007/s40261-017-0503-6.

57. Evens,A.M., Jovanovic,B.D., Su,Y.C. *et al.* (2011) Rituximab-associated hepatitis B virus (HBV) reactivation in lymphoproliferative diseases: meta-analysis and examination of FDA safety reports. *Ann. Oncol.*, **22**, 1170–1180. 10.1093/annonc/mdaq583.

58. Harpaz,R., DuMouchel,W., Shah,N.H. *et al.* (2012) Novel data-mining methodologies for adverse drug event discovery and analysis. *Clin. Pharmacol. Ther.*, **91**, 1010–1021. 10.1038/clpt.2012.50.

59. Golomb,B.A., McGraw,J.J., Evans,M.A. *et al.* (2007) Physician response to patient reports of adverse drug effects: implications for patient-targeted adverse effect surveillance. *Drug Saf.*, **30**, 669–675. 10.2165/00002018-200730080-00003.

60. Goethe,J.W., Woolley,S.B., Cardoni,A.A. *et al.* (2007) Selective serotonin reuptake inhibitor discontinuation: side effects and other factors that influence medication adherence. *J. Clin. Psychopharmacol.*, **27**, 451–458. 10.1097/jcp.0b013e31815152a5.

61. Bull,S.A., Hu,X.H., Hunkeler,E.M. *et al.* (2002) Discontinuation of use and switching of antidepressants: influence of patient-physician communication. *JAMA*, **288**, 1403–1409. 10.1001/jama.288.11.1403.

62. Fakhoury,W.K., Wright,D. and Wallace,M. (2001) Prevalence and extent of distress of adverse effects of antipsychotics among callers to a United Kingdom National Mental Health Helpline. *Int. Clin. Psychopharmacol.*, **16**, 153–162. 10.1097/00004850-200105000-00004.

63. Wong,M.M., Chen,E.Y., Lui,S.S. *et al.* (2011) Medication adherence and subjective weight perception in patients with first-episode psychotic disorder. *Clin. Schizophr. Relat. Psychoses*, **5**, 135–141. 10.3371/CSRP.5.3.3.

64. Wittich,C.M., Burkle,C.M. and Lanier,W.L. (2012) Ten common questions (and their answers) about off-label drug use. *Mayo Clin. Proc.*, **87**, 982–990. 10.1016/j.mayocp.2012.04.017.

65. Osterberg,L. and Blaschke,T. (2005) Adherence to medication. *N. Engl. J. Med.*, **353**, 487–497. 10.1056/NEJMra050100.

66. Viswanathan,M., Golin,C.E., Jones,C.D. *et al.* (2012) Interventions to improve adherence to self-administered medications for chronic diseases in the United States: a systematic review. *Ann. Intern. Med.*, **157**, 785–795. 10.7326/0003-4819-157-11-201212040-00538.

67. Cramer,J.A., Roy,A., Burrell,A. *et al.* (2008) Medication compliance and persistence: terminology and definitions. *Value Health*, **11**, 44–47. 10.1111/j.1524-4733.2007.00213.x.

68. Golder,S., Norman,G. and Loke,Y.K. (2015) Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *Br J Clin Pharmacol*, **80**, 878–888. 10.1111/bcp.12746.

69. Golder,S., O'Connor,K., Hennessy,S. *et al.* (2020) Assessment of beliefs and attitudes about statins posted on Twitter: a qualitative study. *JAMA Netw. Open*, **3**, e208953. 10.1001/jamanetworkopen.2020.8953.

70. Magge,A., Tutubalina,E., Miftahutdinov,Z. *et al.* (2021) DeepADEMiner: a deep learning pharmacovigilance pipeline for extraction and normalization of adverse drug event mentions on Twitter. *J. Am. Med. Inform. Assoc.*, **28**, 2184–2192. 10.1093/jamia/ocab114.

71. Gonzalez-Agirre,A., Marimon,M., Intxaurrondo,A. *et al.* (2019) Pharmacological substances, compounds and proteins named entity recognition track. In: *Proceedings of the 5th Workshop on BioNLP Open Shared Tasks*, Hong Kong, China, 1–10.

72. Wisniewski,A.F., Bate,A., Bousquet,C. *et al.* (2016) Good signal detection practices: evidence from IMI PROTECT. *Drug Saf.*, **39**, 469–490. 10.1007/s40264-016-0405-1.