Journal of Biomedical Informatics 52 (2014) 293–310

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Identifying plausible adverse drug reactions using knowledge extracted from the literature

Ning Shang [a,*], Hua Xu [a], Thomas C. Rindflesch [b], Trevor Cohen [a]

[a] School of Biomedical Informatics, The University of Texas Health Science Center, Houston, TX, United States
[b] National Library of Medicine, Bethesda, MD, United States

ARTICLE INFO

Article history:
Received 24 February 2014
Accepted 10 July 2014
Available online 19 July 2014

Keywords:
Pharmacovigilance
Distributional semantics
Literature-based discovery
Predication-based semantic indexing
Reflective random indexing

ABSTRACT

Pharmacovigilance involves continually monitoring drug safety after drugs are put to market. To aid this process; algorithms for the identification of strongly correlated drug/adverse drug reaction (ADR) pairs from data sources such as adverse event reporting systems or Electronic Health Records have been developed. These methods are generally statistical in nature, and do not draw upon the large volumes of knowledge embedded in the biomedical literature. In this paper, we investigate the ability of scalable Literature Based Discovery (LBD) methods to identify side effects of pharmaceutical agents. The advantage of LBD methods is that they can provide evidence from the literature to support the plausibility of a drug/ADR association, thereby assisting human review to validate the signal, which is an essential component of pharmacovigilance. To do so, we draw upon vast repositories of knowledge that has been extracted from the biomedical literature by two Natural Language Processing tools, MetaMap and SemRep. We evaluate two LBD methods that scale comfortably to the volume of knowledge available in these repositories. Specifically, we evaluate Reflective Random Indexing (RRI), a model based on concept-level co-occurrence, and Predication-based Semantic Indexing (PSI), a model that encodes the nature of the relationship between concepts to support reasoning analogically about drug-effect relationships. An evaluation set was constructed from the Side Effect Resource 2 (SIDER2), which contains known drug/ADR relations, and models were evaluated for their ability to "rediscover" these relations. In this paper, we demonstrate that both RRI and PSI can recover known drug-adverse event associations. However, PSI performed better overall, and has the additional advantage of being able to recover the literature underlying the reasoning pathways it used to make its predictions.

1. Introduction

An adverse drug reaction (ADR) is an "appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medical product" [1]. ADRs were reported to be between the fourth and sixth leading cause of death in the United States in 1994 [2], accounting for 3–7% of medical hospital admissions [3,4] and a substantial number of health care visits [5]. They have a considerable negative impact on health and the healthcare system, despite the fact that extensive pre-marketing clinical trials are designed to test drug safety and efficacy. For example Phase III clinical trials have been estimated to cost 86.3 million U.S. dollars and last 30.5 months on average [6]. Nonetheless, rare ADRs may not be detected due to the limited duration and sample size of such trials, and others may occur on account of idiosyncratic characteristics of individuals excluded from the evaluated sample. The continued monitoring for ADRs after drugs are released into the market, called pharmacovigilance (PV), is therefore an important tool to monitor and improve drug safety.

Over the last decade, drug safety data obtained from spontaneous reporting systems (SRSs) have been analyzed using quantitative data mining procedures to retrieve strongly associated drug/ADR pairs [7–9]. These highlighted associations are subsequently reviewed and scrutinized by domain experts. Unfortunately, research suggests data collected by SRS are limited by long time latency, incorrect or incomplete clinical information, underreporting and reporting bias [10,11]. Consequently, clinicians and researchers have also utilized existing healthcare data sources such as Electronic Health Records (EHRs) to attempt to identify previously unreported ADRs [12–15]. However, these data are inherently noisy as drugs and potential side effects may co-occur in the EHR for many reasons. In addition, the EHR often contains free-text data, and the accuracy of Natural Language Processing

* Corresponding author.
E-mail address: sunnyshang001@gmail.com (N. Shang).

http://dx.doi.org/10.1016/j.jbi.2014.07.011
1532-0464/

(NLP) tools is not perfect. New methods are required to selectively identify potentially hazardous drug/ADR associations. Consequently, the development of computational approaches to more accurately detect potential side effects is currently an active area of research [16–20]. These approaches have predominantly focused on improving signal detection using statistical methods, machine learning (ML) or some combination thereof.

In this paper, we develop an approach that is conceptually different than, and complementary to, such efforts. Methods of literature-based discovery (LBD) are used to detect potential drug/ADR associations and to retrieve literature that supports their plausibility. The paper proceeds as follows. First we discuss the significance and challenges of pharmacovigilance, and how LBD methods might address these. We then provide relevant background on recent developments in LBD; and introduce the NLP tools that were used to extract knowledge from the literature for our experiments. We then discuss these experiments, in which we attempt to identify known ADRs using knowledge from the biomedical literature, and discuss their implications for pharmacovigilance practice.

2. Background

2.1. Pharmacovigilance: post-marketing drug surveillance

Vioxx (Rofecoxib) was withdrawn voluntarily from market by Merck in 2004, after it was found that the use of this agent increased the risk of myocardial infarction [21]. Avandia (Rosiglitazone) was suspended from the European market in 2010 [22–24] on account of an increased risk of cardiovascular complications. These high-profile examples illustrate that PV is an important supplement to existing drug safety profiles because clinical drug trials cannot be large or long enough to identify all problems related to a new drug [7]. Additionally, subjects are pre-selected by eligibility criteria and therefore may not fully represent the patient population after the drugs are put to market [25]. Consequently, it is highly unlikely that instances of all possible ADRs will be detected during pre-marketing clinical trials.

The fact that more than 75 drug products were removed from the market due to safety problems between 1964 and 2002 further emphasizes the importance of post-marketing drug monitoring, known as PV – "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem after drugs are on market" [26]. PV is designed to detect any rare or long-term adverse effects over a very large population and a long period of time. To advance this aim, health departments and organizations (such as the World Health Organization (WHO), U.S. Food and Drug Administration (FDA), and European Medicines Agency (EMA)) encourage physicians, other health care professionals, and patients to report voluntarily about any observed ADRs. In addition to voluntarily reporting, pharmaceutical companies are required to report serious adverse events [27]. These bodies have Spontaneous Reporting Systems (SRSs) to enable the efficient submission of reports electronically [28,29].

In general, the PV process proceeds as follows [30,31]:

(1) Reported drug-related problems are collected in SRSs nationally or internationally.
(2) Quantitative data mining procedures are used to analyze these data and retrieve relatively strongly correlated drug/ADR pairs (drug/ADR associations).
(3) These highlighted associations are then reviewed and evaluated by domain experts making up an expert clinical review panel.
(4) Associations considered to be of clinical interest are then annotated as signals.

Specifically, signal is defined as "reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously" [32]. Overall, the PV process includes two components – a statistical component (quantitative signal detection, steps (1) and (2)) and a qualitative component (expert clinical review, steps (3) and (4)) [31].

Through PV, international and national health institutions gather large amounts of data from SRS for further analysis. In addition, researchers have leveraged the opportunity provided by broader availability of EHRs by utilizing EHR data for signal detection [12,33]. These authors argue that EHR data can compensate for some of the deficiencies of SRS, such as under-reporting, misclassification, a long lag time between observation and reporting, reporting bias and the provision of incomplete clinical information [7,8]. Regardless of source, statistical algorithms are applied to both SRS [34–39] and EHRs [12] to measure the strength of observed drug-event associations.

It has been argued, though, that causality assessment is lacking in pharmacovigilance practice [25]. While expert clinical review is designed to verify potential ADRs, it is a human-intensive and time-consuming process. The available human resources are inadequate to review the large amount of noisy signal detected in SRS and EHR data, creating a bottleneck in the PV process. More research is needed to develop methods to automate, or assist with, the knowledge-intensive task of expert clinical review.

2.2. Assessment of causality

To address the issue of causality assessment, general principles exist that can be applied to evaluate the causality of potential ADRs [40]. The theoretical basis for these principles was proposed by Sir Austin Bradford-Hill in 1965 [41]. Bradford-Hill, an English epidemiologist and statistician, was the first to demonstrate that cigarette smoking contributes toward lung cancer using what are now referred to as the "Bradford-Hill criteria" [42]. The Bradford-Hill criteria provide viewpoints from which to evaluate evidence indicative of causality. These criteria are named 'strength', 'consistency', 'specificity', 'temporality', 'biological gradient' (referring to dose–response relationships), 'plausibility', 'coherence', 'experimental evidence', and 'analogy' [41,43,44]. Since then, the criteria have been widely used in epidemiology and may be applied to assess the causality of drug/ADR relationships [25,40,45]. Three of these criteria seem particularly pertinent to the development of pharmacovigilance methods:

- The *strength* criterion reflects that strong associations are more likely to be causal than weak associations [40]. Quantitative statistical data mining methods evaluate adverse drug reaction signal from the strength of association point of view.
- The *plausibility* criterion relates to evidence about mechanisms that may be involved to support a causal relationship.
- The *coherence* criterion relates to the consistency of the hypothesis in question with contemporary medical knowledge.

Review by domain experts is required to evaluate a signal from the above points of view using their knowledge and judgment to find a signal with clinical significance. However, on account of the human-intensive nature of this task, automated assistance is desirable. In this study, we attempt to partially automate this