Computational Biology and Chemistry xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Research Article

Pharmacoepidemiological characterization of drug-induced adverse reaction clusters towards understanding of their mechanisms

Sayaka Mizutani, Yousuke Noro, Masaaki Kotera, Susumu Goto*

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho Uji, Kyoto 611-0011, Japan

ARTICLE INFO

Article history:
Accepted 23 December 2013

Keywords:
Adverse drug reaction
Biclustering
Drug indication

ABSTRACT

A big challenge in pharmacology is the understanding of the underlying mechanisms that cause drug-induced adverse reactions (ADRs), which are in some cases similar to each other regardless of different drug indications, and are in other cases different regardless of same drug indications. The FDA Adverse Event Reporting System (FAERS) provides a valuable resource for pharmacoepidemiology, the study of the uses and the effects of drugs in large human population. However, FAERS is a spontaneous reporting system that inevitably contains noise that deviates the application of conventional clustering approaches. By performing a biclustering analysis on the FAERS data we identified 163 biclusters of drug-induced adverse reactions, counting for 691 ADRs and 240 drugs in total, where the number of ADR occurrences are consistently high across the associated drugs. Medically similar ADRs are derived from several distinct indications for use in the majority (145/163 = 88%) of the biclusters, which enabled us to interpret the underlying mechanisms that lead to similar ADRs. Furthermore, we compared the biclusters that contain same drugs but different ADRs, finding the cases where the populations of the patients were different in terms of age, sex, and body weight. We applied a biclustering approach to catalogue the relationship between drugs and adverse reactions from a large FAERS data set, and demonstrated a systematic way to uncover the cases different drug administrations resulted in similar adverse reactions, and the same drug can cause different reactions dependent on the patients' conditions.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Adverse drug reactions (ADRs) are one of the main causes of morbidity and mortality worldwide (Giacomini et al., 2007). Since not all ADRs are detected during drug development processes or clinical trials, many of them become a serious public health concern after the drugs are launched in the market. Post-marketed drugs, however, are continuously surveilled for adverse events and medication errors by regulatory agencies, pharmaceutical companies, and hospitals.

A big challenge in pharmacology is the understanding of the underlying mechanisms that cause these ADRs, which are in some cases similar to each other regardless of different drug indications. For example, a number of psychiatric disorders are expressed as a result of taking anti-psychotics, but some are also reported as side effects of unrelated drugs such as tamiful, an anti-influenza drug. As another example, varenicline, used to treat smoking addiction, and gabapentin, used to relieve neuropathic pain, both have showed risks of causing severe adverse events such as suicidal behaviors. It is suggested the similar ADRs by different drugs are the result of their common mechanisms.

In other cases, even when the same drug is used, the expression of ADRs varies from patients to patients. Types and severity of ADRs may vary depending on the patients' physiological conditions such as age, sex, body weight, as well as their health states, and such information is provided in drug package inserts. Kuhn et al. (2010) collected drug side-effects by text-mining package inserts of marketed drugs, and listed them in the SIDER database. However, their approach has not yet covered these variabilities in the population of drug responses. In order to design personalized medicines, it is important to divide populations of patients according to their reaction expression patterns and characterize the populations with patients physiological backgrounds.

Spontaneous ADR reporting systems monitor unexpected drug outcomes over a large population of patients, and are extremely valuable for pharmacoepidemiology, the study of the effects of drugs focusing on the variabilities of drug responses. The FDA Adverse Event Reporting System (FAERS) (http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm) has collected incidence of ADRs in the form of adverse event reports, where each report lists ADRs and the associated drugs along with the drug indications for use

* Corresponding author. Tel.: +81 774 38 3271.
E-mail addresses: smizutan@kuicr.kyoto-u.ac.jp (S. Mizutani), noro@kuicr.kyoto-u.ac.jp (Y. Noro), kot@kuicr.kyoto-u.ac.jp (M. Kotera), goto@kuicr.kyoto-u.ac.jp (S. Goto).

1476-9271/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compbiolchem.2014.01.006

Please cite this article in press as: Mizutani, S., et al., Pharmacoepidemiological characterization of drug-induced adverse reaction clusters towards understanding of their mechanisms. Comput. Biol. Chem. (2014), http://dx.doi.org/10.1016/j.compbiolchem.2014.01.006

(the patients' disease symptoms for which the drugs are prescribed). An event report also provides the patients' demographic data. These resources provide us with profiles of ADRs and drugs from which knowledge may be mined to help characterize the adverse events.

A number of research studied the FAERS data. For example, Bate and Evans (2009) and Harpaz et al. (2013) detected possible relationships between ADRs and drugs that are unknown or incompletely documented in drug package inserts, but they did not systematically take the effects of the drug indications into consideration. Other studies detected ADR-drug pairs in consideration of only particular indications, e.g., anticoagulants (Tamura et al., 2012). Tatonetti et al. (2012) used information on drug indications or the patients' physiological data to improve the detection method, whereas they did not consider the varieties of indications that cause the same ADR in a mechanistic point of view. In this study, we conducted a comprehensive investigation of the relationships between ADRs and the drug indications to identify drug administration patterns which lead to certain groups of ADRs, aiming at inferring the common mechanisms of reaction occurrences.

Clustering ADRs from the FAERS event reports may give us an inference of the common characteristics of ADR incidences. However, in contrast to the manually curated data such as drug package inserts, the spontaneous reporting system such as FAERS contains more noise, which deviates the application of conventional one-way clustering approaches. Specifically, considering a subset of drugs, rather than using all drugs, may help collecting ADRs whose event frequency profiles show high correlations. Accordingly, we applied a biclustering approach to the event frequency data from FAERS to extract clusters of drugs and ADRs in this study. A biclustering method, one of data mining approaches, is a useful tool to extract knowledge from such noisy data. The advantage of biclustering over one-way clustering is that it allows to focus on a subset of conditions over which features of correlated profiles are grouped in a cluster, while omitting other conditions of less correlation.

It was shown that our biclustering approach extracted clusters with different but possibly overlapping drugs and ADRs. Then we characterized the extracted clusters in the points of drug indications, using event reports in each cluster to infer how the drugs with different indications can cause the incidence of medically similar ADRs. Characterizing the extracted adverse event groups revealed the differences in patients' populations (e.g., age, sex, and body weights) which showed expressions of different ADRs from the same drugs. Our findings may help understanding how an ADR can occur from unrelated drugs and how the patients' physiological backgrounds may be an important factor for increased risk of the ADR.

2. Results and discussion

2.1. Extraction of drug-induced adverse reaction clusters

In order to find the biclusters of ADRs and associated drugs, we obtained 425 273 adverse event reports from the FAERS, which contained the total of 5802 ADRs and 1909 drugs. We coded the drugs in "D numbers", unique identification numbers of drug entries in the KEGG DRUG database (Kanehisa et al., 2012), where drugs are characterized in a number of annotating systems including the Anatomical Therapeutic Chemical (ATC) classification system, and are cross-liked by LinkDB (Fujibuchi et al., 1998) to other pharmacological databases such as DrugBank (Knox et al., 2011), SIDER, and DailyMed (http://dailymed.nlm.nih.gov/dailymed/about.cfm). This yielded 1374 D numbers. Medically similar ADR terms were grouped using the MedDRA dictionary at the "*High Level Terms* (HLTs)", which yielded 1317 HLTs. Based on the data we constructed an event report frequency matrix whose rows are the ADRs and columns are the drugs. Then we performed an unsupervised biclustering to identify subsets of the ADRs (referred to as "REACs" hereafter) and subsets of drugs (referred to as "DRUGs" hereafter), where the frequencies of the REACs are consistently high across the DRUGs, using the Iterative Signature Algorithm (ISA) (Ihmels et al., 2004). In general, biclustering approaches pose a tradeoff between the overall coverage of row and column members, and the stringency of the extracted biclusters. Therefore, we optimized the bicluster extraction by setting a parameter range for the ISA workflow, following the strategy employed by Iskar et al. (2013). Detailed optimization process is described in Section 4. As a result, we extracted 163 biclusters (or "clusters" hereafter) containing 691 REACs and 240 DRUGs, and 182 000 event reports (43% of all event reports counted in the matrix) in total. Some DRUGs and REACs appeared in more than one cluster. Fig. 1 shows the distribution of the FAERS adverse report entities, in which the ADRs were classified by the MedDRA "*System Organ Class* (SOC)" level, and the drugs were classified by the ATC anatomical level. Although the number of REACs and DRUGs extracted in the 163 clusters is not large, biclusering process retained the variations of the classifications. Fig. 2 shows the number of REACs and DRUGs extracted in each of the 163 clusters. Detailed information of the cluster entities are found in Table S1.

2.2. Referring for known ADRs in drug package inserts

In order to check how many REACs in the 163 clusters are documented in drug package inserts, we referred the SIDER database. In the original data matrix, 80 019 drug-ADR pairs appeared at least in one event report, of which 34 315 (43 %) were listed in SIDER. In the 163 clusters, the number of drug-ADR pairs that were extracted at least in one cluster was 5159, of which 2481 pairs (48%) were listed in SIDER. This points out that the biclustering did not significantly affect the ratio of ADRs written in drug package inserts. It is also suggested that, even though the spontaneous ADR reports may contain many noises, the 2678 drug-ADR pairs that were found in the 163 clusters but not in SIDER may be worth investigating in order to enhance drug package inserts.

2.3. Characterization of drug-induced adverse reaction clusters by related indication profiles

Each resulting cluster consists of REACs whose overall event report frequencies were consistently high across associated DRUGs in the cluster. Studying REACs and DRUGs within each cluster is expected to reveal physiological or molecular regulations common to take place in the expressions of REACs in the cluster. First we checked whether REACs in a cluster are medically similar to each other. Since REACs are written in the MedDRA HLT level, they were mapped to the HLGT level, one level higher to the HLT in the MedDRA hierarchy. Medical similarity of REACs was measured by how inequally the REACs were distributed across different HLGT terms, using the Gini coefficient (see Section 4). Gini coefficient measures the inequality among values of a frequency distribution, and high Gini coefficient represents high homogeneity in the population. Fig. 3 (right histogram) shows the distribution of 163 clusters as a function of Gini coefficient. In clusters with higher Gini coefficients, REACs were inequally distributed over a group of HLGTs, i.e., the majority of REACs was categorized only in particular HLGT(s), and were not categorized in any other HLGTs, which means the REACs in these clusters showed medical similarity in the MedDRA hierarchical classification. Similar REACs may be derived from administrations of unrelated drugs. In order to measure the degree of variation in DRUGs we constructed drug indication (INDI, hereafter) profiles within each cluster. The drug INDI profiles are

Please cite this article in press as: Mizutani, S., et al., Pharmacoepidemiological characterization of drug-induced adverse reaction clusters towards understanding of their mechanisms. Comput. Biol. Chem. (2014), http://dx.doi.org/10.1016/j.compbiolchem.2014.01.006

Download English Version:

https://daneshyari.com/en/article/6487119

Download Persian Version:

https://daneshyari.com/article/6487119

Daneshyari.com