Artificial Intelligence in Medicine 76 (2017) 7–15

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports

Ruichu Cai[a,*,1], Mei Liu[b,*,1], Yong Hu[c,1], Brittany L. Melton[d], Michael E. Matheny[e,f], Hua Xu[g], Lian Duan[h], Lemuel R. Waitman[b]

[a] Faculty of Computer Science, Guangdong University of Technology, Guangzhou, People's Republic of China
[b] Department of Internal Medicine, Division of Medical Informatics, University of Kansas Medical Center, Kansas City, 66160, USA
[c] Big Data Decision Institute, Jinan University, Guangzhou, People's Republic of China
[d] School of Pharmacy, University of Kansas, Lawrence, USA
[e] Geriatric Research Education & Clinical Care, Tennessee Valley Healthcare System, Veteran's Health Administration, Nashville, USA
[f] Department of Biomedical Informatics, Department of Medicine, Division of General Internal Medicine, & Department of Biostatistics, Vanderbilt University, Nashville, USA
[g] School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, USA
[h] Departent of Information Systems and Business Analytics, Hofstra University, Hempstead, USA

ARTICLE INFO

Article history:
Received 3 December 2016
Received in revised form 29 January 2017
Accepted 31 January 2017

Keywords:
Drug-drug interaction
Adverse drug reaction
Causality
Association rule

ABSTRACT

Objective: Drug-drug interaction (DDI) is of serious concern, causing over 30% of all adverse drug reactions and resulting in significant morbidity and mortality. Early discovery of adverse DDI is critical to prevent patient harm. Spontaneous reporting systems have been a major resource for drug safety surveillance that routinely collects adverse event reports from patients and healthcare professionals. In this study, we present a novel approach to discover DDIs from the Food and Drug Administration's adverse event reporting system.

Methods: Data-driven discovery of DDI is an extremely challenging task because higher-order associations require analysis of all combinations of drugs and adverse events and accurate estimate of the relationships between drug combinations and adverse event require cause-and-effect inference. To efficiently identify causal relationships, we introduce the causal concept into association rule mining by developing a method called Causal Association Rule Discovery (CARD). The properties of V-structures in Bayesian Networks are utilized in the search for causal associations. To demonstrate feasibility, CARD is compared to the traditional association rule mining (AR) method in DDI identification.

Results: Based on physician evaluation of 100 randomly selected higher-order associations generated by CARD and AR, CARD is demonstrated to be more accurate in identifying known drug interactions compared to AR, 20% vs. 10% respectively. Moreover, CARD yielded a lower number of drug combinations that are unknown to interact, i.e., 50% for CARD and 79% for AR.

Conclusion: Evaluation analysis demonstrated that CARD is more likely to identify true causal drug variables and associations to adverse event.

1. Introduction

Each time a person uses a prescription medication, there is a potential for an adverse drug reaction (ADR). That potential increases with an increase in the number of concurrent medications being used. Between 2009–2012, about 47% of the United States population reported they had used at least one prescription medication in the past 30 days and almost 11% reported using at least five prescription medications in the same period. This equated to approximately $270 billion spent on prescription medications in 2013 [1]. ADRs have been attributed to cause over 770,000 injuries and 100,000 deaths each year in US [2], resulting in an annual cost of more than $136 billion [3]. Studies have estimated that drug-drug interactions (DDIs) may account for up to 30% of all unexpected adverse drug reactions (ADRs) [4].

Adverse DDIs can be preventable if discovered early. Unfortunately, it is extremely difficult to study DDIs before market

* Corresponding authors.
E-mail addresses: cairuichu@gmail.com (R. Cai), meiliu@kumc.edu (M. Liu).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.artmed.2017.01.004
0933-3657/

approval. During premarketing surveillance, new drugs can only be tested for interactions with existing drugs using *in vivo* and *in vitro* methods [5]. However, drugs can interact in many different ways [6], it is infeasible to examine every possible type of interaction for all drug combinations [7]. Additionally, many DDIs require certain amount of exposure to manifest and rare DDIs may take several exposures to occur [8]. Therefore, postmarketing surveillance becomes necessary for the early detection of unexpected adverse DDIs in the general population.

To facilitate postmarketing drug safety surveillance, the United States Food and Drug Administration (FDA) established the FDA Adverse Event Reporting System (FAERS) to collect ADR reports from healthcare professionals, patients, and pharmaceutical manufacturers [9]. Similarly at international level, the World Health Organization (WHO) is maintaining a large database of ADR reports, VigiBase [10]. These data provide a significant opportunity to study ADRs computationally. Numerous signal detection algorithms were designed for identifying relationships between drugs and adverse events (AEs) in spontaneous reports. Despite the number of existing algorithms, all have drawbacks that limit their effectiveness, such as noisy results with diminishing accuracy and low robustness due to low signal-to-noise ratio, high dimensionality of the data, and limited sample sizes. The algorithms are largely based on the statistical disproportionality theory such as relative reporting ratio (RR) to quantify the degree of unexpectedness of a relationship [9,11,12]. Both the FDA and WHO utilize adjusted versions of RR for flagging potential ADRs [13,14]. These algorithms primarily focus on binary relationships consisting of one drug and one AE, e.g., *cerivastatin* → *muscle injury* [15–18].

To find higher-order relationships raised by DDIs, e.g., *aspirin* + *warfarin* → *bleeding*, methods have been developed and evaluated on subsets of drugs and specific AEs in the spontaneous reporting systems [19–21]. For instance, van Puijenbroek et al. used logistic regression analysis to examine the influence of combined use of *itraconazole* and oral contraceptives on delayed withdrawal bleeding [19] and diuretics and non-steroidal anti-inflammatory drugs on symptoms indicating decreased efficacy of diuretics [21]. Thakrar et al. [20] investigated two statistical models in detecting four known DDIs and Tatonetti et al. [22,23] built profiles for eight clinically important AEs based on side effects of drugs known to produce them and predicted potential interactions by searching for drug pairs that match the profiles. Furthermore, Harpaz et al. [24,25] employed association rule mining and bi-clustering algorithms to infer associations between drug combinations and adverse event.

Among the existing pharmacovigilance studies, associations are the most studied relationships. However, an association does not necessarily imply causality [26]. Causal relationships do not only indicate two variables are related, but also how they are related and interacted. One step further, the causal mechanism ensures changes in causal variable directly caused changes in the effect variable. For example, we do not only want to know a particular drug is associated with renal failure, but also we want to know definitively whether the association is due to an adverse reaction or a disease. Without knowing the true relationship, plain associations can lead to false conclusions, e.g., the drug causes renal failure.

Causality is at the center stage of biomedical research and work on how to identify it has primarily taken a pragmatic approach with randomized controlled trials (RCTs) being treated as the gold standard for causal inference; however RCT methods have many well-known limitations [27–29]. Experiments are also frequently infeasible due to legal, ethical and practical constraints – no one would randomly assign individuals to smoke to assess its health risks. Instead, large-scale observational datasets on a population can be an indispensable resource for causal inference. Despite significant advances in causal discovery theory that have enabled the computational modeling and learning of causal structures from data [30–33], application of causal discovery algorithms is seriously hindered by its high computational cost [34] and its strict causal mechanism assumption [31,32]. In addressing the challenge, researchers have proposed constraint-based causal discovery methods, instead of searching for a complete Bayesian network they aim to learn local casual structures such as the Markov blanket [35–37], V-structures [38], causal cut [39]. As association rule mining (AR) has been demonstrated to be versatile in exploring relationships in large datasets, researchers have attempted to take advantage of AR for causal discovery by manipulating observational data as in retrospective cohort studies [40,41].

In this study, we aim to investigate whether it is feasible to augment AR to discover causal rules for DDI identification. Our study differs from prior related studies in the following aspects: (1) we propose a Causal Association Rule Discovery (CARD) method by exploiting the properties of V-structures in CBN to guide the causal association rule search; (2) to our knowledge, this is the first CBN-based causal discovery framework for identifying DDIs and their causal relationships to adverse events.

Table 1
Serious outcome code of FAERS reports utilized in this study.

Outcome Code	Description
DE	Death
LT	Life-threatening
HO	Hospitalization – initial or prolonged
DS	Disability
CA	Congenital Anomaly
RI	Required intervention to prevent permanent impairment/damage

2. Materials and methods

2.1. Preparation of datasets

We collected a large sample of spontaneous reports published between the years of 2004–2012 from FAERS that are categorized as having a "serious" patient outcome. In other words, only reports with outcome codes listed in Table 1 are included in the study. Additionally, we restricted our mining process to reports with mentions of at least two drugs; focusing our study on detecting adverse effects corresponding to drug combinations. Furthermore, we excluded duplicated reports and limited the analysis to drugs and AEs that occurred in at least five "serious" reports. The overall process is depicted in Fig. 1 and following sections describe each key step in detail.

2.1.1. Entity standardization

Drug names in FAERS are entered as free-text in a variety of forms. For example, the antidiabetic drug *metformin* can be entered as *Fortamet*, *Diaben/metformin*, *Diabex metformin hydrochloride*, and etc. The terms can also contain dose or route information, e.g., "*500 mg metformin*", and active ingredient information, e.g., "*Competact (pioglitaxone/metformin hydrochloride)*". To codify the medications, we used MedEx [42] to extract drug names out of the free-text terms and mapped each to a generic drug concept in RxNorm. When textual terms contain multiple active ingredients, the mapping was done separately for individual ingredient. Drug names that could not be mapped were still included in the analysis in their original form. Drug names can also be entered as a specific drug or a drug class, for example, *aspirin* vs. *analgesics*. Although *aspirin* falls under the *analgesic* class, we did not attempt to make the translation, so they are included as separate entities. Since adverse outcome reported in FAERS are already coded using the MedDRA terminology [43] (a terminology developed for

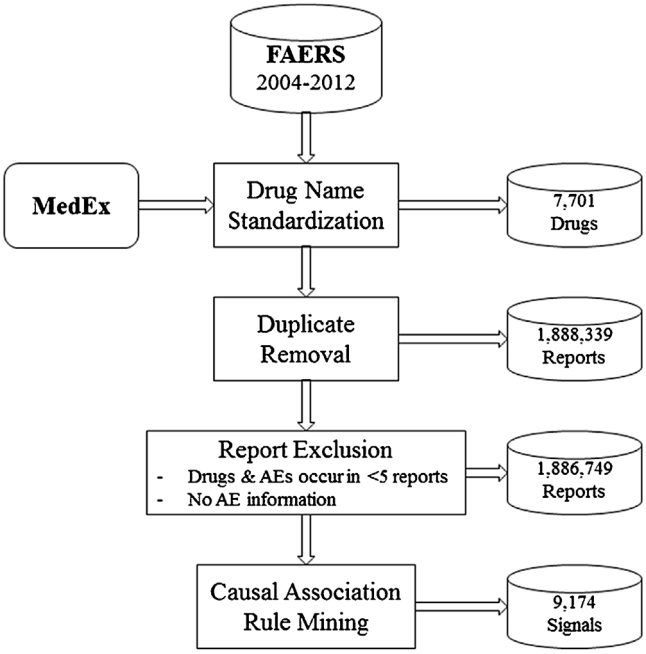

Fig. 1. Overview of the mining process.

adverse drug event applications), we treated each coded AE as a unique entity.

2.1.2. Duplicate report removal

Studies [44,45] have suggested that FAERS may contain 5%–20% duplicated reports, introducing sample bias that may result in spurious signals. FDA utilizes a proprietary algorithm to remove duplicates for analysis. As an alternative, other studies have instead searched for reports with at least eight drugs or AEs and determined reports as duplicate if they match on all the reported drugs, AEs, and patient demographic information [24,25]. Setting the threshold at eight is to minimize the probability of a match by chance. Our study adopted this approach in removing duplicated reports.

2.1.3. Dataset statistics

Statistics on FAERS reports published between 2004 and 2012 are shown in Table 2. The full set contains 4,509,229 reports. After excluding reports with non-serious outcome and duplicates, only 41.88% of the original set remained. Additionally, we removed reports that have no AE information and those only contain drugs and AEs appearing in less than 5 reports, reducing the sample size to 1,886,749 reports; covering 7701 unique generic drug names (reduced from 351,980 unique textual entries) and 11,569 MedDRA coded AEs.

Table 2
FAERS data statistics by year.

Year	Reports	Reports with serious outcome	Reports after duplicate removal
2004	272,295	141,989	135,463
2005	325,674	164,509	157,439
2006	323,791	168,401	160,793
2007	378,176	169,171	157,759
2008	441,009	198,114	180,244
2009	491,305	234,268	209,854
2010	673,170	285,939	253,287
2011	782,795	357,499	313,684
2012	821,014	352,888	319,816
Total	4,509,229	2,072,778	1,888,339

2.2. Association rule mining (AR)

Association rule mining is a well-established data mining method for discovering relationships in data and its algorithmic variations have been developed for ADR detection [24,40,46]. An association rule is an implication expression of the form $A \rightarrow B$, where A and B are two event sets that do not share any common events. In the case of ADR detection, A denotes a set of drugs and B denotes a set of AEs, e.g., $A = (cerivastatin) \rightarrow B = (muscle\ injury)$. An event set can contain one or more items, multi-item associations. For example, $A = (aspirin + warfarin) \rightarrow B = (Bleeding)$ indicates taking *aspirin* and *warfarin* together is associated with the bleeding.

To assess the strength of an association rule, the best-known measures are support and confidence. Support of an association rule $S(A \rightarrow B)$ is the proportion of records in which A and B both appear. Confidence of an association rule $C(A \rightarrow B)$ is the probability $P(B|A)$ of finding the consequent B given the antecedent A. Therefore, support measures the unexpectedness of the rule and confidence measures the reliability of the rule. These two measures allow the screening of interesting rules from a set of all possible rules. An interesting association rule is required to satisfy user-specified minimum thresholds on support and confidence at the same time. The AR method described in Harpaz et al. [24] is implemented in this study as the baseline algorithm for comparison and the relative reporting ratio (RR) is used accordingly as the proxy for measuring rule strength rather than the traditional confidence.

2.3. Causal association rule discovery (CARD)

As associations may not indicate causal relationships, our study aims to detect true causal relationships between drug combinations (drug-drug interactions) and adverse events by proposing a new method called causal association rule discovery (CARD). Given a set of spontaneous reports in the FAERS database, $\boldsymbol{S} = [s_1, s_2, \cdots, s_m]$, let $\boldsymbol{D} = [d_1, d_2, \cdots, d_n]$ be the binary indicators of the drugs taken, where d_i represents whether a patient has taken the i^{th} drug or not. Let y be a binary indicator of the occurrence of an AE and $s_j = \left[d_1^j, d_2^j, \cdots, d_n^j, y^j\right]$ represents an ADR report j. Let I_i denote the event that i^{th} drug is taken ($d_i = 1$) and A denote the event that an AE has occurred ($y = 1$), then an AE triggered by a drug-drug interaction (DDI) can be formulated as an association rule, $I_{i1}I_{i2}\cdots I_{ik} \rightarrow A$. As the study focuses on DDIs, only rules with $k \geq 2$ are considered. To formalize the concept of causal association rules, we derive important properties from Causal Bayesian Network (CBN) and develop an efficient search method.

In the simplest DDI case, a CBN of three variables $\{d_1, d_2, y\}$ with $\{d_1, d_2\}$ being the possible causes and y being the effect has four primitive local structures as shown in Fig. 2. Fig. 2a–c are independent-equivalent BNs because they entail the same set of conditional independence relationships. In other words, variable d_1 is conditionally independent of variable d_2 given variable y (i.e. $d_1 \perp d_2|y$). However, Fig. 2d implies a different assertion that d_1 is conditionally dependent of d_2 given y, referred to as the V-Structure in BN, which is well known and studied in inductive causation methods [26].

In contrast to other CBN local structures, V-structure is more robust and discriminating in causality identification problems because it is not statistically equivalent to any other structures involving the same variables. It is the only local structure that can be used to confirm the direction of causal relationships. If a true V-structure forms as in Fig. 2d, the two drugs involved must be interacting to cause AE. From the statistical aspect, the following statistical independence relation holds for all the V-structures, (1) $d_1 \perp d_2|\boldsymbol{D}'$ holds and (2) $d_1 \perp d_2|\boldsymbol{D}' \cup \{\boldsymbol{y}\}$ *does not* hold, where

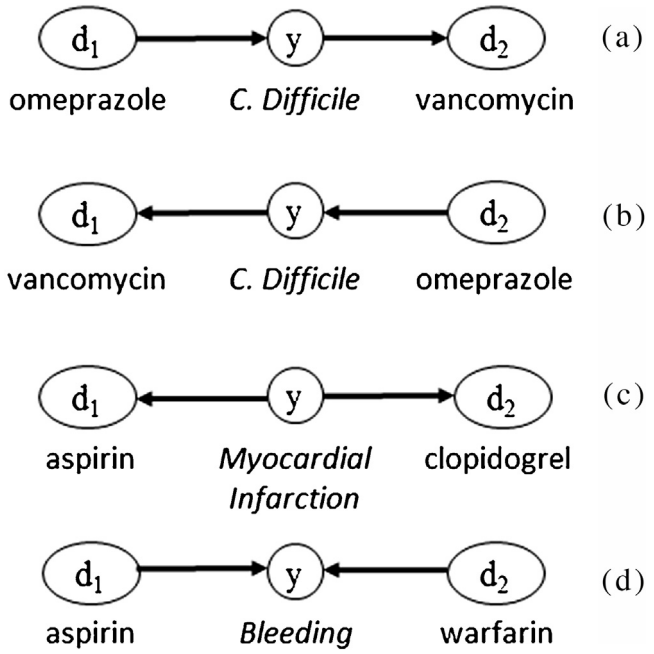

Fig. 2. Basic BN structures: (a–c) conditional independence equivalent; (d) V-structure.

$\boldsymbol{D}' \subset \boldsymbol{D} - \{d_{i1}, d_{i2}\}$. Based on the above essential properties of V-structure, we derived the following causal association interesting measure (CAIM) for $I_{i1}I_{i2} \rightarrow A$.

$$\text{CAIM}(I_{i1}I_{i2} \rightarrow A) = \text{N}(d_{i1}, y) + \text{N}(d_{i2}, y) - \text{N}(d_{i1}, d_{i2}) + \text{N}(d_{i1}, d_{i2}|y) \quad (1)$$

Here, $\text{N}(d_{i1}, y)$ denotes the normalized mutual information [47]. A high CAIM score indicates that both d_1, d_2 are highly associated with y, i.e. the role of $\text{N}(d_{i1}, y) + \text{N}(d_{i2}, y)$, and previously non-associated d_1, d_2 became highly associated given y, i.e. the role of $-\text{N}(d_{i1}, d_{i2}) + \text{N}(d_{i1}, d_{i2}|y)$.

For the rule $\boldsymbol{I} \rightarrow A$ with more than three antecedents, any sub-rules containing two antecedents must also form a V-structure with the adverse event, $d_{i1} \rightarrow y \leftarrow d_{i2}$. Because the interestingness of rule $\boldsymbol{I} \rightarrow A$ is dependent on the weakness of its sub-rules, the generalized interestingness measure is defined as follows.

$$\text{CAIM}(\boldsymbol{I} \rightarrow A) = \min_{\{I_{i1}, I_{i2}\} \subset \boldsymbol{I}} \text{CAIM}(I_{i1}I_{i2} \rightarrow A) \quad (2)$$

To efficiently search through the large space of all possible rules and estimate interestingness of the rules, we derived a pruning strategy based on properties of V-structures and an incremental updating strategy for CAIM. Algorithm details are available in Supplementary Appendix.

3. Results

3.1. Performance evaluation

Following common practice in pharmacovigilance studies in evaluating clinical validity of DDIs [15,16,24], we randomly selected 100 rules identified by CARD and presented them to a pharmacist for manual review. The pharmacist used both Micromedex and Epocrates [48,49] as the clinical reference and settled any disagreements by checking with UpToDate [50]. The pharmacist and a physician also helped us in characterizing the identified rules with a taxonomy developed based on observations. For algorithm performance comparison, we implemented the AR method as presented in Harpaz et al. [24] that was shown to be effective in identifying higher-order associations, and performed the same analysis on its 100 randomly selected findings.

3.2. Higher-order drug-event relationship identification

For the baseline AR algorithm, we aligned this study with the study in [24] by setting AR thresholds at minimum support of 100 and RR of 2. The minimum support and RR thresholds were observed to be a balancing point between the number of rules generated and variation in content (e.g., drugs and AEs in the rules). Low thresholds will result in large set of rules with many false positives. High thresholds, on the other hand, will result in less variation in content. Using the proposed thresholds, AR in total produced 424 higher-order association rules with combinations containing at least 2 drugs and among them 57 contained 3 or more drugs. Analogously, CARD also used minimum support of 100 and the CAIM score threshold was set at $\geq$0.04 to produce similar number of rules (i.e., 457) and among which 10 contained 3 or more drugs (e.g., an identified combination included doxorubicin hydrochloride, prednisone, and rituximab).

From the identified drug-event relationships, we randomly selected 100 for validation and developed a taxonomy from the analysis. The taxonomy characterizing drugs and associations along with the proportions of each category in the 100 randomly sampled findings is presented in Table 3. It contains observed proportions for both AR and CARD. Table 4 provides the taxonomy that characterizes the rules along with its proportion in each category and representative examples.

3.2.1. Known drug interactions

Among the 100 randomly sampled findings, CARD identified more drug combinations known to interact compared to AR, 20% vs. 10%. Further analysis suggested that 6 of the 20 CARD identified known interacting drug combinations cause the indicated AE (i.e., true positive DDI $\rightarrow$ AE associations). For example, CARD identified (*amiodarone* + *warfarin*) increases the international normalized ratio (INR) which is supported by evidence. This interaction is highly clinically significant and the risk of interaction often outweighs the benefit. As another example, (*nitroglycerin* + *rosiglitazone*) $\rightarrow$ *myocardial infarction* was identified by CARD and literature evidence also indicates that the combination should be avoided as it may increase the risk of myocardial ischemia.

The remaining 14 of the 20 interacting combinations identified by CARD are not currently known to cause the indicated AE as an interaction effect but one or both drugs in the combination can still relate to the AE as known side-effects (i.e., Table 4 – taxonomy 1a). For example, the combination (*prednisone* + *rosiglitazone*) was identified to be associated with cardiac failure congestive. Existing evidence indicates that the combination have antagonistic effects where efficacy of the hypoglycemic agent, *rosiglitazone*, is decreased because *prednisone* can cause hyperglycemia. Thus cardiac failure congestive is not known to be a direct effect of the interaction; however it is a cardiovascular side-effect of prednisone due to long-term fluid retention and other direct vascular effects.

Another category of association for known interacting combinations is illustrated by the taxonomy 1b where the identified AEs were unknown. CARD did not retrieve any of these pairs, but AR found 3. For example, prior clinical knowledge suggests that using *cilostazol* together with analgesics like *aspirin* may increase risk of bleeding, but not dyspnea as shown in Table 4.

3.2.2. Unknown drug interactions

CARD yielded a lower number of drug combinations unknown to interact, i.e., 50% for CARD vs. 79% for AR. Among the 50 false

Table 3
Taxonomy of drugs and associations in higher-order association rules.

		AR	CARD
Drugs			
1	Drug combinations known to interact	10%	20%
2	Drug combinations known to be given together or treat the same disease	4%	13%
3	Drug combinations that seem to be due to other confounding issues	7%	17%
4	Drug combinations that are unknown to interact	79%	50%
Associations			
a	One/more drugs in antecedent can cause the adverse event in consequent	61%	77%
b	No drug in antecedent can cause the adverse event in consequent	39%	23%

Table 4
Taxonomy of higher-order association rules.

Taxonomy	AR	CARD	Examples from CARD Findings
1 – a	7%	20%	(amiodarone + warfarin) → INR increased
2 – a	1%	10%	(alendronate + esomeprazole) → femur fracture
3 – a	5%	10%	(simvastatin + ramipril) → rhabdomyolysis
4 – a	48%	37%	(celecoxib + metformin) → myocardial infarction
1 – b	3%	0%	(analgesics + cilostazol) → dyspnea (example from AR)
2 – b	3%	3%	(ranitidine + alendronate) → fall
3 – b	2%	7%	(aspirin + rosiglitazone) → coronary artery disease
4 – b	31%	13%	(acetaminophen + isotretinoin) → intestinal hemorrhage

positives by CARD, 37 (74%) are related to the identified AE through one or both drugs (Table 4 – taxonomy 4-a). For instance, no evidence supports interaction between *celecoxib* and *metformin*, but the identified AE myocardial infarction is related to *celecoxib*. In contrast, 48 out of 79 (61%) of the AR false positives are related to its identified AE. There are also unknown drug combinations with no connection to its identified AE. For example, there is no known interaction evidence between *acetaminophen* and *isotretinoin* and no information on either of the drugs causing intestinal hemorrhage.

3.2.3. Overlapping findings

There were only 4 drug combinations identified in common between CARD (Support = 100, CAIM > 0) and AR (Support = 100, RR ≥ 2), which is an unexpected but interesting finding. The four common drug combinations identified are (Table S5): (*docetaxel*, *carboplatin*), (*pemetrexed*, *dexamethasone*), (*alendronate*, *docetaxel*), and (*dexamethasone*, *potassium*). However, none of the pairs is indicated as interacting by Epocrates. Moreover, only 1 rule in common was found: (*pemetrexed*, *dexamethasone*) → *pneumonia*, but neither drug is known to be associated with pneumonia. The small number of overlapping is because AR and CARD have different preferences of the rules and different mining thresholds. In detail, CARD employs the CAIM to select the rule with threshold CAIM > 0, while AR uses the RR to select the rule with RR ≥ 2.

3.2.4. Confounded drug pairs

As shown in Tables 3 and 4, 30% of CARD identified drug combinations are due to confounding factors such as concomitant (frequently co-administered medications) or indication (symptom of underlying disease, not effect of treatment) biases. As an illustration of taxonomy 2-a, *alendronate* was found to interact with *esomeprazole* to cause femur fracture. *Alendronate* is a bisphosphonate drug used for osteoporosis and other bone diseases. *Esomeprazole* is a proton pump inhibitor that reduces stomach acid secretion, used to treat dyspepsia, peptic ulcer disease, and gastroesophageal reflux disease. Some common adverse effects associated with *alendronate* include acid regurgitation and dyspepsia, which can be treated with *esomeprazole*, but the combination is not known to interact. This implies that our method may be finding BN structures as depicted in Fig. 2a or b rather than d, but conditional independence between the two medications can be wrongly estimated with missing *y* node (indications of *alendronate*). Atypical femur fracture, identified as an AE for (*alendronate* + *esomeprazole*), is among the list of serious adverse reactions for *alendronate* and fracture is also one serious reaction of *esomeprazole*. Similarly for taxonomy 3-a, *ramipril* and *simvastatin* also seem to be a combination found due to confounding variables and not known to interact. The identified AE, rhabdomyolysis is simply a serious side-effect of *simvastatin*.

Furthermore for taxonomy 2-b, *ranitidine* is used to treat acid regurgitation and dyspepsia associated with *alendronate*, but not known to interact. In addition, neither drug is related to fall as indicated by CARD. Finally for taxonomy 3-b, *aspirin* and *rosiglitazone* also seems to be found due to confounding issues; *aspirin* should be a drug for treating coronary artery disease, not causing.

4. Discussion

Accurate identification of causal relationships between drug combinations and adverse events require not only analysis of all drug-event combinations but also cause-and-effect estimation of the relationships. Through the adoption of the essential properties of CBN V-structures in association rule mining, our proposed CARD method can efficiently identify causal associations between drug combinations and adverse events. A comparison to the traditional association rule mining method, CARD is demonstrated to identify more known drug interactions and yield a lower number of unknown drug combinations. Nonetheless, our study has limitations:

(1) CARD is not designed to identify causal relationships that cannot be represented with V-structures. Theoretically in an ideal world, DDIs should form V-structures with its adverse effect. Thus, CARD is suitable for DDI discovery. In future work, we will explore other CBN structures and bootstrapping to identify persistently supported rules to increase confidence.
(2) Those identified rules unsupported by prior clinical knowledge could be spurious or warrant further pharmacologic analysis regarding mechanism of action. From an algorithm perspective, one not only needs to ensure precision on finding known associations is sufficiently high but also the method is able to discover new associations with high confidence. A potential approach to

investigate the unknown associations in the future is to conduct retrospective studies using patient medical records.

(3) The overlapping finding between CARD and AR is low. One explanation is that the two algorithms theoretically favor different rule sets. AR prefers highly relevant rules measured by confidence and interestingness measures based on co-occurrence of item-sets and label. CARD conversely prefers item-sets with larger difference but higher predictability when combined, using mutual information as the penalty for discarding false rules generated from random combination of frequent item-sets. Another possible explanation for the little overlap may be the super large search space. For instance, our final dataset contained 7701 drugs and only considering combination of 2 drugs would yield 29,648,850 pairs; hence CARD and AR may be identifying combinations from different parts of the actual DDI space. This interesting observation made us more aware of the data granularity problem. Although free text drug names were codified to generic names with a NLP tool, it is not perfect; many written forms of drugs are not recognized and must be used in analysis as a unique concept. Furthermore, AE granularity is a challenge as well (11,569 MedDRA coded AEs). Our future work needs to explore analyses with meta-categorization of drugs and AEs by leveraging their hierarchical and ontological structures.

(4) Many of the CARD identified drug combinations are still due to confounding factors such as concomitant usage or indication biases. By design, CARD may identify more interrelated patterns than existing methods but can also inadvertently find more undesired confounded patterns. Since confounding variables correlate with both the dependent and independent variables, incorrect estimate of the relationship between variables may occur when confounding factors are unaccounted for. However, limited by the nature of the spontaneous reporting data, this study only considered drug events; missing essential phenotypic information such as indications and comorbidities. As our future work, a promising approach is to explore larger set of clinical variables from patient medical records standardized through our participation in PCORnet [51] and FDA's Sentinel initiative [52].

Competing financial interests

The authors declare no competing financial interests.

Author contributions

R.C., M.L. and Y.H. conceived the overall design, development, and evaluation of this study. M.L. and Y.H. prepared the dataset and designed the experiment for this study. R.C. implemented the method and conducted the experiments. M.E.M. and B.L.M. conducted the clinical evaluation of the results. H.X. contributed in the data preparation process by running MedEx for drug name mapping. L.D. and L.R.W. contributed in the writing of this manuscript. All authors reviewed the manuscript critically for scientific content, and all authors gave final approval of the manuscript for publication.

Acknowledgements

This work was partially funded through a number of research grants. R.C. was supported by the Natural Science Foundation of China (61472089, 61572143, U1501254), the Natural Science Foundation of Guangdong Province, China (2014A030306004, 2014A030308008), Science and Technology Planning Project of Guangdong (2013B051000076, 2015B010108006, 2015B010131015), Guangdong High-level personnel of special support program (2015TQ01X140) and Pearl River S&T Nova Program of Guangzhou (201610010101). Y.H. was supported by the National Science Foundation of China (71271061, 70801020), Science and Technology Planning Project of Guangdong Province, China (2010B010600034, 2012B091100192). M.E.M was supported by Veterans Administration HSR&D Career Development AwardCDA-08-020 and Veterans Administration HSR&D Investigator Initiated ResearchIIR-11-292. H.X. was supported by NIGMS1R01-GM103859, NLM1R01LM011563, and CPRITR1307. L.R.W was supported in part by the National Institutes of Health grant the Heartland Institute for Clinical and Translational Research (UL1TR000001).

Appendix A.

Association rule mining is a well-researched method for discovering interesting relationships between variables in large databases. The relations discovered are typically in the form of rules $A \rightarrow B$, where A and B in this case denote a set of drugs and a set of adverse events, respectively (e.g., *(aspirin, warfarin)* $\rightarrow$ *bleeding*). A major limitation of the traditional association rule mining method is that the strength of rules is measured based on correlations, which does not imply causality. To address the limitation in this study, we introduce the concept of causality into association rule mining by proposing a Causal Bayesian network based method called Causal Association Rule Discovery (CARD). The CARD algorithm is described in detail as follows.

CARD – Causal association rule discovery

Problem definition

Let $\boldsymbol{D} = [d_1, d_2, \cdots, d_n]$ denote the indicators of the drugs taken, where d_i is a binary indicator of the action whether a patient has taken the i^{th} drug or not. Let y be a binary indicator of the occurrence of an adverse drug reaction (ADR) and $s_j = \left[d_1^j, d_2^j, \cdots, d_n^j, y^j\right]$ represents an ADR report j. Let I_i denote the event that i^{th} drug is taken ($d_i = 1$) and A denote the event that an ADR has occurred ($y = 1$), then an adverse reaction triggered by drug-drug interaction (DDIs) can be formulated as an association rule, $I_{i1} I_{i2} \cdots I_{ik} \rightarrow A$. As DDI involves more than one drug, we only focus on the rules with $k \geq 2$. Given a set of ADR reports, $\boldsymbol{S} = [s_1, s_2, \cdots, s_m]$, we aim to discover drug combinations that cause adverse reactions through causal association rule discovery.

In the following sections, we first formalize the concept of causal association rules, then derive important properties of casual association rules, and finally propose an efficient causal association rule discovery method.

Concept of causal association

Causal faithfulness condition is a commonly used assumption in causal discovery problems. According to the causal faithfulness condition, there exists a causal Bayesian network N faithful to the joint probability distribution P defined on $[d_1, d_2, \cdots, d_n, y]$. For example, Fig. 1 illustrates a causal Bayesian network with the drug and reaction variables as nodes. In the Bayesian network, each directed edge indicates the direct causal influence between the parent node and child node. Thus, d_2, d_3 in Fig. 3 denote the direct causes of ADR y.

Based on the direct causes, we can define the causal association rule as in Definition 1. The causal association rules in the form of $I_2 I_3 \rightarrow A$ are the focus of this study, where the corresponding variables I_2 and I_3, d_2 and d_3 are all the direct causes of the ADR event A (i.e., y = 1).

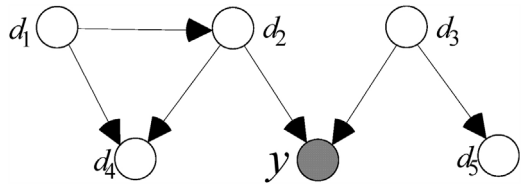

Fig. 3. Example causal Bayesian network (d = drugs, y = reaction).

Definition 1. Causal Association Rule: an association rule, $I_{i1}I_{i2}\cdots I_{ik}\rightarrow A$, defined on $\boldsymbol{D}\cup y$, is a causal association rule if $d_{i1}d_{i2}\cdots d_{ik}$ are the direct causes of y.

Properties of causal association

Although a lot of work has been conducted on the concept of causality, such as Dr. Judea Pearl's inductive causality method [26], Dr. Peter Spirtes' causal Bayesian network based methods [53], additive noise model [54], and a hybrid approach [39], causal discovery on high dimensional and sparse adverse drug-drug interaction data is still an open problem. Instead of reconstructing a causal Bayesian network, we propose to use the properties of causal Bayesian network to derive important properties of causal association rules, and apply such properties to guild the search of causal association rules.

In the simplest case, a causal Bayesian network with three variables $\{d_1, d_2, y\}$ can have four primitive structures as shown in Fig. 2. The structures in Fig. 2a–c are independent-equivalent because they entail the same set of conditional independence relationships. In other words, variable d_1 is conditionally independent of variable d_2 given variable y (i.e. $d_1 \perp d_2|y$). However, Fig. 2d forms a V-structure and implies a different assertion that there exists a variable set $\boldsymbol{D}\prime \subset \boldsymbol{D} - \{d_{i1}, d_{i2}\}$ and variable d_1 is conditionally independent of variable d_2 given $\boldsymbol{D}'$ (i.e. $d_1 \perp d_2|\boldsymbol{D}'$ holds), but dependent given $\boldsymbol{D}\prime \cup \{y\}$ (i.e. $d_1 \perp d_2|\boldsymbol{D}\prime \cup \{\boldsymbol{y}\}$ *does not* hold).

As defined by Pearl [26],"two causal models are equivalent if and only if their direct acyclic graphs have the same links and the same set of uncoupled head-to-head nodes", in which the uncoupled head-to-head nodes indicate a V-structure. This suggests that causal associations can be inferred from V-structures.

Approach to causal association rule discovery

Accordingly, we propose a causality interesting measure for the rule $I_{i1}I_{i2}\cdots I_{ik}\rightarrow A$ based on the properties of V-structure. Consider the simplest case, $I_{i1}I_{i2}\rightarrow A$, if both I_{i1} and I_{i2} are causes of A, then the corresponding three variables d_{i1}, d_{i2} and y of I_{i1},I_{i2} and A forms a V-structure, $d_{i1}\rightarrow y \longleftarrow d_{i2}$. Thus, the following conditions must hold:

1. There does not exist $\boldsymbol{D}'\subset \boldsymbol{D} - \{d_{i1}, d_{i2}\}$ satisfying $d_{i1}\perp y|\boldsymbol{D}'$ or $d_{i2}\perp y|\boldsymbol{D}'$
2. There exists $\boldsymbol{D}'\subset \boldsymbol{D} - \{d_{i1}, d_{i2}\}$ satisfying $d_{i1}\perp d_{i2}|\boldsymbol{D}'$ and $d_{i1}\perp d_{i2}|\{\boldsymbol{D}', A\}$

where $\boldsymbol{D}'$ can be an empty set.

To measure the strength of association and independence relation, we use the normalized mutual information [47]. The above properties can be transferred to the following three heuristic rules of the causal association interestingness measure:

- d_{i1} and d_{i2} should be highly associated with A, thus the rule with higher $\mathrm{N}(d_{i1}, y)+\mathrm{N}(d_{i2}, y)$ is preferred. Here $\mathrm{N}(d_{i1}, y)$ is the normalized mutual information;
- d_{i1} and d_{i2} should be independent of each other given some variable set $\boldsymbol{D}'$, thus the rule with lower $\mathrm{N}(d_{i1}, d_{i1})$ is preferred;
- d_{i1} and d_{i2} should be dependent of each other given some variable set $\{\boldsymbol{D}\prime, y\}$, thus the rule with higher $\mathrm{N}(d_{i1}, d_{i1}|y)$ is preferred

Combining the above three heuristic rules, we obtain the following causal association interestingness measure (CAIM):

$$\mathrm{CAIM}(I_{i1}I_{i2}\rightarrow A)=\mathrm{N}(d_{i1}, y)+\mathrm{N}(d_{i2}, y)-\mathrm{N}(d_{i1}, d_{i2})+\mathrm{N}(d_{i1}, d_{i2}|y) \quad (1)$$

For the rule $\boldsymbol{I}\rightarrow A$ with more than three antecedents, any sub-rules containing two antecedents must also form a V-structure with the adverse drug reaction, $d_{i1}\rightarrow y \longleftarrow d_{i2}$. Because the interestingness of rule $\boldsymbol{I}\rightarrow A$ is dependent on the weakness of its sub-rules, the generalized interestingness measure is defined as follows.

$$\mathrm{CAIM}(\boldsymbol{I}\rightarrow A)=\min_{\{I_{i1}, I_{i2}\}\subset \boldsymbol{I}}\mathrm{CAIM}(I_{i1}I_{i2}\rightarrow A) \quad (2)$$

The interestingness defined in Formula (2) has a good monotonic property with the increasing number of antecedents (Theorem 1), which can be used in the mining procedure to accelerate interesting rule mining process. The proof of the theorem is straightforward and skipped here.

Theorem 1. *Monotonic property: given any two association rules $\boldsymbol{I}_1\rightarrow A$ and $\boldsymbol{I}_2\rightarrow A$, if $\boldsymbol{I}_1\subset \boldsymbol{I}_2$, then holds.*

Based on the definition of $\mathrm{CAIM}(\boldsymbol{I}\rightarrow A)$ given in Formula (2), we can derive the following incremental updating strategy for CAIM. The incremental updating strategy provides an efficient way to estimate the interestingness when combined with incremental association rule mining approach.

Theorem 2. *Incremental updating strategy for CAIM: given two association rules $\boldsymbol{I}\rightarrow A$ and an item $I\notin \boldsymbol{I}$, .*

In addition to the interestingness measure, we can also derive the following pruning strategy of rules based on the properties of V-structures.

Theorem 3. *Pruning strategy: given a causal association rule $\boldsymbol{I}\rightarrow A$, there do not exist two variables I_{i1}, $I_{i2}\in \boldsymbol{I}$ satisfying $d_{i1}\perp d_{i2}|y$, where d_{i1}, d_{i2} and y are the corresponding variables of I_{i1},I_{i2} and A respectively.*

Proof. Because $\boldsymbol{I}\rightarrow A$ is a causal association rule and I_{i1}, $I_{i2}\in \boldsymbol{I}$, corresponding variables on d_{i1}, d_{i2} and y form a V-structure $d_{i1}\rightarrow y \longleftarrow d_{i2}$. Based on the properties of V-structure, we know that $d_{i1}\perp d_{i2}|y$ does not hold.

Based on the CAIM, the proposed CARD algorithm is as follows.

Input: adverse drug reaction sample set $\boldsymbol{S}$, minimal support threshold t, minimal CAIM threshold c;

Output: association rule set $\boldsymbol{R}$;

Initial $\boldsymbol{R}$ as an empty set;

For each item I with $\sup(I \rightarrow A) > t$

  Set $\boldsymbol{R}' = \boldsymbol{R} \cup \{I \rightarrow A\}$;

  For $i = 2$ to the max length of the rules in $\boldsymbol{R}$

    For each rule $\boldsymbol{I} \rightarrow A \in \boldsymbol{R}$ with i items

      Generate a new rule r as $\boldsymbol{I} \cup \{I\} \rightarrow A$;

      If $\sup(r) > t$

        Set CAIM(r) based on Theorem 2;

        If CAIM(r) $\geq c$ and r passed the pruning in Theorem 3

          $\boldsymbol{R}' = \boldsymbol{R} \cup \{r\}$;

        Endif

    EndFor

  EndFor

EndFor

Set $\boldsymbol{R} = \boldsymbol{R}'$;

EndFor

The algorithm shown above is adopted from the incremental rule mining algorithm proposed in Ref. [56]. It takes three input variables including the ADR sample set S and two user specified parameters, minimal support threshold t and minimal CAIM threshold c, and returns a set of potential causal association rules by conducting a mixture of breadth-first and depth-first search in the association rule lattice space. More specifically, the items are sequentially added to the lattice maintained as rule set $\boldsymbol{R}$ in a breadth-first manner. To process item I, existing lattice is partitioned and processed in the depth manner, i.e. the item is added to the rules with 1 item, then the rules with 2 items, and so on and so forth. For each newly generated rule, the support and CAIM are calculated, only the ones that passed the support threshold, CAIM threshold, and the pruning process can be added to the rule set $\boldsymbol{R}$. More details about the incremental mining method can be found in Ref. [47].

References

[1] Health. United States 2014: with Special Feature on Adults Aged 55–64. Hyattsville, MD: National Center for Health Statistics; 2015.
[2] Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279:1200–5.
[3] Johnson JA, Bootman JL. Drug-related morbidity and mortality. A cost-of-illness model. Arch Intern Med 1995;155:1949–56.
[4] Pirmohamed M. Drug interactions of clinical importance. Chapman & Hall; 1998.
[5] Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. Nephron Clin Pract 2012;120:c179–84.
[6] Duke JD, Li X, Dexter P. Adherence to drug-drug interaction alerts in high-risk patients: a trial of context-enhanced alerting. J Am Med Inform Assoc 2013;20:494–8.
[7] Banda JM, Callahan A, Winnenburg R, Strasberg HR, CamiBen A, Reis Y, et al. Feasibility of prioritizing drug–drug-event associations found in electronic health records. Drug Saf 2016;39(1):45–57.
[8] Goldman JL, Sullins A, Sandritter S, Leeder JS, Lowry J. Pediatric pharmacovigilance enhancing adverse drug reaction reporting in a tertiary care children's hospital. Ther Innov Regul Sci 2013;47:566–71.
[9] Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf 2009;18:427–36.
[10] de Abajo FJ, Montero D, Madurga M, Garcia Rodriguez LA. Acute and clinically relevant drug-induced liver injury: a population based case-control study. Br J Clin Pharmacol 2004;58:71–80.
[11] Almenoff JS, DuMouchel W, Kindman LA, Yang X, Fram D. Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting. Pharmacoepidemiol Drug Saf 2003;12:517–21.
[12] Puijenbroek E, Bate A, Leufkens HGM, Lindquist M, Orre R, Egberts ACG, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf 2002;11:3–10.
[13] Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. Drug Saf 2002;25:381–92.
[14] Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, Freitas RM, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 1998;54:315–21.
[15] Hauben M, Madigan D, Gerrits CM, Walsh L, Van Puijenbroek EP. The role of data mining in pharmacovigilance. Expert Opin Drug Saf 2005;4:929–48.
[16] Almenoff J, Tonning JM, Gould AL, Szarfman A, Hauben M, Ouellet-Hellstrom R, et al. Perspectives on the use of data mining in pharmaco-vigilance. Drug Saf 2005;28:981–1007.
[17] Almenoff JS, Pattishall EN, Gibbs TG, DuMouchel W, Evans S, Yuen N. Novel statistical tools for monitoring the safety of marketed drugs. Clin Pharmacol Ther 2007;82:157–66.
[18] Ahmed I, Thiessard F, Miremont-Salame G, Begaud B, Tubert-Bitter P. Pharmacovigilance data mining with methods based on false discovery rates: a comparative simulation study. Clin Pharmacol Ther 2010;88:492–8.
[19] Van Puijenbroek EP, Egberts AC, Meyboom RH, Leufkens HG. Signalling possible drug-drug interactions in a spontaneous reporting system: delay of withdrawal bleeding during concomitant use of oral contraceptives and itraconazole. Br J Clin Pharmacol 1999;47:689–93.
[20] Thakrar BT, Grundschober SB, Doessegger L. Detecting signals of drug-drug interactions in a spontaneous reports database. Br J Clin Pharmacol 2007;64:489–95.
[21] van Puijenbroek EP, Egberts AC, Heerdink ER, Leufkens HG. Detecting drug-drug interactions using a database for spontaneous adverse drug reactions: an example with diuretics and non-steroidal anti-inflammatory drugs. Eur J Clin Pharmacol 2000;56:733–8.

[22] Tatonetti N, Denny J, Murphy S, Fernald G, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther 2011;90:133–42.
[23] Tatonetti NP, Fernald GH, Altman RB. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. J Am Med Inform Assoc 2012;19:79–85.
[24] Harpaz R, Chase HS, Friedman C. Mining multi-item drug adverse effect associations in spontaneous reporting systems. BMC Bioinformatics 2010;11(Suppl. 9):S7.
[25] Harpaz R, Perez H, Chase HS, Rabadan R, Hripcsak G, Friedman C. Biclustering of adverse drug events in the FDA's spontaneous reporting system. Clin Pharmacol Ther 2011;89:243–50.
[26] Pearl J. Causality: models, reasoning and inference. first edition Cambridge Univ Press; 2000.
[27] Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–12.
[28] Rothman KJ. Causes. Am J Epidemiol 1976;141:90–5, discussion 89 (1995).
[29] Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol 2010;39:89–94.
[30] Heckerman D, Geiger D, Chickering DM. Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn 1995;20:197–243.
[31] Hoyer PO, Janzing D, Mooij J, Peters M, Scholkopf B. Nonlinear causal discovery with additive noise models. Neural Information Processing Systems (NIPS); 2008. p. 689–96.
[32] Zhang K, Hyvärinen A. On the identifiability of the post-nonlinear causal model. UAI 2009;64:7–655.
[33] Pearl J. From Bayesian network to causal networks. Bayesian Netw Probab Reason 1994:1–31.
[34] Chickering DM, Heckerman D, Meek C. Large-sample learning of Bayesian networks is NP-hard. J Mach Learn Res 2004;5:1287–330.
[35] Cooper GF. A simple contraint-based algorithm for efficiently mining observational databases for causal relationships. Data Min Knowl Discov 1997;1:203–24.
[36] Pellet JP. Using Markov blankets for causal structure learning. J Mach Learn Res 2008;9:1295–342.
[37] Aliferis CF, Statnikov A, Tsannardinos I, Mani S, Koutsoukos XD. Local causal and Markov blanket induction for causal discovery and feature selection for classification Part I: algorithms and empirical evaluation. J Mach Learn Res 2010;11:171–234.
[38] Cai R, Zhang Z, Hao Z. Causal gene identification using combinatorial V-structure search. Neural Netw 2013;43:63–71.
[39] Cai R, Zhang Z, Hao Z. SADA: a general framework to support robust causation discovery. ICML 2013;20:8–216.
[40] Ji Y, Ying H, Dews P, Mansour A, Tran J, Miller RE, et al. A potential causal association mining algorithm for screening adverse drug reactions in postmarketing surveillance. IEEE Trans Inf Technol Biomed 2011;15:428–37.
[41] J. Li, et al., in IEEE 13th International Conference on Data Mining Workshops (ICDMW) 114–23 (Dallas, TX, 2013).
[42] Xu H, Stenner SP, Doan S, Johnson KB, Waitman LR, Denny JC. MedEx: a medication information extraction system for clinical narratives. J Am Med Inform Assoc 2010;17:19–24.
[43] Welcome to MedDRA and the MSSO, <http://www.meddramsso.com/MSSOWeb/index.htm>.
[44] Noren GN, Orre R, Bate A, Edwards IR. Duplicate detection in adverse drug reaction surveillance. Data Min Knowl Discov 2007;14:305–28, http://dx.doi.org/10.1007/s10618-006-0052-8.
[45] Reich L. 'Extreme duplication' in the USFDA adverse events reporting system database. Drug Saf 2007;30:551–4.
[46] Jin H, Chen J, He H, Williams GJ, Kelman C, et al. Mining unexpected temporal associations: applications in detecting adverse drug reactions. IEEE Trans Inf Technol Biomed 2008;48:8–500.
[47] Vinh NX, Epps J, Bailey J. Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. J Mach Learn Res 2010;11:2837–54.
[48] KDIGO. KDIGO clinical practice guideline for acute kidney injury. Kidney Int 2012;(Suppl. 2):1–141.
[49] Epocrates, <http://www.epocrates.com>.
[50] UpToDate, <http://www.uptodate.com>.
[51] Hou SH, Bushinsky DA, Wish JB, Cohen JJ, Harrington JT. Hospital-acquired renal insufficiency: a prospective study. Am J Med 1983;74:243–8.
[52] McCullough PA, Adam A, Becker CR, Davidson C, Lameire N, Stacul F. Risk prediction of contrast-induced nephropathy. Am J Cardiol 2006;98:27K–36K, http://dx.doi.org/10.1016/j.amjcard.2006.01.022.
[53] Spirtes P, Glymour C, Scheines R. Cusation, prediction, and search. second edition; 2001.
[54] Hoyer PO, Janzing D, Mooij J, Peters M, Scholkopf B. Nonlinear causal discovery with additive noise models. Vancouver, Canada: Neural Information Processing Systems (NIPS); 2008. p. 689–696.
[56] Cai R, Tung AKH, Zhang ZJ, Hao ZF. What is unequal among the equals? Ranking equivalent rules from gene expression data. IEEE Trans Knowl Data Eng 2011;23(11):1735–47.