Mining directional drug interaction effects on myopathy using the FAERS database

Danai Chasioti[1,2]*, Xiaohui Yao[3]*, Pengyue Zhang[4], Samuel Lerner[5], Sara K. Quinney[6], Xia Ning[4], Lang Li[4], Li Shen[3]**

Abstract—**Mining high-order drug-drug interaction (DDI) induced adverse drug effects from electronic health record (EHR) databases is an emerging area, and very few studies have explored the relationships between high-order drug combinations. We investigate a novel pharmacovigilance problem for mining directional DDI effects on myopathy using the FDA Adverse Event Reporting System (FAERS) database. Our work provides information on the risk of myopathy associated with adding new drugs on the already prescribed medication, and visualizes the identified directional DDI patterns as user-friendly graphical representation. We utilize the Apriori algorithm to extract frequent drug combinations from the FAERS database. We use odds ratio (OR) to estimate the risk of myopathy associated with directional DDI. We create a tree-structured graph to visualize the findings for easy interpretation. Our method confirmed myopathy association with previously reported HMG-CoA reductase inhibitors like rosuvastatin, fluvastatin, simvastatin and atorvastatin. New, previously unidentified but mechanistically plausible associations with myopathy were also observed, such as the DDI between pamidronate and levofloxacin. Additional top findings are gadolinium-based imaging agents, which however are often used in myopathy diagnosis. Other DDIs with no obvious mechanism are also reported, such as that of sulfamethoxazole with trimethoprim and potassium chloride. This study shows the feasibility to estimate high-order directional DDIs in a fast and accurate manner. The results of the analysis could become a useful tool in the specialists' hands through an easy-to-understand graphic visualization.**

Index Terms—**Directional effect, high-order drug interaction, FAERS, Apriori, frequent itemsets**

I. INTRODUCTION

Drug interactions occur when the substances of multiple medicines affect one another changing the effect of one or more co-administered drugs. One emerging topic in drug safety research is the investigation of possible drug-drug interactions (DDIs) causing adverse drug events (ADEs). Acquiring knowledge about the possible adverse effects due to drug co-administration can provide useful information for clinical applications, drug development, the control of the ADE-associated medical cost, and safeguarding the patients' wellbeing. In polypharmacy situations, the information about potential DDIs becomes crucial since the harm caused by the concomitant drug administration could outweigh the benefits. Polypharmacy is quite common with 25-50% of patients aged 75 years or older being exposed to at least five drugs [1].

Research on ADE detection can be broadly classified into two categories. The first category discovers ADE patterns via analyzing health record databases [2]–[13]. For example, Harpaz et al. [4] performed a bi-clustering analysis on the data from the Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) to identify drug groups that share common ADEs. Tatonetti et al. [8] analyzed the FAERS data and identified 171 DDIs for eight ADEs. The second category predicts DDI patterns via integrating biochemical and drug target data [14], [15]. For example, Zhang et al. [14] presented an effective ensemble learning method for DDI prediction using a collection of drug data including drug substructure, target, enzyme, transporter, pathway and other relevant data. Ning et al. [15] developed a classification method to predict DDIs associated with an ADE using drug similarity measures constructed from drug properties and DDI patterns. In this work, we focus on the first category of the ADE research. In this category, most prior studies examined the effects of single drugs (e.g., [2]–[5]) or two-way DDIs (e.g., [5]–[13]).

The investigation of DDIs involving three or more drugs (called **high-order DDIs** in this work) is still an under-explored area, partially due to the computational challenge it is facing. Since studying high-order DDIs is important, especially in situations like polypharmacy, this topic has attracted a few recent attentions [5], [16], [17]. For example, Harpaz et al. [5], [16] applied an association rule mining algorithm to the FAERS data for revealing associations of multiple drugs to multiple ADEs. Similar to conventional DDI studies, they estimated the overall effect of a drug set. However, this approach does not allow the exploration of the relative risk between drug combinations. With this observation, we introduce the concept of *directional DDI effects*, defined as the ADE risk induced by adding one or more drugs into an existing drug combination. In a prior study [17], we developed an approach to identify directional effects of high-order DDIs, which was restricted to the study of adding one drug at a time

[1]Department of Radiology and Imaging Sciences, School of Medicine, Indiana University, Indianapolis, IN, USA

[2]Department of BioHealth Informatics, School of Informatics and Computing, Indiana University, Indianapolis, IN, USA

[3]Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[4]Department of Biomedical Informatics, College of Medicine, Ohio State University, Columbus, OH, USA

[5]Department of Computer Science and Engineering, Ohio State University, Columbus, OH, USA

[6]Department of Obstetrics and Gynecology, School of Medicine, Indiana University, Indianapolis, IN, USA

*Equal contribution by Danai Chasioti (dhasioti.@umail.iu.edu) and Xiaohui Yao (xiaohui.yao@pennmedicine.upenn.edu).

**Please address correspondence to: Li Shen, Ph.D., Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, 423 Guardian Ave, 404 Blockley Hall, Philadelphia, PA 19104, USA. Phone: 1-215-573-2956, E-mail: li.shen@pennmedicine.upenn.edu.

to an existing drug combination. Clearly, this method does not apply to general cases of prescribing multiple additional drugs. To bridge this gap, in this work, we aim to extend the above method to estimate the risk of an ADE associated with adding one or more drugs to an existing drug combination.

II. OBJECTIVE

Our objective is to examine the DDI induced ADE risk, and to investigate a novel pharmacovigilance topic where one or more drugs are added concurrently on top of the current medication. Given two sets of drug combinations DC_1 and DC_2 (superset of DC_1), we introduce a new concept of the directional DDI effect from DC_1 to DC_2, which is defined as the altered ADE risk associated with the change from taking DC_1 to taking DC_2. Although DC_2 can contain any number of drugs, in this work, we restrict DC_2 to contain up to three drugs while DC_1 up to two drugs. Thus, we are able to observe the directional effect caused to a patient's medication, by the addition of not only one but up to three drugs at a time. To measure the ADE risk, we use the odds ratio (OR). As seen in previous studies [17], [18], a restriction of this analysis is the low frequency drug combinations. Including such drug combinations in the analysis would result in misleading OR estimations. To avoid such inaccuracies, we employ a data mining technique to exclude the infrequent drug combinations from our analysis and only focus on the analysis of frequent drug combinations. We calculate the OR of the ADE for adding one, two, or three drugs in the patients' medication and visualize the results as an intuitive tree.

To demonstrate the proposed strategy, we analyze data from the publicly available Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) [19]. The widely used FAERS, is a valuable source of information about ADEs. Specifically, we use FAERS to examine the risk of DDI induced myopathy.

Myopathy is unexpected muscle toxicity with primary symptoms that of muscle weakness, pain, and even tissue break down. The prevalence of drug-induced myopathies, one of the most common causes of muscle diseases, has been estimated to be more than 2,000 per 100,000 individuals in the Western Hemisphere [20]. Myopathy has been mainly associated with statin medication, and is one of the major reasons of statin-medication discontinuation [21]–[24]. Based on the Side Effect Resource (SIDER) database [25], among the 99 registered myopathy-related drugs, 8% are statins. In addition to statins, glucocorticoid-induced myopathy due to decreased protein synthesis and increased rate of protein catabolism, has also been observed [26]. We study the proposed novel pharmacovigilance problem, using myopathy as a testbed.

III. MATERIALS AND METHODS

A. Data preparation

Our analysis is based on the publicly available FAERS data collected between Q1 2004 and Q3 2012. Before its structure update in 2013, FAERS was based on case reports, containing one or more reports for each individual, submitted at different time points. In our analysis, we only retained the most recent

report for each individual. The processed FAERS data contain 4,077,447 ADE reports, and 1,763 FDA approved drugs. After processing the original FAERS dataset, we created a dataset where each record is a report containing the administered drugs and the ADEs. To avoid the confusion between a cause and a bystander, in our analysis, only the primary and secondary suspects were included, and drugs labeled as "concomitant" or "interacting" were excluded. A record containing myopathy as an ADE was labeled as *myopathy*; and otherwise as *non-myopathy*. Out of 4,077,447 events, there are 136,860 myopathy events. In our data, the number of drugs in a record ranges from 1 to 103. We only extracted the frequent drug combinations, based on which, we calculated directional DDI effects. Fig. 1 shows the flowchart of our analysis.
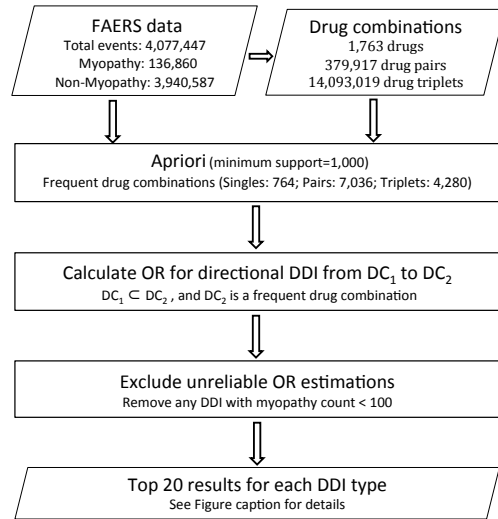

Fig. 1. Flowchart of the analysis process: Odds Ratios (ORs) are estimated for each of the following directional drug interaction types: (1) baseline to 1-drug, (2) baseline to 2-drugs, (3) baseline to 3-drugs, (4) 1-drug to 2-drugs, (5) 1-drug to 3-drugs, and (6) 2-drugs to 3-drugs.

B. Directional drug interaction effects

We define the directional DDI effect as the ADE risk change introduced by adding new drugs to an existing drug combination. We use the odds ratio (OR) as a measure of the ADE risk (see Table I). Here, individuals with myopathy are considered as cases whereas, individuals without myopathy are considered as controls. The OR is calculated as the ratio of odds of the exposed individuals with myopathy over the odds of the exposed individuals without myopathy. Exposed individuals are defined as those who take a specific drug combination $D_1, \ldots, D_i, D_{i+1}, \ldots, D_N$; while unexposed individuals are those taking $D_1, \ldots, D_i$ but not taking at least one of the $D_{i+1}, \ldots, D_N$. For example, the odds of taking ABC over the odds of taking C gives the directional effect of adding AB to C. To calculate the OR of myopathy for individuals taking a drug combination ABC versus those taking only C, we set the exposed group to be those who take any drug combination containing ABC and the unexposed group to be those who take any drug combination containing C but not ABC together (i.e., C, CA, or CB).

Exposure to drug Combination	Myopathy	Non-Myopathy	Total
$D_1, \ldots, D_i, D_{i+1}, \ldots, D_N$	a	b	a+b
$D_1, \ldots, D_i$ without taking at least one of the $D_{i+1}, \ldots, D_N$	c	d	c+d
Total	a+c	b+d	a+b+c+d

C. Extraction of frequent itemsets

The existing low-frequency drug combinations would result in misleading OR results. To exclude these infrequent combinations, we applied the Apriori algorithm [27]. Apriori is a data mining technique used for identifying frequent itemsets and association rules in transactional data sets. Using a bottom-up approach, the algorithm starts by identifying all frequent one-element itemsets. All the frequent itemsets are extended one item at a time to form new candidate itemsets, given that the subsets of a frequent itemset must also be frequent. The occurrence of each candidate itemset is counted using the database to identify frequent ones. This procedure continues until no new extension can be found.

In the processed FAERS data set, a record can be considered as a transaction and a drug combination as an itemset. In Fig. 2, a simple flowchart of the algorithm is presented. For each of the drug combinations (itemsets), the number of its occurrences (support) is calculated. Given a user-specified *minimum support* (MinSup), a drug combination is frequent, if it appears at least as many times as the MinSup value. Apriori was applied using the *arules* package in R. In this work, Apriori identified 764 frequent 1-drug combinations, 7,036 frequent 2-drug combinations and 4,280 frequent 3-drug combinations.
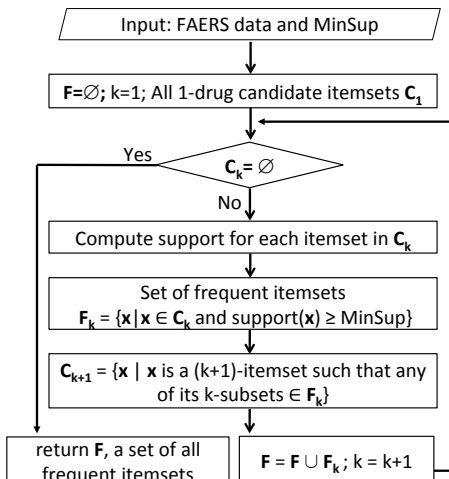

Fig. 2. Flowchart of the main steps of the Apriori algorithm [27].

D. Identifying drug interactions

For each frequent drug combination DC_2, containing up to three drugs, and each of its subset DC_1 (i.e., $DC_1 \subset DC_2$), we calculate the OR of the myopathy risk associated with the directional DDI from DC_1 to DC_2. For convenience, we call DC_1 as *baseline* if DC_1 is empty (i.e., none of the DC_2 drugs are in DC_1). Specifically, the OR is calculated based on the formula $OR = ad/bc$ (see Table I) for the following directional effects: 1) baseline to 1-drug, 2) baseline to 2-drugs, 3) baseline to 3-drugs, 4) 1-drug to 2-drugs, 5) 1-drug to 3-drugs, and 6) 2-drugs to 3-drugs. Here, a represents the count of frequent myopathy itemsets containing all the drugs in DC_2, b is the count of frequent non-myopathy itemsets containing all the drugs in DC_2, c represents the count of frequent myopathy itemsets containing all the drugs in DC_1, and d is the count of frequent non-myopathy itemsets containing all the drugs in DC_1.

Because the myopathy events in our data constitute only 3.5% of all the ADEs, there are drug combination instances where the number of non-myopathy events far exceeds the number of the myopathy events. A big disproportionality between the case and control number could be another source of unreliable OR results. To address this issue, we set an additional constraint regarding the minimum acceptable number of cases (exposed and unexposed myopathy events). Under these two constraints, we estimate all possible directional effects related to the subsets of each frequent drug set containing up to 3 drugs. For example, for a three-drug set ABC, we calculate the OR for each of the following directional effects:

1) Baseline $\rightarrow ABC$
2) $A \rightarrow ABC$
3) $B \rightarrow ABC$
4) $C \rightarrow ABC$
5) $AB \rightarrow ABC$
6) $AC \rightarrow ABC$
7) $BC \rightarrow ABC$

The results are corrected for multiple comparisons (Bonferroni correction) and ordered by their ORs.

E. Visualization of directional drug interactions

Let $\mathbf{S}$ be a frequent drug set and $\mathbf{C} = \{\mathbf{x} \mid \mathbf{x} \subseteq \mathbf{S}\}$ be the set of all of its subsets. Given $\mathbf{S}$, we can generate a tree to organize and visualize all possible directional DDIs between relevant sets in $\mathbf{C}$. Specifically, each node represents a drug combination. An edge from the node DC_1 to the node $DC_2 \setminus DC_1$ indicates the directional DDI of $DC_1 \rightarrow DC_2$, where $DC_1 \subset DC_2 \subseteq \mathbf{S}$. D3: Data-Driven Documents (http://d3js.org/) is used to plot such a tree. Examples of tree visualization are shown in Figs. 3-6. The value shown on each node indicates the OR of the myopathy risk associated with taking the drug(s) on the current node in addition to the drug(s) on the previous node. The node size is proportional to the OR, while the node color indicates the significance of the OR, after Bonferroni correction (white for non-significant and blue for significant). The text color indicates the effect type (green for OR<1 and red for OR>1).
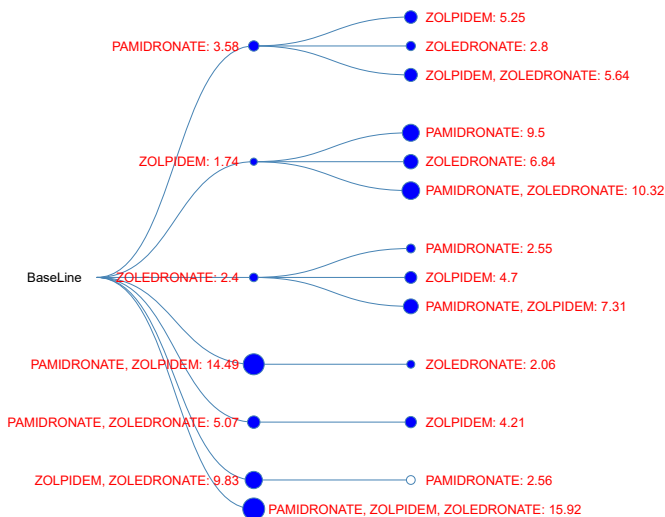
Fig. 3. Tree visualization for all possible directional DDI effects related to the subsets of {pamidronate, zolipidem, zoledronate}. The value shown on each node indicates the OR of the altered myopathy risk associated with taking the drug(s) on the current node in addition to the drug(s) on the previous node. A larger node size corresponds to larger OR. The node color indicates the significance of the OR, after Bonferroni correction (white for non-significant and blue for significant). The text color indicates protective or risk effect (green for OR<1 and red for OR>1).

IV. RESULTS

A. Data summary

The processed FAERS dataset used, contains 4,077,447 records, among which 136,860 are myopathy events, and 3,940,587 are non-myopathy ones. From the 1,763 different drugs in the dataset, 75 have myopathy-related effects. We applied Apriori [27] to our data by setting MinSup=1,000, and we identified 764 frequent single drugs, 7,527 frequent drug pairs, and 4,280 frequent drug triplets.

B. Drug Interactions

The top 100 results for single drug versus baseline, are shown in Table S1, with OR values ranging from 5.57 to 2.18. The top drug is gadoteridol with OR=5.57, meaning that the odds of myopathy in patients taking gadoteridol is 5.57 times higher than in patients not taking it. Nicorandil, ranked ninth in the list, is a vasodilator used for treating angina pectoris. In 2014, specific batches of this drug were withdrawn from the market after FDAs decision. Although, myopathy is not one of its known side effects, cases of weakness have been reported.

For two-drugs versus the baseline, the top 100 results are shown in Table S2. The OR values vary from 21.56 to 7.63. The top result here is the combination of simvastatin and clarithromycin with OR=21.56. Second and third in the list are the simvastatin-cyclospirine and the pamidronate-levofloxacin combinations, with OR values 17.69 and 16.36 respectively. Also, the combination of azithromycin and zole-dronate (ranked 4th), shows an OR of 15.86. All these results are statistically significant, after correcting for multiple comparisons.

Drug	OR	95%CI	p-value
Oxycodone, Zoledronate, Levofloxacin	19.51	(17.16, 22.16)	4.89E-315
Pamidronate, Acetaminophen, Gabapentin	18.32	(16.14, 20.78)	0.00E-00
Acetaminophen, Zoledronate, Levofloxacin	17.01	(15.09, 19.16)	0.00E-00
Oxycodone, Zoledronate, Gabapentin	16.84	(15.07, 18.8)	0.00E-00
Zoledronate, Dexamethasone, Gabapentin	16.75	(14.77, 18.97)	9.31E-293
Pamidronate, Zoledronate, Zolipidem	15.92	(14.05, 18.02)	4.63E-282
Pamidronate, Zoledronate, Gabapentin	15.70	(14.01, 17.57)	0.00E-00
Acetaminophen, Zoledronate, Gabapentin	15.53	(13.93, 17.28)	0.00E-00
Pamidronate, Acetaminophen, Furosemide	14.98	(13.11, 17.10)	3.41E-237
Pamidronate, Acetaminophen, Warfarin	14.85	(13.03, 16.88)	3.54E-247

Table II presents the top 10 findings of comparing drug triplets with the baseline. The OR estimates range from 19.51 to 14.85. Most of the results here are similar to those of Table S2 with the addition of pain reliever (oxycodone or acetaminophen) or gabapentin (which may also be used to treat neuropathic pain). For example, the top result is the combination of oxycodone, zoledronate and levofloxacin (OR=19.51). A similar combination also appears in the comparison of drug pairs to the baseline. Specifically, it was estimated that individuals taking both zoledronate and levofloxacin have an OR=12.61, which is the 11th highest OR (Table S2). We also observe a statistically significant risk of myopathy when zoledronate (belonging to the bisphosphonates drug class and used to treat bone diseases) is administered with pamidronate and zolpidem (Fig. 3). Based on the FAERS data, the risk of myopathy when co-administering these drugs, is 15.92 (p-value<1.17e-05; right hand side of inequality is the Bonferroni p-value threshold. See Table II). To the best of our knowledge, there is neither any previous study nor any report indicating an interaction between these drugs (mild myalgia is an ADE of zoledronate). Also, in our results, dexamethasone is frequently implicated in the risk of myopathy by interacting with drugs like gabapentin-zoledronate (Fig. 4), levofloxacin-zoledronate, and lorazepam-zoledronate. Again, to our knowledge, there are not known interactions between these individual drugs and dexamethasone. The ORs of myopathy when comparing each of these 3-drug combinations to the risk of only taking dexamethasone are 16.76, 13.34 and 12.12 respectively.

For two-drug versus one-drug combinations (Table S4), some results confirm the well-studied association between the CYP3A and statins. Specifically, we observe a statistically significant OR of 20.21 when adding simvastatin on top of cyclosporine (ranked 4th with p-value < the Bonferroni threshold of 3.59e-06) and when adding simvastatin on top of clarithromycin (ranked 9th with OR=13.50 and p-value < the Bonferroni threshold of 3.59e-06). The OR values of this

3-Drug Combination	1-Drug	OR	95%CI	p-value
Gadobenate Dimeglumine, Gadodiamide, Gadoversetamide	Gadobenate Dimeglumine	20.31	(12.98, 33.40)	4.81E-84
Gadobenate Dimeglumine, Gadodiamide, Gadopentetate Dimeglumine	Gadobenate Dimeglumine	19.63	(12.54, 32.27)	7.43E-82
Gadobenate Dimeglumine, Gadopentetate Dimeglumine, Gadoteridol	Gadobenate Dimeglumine	18.71	(12.27, 29.79)	8.67E-85
Gadobenate Dimeglumine, Gadodiamide, Gadoteridol	Gadobenate Dimeglumine	17.91	(11.83, 28.22)	1.48E-83
Zoledronate, Dexamethason, Gabapentin	Dexamethason	16.76	(14.67, 19.12)	7.09E-280
Gadobenate Dimeglumine, Gadoteridol, Gadoversetamide	Gadobenate Dimeglumine	16.18	(10.90, 24.88)	1.11E-81
Gadobenate Dimeglumine, Gadopentetate Dimeglumine, Gadoversetamide	Gadobenate Dimeglumine	16.16	(10.75, 25.25)	2.61E-78
Zoledronate, Dexamethasone, Doxorubicin	Doxorubicin	16.05	(13.54, 18.98)	3.53E-169
Zoledronate, Dexamethasone, Levofloxacin	Dexamethasone	13.34	(11.59, 15.34)	1.08E-207
Gadobenate Dimeglumine, Gadopentetate Dimeglumine Gadoteridol	Gadoteridol	13.06	(6.72, 29.02)	1.13E-28

3-Drug Combination	2-Drug Combination	OR	95%CI	p-value
Potassium Chloride, Sulfamethoxazole, Trimethoprim	Potassium Chloride, Trimethoprim	34.43	(6.02, 1365.53)	9.22E-12
Acetaminophen, Chlorhexidine, Hydrocodone	Chlorhexidine, Hydrocodone	23.86	(6.37, 199.86)	5.34E-14
Zoledronate, Dexamethason, Furosemide	Dexamethason, Furosemide	12.73	(10.26, 15.84)	9.02E-141
Prochlorperazine, Zoledronate, Dexamethasone	Prochlorperazine, Dexamethasone	12.56	(9.86, 16.08)	2.16E-113
Omeprazole, Potassium, Clavulanate	Omeprazole, Clavulanate	12.10	(4.00, 59.79)	3.70E-10
Sulfamethoxazole, Acetaminophen, Hydrocodone	Sulfamethoxazole Hydrocodone	12.01	(3.99, 59.23)	2.36E-10
Lorazepam, Zoledronate, Dexamethasone	Lorazepam, Dexamethasone	11.53	(9.29, 14.34)	1.19E-130
Salmeterol, Valsartan, Fluticasone Propionate	Salmeterol, Valsartan	11.49	(3.08, 96.46)	4.05E-07
Zoledronate, Dexamethason, Doxorubicin	Dexamethason, Doxorubicin	11.33	(9.08, 14.15)	1.99E-107
Morphine, Zoledronate, Dexamethason	Morphine, Dexamethason	11.22	(8.86, 14.28)	3.38E-111

table range from 28.21 to 6.73, with the top result (OR=28.21) being that of gadobenate dimeglumine and gadopentetate dimeglumine versus gadobenate dimeglumine.

Tables III and IV present the top 10 results of drug triplets versus one drug and two drugs respectively. In Table III, the OR values range from 20.31 to 13.06, with the gadolinium contrast agents occupying the top positions, while in Table IV the OR values range from 34.43 to 11.22.

C. Visualization of directional drug interactions

An important aspect of this study is the extended visualization of the DDI that allows tracking down the whole DDI path of the final result. For an n-drug combination $(D_1, \ldots, D_n)$, we start from a baseline node (none of $D_1, \ldots, D_n$ are taken) and we generate all possible paths up to n-drugs along with the corresponding ORs. For example, for the zoledronate-pamidronate-zolpidem combination, a tree visualization of all

the directional DDIs is presented in Fig. 3. For patients at the baseline (none of the 3 drugs are taken), the OR of myopathy when adding zoledronate is 2.4, indicating that zoledronate is a risk factor (red color). Notice that, for the baseline individuals who are prescribed simultaneously zolpidem and zoledronate, the OR is higher (OR=9.83), as indicated by a larger node in the graph. If now, on top of zoledronate, we prescribe pamidronate, the OR becomes 2.55 (baseline $\rightarrow$ zoledronate: 2.4 $\rightarrow$ pamidronate: 2.55). A similar example is also presented in Fig. 4. Given the path baseline $\rightarrow$ dexamethasone: 1.17 $\rightarrow$ zoledronate, gabapentin: 16.76, we observe 17% increase in the risk of myopathy for patients who take dexamethason compared to those in the baseline, while the OR is 16.76 when, on top of dexamethasone, we add zoledronate and gabapentin. Similarly, we can follow any path in the graph to gradually understand the impact of each drug (or drug combination) on the myopathy risk.

V. DISCUSSION

A. Novelty and Contributions

In general, the majority of the studies published in this area, are focusing on examining interactions between two drugs.
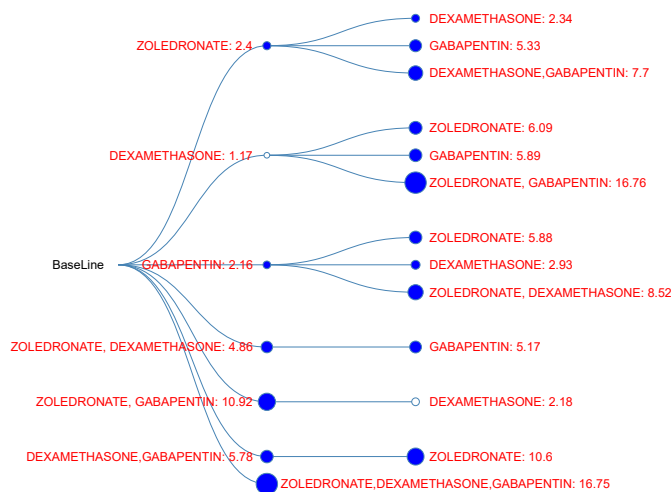
Fig. 4. Tree visualization for all possible directional DDI effects related to the subsets of {dexamethasone, gabapentin, zoledronate}. The value shown on each node indicates the OR of the altered myopathy risk associated with taking the drug(s) on the current node in addition to the drug(s) on the previous node. A larger node size corresponds to larger OR. The node color indicates the significance of the OR, after Bonferroni correction (white for non-significant and blue for significant). The text color indicates protective or risk effect (green for OR<1 and red for OR>1).

However, in cases of polypharmacy, we need to get estimations of higher order interactions. Here, we have estimated a newly proposed directional risk of myopathy as a result of high order DDI, based on the FAERS dataset.

We have extended our previous work [17], by removing the restriction imposed on the risk estimation of adding one drug at a time to an existing combination. We illustrated our methodology by estimating the OR of directional effects resulting from adding up to three drugs at a time. We also expanded the user-friendly visualization to accommodate general DDI patterns. In addition, we accounted for multiple comparison correction, which was neglected in [17]. Of note, although our prior study investigated a similar problem, the analysis was done on different data set with a smaller number of drugs.

Furthermore, we substituted an in-house R implementation of Apriori used in our previous study [17], by a more efficient Apriori implementation in R package "arules", which significantly improved the computational time. Specifically, in Du et al. [17], the experiments were performed on a Linux Xeon 64-bit dual CPU workstation, and the computation of all frequent 3-drug combinations using a dataset of 6,388,674 events and 212 drugs, took 312.9 hours for MinSup=1, 36.4 hours for MinSup=200, and 6.5 hours for MinSup=1,000. In this work, with the new implementation, our experiments were performed on a Windows 10 Enterprise 64 bit desktop with an Intel(R) Core(TM) i5-7500 CPU. We were able to obtain all frequent 3-drug combinations for a dataset of 4,077,447 events and 1,763 drugs, in 17.2 seconds for MinSup=1, 8.9 seconds for MinSup=200, and 8.4 seconds for MinSup=1,000. To obtain all frequent 2-drug combinations, it took 8.2 seconds for any MinSup ranging from 1 to 1,000.

Ibrahim et al. [28] and Li et al. [29] also analyzed FAERS data but focused on two-way drug interactions. Cai et al. [30] examined the adverse effect induced by three co-administered drugs (similar to our baseline versus 3-drug combination effect) in the FAERS data, while they did not investigate the directional DDIs. To the best of our knowledge, our work is the first study offering estimates on high-order directional DDIs using the FAERS data.

Below, we discuss our findings, grouped into several categories: (1) gadolinium-based imaging agents and drugs associated with diagnosis or treatment of myopathy, (2) known myotoxic drugs or drugs with interactions known to increase the risk of myopathy, (3) previously unidentified but mechanistically plausible interactions, and (4) previously unidentified interactions with no obvious mechanism for increasing the risk of myopathy.

B. Drugs associated with diagnosis and treatment of myopathy

The data in the FAERS reports do not provide the ability to evaluate the timing of drugs with respect to the ADE manifestation. Therefore, drugs used for the diagnosis and treatment of myopathy, not just those that could be potentially causative, may be highly associated with the ADE. The top two agents (both gadolinium-based contrast agents) associated with myopathy were gadoteridol and gadoversatimide (Table S1). Two other gadolinium compounds (gadodiamide and gadobenate dimuglumine) were also in the top 10 and top 20 myopathy-associated drugs respectivelly (Table S1). These agents have been reported to cause muscular calcification and weakness, which may present as myopathy [31]. However, MRI is often a part of the diagnostic work-up of patients presenting with myopathy. Therefore, it is likely that the high association between myopathy and gadolinium-based contrast agents is due to their use in the diagnosis of myopathy. However, this does not necessarily explain why a myopathy report would be associated with multiple gadolinium-based contrast agents nor why individuals exposed to multiple contrast agents are at increased risk of myopathy (Tables S4 and III).

This same conclusion may be made for other drugs with high OR for myopathy, such as chlorhexidine (Table S1 and S2), a topical antiseptic that may have been applied prior to muscle biopsy. We also note a number of pain relievers (e.g. oxycodone, acetaminophen, hydrocodone) among the 2-drug and 3-drug combinations associated with increased myopathy risk. While there is some evidence of myopathy with high doses of these drugs [32], the use of these agents to relieve myopathy symptoms cannot be ruled out.

C. Drugs or drug interactions previously reported to increase myopathy risk

The presence of drugs or drug interactions that have been previously reported to increase myopathy risk substantiates. Most noticeably, a number of HMG-CoA reductase inhibitors (Table S1) were observed to increase the risk of myopathy over baseline, including rosuvastatin (OR=4.59 Table S1), fluvastatin (OR=4.29, Table S1), simvastatin (OR=3.98, Table S1) and atorvastatin (OR=3.85, Table S1). We also identified increased OR for myopathy with other lipid lowering agents,

such as gemfibrozil, ezetimibe, and colesevelam. Muscle pain, weakness, and myopathy have been reported with these drugs, and they are also known to increase the risk of statin-induced myopathy. Our analyses also identified a number of DDIs known to increase the risk of myopathy through well-defined mechanisms. For instance, clarithromycin and cyclosporine are known to inhibit metabolism and transport, respectively, of many statin drugs, including simvastatin [33] (Table S2). This leads to increased plasma concentrations of the statin and higher risk for myopathy [34], [35]. Additional DDIs previously linked to myopathy, which were captured by our analysis and were highly ranked, include atorvastatin with rosuvastatin, simvastatin with gemfibrozil and rosuvastatin with simvastatin (7th, 12th and 16th highest OR respectively, Table S2).

In order to investigate how many of the known myopathy-related findings coincide with our results, we searched the OFFSIDES and TWOSIDES databases (both developed by Tatonetti et al. [7]). Our analysis captured almost 92% of the drugs in OFFSIDES, that were linked to the events of myopathy toxic and myopathy steroids (all have significant p-values). In the TWOSIDES database, no events were found under the term "myopathy". Hence, we focused on myopathy-related events. Specifically, we searched for 2-drug interactions linked to muscle weakness, rhabdomyolysis, muscle disorder, muscle paresis, muscle spasm, muscle inflammation, musculoskeletal pain, myasthenia gravis, muscle strain, and muscle rupture. TWOSIDES contains 32,304 unique 2-drug combinations related to the aforementioned events. From these drug combinations, 2,906 are among our 2-drug vs. baseline results. The ORs of myopathy in our analysis for these drug combinations range from 21.56 (i.e., for (simvastatin, clarithromycin)) to 0.08 (i.e., for (capecitabine, cisplatin)).

While the statin interactions are widely known to be implicated to myopathy events and are the top interactions identified among our results, it warrants further investigation to build a complete collection of every known DDI related to myopathy. An important future direction is to perform a comprehensive curation of relevant literature, based on which an unbiased estimate of our method's sensitivity can be obtained.

D. Previously unidentified but mechanistically plausible interactions

Although myalgia has been reported as an adverse event for the bisphosphonate pamidronate, there is no literature evidence regarding the mechanism of the effect. However, renal toxicity has been associated with pamidronate. In turn, renal failure is associated with an increased risk of myopathy. Thus, it may be hypothesized that the increased risk of myopathy with pamidronate (OR=3.59 compared to baseline; Table S1), is likely secondary to renal failure. This hypothesis is further strengthened by the increased risk of myopathy observed when pamidronate is found in combination with other nephrotoxic drugs (e.g. OR of pamidronate with levofloxacin is 16.36 over baseline; Table S2; Fig. 5).
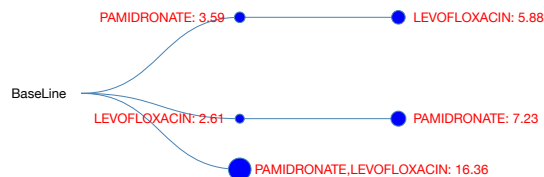

Fig. 5. Tree visualization for all possible directional DDI effects related to the subsets of {pamidronate, levofloxacin}. The value shown on each node indicates the OR of the altered myopathy risk associated with taking the drug(s) on the current node in addition to the drug(s) on the previous node. A larger node size corresponds to larger OR. The node color indicates the significance of the OR, after Bonferroni correction (white for non-significant and blue for significant). The text color indicates protective or risk effect (green for OR<1 and red for OR>1).

E. Previously unidentified interactions with no obvious mechanism for increased risk of myopathy

Our analysis identified several drugs and drug combinations with high OR for myopathy not supported by previous studies, e.g. sulfamethoxazole with trimethoprim and potassium chloride (Table IV; Fig. 6). Another one is that of azithromycin with zoledronate (Table S2). The mechanism for increased risk of myopathy for these drugs or drug pairs is unclear and requires further investigation.
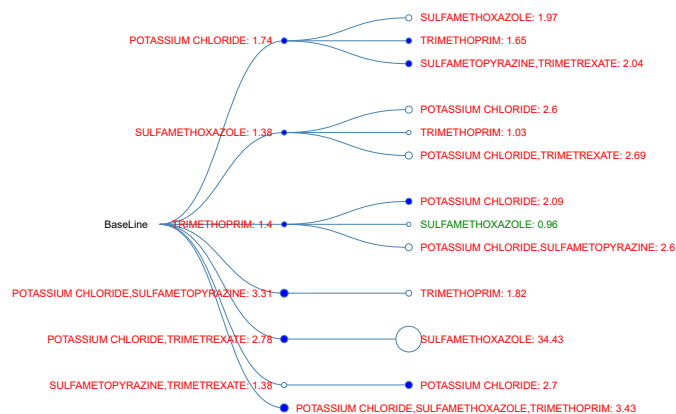

Fig. 6. Tree visualization for all possible directional DDI effects related to the subsets of {Sulfamethoxazole, Potassium Chloride, Trimethoprim}. The value shown on each node indicates the OR of the altered myopathy risk associated with taking the drug(s) on the current node in addition to the drug(s) on the previous node. A larger node size corresponds to larger OR. The node color indicates the significance of the OR, after Bonferroni correction (white for non-significant and blue for significant). The text color indicates protective or risk effect (green for OR<1 and red for OR>1).

VI. CONCLUSIONS

Through this work, we were able to get directional high-order drug interactions in a computationally efficient way, by applying the Apriori algorithm for frequent itemset identification. We calculated the OR for the directional adverse effect of myopathy for up to three drugs at a time. Our findings confirmed well-known adverse effects of specific drug combinations, revealed new but plausible associations, but also identified drug combinations with no profound mechanism but warranting further investigation. Despite the valuable knowledge hidden in the FAERS data, there is no information

about the elapsed time from the drug administration until the adverse effect manifestation. The absence of temporal relation information, between the exposure and the event, complicates the inference process, by introducing higher uncertainty. Thus, it is more difficult to decide whether an observed association could potentially be causal or due to pre-existing myopathy. Examples of this phenomenon are the findings related to gadolinium compounds, and a series of pain relievers. Despite this limitation, we were able to confirm drug interactions previously associated to myopathy and also reveal new drug combinations with a plausible underlying mechanism for causing myopathy. The current methodology can be extended to higher order drug combinations (e.g., >3 drugs), with affordable computational cost and accurate risk estimation. That, combined with the interactive visual representation could be a valuable tool in the specialists hands, not only for ensuring patients health but also for promoting drug research.

References

[1] A. Sonnichsen, U. S. Trampisch, A. Rieckert et al., "Polypharmacy in chronic diseases-reduction of inappropriate medication and adverse drug events in older populations by electronic decision support (prima-eds): study protocol for a randomized controlled trial," Trials, vol. 17, p. 57, 2016.

[2] W. DuMouchel, "Bayesian data mining in large frequency tables, with an application to the fda spontaneous reporting system," The American Statistician, vol. 53, no. 3, pp. 177–190, 1999.

[3] L. Lian Duan, M. Khoshneshin, W. N. Street, and M. Liu, "Adverse drug effect detection," IEEE J Biomed Health Inform, vol. 17, no. 2, pp. 305–11, 2013.

[4] R. Harpaz, H. Perez, H. S. Chase et al., "Biclustering of adverse drug events in the fda's spontaneous reporting system," Clin Pharmacol Ther, vol. 89, no. 2, pp. 243–50, 2011.

[5] R. Harpaz, W. DuMouchel, N. H. Shah et al., "Novel data-mining methodologies for adverse drug event discovery and analysis," Clin Pharmacol Ther, vol. 91, no. 6, pp. 1010–21, 2012.

[6] N. P. Tatonetti, P. P. Ye, R. Daneshjou, and R. B. Altman, "Data-driven prediction of drug effects and interactions," Sci Transl Med, vol. 4, no. 125, p. 125ra31, 2012.

[7] N. P. Tatonetti, J. C. Denny, S. N. Murphy et al., "Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels," Clin Pharmacol Ther, vol. 90, no. 1, pp. 133–42, 2011.

[8] N. P. Tatonetti, G. H. Fernald, and R. B. Altman, "A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports," J Am Med Inform Assoc, vol. 19, no. 1, pp. 79–85, 2012.

[9] Y. Ji, H. Ying, J. Tran et al., "A functional temporal association mining approach for screening potential drug-drug interactions from electronic patient databases," Inform Health Soc Care, vol. 41, no. 4, pp. 387–404, 2016.

[10] H. Yang and C. C. Yang, "Discovering drug-drug interactions and associated adverse drug reactions with triad prediction in heterogeneous healthcare networks," in Healthcare Informatics (ICHI), 2016 IEEE International Conference on. IEEE, Conference Proceedings, pp. 244–254.

[11] G. N. Noren, R. Sundberg, A. Bate, and I. R. Edwards, "A statistical methodology for drug-drug interaction surveillance," Stat Med, vol. 27, no. 16, pp. 3057–70, 2008.

[12] S. V. Iyer, R. Harpaz, P. LePendu et al., "Mining clinical text for signals of adverse drug-drug interactions," J Am Med Inform Assoc, vol. 21, no. 2, pp. 353–62, 2014.

[13] X. Han, S. K. Quinney, Z. Wang et al., "Identification and mechanistic investigation of drug-drug interactions associated with myopathy: A translational approach," Clin Pharmacol Ther, vol. 98, no. 3, pp. 321–7, 2015.

[14] W. Zhang, Y. Chen, F. Liu, F. Luo et al., "Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data," BMC Bioinformatics, vol. 18, no. 1, p. 18, 2017.

[15] X. Ning, L. Shen, and L. Li, "Predicting high-order directional drug-drug interaction relations," in 2017 IEEE International Conference on Healthcare Informatics (ICHI), Aug 2017, pp. 556–561.

[16] R. Harpaz, H. S. Chase, and C. Friedman, "Mining multi-item drug adverse effect associations in spontaneous reporting systems," BMC Bioinformatics, vol. 11 Suppl 9, p. S7, 2010.

[17] L. Du, A. Chakraborty, C. W. Chiang et al., "Graphic mining of high-order drug interactions and their directional effects on myopathy using electronic medical records," CPT Pharmacometrics Syst Pharmacol, vol. 4, no. 8, pp. 481–8, 2015.

[18] P. Zhang, L. Du, L. Wang et al., "A mixture dose-response model for identifying high-dimensional drug interaction effects on myopathy using electronic medical record databases," CPT Pharmacometrics Syst Pharmacol, vol. 4, no. 8, pp. 474–80, 2015.

[19] "FDA adverse event reporting system," 03/16/2017. [Online]. Available: https://open.fda.gov/data/faers/

[20] S. Peters and R. A. Kley, "Toxic and drug-induced myopathies," in Neuromuscular Imaging. Springer, 2013, pp. 335–348.

[21] M. Y. Wei, M. K. Ito, J. D. Cohen et al., "Predictors of statin adherence, switching, and discontinuation in the usage survey: understanding the use of statins in america and gaps in patient education," J Clin Lipidol, vol. 7, no. 5, pp. 472–83, 2013.

[22] D. Rosenbaum, J. Dallongeville, P. Sabouret, and E. Bruckert, "Discontinuation of statin therapy due to muscular side effects: a survey in real life," Nutr Metab Cardiovasc Dis, vol. 23, no. 9, pp. 871–5, 2013.

[23] R. Rosenson, S. Baker, T. Jacobson et al., "The national lipid association's muscle safety expert panel," An assessment by the statin muscle safety task force, pp. S58–S71, 2014.

[24] H. Zhang, J. Plutzky, S. Skentzos, F. Morrison, and others., "Discontinuation of statins in routine care settings: a cohort study," Ann Intern Med, vol. 158, no. 7, pp. 526–34, 2013.

[25] C. C. Hays, Z. Z. Zlatar, and C. E. Wierenga, "The utility of cerebral blood flow as a biomarker of preclinical alzheimer's disease," Cellular and Molecular Neurobiology, vol. 36, no. 2, pp. 167–179, 2016.

[26] R. M. Pereira and J. Freire de Carvalho, "Glucocorticoid-induced myopathy," Joint Bone Spine, vol. 78, no. 1, pp. 41–4, 2011.

[27] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," pp. 487–499, 1994.

[28] H. Ibrahim, A. Saad, A. Abdo, and A. S. Eldin, "Mining association patterns of drug-interactions using post marketing fda's spontaneous reporting data," Journal of biomedical informatics, vol. 60, pp. 294–308, 2016.

[29] Y. Li, P. Zhang, Z. Sun, and J. Hu, "Data-driven prediction of beneficial drug combinations in spontaneous reporting systems." AMIA Annual Symposium Proceedings, p. 808817, 2016.

[30] R. Cai, M. Liu, Y. HU, B. Melton et al., "Identification of adverse drug-drug interactions through causal association rule discovery from spontaneous adverse event reports," Artificial Intelligence in Medicine, vol. 76, pp. 7–15, 2017.

[31] E. Edgar, R. Woltjer, R. Whitham et al., "Nephrogenic systemic fibrosis presenting as myopathy: a case report with histopathologic correlation," Neuromuscul Disord, vol. 20, no. 6, pp. 411–3, 2010.

[32] E. B. Larbi, "Drug induced rhabdomyolysis: case report," East Afr Med J, vol. 74, no. 12, pp. 829–31, 1997.

[33] R. Amundsen, H. Christensen, B. Zabihyan, and A. Asberg, "Cyclosporine a, but not tacrolimus, shows relevant inhibition of organic anion-transporting protein 1b1-mediated transport of atorvastatin," Drug Metab Dispos, vol. 38, no. 9, pp. 1499–504, 2010.

[34] A. M. Patel, S. Shariff, D. G. Bailey et al., "Statin toxicity from macrolide antibiotic coprescription: a population-based cohort study," Ann Intern Med, vol. 158, no. 12, pp. 869–76, 2013.

[35] W. Yang, Z. Zeng, X. Ren, Y. Li et al., "Simvastatin-induced myopathy with concomitant use of cyclosporine: case report," International journal of clinical pharmacology and therapeutics, vol. 49, no. 12, pp. 772–777, 2011.