

SCIENTIFIC REPORTS



OPEN

Label Propagation Prediction of Drug-Drug Interactions Based on Clinical Side Effects

Received: 27 January 2015

Accepted: 22 June 2015

Published: 21 July 2015

Ping Zhang¹, Fei Wang², Jianying Hu¹ & Robert Sorrentino¹

Drug-drug interaction (DDI) is an important topic for public health, and thus attracts attention from both academia and industry. Here we hypothesize that clinical side effects (SEs) provide a human phenotypic profile and can be translated into the development of computational models for predicting adverse DDIs. We propose an integrative label propagation framework to predict DDIs by integrating SEs extracted from package inserts of prescription drugs, SEs extracted from FDA Adverse Event Reporting System, and chemical structures from PubChem. Experimental results based on hold-out validation demonstrated the effectiveness of the proposed algorithm. In addition, the new algorithm also ranked drug information sources based on their contributions to the prediction, thus not only confirming that SEs are important features for DDI prediction but also paving the way for building more reliable DDI prediction models by prioritizing multiple data sources. By applying the proposed algorithm to 1,626 small-molecule drugs which have one or more SE profiles, we obtained 145,068 predicted DDIs. The predicted DDIs will help clinicians to avoid hazardous drug interactions in their prescriptions and will aid pharmaceutical companies to design large-scale clinical trial by assessing potentially hazardous drug combinations. All data sets and predicted DDIs are available at <http://astro.temple.edu/~tua87106/ddi.html>.

Drug-drug interaction (DDI) may happen unexpectedly when more than one drugs are co-prescribed, causing serious side effects. DDI is a serious health and safety issue which draws great attention from both academic and industry¹. As the number of approved drugs increases, the number of potential interactions between prescribed medications rapidly rises. Moreover, elderly patients and cancer patients are typically administered numerous medications^{2,3}, exposing them to a high risk of adverse DDIs. Discovering and predicting DDIs will not only prevent life-threatening consequence in clinical practice, but also prompt safe drug co-prescriptions for better treatments.

Most DDIs are discovered by accident in the clinic or during phase IV clinical trials that take place once a drug is already on the market¹. In order to effectively detect DDIs, a number of statistical methods were developed for discovering DDIs from scientific literature^{4,5}, electronic medical records⁶, insurance claim databases⁷, and the FDA Adverse Event Reporting System^{8,9}. However, these methods still rely on the accumulation of sufficient clinical evidence in the post-marketing surveillance. As such, they are insufficient for detecting all DDIs and cannot alert the public to potentially dangerous DDIs before a drug enters the market.

In recent years, a new direction is to predict novel DDIs based on mechanistic and structural information of the drugs themselves and their interactions with proteins. For example, computational methods were developed for predicting DDIs by analyzing chemical structure similarity¹⁰, implementing the chemical-protein interactome¹¹, modelling interaction profile fingerprints¹², and exploiting pharmacoin-teraction network structure¹³. There are also some efforts on predicting DDIs by integrating multiple molecular and pharmacological data^{14,15}. The advantage of these methods lies in the fact that they rely

¹Healthcare Analytics Research, IBM T. J. Watson Research Center, Yorktown Heights, USA. ²Department of Computer Science and Engineering, University of Connecticut, Storrs, USA. Correspondence and requests for materials should be addressed to P.Z. (email: pzhang@us.ibm.com)

mainly on chemical and bioactivity data from laboratory studies rather than clinical records. As a result, they could potentially be used to predict DDIs years in advance, enabling drug safety professionals to better prioritize their limited investigative resources and take appropriate regulatory action.

Similarity-based approach is a representative strategy for DDI prediction^{10,12,15}. However, most of the existing similarity-based DDI prediction algorithms only utilize first-order similarity (i.e., use immediate similarities for prediction) and doesn't consider transitivity of similarity. To address this deficiency, we proposed a label propagation approach for predicting DDIs by considering high-order similarity. Furthermore, clinical phenotypic information has not been adequately investigated for its power in predicting DDIs. The advantages of leveraging clinical phenotypic information lies on two aspects: (1) clinical phenotypic information could serve as biomarkers for both therapeutic effects^{16,17} and toxic effect¹⁸, thus has potential to be used in DDI predictions. (2) Clinical phenotypic information is derived from direct observations from human, thus has better translational power when comparing with molecular or animal models¹⁹ in DDI predictions. In this study we also investigated the usage of clinical side effects (SEs) to predict DDIs.

To facilitate the use of different SEs and chemical structures information sources, we propose an integrative label propagation framework to predict DDIs by integrating SEs extracted from package inserts of prescription drugs, SEs extracted from FDA Adverse Event Reporting System, and chemical structures from PubChem. The proposed framework is also extensible, thus our method can incorporate additional types of drug information sources.

To summarize, our study differs from prior related studies in the following aspects: (1) we investigate the use of clinical SEs as key features to predict DDIs. To our knowledge ours is the first study to do so. While Gottlieb *et al.*¹⁵ used SEs as one of the sources to build predictive models, they did not consider any off-label SEs extracted from FDA Adverse Event Reporting System. (2) we use a label propagation approach for DDI prediction by considering high-order similarity, and propose an integrative label propagation framework by considering drug information from multiple sources for better solutions. The new method not only provides DDI prediction but also ranks and prioritizes multiple drug information sources.

Materials and Methods

Preparation of datasets. *FAERS DDI database.* FDA Adverse Event Reporting System (FAERS) is a database that contains information on adverse events submitted to FDA, which is designed to support FDA's post-marketing safety surveillance program for drugs and therapeutic biological products. Mined from FAERS, TWOSIDES²⁰ is a dataset containing only SEs caused by the combination of drugs rather than by any single drugs. In this study, we used the unsafe co-prescriptions from TWOSIDES as known set of DDIs. There are 645 drugs and 63,473 distinct pairwise DDIs in the dataset.

Side effect datasets. SIDER is a side effect database of drugs containing information on market medicines and their recorded adverse drug reactions²¹. The information is extracted from package inserts (i.e., drug labels). In this study, we downloaded the entire database from <http://sideeffects.embl.de/>. There are 996 drugs and 4,192 side effects in the dataset. We called side effects extracted from SIDER as "Label Side Effect".

SIDER is an important source of known side effects, but the knowledge is limited: (1) Clinical trials are conducted on relatively small patient populations, only common effects can be detected with sufficient confidence to be listed on a drug's package insert. (2) Effect observed during the clinical trials may be incidental and not actually caused by the drug. OFFSIDES²⁰ is a side effect dataset built by mining FAERS system while controlling confounding factors such as concomitant medications, patient demographics, and patient medical histories. There are 1,332 drugs and 10,093 side effects in the dataset. We called side effects extracted from OFFSIDES as "Off-Label Side Effect".

Merging the drugs from SIDER, OFFSIDES, and TWOSIDES, we obtained 569 drugs with label side effect, off-label side effect, and DDI information. Among 569 drugs, there are 52,416 distinct pairwise DDIs.

Chemical structure dataset. Also we extracted chemical structures of the 569 drugs from PubChem²², thus we can compare chemical structures information with side effect information.

Similarity measures. For side effect information, each drug was represented by a binary side effect profile (4,192-dimensional for label side effect and 10,093-dimensional for off-label side effect) whose elements encode for the presence or absence of each of the side effect key words by 1 or 0 respectively.

For chemical structure information, we used a chemical structure fingerprint corresponding to the 881 chemical substructures²³ defined in the PubChem. Each drug was represented by an 881-dimensional binary profile whose elements encode for the presence or absence of each PubChem substructure by 1 or 0, respectively.

We used Tanimoto coefficient (TC), also known as the Jaccard index, to compute similarities between all the fingerprints. The TC between two fingerprints A and B is defined as the ratio between the number of features in the intersection to the union of both fingerprints: $TC(A,B) = |A \cap B| / |A \cup B|$. Then we

created matrices so that the rows and columns represent drugs and each cell represents the TC between fingerprint pair of drugs. Drug information from chemical structure, prescription package inserts, and FAERS was thus transferred into chemical similarity, label side effect similarity, and off-label side effect similarity matrices.

Label propagation algorithm. Label propagation algorithms address the following problem: given an undirected weighted network with n nodes where a small portion of them are labeled (e.g., as positive), estimate the labels of the rest unlabeled nodes. In our case, we treated different drugs as nodes on the network, and computed the edge weights on the network with drug similarities evaluated using the method in last section. For each drug, we labelled all other drugs in the network as positive if they are known to have DDI with this drug, and utilized label propagation on such drug network to estimate the possibility that the unlabeled drugs will have DDI with this drug. We didn't consider negative examples, as in the DDI prediction problem negatives (i.e., the known safe co-prescriptions) are rarely available.

More concretely, we represented the input drug network using an $n \times n$ affinity matrix A , where $A_{ij} \geq 0$ is the similarity between drug i and j . For each drug, we constructed a label vector $y \in \{0,1\}^n$ over the network, where $y_i = 1$ if drug i is known to have DDI with the reference drug, and $y_i = 0$ otherwise. With those notations, label propagation assigns scores (which indicate the possibility that each drug will have DDI with the reference drug) to drug nodes by an iterative procedure which propagates evidence out from positive nodes through the edges in the network. To ensure convergence of the updates, the original affinity matrix A needs to be normalized so that the row sum is one. In this study, we used Bregmanian Bi-Stochasticity (BBS) algorithm²⁴ for such normalization and denote the normalized matrix as W . One nice thing of BBS is that the resultant normalized matrix is still symmetric.

Using W , we propagated labels from the labeled drug nodes to the unlabeled nodes. In each propagation iteration, the estimated score of each drug node "absorbs" a portion (μ) of the label information from its neighborhood, and retains a portion $(1 - \mu)$ of its initial label information. The updating rule for node i is given by $f_i^{after} \leftarrow \mu \sum_{j=1}^n W_{ij} f_j^{before} + (1 - \mu) y_i$. In this formula, $0 < \mu < 1$ is a parameter that determine the influence of a node's neighbors relative to its provided label. By concatenating the predicted scores for all drug nodes, we can obtain the matrix form of the updating rule²⁵ as $f^{after} \leftarrow \mu W f^{before} + (1 - \mu) y$. It can be shown that after t iterations, the predicted score vector f^t can be written as $f^t = (\lambda W)^{t-1} y + (1 - \lambda) \sum_{i=0}^{t-1} (\lambda W)^i y$. Since $W_{ij} \geq 0$, and $\sum_j W_{ij} = 1$, the spectral radius of W , or $\rho(W) \leq 1$. In additional, $0 < \mu < 1$, thus $\lim_{t \rightarrow \infty} (\mu W)^{t-1} = 0$ and $\lim_{t \rightarrow \infty} \sum_{i=0}^{t-1} (\mu W)^i = (I - \mu W)^{-1}$ where I is the identity matrix of order n . Therefore, f^t will eventually converge to $f = \lim_{t \rightarrow \infty} f^t = (1 - \mu)(I - \mu W)^{-1} y$.

In our scenario, there are n tasks, i.e., we want to predict the DDI profile for each drug. To achieve this, we can first concatenate the initial label vector y for each drug into an initial label matrix Y , where its (i, j) -th entry is 1 if drug i interacts with drug j , and 0 means there is no known interaction between drug i and drug j . Then we can get all the DDI predictions in one-shot as $F = (1 - \mu)(I - \mu W)^{-1} Y$ ²⁶.

Actually the converged solution for label propagation can also be obtained by minimizing the following objective:

$$J = \mu \text{tr}(F^T (I - W) F) + (1 - \mu) \|F - Y\|_F^2 \quad (1)$$

In this formula, $\text{tr}(\bullet)$ denotes the trace of a matrix, $\|\bullet\|_F$ denotes Frobenius norm of a matrix. The first term of (1) is a smoothness term, which assumes that the prediction should not vary too much on the intrinsic network. In our case, it means that the predicted DDI score for any reference drugs should change smoothly over the drug network. This coincides with our drug similarity assumption: similar drugs tend to have similar DDI effects. The second term of (1) is the fitting term, which restricts the predicted DDI scores to be close to their initial values. The trade-off between those two competing terms is captured by a nonnegative parameter μ between 0 and 1. As formula (1) is convex with respect to F , we can get its global optimum by setting the first order derivative of J with respect to F to zero. As $\partial J / \partial F = 2(\mu(I - W)F + (1 - \mu)(F - Y))$. By setting $\partial J / \partial F = 0$, we can still get $F = (1 - \mu)(I - \mu W)^{-1} Y$.

The difference between label propagation and nearest neighbor strategies. In last section, we introduced a label propagation approach for predicting DDIs. Here we compare it to the nearest neighbor strategy. The motivating hypothesis for similarity based DDI prediction is: if drug i and drug j are similar according to some criteria and drug k interacts with drug i , then drug k will be very likely to interact with drug j as well. We first constructed a drug similarity network according to various drugs' characteristics (chemical structures, label side effects, or off-label side effects), and then spread DDI labels of the drugs on such network.

To better depict this idea, we constructed a synthetic example in Fig. 1. Figure 1(a) shows a drug network structure, where any two drugs whose similarity is larger than a specific threshold are connected. We colored the nodes according to whether the drugs in the network are known to interact with the reference drug. In this case, the two green drugs are known to interact with the query drug, and the

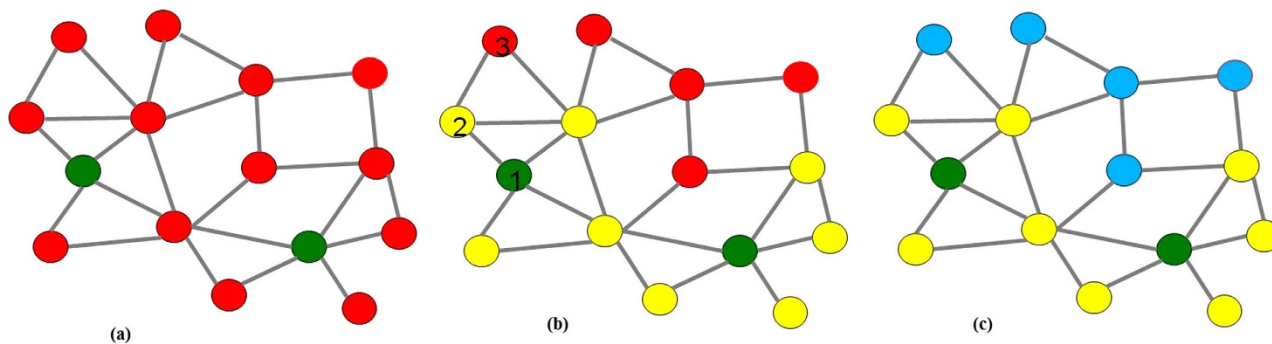


Figure 1. Comparison of nearest neighbor and label propagation strategies. (a) toy training data with two known drugs (in green) interacting with the input; (b) nearest neighbor method only searches drugs similar to the training drugs (in yellow); (c) label propagation method searches the whole drug similarity graph, i.e., search yellow drugs' nearest neighbors as well (in blue).

red drugs are the unknown ones that we want to predict. The nearest neighbor algorithms will act as in Fig. 1(b): they search the whole drug space, find the most similar drugs to the green ones (these drugs are plotted in yellow), and predict the yellow ones will interact with the input drug. For example, drug 1 is similar to drug 2, but drug 1 is not similar enough to drug 3. Therefore, only drug 2 is predicted to interact with the input drug. However, if drug 3 is very similar to drug 2 (and also another yellow drug in the figure), it may also have high likelihood of interacting with the input drug. Our proposed approach goes further, acting as Fig. 1(c): it will continue to search yellow drugs' nearest neighbors as well (in blue), and iterate this process until convergence.

Label propagation with multiple similarities. In previous sections we discussed the case when we have only one drug similarity matrix. In the real-world applications, we can have multiple drug similarities derived from different data sources (e.g., chemical, label side effect, and off-label side effect). In this case if we only use one drug similarity matrix then the resultant predictions could be biased. It may also suffer from the noise in the focused drug information source. Because of the complex mechanism of the DDIs, their predictions clearly need methods to integrate drug information from multiple sources for better solutions.

Let Y be the initial DDI matrix and suppose we have K symmetrically normalized drug similarity matrices W_1, W_2, \dots, W_K , where each of them captures the drug similarity from one specific perspective. We would like to obtain DDI predictions F by solving the following composite optimization problem:

$$\begin{aligned} \min_{F, \alpha} \mu \operatorname{tr} \left(F^T \sum_{k=1}^K \alpha_k (I - W_k) F \right) + (1 - \mu) \|F - Y\|_F^2 + \delta \|\alpha\|_2^2 \\ \text{s.t.}, \forall k, \alpha_k \geq 0, \sum_{k=1}^K \alpha_k = 1 \end{aligned} \quad (2)$$

The objective function (2) is derived from objective function (1) by linear combination of individual drug similarity matrices from different data sources. In this formula, $\operatorname{tr}(\bullet)$ denotes the trace of a matrix, $\|\bullet\|_F$ denotes Frobenius norm of a matrix, and $\|\bullet\|_2$ denotes l_2 norm of a vector. Since we have K drug similarity matrices W_1, W_2, \dots, W_K , we defined and constrained a weight coefficients $\alpha = \{\alpha_k\}_{k=1}^K$ in a simplex, whose k -th element is the importance of drug similarity matrix W_k , and the sum of all element of α is 1. Similar to the definitions in formula (1), the first term of the objective in problem (2) is the prediction smoothness, which means that a good classification function should not change too much between nearby points. The second term of (2) is the fitness term, which means a good classification function should be consistent with the initial label assignment. The trade-off between those two terms is captured by a positive parameter μ which between 0 and 1. In the third term of (2), $\|\alpha\|_2^2$ is a regularizer to avoid trivial solutions, and $\delta > 0$ is a regularization parameter.

There are two groups of variables, F and α , in problem (2). Although the objective of problem (2) is not jointly convex with respect to F and α , it is convex with respect to one group of variables with the other group fixed. Thus, we adopted *Block Coordinate Descent (BCD)* schema, which starts by initializing both groups' variables, and then alternatively solves the optimization problem with respect to one group of variable with the other group fixed. The resultant two sub-problems are:

Fix α , solve F In this case, the weight vector α is given: the k -th element of α corresponds to the importance of the k -th similarity matrix, and the sum of all elements in α is 1. This is the same problem

as the single similarity matrix case if we make $\bar{W} = \sum_{k=1}^K \alpha_k W_k$. Thus the solution of this iterative step is $F = (1 - \mu)(I - \mu\bar{W})^{-1}Y$. In our experiment, we don't have any prior knowledge on different sources. Therefore, we initialize α from a uniform distribution (i.e., assuming all similarity matrices are equally weighted in the beginning).

Fix F, solve α . When F is fixed, the problem in formula (2) becomes

$$\begin{aligned} \min_{\alpha} \quad & \mu \sum_{k=1}^K \alpha_k c_k + \delta \|\alpha\|_2^2 \\ \text{s.t.}, \quad & \forall k, \alpha_k \geq 0, \sum_{k=1}^K \alpha_k = 1 \end{aligned} \quad (3)$$

where $c_k = \text{tr}(F^T(I - W_k)F)$, which is a standard *Quadratic Programming (QP)* problem.

Here, we reformulated this problem to facilitate more efficient solutions. For notational convenience, we denoted $C = (c_1, c_2, \dots, c_k)^T$. we first rewrote the objective of problem (3) as

$$\mu \alpha^T C + \delta \|\alpha\|_2^2 = \delta \left\| \alpha - \frac{-\mu}{2\delta} C \right\|_2^2 - \frac{\mu^2}{4\delta} \|C\|_2^2$$

As $\frac{\mu^2}{4\delta} \|C\|_2^2$ is a constant with respect to α and $\delta > 0$, (3) can be rewritten as

$$\begin{aligned} \min_{\alpha} \quad & \left\| \alpha - \frac{-\mu}{2\delta} C \right\|_2^2 \\ \text{s.t.}, \quad & \forall k, \alpha_k \geq 0, \sum_{k=1}^K \alpha_k = 1 \end{aligned}$$

where μ is the influence parameter and δ is the regularization parameter as in (2) and (3). This is a Euclidean projection problem under the simplex constraint and can be solved by the algorithm in^{27,28} (as in our experiments).

After the alternating optimization procedure converges, we obtained both the predicted DDI matrix F and the weight coefficient vector α that can be used to rank and prioritize multiple similarity matrices.

Results

Performance evaluation. A hold-out validation was carried out by dividing the initial DDI dataset into training and testing subsets. To ensure the validity of the test cases, we held out all the DDIs associated with a fixed percentage of the drugs, rather than holding out DDIs directly. This validation setting mimics a real-world situation of the drug discovery: once drugs without any interaction information come, a computational method should provide the DDI prediction based on properties of both these drugs and training drugs which exist in the current system. To be specific, we randomly selected a fixed percentage (i.e., 15%, 25%, 50%, 75%, and 85%) of drugs for testing, and moved all DDIs associated with these drugs to the testing set. Then we constructed the models with the remaining DDIs as the training set. The model parameters were tuned with cross validation based on the training set. Models were tested on the testing set only after all model parameter tuning has been done. For each testing, we repeated the hold-out validation experiment 50 times with different random divisions of the DDI dataset, and computed the mean and the standard deviation of the *Area Under the Receiver Operating Characteristic Curve (AUROC)* as well as the *Area Under the Precision-Recall Curve (AUPR)* over the 50 repetitions. In the ROC and PR analytics, we utilized DDI interactions from TWOSIDES as reference positives, and the complement set of TWOSIDES DDI interactions as reference negatives.

In the experiment, we compared seven DDI prediction methods: (1) **Nearest Neighbor with Chemical Similarity (NN-Chemical)** that identifies novel DDIs by using the nearest neighbor similarity to drugs involved in established DDIs. The similarity metric used in this method is based on chemical structures. A similar method was published in reference¹⁰. (2) **Nearest Neighbor with Label Side Effect Similarity (NN-LabelSE)** that identifies novel DDIs by using the nearest neighbor similarity to drugs involved in established DDIs. The similarity metric used in this method is based on drug label side effect information. (3) **Nearest Neighbor with Off-Label Side Effect Similarity (NN-OffLabelSE)** that identifies novel DDIs by using the nearest neighbor similarity to drugs involved in established DDIs. The similarity metric used in this method is based on FAERS off-label side effect information. (4) **Label Propagation with Chemical Similarity (LP-Chemical)** that predicts novel DDIs by propagating established DDI information iteratively through the network (details described in "Label propagation algorithm" section). Weight between two nodes (i.e., drugs) in the network is measured by chemical structure similarity. (5) **Label Propagation with Label Side Effect Similarity (LP-LabelSE)** that predicts novel DDIs by propagating established DDI information iteratively through the network. Weight between two nodes in the network is measured by drug label side effect similarity. (6) **Label Propagation with Off-Label Side Effect Similarity (LP-OffLabelSE)** that predicts novel DDIs by propagating established DDI information

Methods	15%	25%	50%	75%	85%
NN-Chemical	0.6951+/-0.0031	0.6949+/-0.0021	0.6910+/-0.0016	0.6838+/-0.0017	0.6805+/-0.0017
NN-LabelSE	0.7359+/-0.0018	0.7348+/-0.0019	0.7304+/-0.0016	0.7288+/-0.0016	0.7241+/-0.0020
NN-OffLabelSE	0.7557+/-0.0017	0.7500+/-0.0023	0.7436+/-0.0015	0.7397+/-0.0015	0.7390+/-0.0019
LP-Chemical	0.8676+/-0.0015	0.8654+/-0.0012	0.8603+/-0.0007	0.8515+/-0.0009	0.8415+/-0.0015
LP-LabelSE	0.8907+/-0.0014	0.8880+/-0.0010	0.8816+/-0.0007	0.8713+/-0.0007	0.8642+/-0.0014
LP-OffLabelSE	0.9219+/-0.0012	0.9194+/-0.0010	0.9115+/-0.0009	0.8994+/-0.0009	0.8888+/-0.0012
LP-AllSim	0.9258+/-0.0011	0.9233+/-0.0009	0.9156+/-0.0008	0.9033+/-0.0008	0.8921+/-0.0012

Table 1. Comparison of DDI prediction methods according to AUROC at different testing percentages.

Methods	15%	25%	50%	75%	85%
NN-Chemical	0.5182+/-0.0018	0.4932+/-0.0020	0.4067+/-0.0015	0.3810+/-0.0013	0.3667+/-0.0014
NN-LabelSE	0.5537+/-0.0024	0.5137+/-0.0019	0.4549+/-0.0016	0.4038+/-0.0016	0.3791+/-0.0015
NN-OffLabelSE	0.5780+/-0.0022	0.5412+/-0.0026	0.4710+/-0.0015	0.4331+/-0.0016	0.3924+/-0.0014
LP-Chemical	0.6128+/-0.0040	0.6026+/-0.0026	0.5876+/-0.0016	0.5545+/-0.0022	0.4726+/-0.0025
LP-LabelSE	0.6567+/-0.0035	0.6481+/-0.0021	0.6066+/-0.0019	0.5826+/-0.0023	0.5323+/-0.0028
LP-OffLabelSE	0.7195+/-0.0031	0.7189+/-0.0019	0.6827+/-0.0022	0.6477+/-0.0026	0.6349+/-0.0031
LP-AllSim	0.7292+/-0.0032	0.7282+/-0.0021	0.7052+/-0.0022	0.6736+/-0.0025	0.6501+/-0.0034

Table 2. Comparison of DDI prediction methods according to AUPR at different testing percentages.

iteratively through the network. Weight between two nodes in the network is measured by FAERS off-label side effect similarity. (7) **Label Propagation by Integrating All Similarities** (LP-AllSim) that predicts novel DDIs by integrating label propagation processes of multiple drug similarity networks (details described in “Label propagation with multiple similarities” section). In this study, we integrated networks derived from chemical structure, label side effect, and off-label side effect information sources. The comparisons of the hold-out validation assessed by AUROC are summarized in Table 1 and by AUPR are summarized in Table 2.

Several observations can be made from Table 1: (1) Our proposed label propagation algorithm boosted the DDI prediction performance. With the similarity from the same information source, label propagation based methods obtained much higher AUROC scores than nearest neighbor based methods (e.g., at testing percentage of 15%, NN-Chemical achieved averaged AUROC of 0.6951 and LP-Chemical achieved average AUROC of 0.8676). (2) Side effect information is an effective source for constructing drug network when applying the label propagation algorithm, as LP-LabelSE and LP-OffLabelSE obtained higher AUROC than LP-Chemical. For example, with 15% data testing, LP-LabelSE and LP-OffLabelSE achieved averaged AUROC of 0.8907 and 0.9219 respectively, and LP-Chemical achieved average AUROC of 0.8676. (3) Our proposed LP-AllSim method effectively integrated multiple similarities. For DDI prediction, if we only use one drug information then the resultant predictions could be biased. It may also suffer from the noise in the focused drug information source. At all percentages for testing data, LP-AllSim outperformed both the label propagation methods that use a single similarity source (i.e., LP-Chemical, LP-LabelSE, and LP-OffLabelSE) and the methods that use nearest neighbor algorithmic framework (i.e., NN-Chemical, NN-LabelSE, and NN-OffLabelSE). (4) Across different testing percentages, the higher the percentage (i.e., the less known DDIs as training) the lower the AUROC scores.

In Table 1, the AUROC scores only decrease less than 5% for all DDI prediction methods, even the testing percentage jumps from 15% to 85%. It does not mean DDI prediction is an easy task and a small DDI training data is enough to build a reliable model. For ROC analysis, a large change in the number of false positives may only lead to a small change in the false positive rate. Therefore, AUROC scores can present an over optimistic view of an algorithm's performance for the highly skewed data²⁹. In this study, we also provided a precision-recall analysis as a supplement to the ROC analysis.

As shown in Table 2, AUPR scores of all methods decrease at least 10% for all DDI prediction methods, when the testing rate increases from 15% to 85%. AUPR provides a quantitative assessment of how well, on average, predicted scores of true interactions are separated from predicted scores of true non-interactions. Thus it holds biological significance in practice: among the best ranked predictions that could potentially be experimentally tested, what proportion of true positives is present. Compare to NN-Chemical, NN-LabelSE, and NN-OffLabelSE, the large improvement in AUPRs suggests that the top ranked DDIs found by our methods are more likely to be correct. Even with testing percentage 50%,

Sources	15%	25%	50%	75%	85%
chemical structure	0.2451+/-0.0035	0.1417+/-0.0021	0.0264+/-0.0012	0.0000+/-0.0003	0.0000+/-0.0000
label side effect	0.3412+/-0.0047	0.3464+/-0.0018	0.3609+/-0.0012	0.3505+/-0.0009	0.2947+/-0.0008
off-label side effect	0.4137+/-0.0041	0.5119+/-0.0019	0.6127+/-0.0016	0.6495+/-0.0006	0.7055+/-0.0010

Table 3. LP-AllSim derived weights for chemical structure, label side effect, and off-label side effect information sources in experiments at different testing percentages.

LP-AllSim still achieves AUPR score that is higher than 0.70. The observation indicates that LP-AllSim is more robust to different performance measures and has higher potential to be useful in real-world pharmaceutical applications.

Table S1 lists the best influence parameter μ we used in the experiments for all label propagation algorithms. LP-AllSim also has a regularization parameter δ , which is not sensitive to test percentage values. Thus we set δ as 1 for all LP-AllSim experiments. Table S1 show that the higher testing percentage value (i.e., the less training data), the larger the best influence parameter μ . This observation indicates an important property of the propagation algorithms: when there is little training data, the predictions depend more on the geometric structure of the entire dataset, rather than just the training data. It is also the reason that label propagation strategy outperforms nearest neighbor strategy.

Data source comparison. Table 1 and Table 2 clearly show that clinical side effect information is more important than chemical structure information. From another perspective, LP-AllSim integrated chemical structure, label side effect, and off-label side effect information sources. Besides DDI prediction, the weight vector α derived from LP-AllSim is interpretable: the i -th element of α corresponds to the importance of the i -th data source, and the sum of all elements of α is 1. The weights of each data source are summarized in Table 3. Table 3 shows that compared to chemical structure source, side effect sources are more important in DDI prediction. The higher the testing percentage values, LP-AllSim depends more on side effect sources.

To further compare the information sources, we compared the worst-case DDI prediction performance by using chemical structure, label side effect, or off-label side effect information. Our hypothesis is to directly use drug-drug similarity scores as the DDI predictions, and evaluate their performance by using all known DDIs as the testing set (i.e., we didn't use any known DDI as training set). In the experiment, chemical structure similarity only achieves AUROC of 0.5366, which indicate that drugs that have interactions do not necessarily have similar chemical structures. On the other hand, label side-effect similarity and off-label side-effect similarity achieve AUROCs of 0.6108 and 0.6433 respectively. The observation seems to indicate that drugs having interactions sometime have overlapping serious side effects. The analysis provides a potential guideline to help clinicians to rule out unsafe co-prescriptions.

Novel predictions and case studies. We applied our LP-AllSim algorithm to predict new DDIs between all 1626 drugs with one or more side effect profiles. Among all the 1626 drugs, 702 have both label side effect information and off-label side effect information, 294 have only label side effect information, and 630 have only the off-label side effect information. We extracted chemical structures of all 1626 drugs from PubChem, thus each drug has at least two data sources. Among the 1,321,125 drug pairs, 63,468 pairs are identified as DDIs from TWOSIDES²⁰. We used all the 63,468 pairwise DDIs as the training data, and provided DDI prediction scores for the remaining 1,257,657 drug pairs. For pairs of drugs with all the chemical structure, label side effect, and off-label side effect information, we used LP-AllSim algorithm to integrate all three data sources. Otherwise, we used LP-AllSim algorithm to integrate two data sources (i.e., chemical structure and label side effect, or chemical structure and off-label side effect). We selected a cutoff for the ranked list of predictions according to the best F1-measure obtained from cross validation, thus obtaining 145,068 predicted DDIs from the 1,257,657 drug pairs. All DDIs prediction results are available in <http://astro.temple.edu/~tua87106/ddi.html>.

We validated our novel predictions by comparing the resulting DDI prediction to known interactions from DrugBank³⁰. DrugBank DDIs are extracted from drug's package inserts (accurate but far from complete), which can be used as independent validation sources. Of all the drug pairs, 1,892 DDIs are found only in DrugBank, but not found in TWOSIDES. The mean and standard deviation of the 1,892 DDI prediction scores are 0.3932 ± 0.1431 , which are significantly larger than those of the other 1,255,765 drug pairs (0.0717 ± 0.0996). Besides that, of the 1,892 DrugBank DDIs, 1,876 (99.15%) DDIs are included in our 145,068 predicted DDIs.

As an example, DDI predictions were analyzed between antihypertensive drugs and anti-inflammatory drugs. Table S2 (a) shows the prediction scores between Nonsteroidal anti-inflammatory drugs (NSAIDs) (e.g., ibuprofen, aspirin, and naproxen) and Angiotensin-Converting Enzyme (ACE) Inhibitors (e.g., benazepril, lisinopril, and ramipril)/Angiotensin II Receptor Blockers (ARBs) (e.g., candesartan, eprosartan, and valsartan) antihypertensive drugs. The mean and standard deviation of the 167 predictions are 0.3131 ± 0.0933 . And of the 167 predictions, 152 (91.02%) drug pairs are determined as DDIs using

the cutoff value described above. Table S2 (b) shows that the prediction scores between NSAIDs and Calcium Channel Blockers (CCBs) (e.g., amlodipine, diltiazem, and felodipine)/Central-Acting Agents (CAAs) (e.g., methyldopa, and clonidine) antihypertensive drugs. The mean and standard deviation of the 88 predictions are 0.1221 ± 0.0647 , which are significantly lower than those of predictions between NSAIDs and ACE/ARBs. Table S2 shows that when co-prescribing NSAIDs, hypertensive patients might take CCBs or CAAs antihypertensive drugs instead of ACE or ARBs to avoid potential adverse DDIs. This could be partially explained with the following reasons. NSAIDs inhibit prostaglandin-mediated vasodilation and promote salt and water retention. Both of these mechanisms contributing to NSAIDs partially reverse the effects ACE and ARBs, whose mechanism depends on modulating prostaglandins, renin, or sodium and water balance. In contrast, NSAIDs do not interact with CCBs and CAAs whose actions are apparently unrelated with renal/extrarenal production of prostaglandin³¹. The mechanism of action (MOA) has also been verified by a recent large-scale clinical study³².

Another example we analyzed is DDI predictions between cholesterol-lowering statin drugs and antibiotics. Table S3 shows the prediction scores between statin drugs (e.g., atorvastatin, lovastatin, and simvastatin) and clarithromycin/erythromycin. Of the 10 drug pairs included in our predictions, 9 drug pairs are predicted as DDIs. In addition, the only drug pairs which have been predicted as non-DDIs (i.e., atorvastatin and erythromycin lactobionate) gain a relatively high prediction score (i.e., 0.184649), showing that they still have some possibilities to interact with each other. Table S3 indicates that doctors might avoid ordering clarithromycin and erythromycin for older patients who take cholesterol-lowering statin drugs. This could be partially explained with the following reasons. Clarithromycin and erythromycin that inhibit the liver enzyme cytochrome P450 isoenzyme 3A4, and the inhibition might increase statin concentration in the blood, which can cause muscle or kidney damage, and even death. The MOA has been well documented and clinically verified by a recent population-based cohort study³³.

As a third example, DDI predictions were analyzed between Selective Serotonin Reuptake Inhibitor (SSRI) antidepressants and hydrocodone (a commonly used antitussive). Table S4 shows that the prediction scores between SSRI medications (including citalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline) and hydrocodone. All the 5 drug pairs were predicted as DDIs. Table S4 indicates that hydrocodone may interact with SSRI medications. Our predictions are in agreement with a clinical study³⁴ which shows that co-prescription of hydrocodone and SSRI medications may result in serotonin syndrome, a potentially life threatening drug reaction, that causes the body to have too much serotonin, a chemical produced by nerve cells.

Discussion

In this study, we have proposed an integrative label propagation framework to predict DDIs by integrating label side effects, off-label side effects, and chemical structures. A systematic comparison of the experimental results shows (1) side effect profiles are more predictive features than chemical structures in DDI prediction. It greatly benefits from the fact that clinical side effects are human phenotypic data obviating translation issues. (2) label propagation algorithm boosted the DDI prediction by considering high-order relationships between drugs. (3) our proposed integrative label propagation algorithm effectively integrated multiple drug properties and outperformed competitors. Furthermore, we applied the proposed algorithm to all known drugs which have one or more side effect profiles and obtained 145,068 predicted DDIs. These predicted DDIs can be leveraged for clinical surveillance and real-world drug discovery.

There are some limitations in our study:

- (1) Although SE information is more predictive than chemical structures in our DDI prediction experiments, SEs are usually unavailable for clinical candidates. For example, label SEs are detected during the phases of the clinical trials, and off-label SEs are accumulated during the post-market surveillance. For clinical candidates which at the early stages of drug development, the only available information is chemical structures. In future work, we may consider to map compound structure of clinical candidates to possible SEs via quantitative structure-activity relationship (QSAR) models and then predict DDIs based on inferred SEs. A similar idea was introduced by DRoSEf⁷, a methodology of drug repositioning based on the side-effectome.
- (2) We used 4,192-dimensional label SE vector and 10,093-dimensional off-label SE vector to quantify each drug. Since some SE terms are similar (e.g., “muscle stiffness” and “muscle tightness”, “ear discomforts” and “ear pain”), multiple synonyms could cause biases in SE profiling. Of the 569 drugs in our evaluation experiments, the chance of two drugs share a group of very similar SE terms is rare; and Tanimoto coefficient (which we used as metric to measure SE similarities) is not sensitive to the dimensionality problem (more details in Supplementary Section 2). Thus, the multiple-name problem doesn't have a significant impact on the SE similarity measurement. In future work, we will investigate the synonyms among SE terms. We may consider to measure semantic similarity between SE terms, remove redundant SE terms, and build more robust SE profiles to support our DDI predictions.
- (3) In this study, we used the unsafe co-prescriptions from TWOSIDES as known set of DDIs. However, TWOSIDES is directly derived from FDA Adverse Event Reporting System (FAERS), which contains some false positives (i.e., drug pairs included in TWOSIDES don't interact). When we utilized TWOSIDES as known interactions to build predictive models, the false-positive noises

may affect the predictions. Fortunately, our proposed label propagation algorithm is robust to these false-positive noises. The objective function (formula (1) in the section “label propagation algorithm”) of our method contains two terms (i.e., smoothness term and fitting term) and these two competing terms is captured by a nonnegative tradeoff parameter μ between 0 and 1. The larger the parameter μ , the more the objective function depends on the structure of the drug network (which built from either chemical structure or side-effect profiles). Thus as long as the data structure is reliable enough (where we use chemical compound and side-effects), the label propagation process can even correct some of the false-positive label noises.

- (4) In the experiment, we only predicted whether two drugs interact or not, without providing the reasons by which the drugs interact. In future work, we may extend our model to link specific reasons to DDI candidates. In this case, the “label” is no longer a simple “interaction”, but interaction with specific reasons. In other words, we have multiple types of labels to propagate, where each type of labels corresponds to interactions with a specific reason. In machine learning, this type of problem is usually referred to as multi-label learning. The proposed label propagation method does have the flexibility of handling multiple labels, where the whole procedure is equivalent to propagation on every specific type of label independently.

References

- Percha, B. & Altman, R. B. Informatics confronts drug-drug interactions. *Trends Pharmacol Sci* **34**, 178–184 (2013).
- Juurlink, D. N., Mamdani, M., Kopp, A., Laupacis, A. & Redelmeier, D. A. Drug-drug interactions among elderly patients hospitalized for drug toxicity. *JAMA* **289**, 1652–1658 (2003).
- van Leeuwen, R. W., Swart, E. L., Boom, F. A., Schuitenmaker, M. S. & Hugtenburg, J. G. Potential drug interactions and duplicate prescriptions among ambulatory cancer patients: a prevalence study using an advanced screening method. *BMC Cancer* **10**, 679; doi: 10.1186/1471-2407-10-679 (2010).
- Kuhn, M. *et al.* STITCH 2: an interaction network database for small molecules and proteins. *Nucleic Acids Res* **38**, D552–D556 (2010).
- He, L., Yang, Z., Zhao, Z., Lin, H. & Li, Y. Extracting drug-drug interaction from the biomedical literature using a stacked generalization-based approach. *PLoS One* **8**, e65814; doi: 10.1371/journal.pone.0065814 (2013).
- Duke, J. D. *et al.* Literature based drug interaction prediction with clinical assessment using electronic medical records: novel myopathy associated drug interactions. *PLoS Comput Biol* **8**, e1002614; doi: 10.1371/journal.pcbi.1002614 (2012).
- Noren, G. N., Sundberg, R., Bate, A. & Edwards, I. R. A statistical methodology for drug-drug interaction surveillance. *Stat Med* **27**, 3057–3070 (2008).
- Tatonetti, N. P. *et al.* Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. *Clin Pharmacol Ther* **90**, 133–142 (2012).
- Tatonetti, N. P., Fernald, G. H. & Altman, R. B. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. *J Am Med Inform Assoc* **19**, 79–85 (2012).
- Vilar, S. *et al.* Drug-drug interaction through molecular structure similarity analysis. *J Am Med Inform Assoc* **19**, 1066–1074 (2012).
- Luo, H. *et al.* DDI-CPI, a server that predicts drug–drug interactions through implementing the chemical–protein interactome. *Nucl. Acids Res* **42**, W46–W52 (2014).
- Vilar, S., Uriarte, E., Santana, L., Tatonetti, N. P. & Friedman, C. Detection of drug-drug interactions by modeling interaction profile fingerprints. *PLoS ONE* **8**, e58321; doi: 10.1371/journal.pone.0058321 (2013).
- Cami, A., Manzi, S., Arnold, A. & Reis, B. Y. Pharmacointeraction network models predict unknown drug-drug interactions. *PLoS ONE* **8**, e61468; doi: 10.1371/journal.pone.0061468 (2013).
- Takarabe, M., Shigemizu, D., Kotera, M., Goto, S. & Kanehisa, M. Network-based analysis and characterization of adverse drug-drug interactions. *J Chem Inf Model* **51**, 2977–2985 (2011).
- Gottlieb, A., Stein, G. Y., Oron, Y., Ruppim, E. & Sharan, R. INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol* **8**, 592; doi: 10.1038/msb.2012.26 (2012).
- Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).
- Yang, L. & Agarwal, P. Systematic drug repositioning based on clinical side-effects. *PLoS ONE* **6**, e28025; doi: 10.1371/journal.pone.0028025 (2011).
- Liu, Z. *et al.* Translating clinical findings into knowledge in drug safety evaluation–drug induced liver injury prediction system (DILIPS). *PLoS Comput Biol* **7**, e1002310; doi: 10.1371/journal.pcbi.1002310 (2011).
- Duran-Frigola, M. & Aloy, P. Recycling side-effects into clinical markers for drug repositioning. *Genome Med* **4**, 3; doi: 10.1186/gm302 (2012).
- Tatonetti, N. P., Ye, P. P., Daneshjou, R. & Altman, R. B. Data-driven prediction of drug effects and interactions. *Sci Transl Med* **4**, 125ra31; doi: 10.1126/scitranslmed.3003377 (2012).
- Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J. & Bork, P. A side effect resource to capture phenotypic effects of drugs. *Mol Syst Biol* **6**, 343; doi: 10.1038/msb.2009.98 (2010).
- Wang, Y. *et al.* PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* **37**, W623–W633 (2009).
- PubChem substructure fingerprint. Available at: ftp://ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprints.pdf (Accessed: 17th January 2015).
- Wang, F., Li, P., Konig, A. C. & Wan, M. Improving clustering by learning a bi-stochastic data similarity matrix. *Knowl Inf Syst* **32**, 351–382 (2010).
- Vanunu, O., Magger, O., Ruppim, E., Shlomi, T. & Sharan, R. Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* **6**, e1000641; doi: 10.1371/journal.pcbi.1000641 (2010).
- Wang, F. & Zhang, C. Label propagation through linear neighborhoods. *International Conference on Machine Learning (ICML)*. 985–992 (2006).
- Duchi, J., Shalev-Shwartz, S., Singer, Y. & Chandra, T. Efficient projections onto the l1-ball for learning in high dimensions. *International conference on Machine learning (ICML)*. 272–279 (2008).
- Chen, Y. & Ye, X. *Projection onto a simplex*. arXiv:1101.6081 (2011).
- Davis, J. & Goadrich, M. The relationship between precision-recall and ROC curves. *International Conference on Machine Learning (ICML)*. 233–240 (2006).

30. Knox, C. *et al.* DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* **39**, D1035–D1041 (2011).
31. Polonia, J. Interaction of antihypertensive drugs with anti-inflammatory drugs. *Cardiology* **88**, 47–51 (1997).
32. Fournier, J. P. *et al.* Non-steroidal anti-inflammatory drugs (NSAIDs) and hypertension treatment intensification: a population-based cohort study. *Eur J Clin Pharmacol* **68**, 1533–1540 (2012).
33. Patel, A. M. *et al.* Statin toxicity from macrolide antibiotic coprescription: a population-based cohort study. *Ann Intern Med* **158**, 869–876 (2013).
34. Gnanadesigan, N., Espinoza, R. T. & Smith, R. L. The serotonin syndrome. *N Engl J Med* **352**, 2454–2456 (2005).

Acknowledgements

The authors would like to thank Shahram Ebadollahi of IBM Research for his support during the progression of this study.

Author Contributions

Conceived and designed the experiments: P.Z. Designed algorithms and performed the experiments: P.Z. and F.W. Analyzed the experimental results and wrote the paper: P.Z., F.W., J.H. and R.S.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zhang, P. *et al.* Label Propagation Prediction of Drug-Drug Interactions Based on Clinical Side Effects. *Sci. Rep.* **5**, 12339; doi: 10.1038/srep12339 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>