ORIGINAL ARTICLE

Integrating Multiple Evidence Sources to Predict Adverse Drug Reactions Based on a Systems Pharmacology Model

D-S Cao[1]*, N Xiao[2], Y-J Li[1], W-B Zeng[1], Y-Z Liang[3], A-P Lu[4], Q-S Xu[2] and AF Chen[1]

Identifying potential adverse drug reactions (ADRs) is critically important for drug discovery and public health. Here we developed a multiple evidence fusion (MEF) method for the large-scale prediction of drug ADRs that can handle both approved drugs and novel molecules. MEF is based on the similarity reference by collaborative filtering, and integrates multiple similarity measures from various data types, taking advantage of the complementarity in the data. We used MEF to integrate drug-related and ADR-related data from multiple levels, including the network structural data formed by known drug–ADR relationships for predicting likely unknown ADRs. On cross-validation, it obtains high sensitivity and specificity, substantially outperforming existing methods that utilize single or a few data types. We validated our prediction by their overlap with drug–ADR associations that are known in databases. The proposed computational method could be used for complementary hypothesis generation and rapid analysis of potential drug–ADR interactions.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC? ☑ One of the main objectives in drug discovery and public health is to predict and monitor a drug's ADRs. Systems pharmacological models are urgently needed to combine various types of data for accurately predicting likely ADRs because of complex mechanisms underlying ADRs. • WHAT QUESTIONS DID THIS STUDY ADDRESS? ☑ We developed a multiple evidence fusion framework for systems pharmacology, and applied it to integrate drug-related and ADR-related data from multiple levels, the network structural data formed by known drug–ADR relationships for predicting likely unknown ADRs. • WHAT THIS STUDY ADDS TO OUR KNOWLEDGE ☑ Our results showed that our proposed model obtains high sensitivity and specificity, substantially outperforming existing methods that utilize single or a few data types. By integrating drug, ADR, and network-related information, we established a high accuracy systems pharmacology model for predicting potential ADRs. • HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS ☑ Our method is simple, applicable on a large scale, useful to predict or monitor drug ADRs, test a complementary hypothesis generation, and realize rapid analysis of potential drug–ADR interactions.

Drug use in medicine is based on a balance between expected benefits (already investigated indications before marketing authorization) and possible risks (i.e., adverse effects). Adverse drug reactions (ADRs) are undesirable effects that occur even when a drug is administered at the proper dose in the correct manner for an appropriate indication.[1] There is a major concern for ADRs in both the drug development and public health fields.[2] In the pharmaceutical industry, ADRs are one of the main causes of failure in the process of drug development and of drug withdrawal once a drug has reached the market. It is also the top reason for drug discontinuation in patients. In the healthcare industry, unrecognized or underreported ADRs not only cause preventable human suffering and costs to the healthcare system, but also

unnecessarily undermine the public's faith in drug therapy. Serious ADRs account for an estimated over two million hospitalizations annually. The fatal serious ADRs have become the 4th–6th leading causes of death annually. Studies in Europe and Australia have yielded similar estimates. It takes many years of study and safety surveillance to identify these ADRs completely. This delay in understanding impedes our ability to identify, evaluate, and use ADRs to optimize drug selection and dose. There is therefore a great need to predict and monitor a drug's ADRs throughout its life cycle, from preclinical screening phase to postmarketing surveillance.

To reduce ADR-related morbidity and mortality, several computational attempts to identification potential ADRs have been made, including: I) generating various drug-related

[1]School of Pharmaceutical Sciences, Central South University, Changsha, P.R. China; [2]School of Mathematics and Statistics, Central South University, Changsha, P.R. China; [3]Research Center of Modernization of Traditional Chinese Medicines, Central South University, Changsha, P.R. China; [4]Institute for Advancing Translational Medicine in Bone & Joint Diseases, School of Chinese Medicine, Hong Kong Baptist University, Hong Kong SAR, P.R. China. *Correspondence: D-S Cao (oriental-cds@163.com)

profiling (e.g., chemical profiling, cellular response profiling) to predict ADRs at different levels.[3–6] For example, ADRs at the level of organ systems are analyzed with screening data from the PubChem BioAssay database.[7] The premise for this is that some of the molecular actors of ADRs involve interactions detectable in compound screening campaigns. A study proposed by Liu *et al.* integrates the profiling from chemical, biological, and phenotypic properties of drugs to establish the classification model.[8] II) Utilizing sophisticated network inference methods such as network diffusion. Atias and Sharan proposed a diffusion process in the ADR similarity matrix to score each ADR by assuming that similar ADRs receive similar scores.[9] A similar approach was proposed by Cheng *et al.*, using a two-step resource allocation process.[10] These two methods only utilize a part of the network structure information, and neglect drug or ADR intrinsic properties. The pharmacological network model developed by Cami *et al.* improves this situation by implicitly introducing several network, taxonomic, and intrinsic covariates.[11] III) Detecting true signals from suspected adverse drug events (ADEs).[12] Currently, systematic evaluation of five signal detection algorithms was performed by Harpaz *et al.*[13] To varying degrees, these methods suffer from low sensitivity or specificity. They also involve well-known limitations, such as the difficulty of detecting rare or delayed-onset ADEs,[14–16] as well as ADEs that are already common in the treatment population.[17] IV) Identifying candidate targets that have a causal connection with ADRs.[18,19] Currently, a large-scale study proposed by Kuhn *et al.* systematically identifies protein-ADR associations by correlating drug–target interaction data with drug–ADR interaction data.[20] Furthermore, the identification of genetic risk factors for ADRs could lead to the safer use of drugs. Some genetic and genomic approaches may facilitate the identification of biological risk markers and reveal novel underlying mechanisms.[21]

Here we propose an approach for predicting novel associations between drugs and ADRs. Our method is logically derived from the following assumptions: I) Drug behavior at different levels provide some clues to understanding ADRs; II) semantic relationships between ADRs help us infer new ADRs for some drugs; III) the global or local information from the drug–ADR interaction network could inform us how to infer unknown drug–ADR associations. These assumptions are not universally true, but their degree of truth determines the accuracy and utility of our method. Our algorithmic framework follows the collaborative filtering system widely used in various electronic commerce websites. Given a query about a drug–ADR pair, we exploited several drug–drug or ADR–ADR similarity measures, and employed three aspects of recommendations based on three assumptions. The score of each drug–ADR pair according to each similarity measure allows us to determine the likelihood that the query concerning the drug–ADR pair interacts. The knowledge from multiple resources is systematically integrated to generate a multiscale model for predicting ADRs, following the idea of systems pharmacology. The prediction process is strictly evaluated and validated by different evaluation schemes. Importantly, the proposed computational method can be used for a comple-

mentary hypothesis generation and rapid analysis of potential drug–ADR interactions.

METHODS

Full details of the method and results are provided in the **Supplementary Data** online.

Data sources
Drugs and their associated ADRs were obtained from SIDER (as of October 2009).[22] This dataset consists of 880 drugs, 1,382 ADRs, and 61,102 drug–ADR associations. The ADRs in the databases were mapped to the MedDRA preferred term (PT). For a very small number of ADR names (less than 1%), we were not able to find a mapping at the MedDRA PT. We excluded those ADR names from our analysis. Moreover, drugs and ADRs vary greatly in their number of associations. Some ADRs are present in almost all drugs, while others are associated with very few drugs, and similarly for drugs. Thus, we filtered from the association data drugs and ADRs that lie at the top 5%, as well as ADRs and drugs having fewer than two associations. The resulting drug–ADR network contained 746 drugs, 817 ADRs, and 24,803 associations. All drug and ADR-related information was collected from different databases.

Similarity measures
We used node attribute-based and network structure-based similarity measures. For node attribute-based similarities, we computed eight drug–drug similarity measures and five ADR–ADR similarity measures, respectively, including chemical-based (ECFP), ATC-based (ATC), sequence-based (ProSeq), closeness in a PPI network (PPI), GO-based (ProGO), pathway-based (Pathway), disease-based (Disease), CMap-based (CMap), UMLS-based (UMLSLin and UMLSJCN), ADR coexist-based (Coexist), MedDRA-based (MedDRA), and ADR-related protein-based (APro). For network structure-based similarity, we computed three drug–drug and ADR–ADR similarity measures, including network neighbor-based (DNN and ANN), SimRank-based (DSimRank and ASimRank), path-based (DKatz and AKatz), and preferential attachment score (PAS). For the definition of each similarity, please see the Materials and Methods in the **Supplementary Materials**.

Generating classification features
The classification features were constructed from drug–drug and ADR–ADR similarity measures, resulting in 13 node attribute-based features and seven network structure-based features. Herein, we extended neighborhood-based collaborative filtering recommendation methods to generate drug/ADR-based recommendation scores as classification features.[23] For a given similarity measure, the score of a given drug–ADR association (d_i-a_j) is calculated by considering the similarity, according to the given pair, of k most similar to known drugs or ADRs to those in this association.

For a drug–ADR pair d_i-a_j, a linkage between d_i and a_j is determined by the following two predicted scores:

$$Score(d_i, a_j) = \frac{\sum_{m \in N\_k(i)} s(d_i, d_m) t_m, j}{\sum_{m \in N\_k(i)} t_{m,j}}$$

where $s(d_i, d_m)$ is similarity between drugs d_i and d_m (i.e., various similarity measures from drugs). $N\_k(i)$ denotes the set of k drugs most similar to drug i. $t_{m,j}$ is equal to 1 if drug m connects to ADR j, otherwise 0. Here, when the number of drugs most similar to drug i in the dataset is less than k, we used all these drugs to calculate the predicted score.

$$Score(d_i, a_j) = \frac{\sum_{n \in N\_k(j)} s(a_j, a_n) t_i, n}{\sum_{n \in N\_k(j)} t_{i,n}}$$

where $s(a_j, a_n)$ is similarity between ADRs a_j and a_n (i.e., various similarity measures from ADRs). $N\_k(j)$ denotes the set of k ADRs most similar to ADR j. $t_{i,n}$ is equal to 1 if drug i connects to ADR n, otherwise 0. Here, when the number of ADRs most similar to ADR j in the dataset is less than k, we used all these ADRs to calculate the predicted score. For each similarity measure, the optimal k value could be selected to generate the best classification feature by maximizing the AUC score (**Supporting Figure S1**).

Performance evaluation and novel predictions

We constructed different performance evaluation strategies to fully evaluate the prediction performance of our method. We predicted all 746 × 817 drug–ADR pairs, and selected those associations with high prediction confidence by setting a cutoff according to the precision-recall curve obtained from cross-validation. These predicted associations were further validated by manually looking up the drug–ADR associations from the SIDER (those associations from 2009 to 2012) and OFFSIDES databases. For detailed description, please see Materials and Methods in the **Supplementary Materials**.

RESULTS AND DISCUSSION
MEF: a multiple evidence fusion algorithmic framework for predicting ADRs

We designed a multiple evidence fusion (MEF) algorithm for predicting ADRs. The algorithm is based on the recommender system using multiple evidence resources. Given a gold standard set of drug–ADR associations, the basic idea behind the method is: If a drug interacts with an ADR, other drugs similar to the drug will be recommended to the ADR, and vice versa. Thus, we computed a number of similarity measures from different resources for drugs and ADRs, each of which represents a kind of evidence. Given a query drug–ADR pair, we made three aspects of recommendations (i.e., drug, ADR, and network) to determine whether this query drug–ADR pair interacts or not. The score of each drug–ADR pair according to each similarity measure allows us to determine the likelihood that the query drug–ADR pair inter-

acts. The accumulated scores were finally fed into a learned classifier that automatically weights different scores to yield a classification outcome. The algorithm works in three successive steps (**Figure 1**): I) construction of drug–drug and ADR–ADR similarity measures from different evidence; II) applying the collaborative filtering algorithm to construct classification features based on these similarity measures, and subsequent learning of a classifier which distinguishes true from false drug–ADR associations; III) applying the classifier to predict new drug–ADR associations.

Assembly of drug–ADR interaction data and evidence resources

We extracted 24,803 drug–ADR interactions from the 2009 SIDER data snapshot as our training set. A total of 746 drugs and 817 ADRs were involved in the extracted interactions (**Figure 2**). This network had 24,803 edges and 584,679 nonedges (proportion of edges in the training set: 4.07%). In the drug–ADR network, each ADR was mapped to the medical dictionary for regulatory activities (MedDRA, v. 16.0) and the unified medical language system (UMLS). Each drug was mapped to different domain-related levels including chemical structure, World Health Organization anatomical therapeutic chemical (ATC) classification system, target proteins, and various phenotypes.

Integrating the drug-related data from different databases, we assembled eight drug–drug similarity measures between the 746 drugs. Likewise, we constructed five ADR–ADR similarity measures between the 817 ADRs by integrating the ADR-related data from MedDRA and UMLS. These similarity measures totally derive from node attributes (i.e., drug or ADR representations), which greatly rely on the specific domain knowledge. We therefore called them node attribute-based similarity measures. Additionally, we defined three similarity measures and one preferential attachment score in terms of various topological characteristics of the drug–ADR network graphs. For convenience, we called them the network structure-based similarity measures.

Construction of classification features from multiple resources

We generated one classification feature for each similarity measure using the collaborative filtering algorithm, as described in the Methods section. Each feature represents a kind of evidence that allows us to infer novel drug–ADR associations. A total of 20 classification features were constructed, including eight drug-related features, five ADR-related features, and seven network-related features. The knowledge from the systematic levels, such as the molecular, cellular, individual, and network levels, represents multiple evidence features to generate a general assumption based on systems pharmacology.[24–28]

In a set of preliminary investigations, we individually checked the contribution of each feature to classification performance, and then evaluated the information overlap between the evidence. We show that each evidence feature has a moderate predictive ability, and the area under the receive operating characteristic curve (AUC) ranges from 0.57 to 0.88 (**Figure 3** and **Table S1**). Network
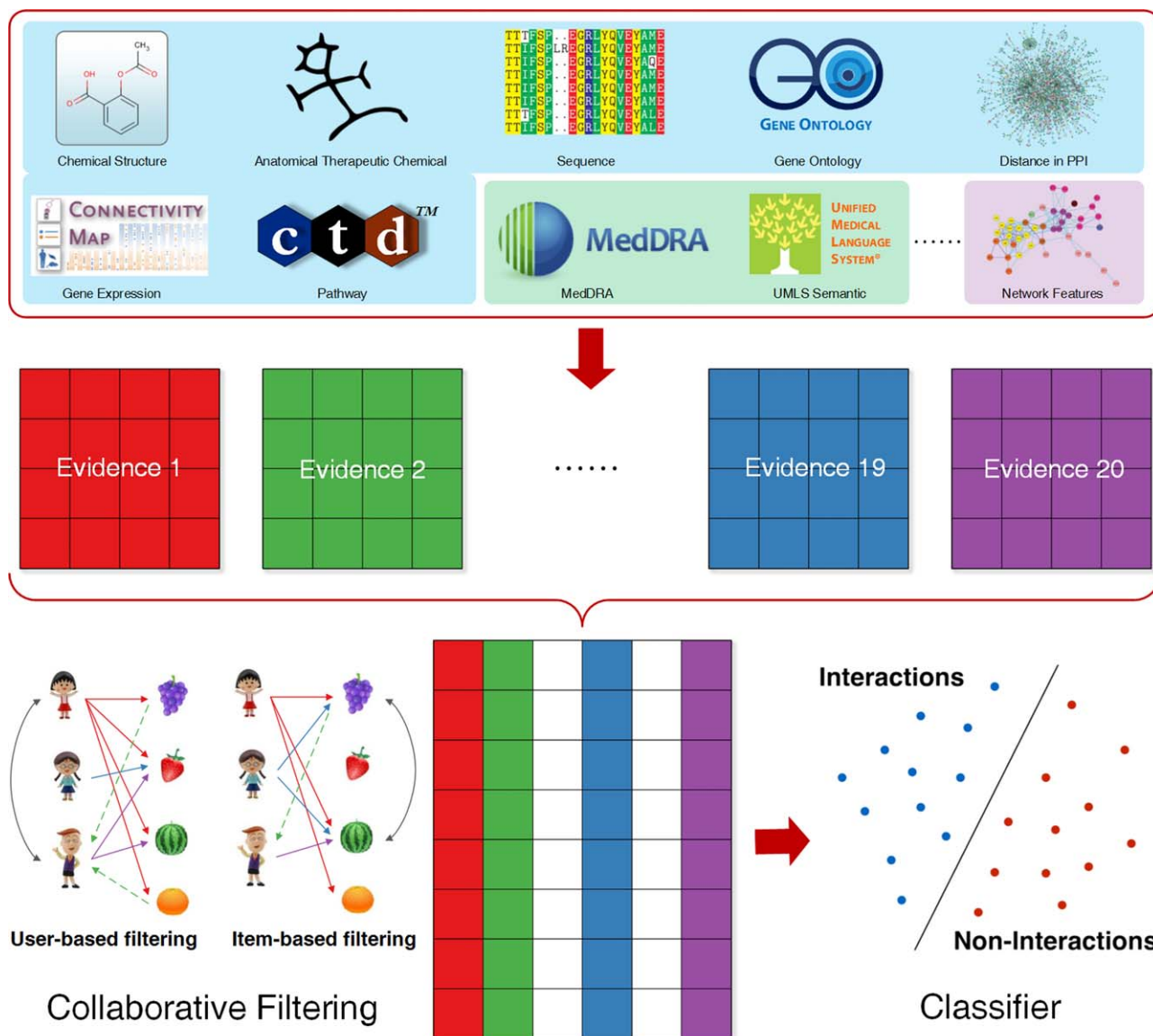
Figure 1 Illustration of MEF algorithm. By integrating multiple evidence resources, we computed a number of similarity measures for drugs and ADRs, including node attribute-based and network-based similarity. Given a query drug–ADR pair, we made three aspects of recommendations (i.e., drug-related, ADR-related, and network-related), based on collaborative filtering recommendation systems, to determine whether this query drug–ADR pair interacts or not. The score of each drug–ADR pair according to each similarity measure allows us to determine the likelihood that the query drug–ADR pair interacts. The accumulated scores were finally fed into a learned classifier that automatically weights different scores to yield a classification outcome.

structure-based features obtain the best performance as a whole. The best prediction is obtained by ANN and DNN (ANN: 0.88, DNN: 0.87), representing the structural equivalence in the drug–ADR network. The next is AKatz and DKatz (AKatz: 0.82, DKatz: 0.78), which consider the path length in the drug–ADR network. Among node attribute-based features, the best drug-related and ADR-related features are ATC and Coexist, resulting in AUCs of 0.77 and 0.84, respectively. However, it should be noted that the information in some evidence sources is incomplete because it is hard to find the complete information for some drugs or ADRs (**Table S2**). Thus, the prediction does

not totally reflect the true contribution to each kind of evidence. For instance, only 45% of drugs could obtain gene expression responses from the Connectivity Map, and only 46% ADRs could obtain ADR-causing proteins from Kuhn *et al.*'s work. Maybe the addition of the lost information for the evidence will continue to improve the prediction performance. To investigate the information overlap between the evidence, we first calculated their correlation for drugs and ADRs, respectively (**Figures S2, S3** and **Tables S3, S4**). The similarity values between drugs and drugs or ADRs and ADRs are relatively low, illustrating the difference of the evidence collected from drugs or ADRs. In
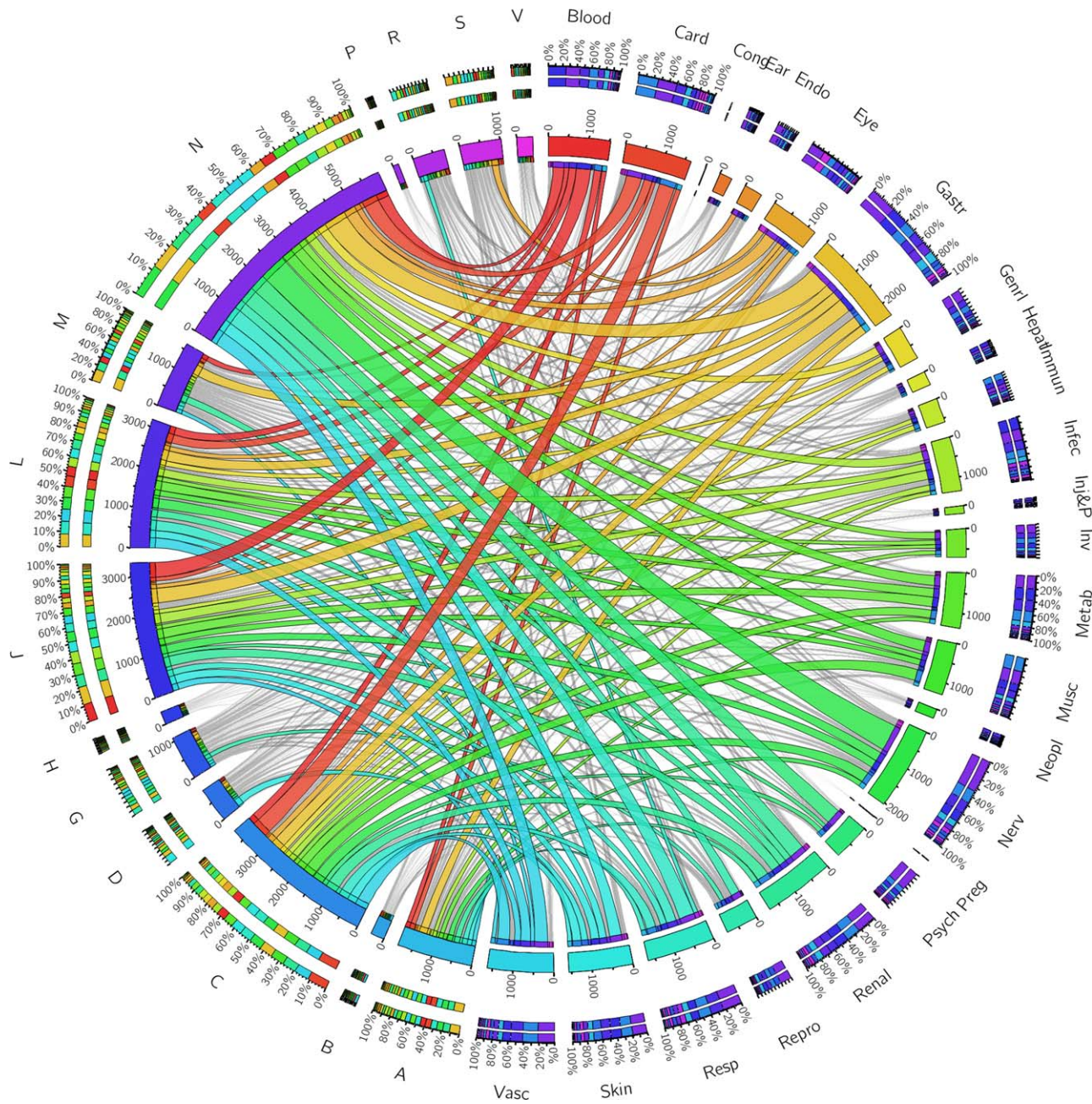
Figure 2 Drug–ADR interaction network. To clearly visualize the network relationships between drugs and ADRs, the drugs with the same top ATC level were bundled together, while the ADRs with the same top SOC level were bundled together, using an edge bundling technique. The length of the bars of the ATC levels on the outer ring represented the percentage of SOC levels at each ATC level, and vice versa. From this figure, nervous system agents have a wide range of ADRs, covering most of SOC categories. Significantly, nervous system agents usually cause nervous system disorders. Cardiovascular drugs usually cause cardiac disorders. The drugs applied to sensory organs usually lead to eye disorders.

order to further verify the situation, we next checked the accurately predicted associations from each kind of evidence (**Figure S4** and **Table S5**). Although the prediction accuracies from some evidence are similar, we found that the accurately predicted associations are different, indicating that different evidence preferred different drug–ADR associations. To a certain extent, integrating multiple evidence features should yield better performance. In summary, these observations are not perfectly predictive, but they adjust probabilities. This is all that our method requires, because the goal is simply to classify drug–ADR pairs on the basis of higher and lower aggregate probabilities. By combining different evidence signals in these datasets with the existing model favoring drug–ADR pairs for which we have more evidence, we can predict the ADRs more reliably.
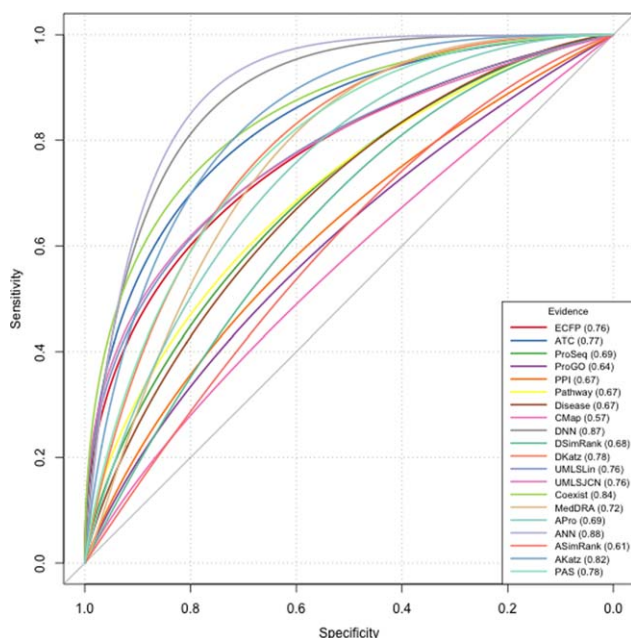
Figure 3 The ROC curves and the AUCs of 20 classification features from different evidence sources.

Training and validating a random forest classifier

We next investigated whether the combination of multiple evidence features improved the prediction accuracy or not. We trained a random forest (RF) classifier in all 20 features using a 10-fold cross-validation setting. Here, 1,000 classification trees are grown to construct the RF classifier. To avoid easy prediction cases, we hid all the associations involved with 10% of the drugs or the ADRs in each fold, rather than hiding 10% of the associations. The resulting models yielded AUC scores of 0.97 ± 0.01 for drug-based

cross-validation, and the same AUC scores of 0.97 ± 0.01 for ADR-based cross validation. The RF models correctly classified about 91.6% of the associations, with a sensitivity of 93.4% and specificity of 89.8%. We noted that no single feature obtained these high AUCs: the highest AUC of 0.88 was obtained using the ANN feature; removing each feature had a marginal effect on the overall AUC ($<$0.03). Distinct data sources complement each other in the prediction of ADRs since the coverage of each feature is incomplete and the overlap between different features is low (**Figures S2–S4**). By integrating these features where available, we improved the coverage of drug–ADRs interactions compared to a single feature.

Network features vs. node features

We investigated the use of node attribute-based features and network-based features in the model. By individually training RF classifiers based on two feature sets, we obtained their prediction performance using the same validation strategy. The RF models using node attributes and network features obtained AUCs of 0.91 ± 0.01 and 0.96 ± 0.01, respectively (**Figure 4a**). The prediction from network features is significantly better than that from node attributes, indicating the potential ability of network information in predicting ADRs. For all this, a part of accurately predicted drug–ADR associations for two types of features are different (**Figure 4b**), illustrating that different types of features preferred different drug–ADR associations again. By combining them into a RF model, reassuringly, the prediction performance again improved, although this improvement was relatively low.

Drug- and ADR-specific prediction performance

In order to investigate any potential variation in performance according to drug and ADR, we carried out a second evaluation. For this evaluation, we generated different
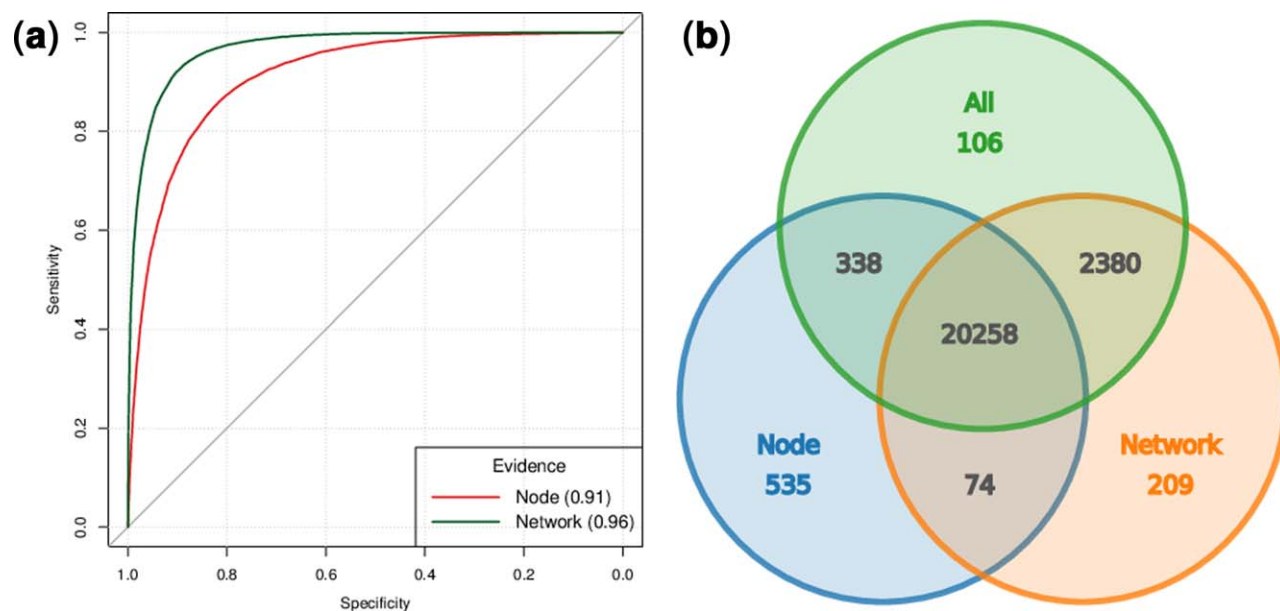

Figure 4 (**a**) The ROC curves and the AUCs using node features and network features, respectively. (**b**) Venn diagram for MEF predictions using node features, network features, and their combinations.

Table 1 The 24 predicted associations associated with six drugs and their corresponding scores

Drug names	ADRs	Scores
Ibuprofen	Bronchitis	0.998
	Peripheral edema	0.990
	Dysphagia	0.990
	Gastroenteritis	0.998
	Hypokalemia	0.986
	Sinusitis	0.974
Goserelin	Breast tenderness	0.960
	Cyst	0.988
	Pulmonary embolism	0.998
	Acne	0.958
	Cystitis	0.952
Prednisone	Cardiomegaly	0.996
	Vasculitis	0.974
	Thrombophlebitis	0.982
	Diabetes mellitus	0.962
Phenytoin	Diabetes mellitus	0.998
	Aspiration pneumonia	0.990
	Liver function tests abnormal	0.996
Atenolol	Weight gain	0.964
	Pneumonia	0.984
	Erythema multiforme	0.962
Methylprednisolone	Vasculitis	0.976
	Neuropathy	0.982
	Cardiomyopathy	0.984

level-specific validation sets and, for each set, computed an AUC statistic on the basis of prediction probabilities. First, we generated 14 ATC top-level validation sets and 24 MedDRA top-level validation sets for drugs and ADRs, respectively. Drug-specific AUCs were plotted; with drugs grouped according to the ATC top-level categories (**Figure S5** and **Table S6**). For all the drugs, the AUC scores were above 0.95, and mean AUC scores did not vary much across drug ATC categories. There are two groups for which the model produced AUC scores above 0.975, such as dermatological and sensory organs. The prediction precisions of 13 ATC categories are above 0.85, except antiparasitic products (**Figure S6** and **Table S6**). ADR-specific AUCs were also plotted for MedDRA top-level groups (**Figure S7** and **Table S8**). For most of ADRs, the AUC scores were above 0.97. Congenital, familial, and genetic disorders got the best AUC score of 0.99. The prediction precisions of 23 MedDRA top-level groups were above 0.85, except pregnancy, puerperium, and perinatal conditions (**Figure S8** and **Table S7**).

To make a further validation, we generated all 746 drug-specific validation sets and 817 ADR-specific validation sets for drugs and ADRs, respectively. The plots of drug-specific or ADR-specific AUCs against each drug or each ADR are provided (**Figure S9** and **Tables S8, S9**). On the whole, the prediction of each drug or each ADR is relatively satisfactory. For drugs, there are 174 drugs for which the model produced a very high AUC score (>0.99). There are also a small number of drugs for which the model produced a small AUC score, including oxandrolone (0.73), prazi-

quantel (0.84), amrinone (0.84), and rifabutin (0.87). Likewise, there were 283 examples of ADRs for which the model produced high AUC scores. Examples of ADRs for which the model did not produce high AUC scores include perirectal abscess (0.78) and infestation (0.79). The prediction from multiple levels demonstrated the reliability and robustness of our proposed algorithm.

Analysis of novel predictions

In order to predict novel drug–ADR pairs, we scanned the entire drug–ADR associations systematically. We found a significant enrichment of drug–ADR associations according to prediction probabilities. By using alternative score thresholds, our method may be tuned to predict a subset of drug–ADR pairs with high likelihood at the cost of false discovery rate (FDR). We can estimate our FDR from the precision-recall curve (**Figure S10**). To trade off precision and recall, we chose a cutoff that corresponds to 30% recall, which has an RF prediction probability of 0.95. At this cutoff, precision is about 99% when recall is 30%, and therefore we estimate our FDR to be about 1%. In other words, at this cutoff (RF >0.95), on the training set, we capture 30% of the drug–ADR interactions. Based on this threshold, we predict 18,629 drug–ADR interactions in the total screening set, and 2,536 new drug–ADR interactions after excluding those appearing in the training set (**Table S10**). We expect about 2,510 of these associations to be true drug–ADR interactions. These associations only take up 0.39% of all cross-linking associations. Further analysis found that nearly 373 drug–ADR pairs have scores above 0.99, suggesting that there are many potential candidate drug–ADR pairs with a relatively high likelihood.

We checked open ADR-related databases, and confirmed some drug–ADR interactions supported our predictions. Among the 2,536 predicted associations, about 70.5% were successfully validated from two databases (SIDER and OFFSIDES). Subsequent targeted survey revealed that 11.24% of our predictions were approved in the SIDER database. 63.52% of our predictions were previously reported as potential ADRs in the OFFSIDES database, although they are not approved yet, corroborating the predictive power of our proposed method. In all, 108 associations related to 73 drugs are commonly confirmed by two databases, whose prediction scores are in the range of 0.95–1.00. **Table 1** lists the 24 predicted associations related to six drugs and their corresponding scores. For example, ibuprofen, a prototypical nonsteroidal antiinflammatory agent with analgesic and antipyretic properties, obtained an AUC score of 0.96 ± 0.016 and prediction precision of 0.94 ± 0.024 based on drug-related prediction (**Table S8**). We detected 30 novel associations, 15, 2, and 6 of which have been identified by OFFSIDES, SIDER, and two databases, respectively. Moreover, six predicted associations are still unidentified, implying that they could be novel potential ADRs (**Table S10**). Goserelin, a synthetic hormone, obtained an AUC score of 0.96 ± 0.016, and prediction precision of 0.89 ± 0.025. Among the 15 predicted associations by our method, all have been successfully identified, five of which are commonly confirmed by two databases. In summary, the large overlap (70.5%) between our predictions and those reported

demonstrates that our proposed method effectively predicts new potential drug–ADR interactions that are still not determined from clinical trials.

In conclusion, in the present study we proposed an MEF algorithmic framework for predicting unknown drug–ADR associations by integrating multiscale evidence sources. The proposed model achieved high rates of specificity and sensitivity in cross-validation. When all features are integrated into a model, we attained a higher AUC of 0.97, surpassing existing methods. Applying the model to the screening of new associations, we confirmed about 70.5% associations by looking up the databases. Our findings suggest that MEF could be useful for predicting future reported drug–ADR relationships.

We compared our current study with three previous studies with similar but intrinsically different ideas. Those studies and the current study are similar in that they integrate various types of information, and all are similarity-based inference processes for classification or clustering.[29–33] Compared to the method proposed by Gottlieb *et al.*,[34,35] the current study used two types of features—node attributes (drug attributes and ADR attributes), network structure—whereas Gottlieb *et al.* only used node attribute-based information, and did not include the network-based information. Additionally, the similarity-based inference process is totally different. The current study used collaborative filtering to generate recommendation scores as classification features, and therefore help the transparent interpretation of drug-specific or ADR-specific information, while Gottlieb *et al.* used the pairwise similarity-based inference scheme to generate classification features, which are not easily interpretable. Compared to the pharmacological network models proposed by Cami *et al.*,[11] the current study proposed a flexible computational framework to integrate arbitrary similarity measures from multiple sources, while Cami *et al.* directly generated several covariates as classification features by covariate definitions. Currently, Wang *et al.* proposed a similarity fusion network model for identifying cancer subtypes and predicting survival.[36] This method is inspired by the theoretical multiview learning framework developed for computer vision and image processing applications, while our method is inspired by the recommendation systems developed for various electronic commerce applications. Furthermore, collaborative matrix factorization methods were also developed for drug repositioning studies by integrating multiple aspects of similarities.[32,33]

The main limitations of our proposed approach are summarized as follows: I) When network features are included, our proposed approach can only be applied to detect new interactions for a drug or an ADR for which at least one interaction has already been established. When predicting new drugs or ADRs with no prior interaction data, we suggest applying drug features and ADR features to construct the classification model, and then make predictions. II) A limitation of our method in predicting ADRs is that it does not take into consideration genetic risk factors of some ADRs. This information is crucial for determining whether an interaction will take place in clinical reality or not. Nevertheless, this situation should be taken into account by the physician in each case individually.

We suggest that our predictions may be beneficial in three areas: I) drug development, especially postmarketing surveillance, aiding in assessment and verification of potential ADRs; II) driving and directing early focus of potentially serious ADRs and cost reduction of large-scale clinical trials; III) assisting in discovery of new mechanisms of drugs by recognizing the group of ADRs targeted by a particular drug, especially in solving problems related to drug repositioning, drug-target selectivity, and polypharmacology.

Author Contributions. D.S.C. and A.F.C. wrote the article; D.S.C., Y.J.L., and Q.S.X. designed the research; D.S.C., N.X., W.B.Z., Y.Z.L., and Q.S.X. performed the research; D.S.C., N.X., W.B.Z., Y.J.L., Y.Z.L., L.A., and A.F.C. analyzed the data; D.S.C., L.A., and A.F.C. contributed reagents/analytical tools.

1. Edwards, I.R. & Aronson, J.K. Adverse drug reactions: definitions, diagnosis, and management. *Lancet* **356**, 1255–1259 (2000).
2. Chiang, A.P. & Butte, A.J. Data-driven methods to discover molecular determinants of serious adverse drug events. *Clin. Pharmacol. Ther.* **85**, 259–268 (2009).
3. Pauwels, E., Stoven, V. & Yamanishi, Y. Predicting drug side-effect profiles: a chemical fragment-based approach. *BMC Bioinform.* **12**, 169 (2011).
4. Hammann, F., Gutmann, H., Vogt, N., Helma, C. & Drewe, J. Prediction of adverse drug reactions using decision tree modeling. *Clin. Pharmacol. Ther.* **88**, 52–59 (2010).
5. Fliri, A.F., Loging, W.T., Thadeio, P.F. & Volkmann, R.A. Analysis of drug-induced effect patterns to link structure and side effects of medicines. *Nat. Chem. Biol.* **1**, 389–397 (2005).
6. Scheiber, J. *et al.* Mapping adverse drug reactions in chemical space. *J. Med. Chem.* **52**, 3103–3107 (2009).
7. Pouliot, Y., Chiang, A.P. & Butte, A.J. Predicting adverse drug reactions using publicly available PubChem BioAssay data. *Clin. Pharmacol. Ther.* **90**, 90–99 (2011).
8. Liu, M. *et al.* Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *J. Am. Med. Inform. Assoc.* **19**, e28–e35 (2012).
9. Atias, N. & Sharan, R. An algorithmic framework for predicting side effects of drugs. *J. Comput. Biol.* **18**, 207–218 (2011).
10. Cheng, F. *et al.* Adverse drug events: database construction and in silico prediction. *J. Chem. Inform. Model.* **53**, 744–752 (2013).
11. Cami, A., Arnold, A., Manzi, S. & Reis, B. Predicting adverse drug events using pharmacological network models. *Sci. Transl. Med.* **3**, 114ra127–114ra127 (2011).
12. Vilar, S. *et al.* Similarity-based modeling applied to signal detection in pharmacovigilance. *CPT Pharmacometrics Syst. Pharmacol.* **3**, e137 (2014).
13. Harpaz, R. *et al.* Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. *Clin. Pharmacol. Ther.* **93**, 539–546 (2013).
14. Frantz, S. Pharma's year of trouble and strife. *Nat. Rev. Drug Discov.* **5**, 7–9 (2006).
15. Ellis, C. FDA alerts asthmatics to drug safety risk. *Nat. Rev. Drug Discov.* **2**, 765–765 (2003).
16. Hughes, B. Spotlight on drug safety. *Nat. Rev. Drug Discov.* **7**, 5–7 (2007).
17. Breckenridge, A., Woods, K. & Raine, J. Monitoring the safety of licensed medicines. *Nat. Rev. Drug Discov.* **4**, 541–543 (2005).
18. Tatonetti, N.P., Patrick, P.Y., Daneshjou, R. & Altman, R.B. Data-driven prediction of drug effects and interactions. *Sci. Transl. Med.* **4**, 125ra131–125ra131 (2012).
19. Campillos, M., Kuhn, M., Gavin, A.-C., Jensen, L.J. & Bork, P. Drug target identification using side-effect similarity. *Science* **321**, 263–266 (2008).
20. Kuhn, M. *et al.* Systematic identification of proteins that elicit drug side effects. *Mol. Syst. Biol.* **9**, 663 (2013).
21. Wilke, R.A. *et al.* Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. *Nat. Rev. Drug Discov.* **6**, 904–916 (2007).

22. 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 (2010).

23. Ricci, F., Rokach, L. & Shapira, B. *Introduction to Recommender Systems Handbook* (New York, Springer, 2011).

24. Ma'ayan, A. *et al.* Lean big data integration in systems biology and systems pharmacology. *Trends Pharmacol. Sci.* **35**, 450–460 (2014).

25. van der Graaf, P.H. CPT: pharmacometrics and systems pharmacology. *CPT Pharmacometrics Syst. Pharmacol.* **1**, e8 (2012).

26. Sorger, P.K. *et al.* An NIH white paper by the QSP workshop group 1-48 (NIH Bethesda, MD, 2011).

27. Tatonetti, N.P., Liu, T. & Altman, R.B. Predicting drug side-effects by chemical systems biology. *Genome Biol.* **10**, 238 (2009).

28. Hopkins, A.L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682–690 (2008).

29. Ding, H., Takigawa, I., Mamitsuka, H. & Zhu, S. Similarity-based machine learning methods for predicting drug-target interactions: a brief review. *Brief. Bioinform.* **15**, 734–747 (2014).

30. Hansen, N.T., Brunak, S. & Altman, R. Generating genome-scale candidate gene lists for pharmacogenomics. *Clin. Pharmacol. Ther.* **86**, 183–189 (2009).

31. Bass, J.I.F. *et al.* Using networks to measure similarity between genes: association index selection. *Nat. Methods* **10**, 1169–1176 (2013).

32. Zheng, X., Ding, H., Mamitsuka, H. & Zhu, S. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1025–1033 (New York, ACM, 2013).

33. Zhang, P., Wang, F. & Hu, J. Towards Drug Repositioning: A Unified Computational Framework for Integrating Multiple Aspects of Drug Similarity and Disease Similarity. *AMIA* (2014).

34. Gottlieb, A., Stein, G.Y., Oron, Y., Ruppin, E. & Sharan, R. INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol. Syst. Biol.* **8**, 592 (2012).

35. Gottlieb, A., Stein, G.Y., Ruppin, E. & Sharan, R. PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* **7**, 496 (2011).

36. Wang, B. *et al.* Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014).

Supplementary information accompanies this paper on the *CPT: Pharmacometrics & Systems Pharmacology* website (http://www.wileyonlinelibrary.com/psp4)