Statistical Learning For Predicting Probable Drug-Drug Interactions Using Machine Learning Classifiers

Prof. Pallavi N[1], Karthik N[2], Kiran V Kulkarni[3], NL Subramanya Tejas[4]
[1, 2] Dept of Computer Science & Engineering
[1, 2] Atria Institute of Technology Bangalore, India

Abstract- *Drug-drug interaction (DDI) is a change in the effect of a drug when patient takes another drug. Characterizing DDIs is extremely important to avoid potential adverse drug reactions. We represent DDIs as a complex network in which nodes refer to drugs and links refer to their potential interactions. Recently, the problem of link prediction has attracted much consideration in scientific community. We represent the process of link prediction as a binary classification task on networks of potential DDIs. We use link prediction techniques for predicting unknown interactions between drugs in five arbitrary chosen large-scale DDI data-bases, namely Drug Bank, KEGG, NDF-RT, SemMedDB, and Two sides. We estimated the performance of link prediction using a series of experiments on DDI networks. We per-formed link prediction using unsupervised and supervised approach including classification tree, k-nearest neighbors, support vector machine, random forest, and gradient boosting machine classifiers based on topological and semantic similarity features. Supervised approach clearly outperforms unsupervised approach. The Two sides network gained the best prediction performance regarding the area under the precision-recall curve (0.93 for both random forests and gradient boosting machine). The applied methodology can be used as a tool to help researchers to identify potential DDIs. The supervised link prediction approach proved to be promising for potential DDIs prediction and may facilitate the identification of potential DDIs in clinical research.*

I. INTRODUCTION

Combined use of multiple drugs at the same time (i.e., polypharmacy) is common in modern pharmacotherapy [1], particularly in older population who has required continuous treatment for one or more chronic diseases [2]. Empirical evidence reported that the percentage of theU.S. population taking three or more drugs increased for 12% in years 1988±1994 to 21% in years 2007±2010 [3]. In such settings drugs may interact; they are not independent from one another. Drug-drug interaction (DDI) is an event in which one drug influences thepharmacologic effect of another drug when both are administered together [4, 5]. Identifying DDIs is a critical process in drug industry and clinical patient care, especially in drug administration [6].

Adverse drug reactions (ADRs) are harmful reactions that are caused by intake of medications [7]. Many ADRs are not identified during clinical studies (i.e., before a drug is approved by a government). Liu [8] recently demonstrated that about 10% of all possible drug pairs may probably induce ADRs through DDIs. Therefore, one of the fundamental aspects in pharmacovigilance research field related to the detection and prevention of ADRs Đis to generate new knowledge about DDIs. Despite several resources for DDIs [10] (e.g., DrugBank,Drugs.com), a study has demonstrated that none of the actual public databases provide a tolerable coverage of all the known DDIs; these databases are either incomplete or they record a large number of irrelevant interactions [11]. Additionally, the great majority of DDIs is hidden in a crowd of unstructured textual data which is expanding at a large scale [12]. For example, as of date of this writing simple PubMed search returns about 150000 bibliographic citations which include MeSH term `Drug Interaction'. Hence, the main motivation behind this study is consideration of computerized approach to identify potential DDIs.

DDIs may be naturally represented as a network in which nodes refer to different drugs and relationships between them designate their interactions [13, 14]. Complex networks fascinate many researchers after the small-world [15] and scale-free [16] features were recognized in numerous real-life networks, such as the Web and large social networks that capture relationships between actors. The network induced can be employed to elucidate the architecture and dynamics of a complex system and assist us in identification of relevant topological properties, interesting patterns, and predicting future trends. Various studies have already been performed in pharmacology with interesting applications of complex networks, including DDIs prediction (e.g., [17, 18]). There are three main benefits of processing DDIs with net-work analysis approach [19]: (i) researcher can predict potential, previously unknown, DDIs; (ii) certain (insignificant) DDIs will be avoided in such knowledge representation; and (iii) relationships which link pharmacodynamic and pharmacokinetic drug characteristics to DDIs can be explored.
A plethora of statistical methods were employed and developed to predict DDIs. An extensive overview of recent

approaches is presented separately in the next section. Existing methods may be categorized into three main approaches to DDI prediction: (i) a similarity-based approach, (ii) classification-based approach, and (iii) text mining approach. A similarity-based methods are based on the assumption that similar drugs may interact with the same drug. For instance, two drugs may interact if they have similar molecular profile. Classification-based techniques mimic the DDI prediction task as a binary classification problem. For example, drug-drug pairs are represented as feature vectors, while target variable is represented by presence or absence of interactions. A particular instance of classification-based methods is link prediction, which aim is to assess the probability that a relation exists between pair of nodes in a network, based on observation of topology of existing nodes and their attributes [20, 21]. Finally, text-mining methods employ natural language processing techniques to extract plausible relations among drugs from unstructured data sources (e.g., from MEDLINE citations). However, Abdelaziz et al. [22] identified several issues that are overlooked by a great majority of DDI prediction studies: (i) inability to predict newly developed drugs, (ii) failure to handle extreme data skewness of DDI pairs, (iii) relying the analysis only on selected data source (mainly DrugBank), and (iv) careless evaluation techniques which is reflected by employing area under the ROC curve as the main evaluation metric to assess the quality of prediction. All these limitations encourage us to perform a new, improved experiment.In this study we examine link prediction from the viewpoint of predicting potential DDIs. The main objectives of this work are: (i) to represent the process of discovering potential DDIs as a binary classification task in which features are represented as topological and semantic measures between drugs, and (ii) to evaluate performance of unsupervised and supervised machine learning methods for predicting potential DDIs. This study is different from other related studies in the following facets: (i) we use broader set of databases for DDIs prediction including DrugBank, KEGG, NDF-RT, SemMedDB, and Twosides; (ii) besides network-based features we also include semantic-based features, for instance chemical information of a drug and assigned Medical Subject Headings (MeSH); (iii) regarding methodological considerations we assume balanced distribution of DDI pairs; (iv) in addition to unsupervised approach we also include supervised statistical learning methods; and (v) last but not least, the study relies on comprehensive statistical evaluation and on manual evaluation performed by trained pharmacist.

II. RELATED WORK

A recent comprehensive review of DDI detection utilizing clinical resources, scientific literature, and social media is given by Vilar et al. [23]. In previous section we defined three approaches to DDIs prediction, namely similarity-based approach, classification-based approach, and text mining approach. We review the most recent literature for each of the approaches in the next paragraphs.

The similarity-based approach exploits the idea of biological profiles which are used to compare drugs and infer new molecular properties [24]. Gottlieb et al. [25] performed statistical validation by considering various types of drug-drug similarities, including chemical-based and side-effect-based similarity. Vilar et al. [26] developed new approach appropriate for large scale data that detects DDIs based on similarity of molecular structural properties. Li et al. [27] presented a Bayesian network model which was combined with a similarity algorithm to predict the drug pairs from drug molecular and pharmacological features. Zhang et al. [28] developed an integrative label propagation framework to model DDIs by integration of ADRs and chemical structures. Sridhar et al. [29] developed a probabilistic approach for predicting DDIs.

They used probabilistic soft logic framework which is highly scalable. The evaluation demonstrated of more than 50% improvement over baselines. Ferdousi et al. [30] reported on a methodology for DDIs modelling based on comparison of functional profiles of drugs, where drug profiles were constructed using carriers, transporters, enzymes, and targets information. They predicted over 250000 potential interactions. Takeda et al. [31] predicted DDIs based on structural similarities and the interaction networks that consist of pharmacokinetics and pharmacodynamics properties.Classification-based approaches mimic the prediction of DDIs as a two-class classification task. Cami et al. [32] defined DDIs as combinations of feature vectors and then employ logistic regression model to predict future interactions. Their model achieves a sensitivity of 48% with a specificity of 90%. Cheng and Zhao [33] used four DDI similarity measures and applied various statistical learning methods (naive Bayes, classification tree, *k*-nearest neighbors, logistic regression, and support vector machine) to learn interactions between pairs of drugs. Jamalet al. [34] studied neurological ADRs. They use various properties of drugs including biological, chemical, phenotypic, and their combinations. They used feature selection based on relief to detect most important variables and then employed advanced statistical techniques to predict side effects. Abdelaziz et al. [22] developed a large-scale similarity-based framework that predicts DDIs using link prediction. The system can predict both novel DDIs among existing drugs as well as newly developed drugs. Similarly, Lu et al. [35] studied whether classicalsimilarity measures provide plausible approach to drug-target interaction prediction, when only information from network topology is available. They compare their method against restricted Boltzmann machines and demonstrated higher precision of the proposed approach.Zhang et al. [18] collected a variety of information sources (i.e., data about substructures, tar-gets, enzymes, transporters, pathways) and build prediction models using neighbor recommender, random walk, and matrix perturbation method. They demonstrated that the methods based on ensemble learning could derive higher prediction performance than individual algorithms. Hameed et al. [36] developed a methodology for DDI prediction that is especially use-able in situations when true negative instances for training are inadequate.Information about DDIs in the research literature is increasing rapidly. Third line of research thus utilizes text mining methods to infer novel DDIs. Duke et al. [37] perform literature discovery approach on large health information exchange data repository to predict and evaluate new DDIs. Their method could identify new clinically significant DDIs and also sup-ports mining for their potential biological roots. Huang et al. [38] presented a method that esti-mates the strength of network connection between drug targets to predict pharmacodynamics DDIs with 82% accuracy. Tari et al. [39] proposed a novel approach that integrates automated reasoning techniques and text mining do derive new enzyme-based DDIs from MEDLINE abstracts. Manual evaluation revealed about 81% accuracy of their approach. Gottlieb at al. [25] introduced an interaction prediction framework that allows the inference of both pharmacokinetic and pharmacodynamic DDIs. They reported high sensitivity and specificity ratesof the proposed approach. Lu et al. [17] recently described an automatic approach for the description of the mechanism of interactions using MEDLINE MeSH descriptors. Authors reported high accuracy for identification of appropriate MeSH headings, including drugs and proteins. Besides scientific literature, social media also provides promising approach that can be useful in detection of DDIs [23]. For example, Hamed et al. [40] presented computational framework that detects DDI patterns from Twitter hashtag-based networks.

III. MATERIALS AND METHODS

We compiled knowledge networks by using DDI data from five public drug databases, including DrugBank, KEGG, NDF-RT, SemMedDB, and Twosides. We formed a pair of drugs if both are involved in one adverse DDI. DDIs are typically represented as directed connections. In this work the direction of the interaction was ignored. DrugBank is an encyclopedic Web repository containing complete biochemical and pharmacological data about drugs, including biological mechanisms and targets information [41]. Most of the information in DrugBank is throughly curated from research literature. Currently, DrugBank lists 10376 drug entries and 577712 directed interactions among them. In this study we used version 5.0 of the DrugBank which was obtained from the Drug-Bank Web page (https://www.drugbank.ca) on August 1, 2017. We parsed the DDI information from the provided XML file and compiled an edgelist of drug identifiers combinations.SemMedDB is a database of semantic predications (i.e., subject-relation-object triples) parsed from MEDLINE bibliographic database abstracts by the SemRep tool. Subject and object arguments of each predication correspond to concepts from the Unified Medical Language System (UMLS) Metathesaurus while relations coincide with links from the UMLS Semantic Network. SemMedDB contains information from about 91 million predications from all of the MEDLINE citations (approximately 27 million bibliographic records as of this writing). We used the version v.30 of the SemMedDB database in this study that processed the MEDLINE up to end of June 2017. In this study, all`INTERACTS_WITH' relationships between pairs of drugs were used as potential DDIs. Pre-processed database contains 1447792 directed interactions among UMLS concepts that refer to drugs. Next we use MRCONSO table from UMLS Metathesaurus to map UMLS concepts to DrugBank identifiers. Final database of interactions contains 1688 compounds and 37287 interactions.Twosides is a comprehensive source of polypharmacy ADRs for combinations of drugs [45]. The version used in this study was obtained from the Twosides Web page on August 1, 2017. Interactions in Twosides database are restricted to only those that cannot be unambiguously ascribed to either drug alone. We parsed the interaction information from the downloaded text file (http://tatonettilab.org/) and build a database of drug identifier pairs for the interacting compounds. We use PubChem (https://pubchem.ncbi.nlm.nih.gov/) identifiers to map Twosides identifiers to DrugBank identifiers. Final database of interactions contains 340 unique compounds and 19020 interactions.

IV. DATA REPRESENTATION

Consider an undirected and unweighted network which is depicted as a simple graph G (V, E) that consists of a set of nodes V referring to drugs and a set of edges E representing interactions between drugs. Let |.| represent the cardinality of the set. Let us first introduce some notation which is essential to understand the basics of the link prediction; for a comprehensive introduction to the technical details of link prediction we refer the reader to excellent reviews by Liben-Nowell and Kleinberg [20] or LuÈ and Zhou [21].

Let U be the universal set containing (|V| |V| − 1)/2 possible edges. By U − E we denote a set of non-existing links (or links that will appear later in time). The problem of link prediction is to predict these missing links. To test prediction algorithms we split the set of observed links E into two partitions: the training partition ET and test partition EP. It follows that ET [EP = E and ET \ EP = ⌀. In this study, we split each data set E into 66% training and 33% test data.

For all pairs of nodes in the training data we calculate similarity measure, which reflects the chance that a pair of nodes will interact in the test data set. In terms of machine learning, each pair of nodes serve as a positive or negative example, depending on whether those node pairsform a link in the test network. We organize the whole network as a list of relations

$$U \hat{} \text{ fh}u_1 ; u_2\text{i}; \text{h}u_1 ; u_3\text{i}; \ldots ; \text{h}u_i; u_j\text{i}; \ldots ; \text{h}u_{n1} ; u_n\text{ig};$$

where$n$ is the number of nodes in the network. Each term of the list comprises a feature vector and a relationship (i.e., class) label. The label is 1 when u_i following u_j and 0 otherwise. A feature vector is composed by the two feature subsets, as described in the next section.

Our basic assumption is that similar nodes more probably form a potential DDI. For each non-existent pair (x, y) in a test data, a link prediction algorithm provides a score $s(x, y)2\ U - E^T$that is an estimate of the existence of link between nodes x and y.

V. FEATURE EXTRACTION

Extracting a relevant set of features is one of the most critical part of any statistical learning algorithm. Traditional link prediction research considers mostly the topological features. In this study we augment the set of topological features with four semantic features.

Topological features. *Common neighbor (CN)*. Due to its simplicity this is one of themost commonly used measure in link prediction [46]. For a node x, let $\Gamma(x)$ denotes a set of neighbors of x. For nodes x and y the CN is defined as the number of nodes that x and y have in common. CN gives the relative similarity between a pair of nodes. CN is formally defined as

$$s^{CN}{}_{x;y} \hat{}\ \text{j}L_{x;y}\text{j} \ \hat{}\ \text{jG}\ldots x\dagger \ \backslash\text{G}\ldots y\dagger\text{j}:$$

Jaccard's coefficient (JC). It is a normalized version of CN. JC assumes higher values of nodepairs (x, y), which have many common neighbors proportionate to the total number of neighbors they have [47].

Adamic/Adar index (AAI). This index was first proposed for measuring similarity betweentwo Web pages [48]. AAI definition is related to JC, with a correction that lower-connected neighbors are weighted more heavily.

Preferential attachment (PA). This is simply the product of the degrees of nodes x and y.

This measure rest on an assumption that new edges more probably connect to higher-degree nodes than to lower-degree ones [49]. PA is defined as

$$s^{PA}{}_{x;y} \hat{}\ \text{jG}\ldots x\dagger \qquad \text{G}\ldots y\dagger\text{j}:$$

Resource allocation (RAI). It is similar to AAI but it penalizes the common neighbors withhigher degree more rigorously. RAI is formally defined as

Common neighbors 1 (CCN). This measure begins with the base score given by $|\Lambda_{x,y}|$ andthen for every neighbor$i$ shared by x and y, CCN receives an additional point for everycommunity that x, b, and i are all in.

Resource allocation 1 (CRA). It is similar to the original resource allocation definition, but itgives extra weight to shared neighbors$i$ that are in at least one community with both x and y, and weight i's contribution toward the total score by the number of communities that i shares with x and y.

Within-inter cluster (WIC). WIC predicts link between a pair of nodes using informationfrom within-cluster (W) and inter-cluster (IC) common neighbors of these nodes. A community detection must be performed on the network before applying this metric. Each vertex belongs to only one community.

*Semantic features.*Drug therapeutic-based similarity (ATC). This type of similarity was evaluated through ATC codes. ATC coding system partitions compounds into different clus-ters according to the biological system or organ on which they act. The first level of the code which was used in this study indicates the anatomical main group. There are 14 main clusters. The ATC codes for all compounds were extracted from the main DrugBank file. There are 3322 unique ATC codes as of this writing in the DrugBank database. Each compound was represented by a binary vector in which elements refer to the presence or absence of the ATC codes.

*MeSH-based similarity (MESH).*MeSH is a controlled vocabulary which is used to indexMEDLINE database. MeSH-based similarity is based on MeSH terms that are associated with DrugBank entries. There are 2072 different MeSH terms in the DrugBank database. As in the case of drug therapeutic-based similarity, each compound was represented by a binary vector whose elements represent the presence of the MeSH terms. The MeSH-based similarity is defined as the cosine similarity between the IDF-weighted MeSH vectors of the two corresponding compounds.

Adverse drug effect-based similarity (ADE). For this type of similarity we use informationprovided by SIDER side effects database of drugs. SIDER provides data on marketed drugs and their known ADRs. The version used in this study (4.1) was obtained from the SIDERWeb page [51]. There are 1430 drugs and 5868 side effects in the database. Each compound was represented by a vector with binary values in which elements represent the presence of the side effect terms. The side effect similarity of two compounds is defined as a cosine similarity between the IDF-weighted side effect vectors of the two compounds.

VI. STATISTICAL LEARNING

From In this study we used unsupervised and supervised learning. Later was performed by using five state-of-the-art classifiers, namely classification tree (DT), *k*-nearest neighbors (*k*NN), support vector machine (SVM), random forest (RF) and stochastic gradient boosting also known as gradient boosting machine (GBM). These classifiers have become mainstream in modern statistical learning. A comprehensive overview of all learning methods is not in scope of this paper. However, in the following lines we will shortly introduce the basic background. For more deep insight please see Friedman et al. [52].

*Unsupervised classification.*For unsupervised classification we use combined similaritymeasure which is derived from standardized similarity scores for pairs of nodes based on topological and semantic properties of the networks. More formally, we define combined similarity measure as

$$s^{Comb}_{x;y}\hat{}\,\mathrm{Avg}\ldots s^{CN}_{x;y}\,;\, s^{JC}_{x;y}\,;\, \ldots\,;\, s^{ADE}_{x;y}\dagger;$$

whereAvg is arithmetic mean. A pair of drugs is predicted to have a link if its score is over a certain threshold *t*. Clearly, a lower threshold predicts more pairs to be links. In our settings we use *t* = 90th percentile as a threshold. For example, value of combined similarity above chosen threshold therefore predicts a link between selected nodes. We use class information as described previously in `Data representation' section.

*Classification tree.*DT is built by partitioning instances into local subsets using a series ofrecursive splits. Each node of a tree is constructed by a logical rule, where instances below a certain threshold fall into one of the two child nodes, and instances above fall into the other child node. Partitioning continues until a terminal node, where data instances are assigned a class label [52]. The prediction for an instance is obtained by a majority vote of the instances reaching the same terminal node. Classifier was constructed using the rpart package in R.

*k-nearestneighbors.kNN*classifier defines the class of a test instance according on themajority vote of its *k* nearest neighbors from training data [52]. We set the value of *k* using internal 5-fold cross-validation. We used the Euclidean metric for calculating distances between data points. *k*NN classifier was implemented using the class package in R.

*Support vector machine.*SVM classifier maps the input data set into a high-dimensionalfeature space and then constructs a hyperplane to separate classes based on a maximum margin principle. We can choose various kernel functions including linear or nonlinear [52]. SVM classifier was implemented using the e1071 package in R. The penalty parameter was determined by an internal 5-fold cross-validation. Our implementation uses the linear kernel.

*Random forest.*RF is a statistical learning methodology that perform ensemble learningfor classification. Ensemble consists of multiple classification trees [53]. We used bootstrap sampling on training data to grow each tree. We split each node using the best among a ran-domly selected subset of given features. Next, we combined class labels predicted by each tree in the forest. Majority vote is finally used to create final prediction. RF classifier was imple-mented using the ranger package in R.

*Gradient boosting machine.*GBM also provides ensemble learning, but the base learnersin a GBM are weak learners [54]. The trees in GBM are not grown to the maximum possible extent as in RF. The GBM starts with an imperfect model (i.e., the base learner that is not grown maximally) and generates a new model by successively fitting the residuals of the cur-rent model, using the same class of base learners as the initial imperfect model. GBM classifier was implemented using the gbm package in R.

VII. EVALUATION METRICS

To estimate the quality of the proposed methodology, we performed two types of analyses: we performed statistical validation on selected DDI data sets as well as qualitative validation on a small subset of DDIs. The performance of algorithms was evaluated by employing train-test schema. First we used ovun.sample() function from the ROSE package in R to create a representative sample of DDI pairs for each network. Models were trained and tuned using the caret package in R utilizing doMC package for parallel processing. We used createDataPartition() function to split the entire data set into training subset containing 66% of examples and a test subset containing 33% of examples. Model selection was carried out using 10-fold cross-validation on training subset, which is known to give the lowest bias and variance [52]. The model with the highest accuracy was selected as the candidate model and used to predict interactions in the testing dataset.
To benchmark the performance of our algorithms we used standard evaluation measures from statistical learning including precision, recall, F1 measure, area under the receiver operating characteristic (ROC) curve (AUROC), and area under the precision-recall curve (AUPR).

Precision refers to the proportion of instances classified as positive that are actually positives, while recall refers to the proportion of true positive instances correctly classified as positives. F1 measure is used to integrate precision and recall into a single measure. ROC curve is a plot of true positive rate (sensitivity) vs. false positive rate (1Ðspecificity). Despite its popularity, the ROC curve has some drawbacks including the inappropriateness for imbalanced data [55].For this reason we also used the AUPR.To evaluate statistically significant differences between classifiers across different networks, we followed the methodology proposed by DemsÏar [56] as implemented in scmamp package.

We used Friedman test, which is a non-parametric alternative of repeated ANOVA design. The test is based on rank comparison that identify an overall effect of the choice of classifier on performance across multiple experiments. The null hypothesis is that all classifiers are equivalent. When the null hypothesis of the Friedman test is rejected ($p < 0.05$), we proceed with theNemeny post-hoc test, which compares classifiers to each other across datasets and finds the statistical significance of differences between their average performance ranks.We used custom AWK and Python scripts for data preprocessing. Similarity measures were implemented using NetworkX package in Python. Other numerical computations, including statistical learning were performed using R programming language for statistical computing and graphics. Complete programming code to reproduce the results of this study is accessible in GitHub repository at URL https://github.com/akastrin/ddi-prediction.

VIII. CONCLUSION

Link prediction is a promising methodological framework for studying complex systems in different scientific disciplines, including pharmacology. We evaluate an approach to potential DDIs prediction using link prediction methodology. We study the prediction performance of unsupervised and supervised link prediction algorithms on several large-scale DDI networks.Although there exist many different approaches and algorithms, reliable prediction of links in a network is still a very challenging problem. Computational approach presented here can be used as tool to help researchers to identify potential DDIs. Overall, our results demonstrated favourable classification performance and suggest appropriateness of the presented methodology for potential DDIs identification.

IX. ACKNOWLEDGMENT

This research was supported by the Department of Computer Science and Engineering, AIT, India. We are grateful to all our friends who provided support and insight which assisted us a lot in carrying out this research. We also thank all of them for their worthy comments & criticism which helped us to refine the content and improvise the paper.

REFERENCES

[1] Lu Y, Shen D, Pietsch M, Nagar C, Fadli Z, Huang H, et al. A novel algorithm for analyzing drug-drug interactions from MEDLINE literature. Scientific Reports. 2015; 5:17357. https://doi.org/10.1038/srep17357 PMID: 26612138

[2] Juurlink DN. Drug-drug interactions among elderly patients hospitalized for drug toxicity. JAMA. 2003; 289(13):1652. https://doi.org/10.1001/jama.289.13.1652 PMID: 12672733

[3] Percha B, Altman RB. Informatics confronts drug-drug interactions. Trends in Pharmacological Sciences. 2013; 34(3):178±184. https://doi.org/10.1016/j.tips.2013.01.006 PMID: 23414686

[4] YanjunWienkers LC, Heath TG. Predicting in vivo drug interactions from in vitro drug discovery data. Nature Reviews Drug Discovery. 2005; 4(10):825±833. https://doi.org/10.1038/nrd1851 PMID: 16224454

[5] Jia J, Zhu F, Ma X, Cao ZW, Li YX, Chen YZ. Mechanisms of drug combinations: Interaction and net-work perspectives. Nature Reviews Drug Discovery.

2009; 8(2):111±128. https://doi.org/10.1038/nrd2683PMID: 19180105

[6] Tatonetti NP. The next generation of drug safety science: Coupling detection, corroboration, and validation to discover novel drug effects and drug-drug interactions. Clinical Pharmacology & Therapeutics. 2018; 103(2):177±179. https://doi.org/10.1002/cpt.949

[7] Edwards IR, Aronson JK. Adverse drug reactions: Definitions, diagnosis, and management. The Lan-cet. 2000; 356(9237):1255±1259. https://doi.org/10.1016/S0140-6736(00)02799-9

[8] MZhu Y, Elemento O, Pathak J, Wang F. Drug knowledge bases and their applications in biomedical informatics research. Briefings in Bioinformatics. 2018; p. bbx169

[9] Ayvaz S, Horn J, Hassanzadeh O, Zhu Q, Stan J, Tatonetti NP, et al. Toward a complete dataset of drug-drug interaction information from publicly available sources. Journal of Biomedical Informatics. 2015; 55:206±217. https://doi.org/10.1016/j.jbi.2015.04.006 PMID: 25917055