

Predicting Drug-Drug Interactions for Pre-market Drug Development Process: A Network Based Approach

S. D. L. Gunawardena^{#1}, A. R. Weerasinghe^{#2}, M. W. A. C. R. Wijesinghe^{#3}

Abstract— Drug-drug interactions (DDIs) are responsible for many serious adverse events; their detection is crucial for the safety of the patient but also it is very challenging. In recent years, several drugs have been withdrawn from the market due to interaction related Adverse Events (AEs).

This study describes a model which can be used to predict novel DDIs based on the similarity of drug interaction candidates to drugs involved in established DDIs which can be used in a large scale to discover novel DDIs. This model is mainly based on the assumption that if drug A and drug B interact to produce a specific biological effect, then drugs similar to drug A (or drug B) are likely to interact with drug B (or drug A) to produce the same effect. We have created a drug network using the 2011 snapshot of a widely used drug safety database which utilizes 352 distinct drugs and contains 3 700 interactions. Then, it was used to develop the proposed model for predicting future DDIs. The target similarities and side effect similarities (P-score) were calculated for all selected pairs of drugs. Then, it was used to develop the proposed model for predicting future DDIs. The proposed model mainly follows two distinct approaches: ‘Which forces the preservation of existing (known) DDIs’ and ‘Without forced to preserve existing DDIs.’ Underneath each of these approaches, three different techniques: target similarity score, side effect similarity (P-score) and resulting score were used to retrieve novel DDIs.

The proposed model was evaluated using the Drugbank 2014 snapshot as a gold standard for the same set of drugs which produce novel DDIs with an average accuracy of 95% and 92%, average AUC (Area Under the Curve) of 0.9834 and 0.8651 under each of these two approaches respectively.

The results presented in this study demonstrate the usefulness of the proposed network based drug-drug interaction methodology as a promising approach. The method described in this article is very simple, efficient, and biologically sound.

Keywords— Drug-drug Interactions, Adverse Events, Target similarity score, P-score.

I. INTRODUCTION

Adverse DDIs are a serious health issue that can result in significant morbidity, mortality and also a leading source of treatment inefficacy. Due to interaction related AEs, several drugs have been withdrawn from the market in the past few years. For example, Terfenadine (Seldane®) in

February 1998 and Cisapride (Propulsid®) in January 2000. According to statistics provided by the National Health and Nutrition Examination Survey, over 76% of elderly Americans are on two or more drugs [2]. Unfortunately, adverse drug events (ADEs) are common and account for 770 000 injuries and deaths in each year, and drug interactions account for as much as 30% of these ADEs [2]. Thus, there is a practical necessity to identify DDIs in the pre-stage of the drug development process. However, detection of DDIs is a very difficult task. Therefore it is more critical to develop effective methods for predicting DDIs years in advance.

DDIs may be categorized into various criteria depending on the task. Mainly, DDIs can be classified into two main categories: severity level and the underlying DDI mechanism [1]. Each of these categories can be further categorized into three fundamental sub categories. In terms of severity, DDIs can be categorized into minor, moderate and severe [1][7]. Minor DDIs are considered to have a slight clinical significance and typically they are only called for routine patient monitoring. Moderate DDIs have a higher clinical significance and may require dosage changes and closer monitoring. Major DDIs can lead to serious adverse effects and should typically be avoided. In terms of underlying mechanism, DDIs can be categorized into Pharmaceutical, Pharmacokinetic (PK) and Pharmacodynamics (PD) [1][3]. Pharmaceutical interactions occur due to the chemical or physical incompatibility. PK DDIs occur when one drug interferes with the absorption, distribution, metabolism, or elimination of another drug, leading to changes in the plasma concentration of the affected drug. PD interactions occur when one drug interferes with a second drug at a target site, leading to additive or subtractive effects on the drugs involved.

Typically, there are two major stages of identifying DDIs: pre-market and post market [1]. These identifications include variety of medical laboratory experiments and computer based simulation techniques. Even though there are plenty of methods exist, when it comes to practical scenario most of the crucial DDIs go undetected in the phase of pre-market stage, according to the evidence of interaction related post market warnings and withdrawals. There are several statistical methods exist to identify whether the combination of two or more drugs are associated with an increased risk of certain AEs. Most of these statistical approaches depend on analysing results of post market data such as insurance claim databases, spontaneous reports and other available electronic medical records [8][9][10][11]. The main weakness of these methods is, they rely on waiting for sufficient post-market evidence to accumulate. It is a process which is highly time consuming. By this time, there is a higher probability that various people could risk their lives due to interaction related AEs. Additionally, these methods also suffer from the problem of vastness and the space of possible drug-drug AE

Manuscript received on 18th Feb 2018. Recommended by Dr. M. G. N. A. S. Fernando on 25th May 2018.

This paper is an extended version of the paper “Network Based Prediction of Drug-Drug Interactions” presented at the ICTer 2017.

S. D. L. Gunawardena is from the University of Colombo School of Computing. (*dinusha_89@yahoo.com*).

Dr. A. R. Weerasinghe and Ms. M.W.A.C.R. Wijesinghe are Senior Lecturers at the University of Colombo School of Computing. (*arw@ucsc.cmb.ac.lk, crw@ucsc.cmb.ac.lk*).

combinations. Thus, there is a practical need for a methodology that could identify adverse DDIs promptly in the arena of the drug development process.

To meet this need, we propose a method based on the network structure. This network is constructed based on already known DDIs, as well as various intrinsic and taxonomic properties of drugs. In this network, nodes represent drugs and edges represent interactions among the drugs. The main objective of the predictive model is to identify missing edges (unknown DDIs) in the constructed network. The process of identifying new DDIs using target similarity and side effect similarity (P-score) is based on the basic idea, if drug A interacts with drug B and drug C is similar to A, then C should also interact with B (the argument also follows if A is replaced with B). Thus, by combining the knowledge of known interactions with target similarity and side effect similarity, it is possible to identify new interactions. We have also integrated the results which were obtained through target similarity and side effect similarity as resulting score to investigate the impact of predicting new interactions based on two distinct approaches: forces the preservation of existing (known) DDIs and without force to preserve existing DDIs.

II. RELATED WORK

The interaction among the drugs play an important role in terms of drug development and the drug discovery. According to the study, many patients are relying on multiple medicines, especially older adults often take more than one drug at a time. Therefore, there is a higher chance that they could suffer from the effect of significant cause of Adverse Drug Reactions (ADRs), which occurs mainly due to DDI.

Prior identification of DDIs is mainly useful in two distinct ways: (1) to prevent occurring adverse event and (2) to take the advantage of the beneficial effects of drug interactions, which provides greater total effect than the sum of the individuals of two (or more) drugs. Thus, development of predictive tools to help behaviour of possible DDIs is of great interest to pharmaceutical companies and regulatory authorities, such as the United States Food and Drug Administration (FDA).

To support this need, scientists have come up with multiple computational based approaches to predict DDIs. These researches are based on two fundamental approaches known as: knowledge based and similarity based. Often, knowledge based approaches are focused on the post-market stage and the similarity based approaches are focused on the pre-market stage and the initial part of the post-market stage in the drug development process. However, both approaches suffer from several limitations such as the necessity to distinguish drug classes and the inability to handle novel drugs for which limited reports exist [12].

Knowledge Based Approaches

Knowledge based methods predict DDIs based on the information from the numerous resources such as scientific literature [3], Drug-specific patient registries, administrative claims databases [14] and an electronic medical record database [4][14]. It also retrieves the information from the spontaneous reporting systems such as US Food and Drug Administration's Adverse Event Reporting System (AERS)

[2][11][13] by using the technologies like the semantic web [15] and linked data.

A. Adverse event Reports

This approach mainly focuses on retrieve DDIs based on the adverse event reports. In the study which was carried out by Nicholas et al., [2], they have used a novel signal detection algorithm to identify hidden DDIs signal from adverse event reports. Fundamental concept of their study is to divide the severe adverse events (SAE) to eight distinct classes based on their clinical significance; cholesterol, renal impairment, diabetes, liver dysfunction, hepatotoxicity, hypertension, depression, and suicide. Then, separate models were created for each class by using supervised machine learning methods. Then, using those models, they have discovered hidden signals for one of the eight AEs. Since there is no standard for drug interaction AEs, they have used two main strategies to label the appropriate drug-pairs. According to the first approach, they labelled drug-pairs as "positive" if at least one of the drugs in the pair was known to be associated with the AE (i.e., the single drug-event associations). In the second strategy, they labelled drug-pairs as "positive" if the pair is known to interact according to a list of clinically significant interactions with selected data.

B. Semantic Web and Linked Data Technologies

A different study was carried out by Pathak et al., [7] to identify DDIs by applying linked data principles and semantic web technologies for the patient electronic health records. The main idea of using semantic web technologies is to facilitate data to be manipulated by machines not just for the display, but for automation, integration, and reuse across various applications. The retrieval process of appropriate DDIs data from the electronic health record databases was conducted by using resource description framework.

The major problem of knowledge based approaches is that they are relying on waiting for sufficient post-market evidence to accumulate. This is a process which is highly time consuming. Therefore, the ultimate DDIs prediction results cannot be generated in the pre-market or in the initial post-market stages in the drug development process. This delay leads to waste of both time and money which are being allocated for the drug development process. In the worst-case scenario this might be a reason for the increase patient's deaths.

Knowledge based approaches are also suffering from the practical problems such as lack of reliable data. Due to privacy, security, ethical, policy and issues with confidentiality, patient data is closely guarded and monitored from unauthorized access within institutional firewall boundaries. Even though the data is found, there is no guarantee whether they are fake or accurate.

Similarity Based Approaches

Similarity based methods are predicting DDIs based on the measurement of the similarity of drug information. To identify similar drug pairs, researchers have used numerous similarity measures like: Chemical-based, Ligand-based, Side effect based, Annotation-based, Sequence-based, Closeness in a PPI network and Gene ontology based. Among these measures, the last three measures are gene related. There are

various researches which were being carried out to retrieve the drug similarity information such as; utilizing multiple drug-drug similarity measures to predict DDI: Inferring Drug Interactions (INDI) [12], target overlap [16], S-score [3], P-score [17], target distance, C-score [18], Indication overlap [19] and the molecular structural similarity [6]. Among these approaches, target based methods (target overlap, target distance and S-score) are comparatively better than using indication overlap and the C-score methods [3].

A. Predictive Pharmacointeraction Networks

A novel approach was developed by Cami et al [1] named Predictive Pharmacointeraction Networks (PPINs) to predict DDIs. This method utilizes known DDIs along with other intrinsic and taxonomic properties of drugs to predict novel DDIs. Based on the known drug-drug interaction data, a drug network was constructed. Then those data is used to construct a set of covariates and to develop predictive logistic regression and generalized mixed model [1]. The main idea of this model is to predict the probabilities for each non-edge in the already constructed drug network. Then, based on the resulted probability values, all the non-edges with higher probability values were considered as a successful edge in the network. Then, those results were used to predict novel DDIs. In simple term, interacting drugs were represented by two end nodes of the newly drawn edges.

B. Bayesian Model Averaging Approach

This is the approach which was developed by Guimera et al. They have introduced a network inference algorithm to predict uncharacterized DDIs [20]. The algorithm that they used is highly unsupervised and parameter free. Since their algorithm is highly abstract, it takes known DDIs as its only input. This algorithm does not require any biochemical or pharmacological information as inputs. The prediction is performed based on the technique stochastic block models [21][22], node partitioning by their cellular function. According to this model, the interaction between any pair of nodes depends only on the groups to which they belong.

C. Interaction Profile Fingerprints

Interaction profile fingerprints (IPFs) [5] is another successful way of predicting DDIs. This model consumes IPFs to measure the similarity of drug pairs and generate presumed DDIs from the non-intersecting interactions of a pair. The concept of IPFs is similar to the molecular structure fingerprint [23]. After identifying IPFs, the similarity between fingerprints are calculated using JC (refer Fig. 1 for example of IPFs calculation).

D. Target Based Methods

This is a technique which is used to identify DDIs based on the targets that it's going to hit. There are several methods which are being used to identify DDIs based on its target.

Target Overlap

Target overlap is a technique with the strategy of connecting two drugs if both drugs share at least one target protein [27]. In simple terms, target overlap is based on the basic idea that if drug A hits target proteins (X, Y and Z) and the drug B also hits the target proteins (V,W and X) drug A and drug B is said to be somewhat similar, because they share at least one target protein (according to this example it is X).

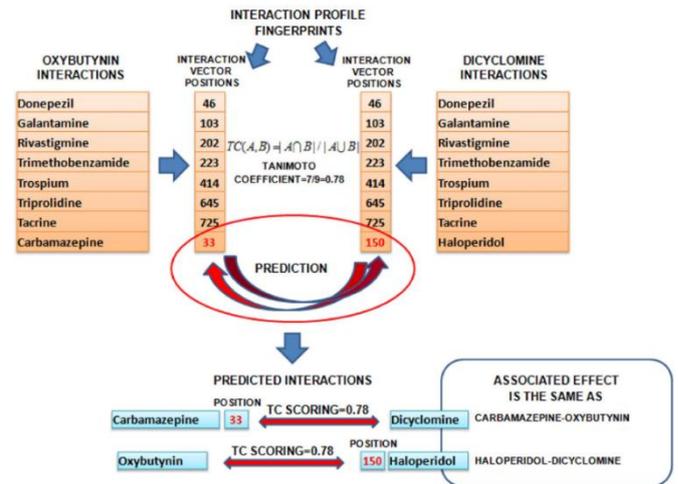


Fig. 1 Examples of interaction profile fingerprints (IPFs) calculated for the drugs oxybutynin and dicyclomine [5]

Jaccard Coefficient

The Jaccard coefficient (JC) (also named as Tanimoto Coefficient (TC)) [28]. Jaccard Coefficient is a statistic used to compare the similarity and diversity of sample sets. Jaccard coefficient measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets. JC is a value that can be varied in a range of 0 to 1 [$0 \leq J(A, B) \leq 1$], where 0 means maximum dissimilarity and 1 means maximum similarity. JC for given target representations A and B can be calculated as follows:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

S-score

The S-score reflects the tightness of a connection between two target-centered systems in the network. This value mainly depends on two parameters: number of edges connecting the genes in these two target-centered systems and the similarity in expression patterns across tissues [3]. The S-score can be calculated as follows:

$$S - score = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$$

where \bar{x} , s and n represent the mean, standard deviation and the number of the cross-tissue expression Pearson's correlation coefficient (PCC) of edges connecting two drugs-centered systems, respectively; μ represents the average PCC of all edges in the network as background. In addition, if two target-centered systems share a gene, an artificial edge with PCC of 1 is added between the two systems.

P-Score

The P - score is a technique with the strategy of connecting two drugs by their side effect similarity. Identifying drug side effect is a complex phenomenological observation that is based on various molecular scenarios, including interaction with the primary or additional targets (off-targets) [17]. These off-targets are being used to identify similar side effects of unrelated drugs.

Bayesian probabilistic model

Bayesian probability is one of the different interpretations of the concept of probability. Bayesian inference derives the posterior probability because of two precursors, a prior probability and a "likelihood function" derived from a probability model for the data to be observed [28]. Bayesian inference calculates the posterior probability based on the Bayes rule,

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

Where, $|$ denotes conditional probability, H denotes any hypothesis whose probability may be affected by the data, E denotes evidence corresponds to new data that were not used in computing the prior probability, $P(H)$ denotes the prior probability, $P(H|E)$ denotes the posterior probability, $P(E|H)$ denotes the probability of observing E for a given H and $P(E)$ denotes the marginal likelihood.

E. Inferring Drug Interactions

Another method of detecting DDIs is by using the Inferring Drug Interactions (INDI) [12]. INDI algorithm is designed based on two main objectives (1) predicting both new CYP-related DDIs and non-CYP-related DDIs and (2) developing a general strategy that allows predicting interactions of novel drugs for which no interaction information is currently available. This algorithm mainly relies on three fundamental steps (i) construction of drug-drug similarity measures; (ii) constructing classification features based on these similarity measures; and (iii) application of the classifier to these features to predict new DDIs (refer Fig. 2 for INDI steps)

This algorithm mainly relies on chemical and side-effect similarity of applicable drugs as its inputs, and based on those results novel prediction for DDIs were generated.

F. Molecular Structure Similarity

Molecular structure similarity is a technique which is used to identify similar drug pairs by considering the structural similarity of the drugs [24][25]. By using the knowledge of known interactions with structural similarity, it is possible to identify new DDIs [6][26]. In simple terms, structural similarity is based on the basic idea that if drug A interacts with drug B, and drug C is structurally similar to A, then C should also interact with B (the argument also follows if A is replaced with B). The identification of structural similarity is a three-step process: (1) collecting and processing drug structures, (2) structural representation and similarity measures, computation, and (3) data representation respectively [6]. Then these structural similarity calculations are used along with the established drug interactions to predict novel DDIs.

The proposed methodology mainly focused on implementing a novel network based model to predict DDIs. The main idea behind this methodology is to utilize the findings of two distinct approaches: target similarities and the P-score values to predict DDIs. This study also focuses on generating a new resulting matrix by integrating results obtained through target similarity matrix and the P-score value matrix.

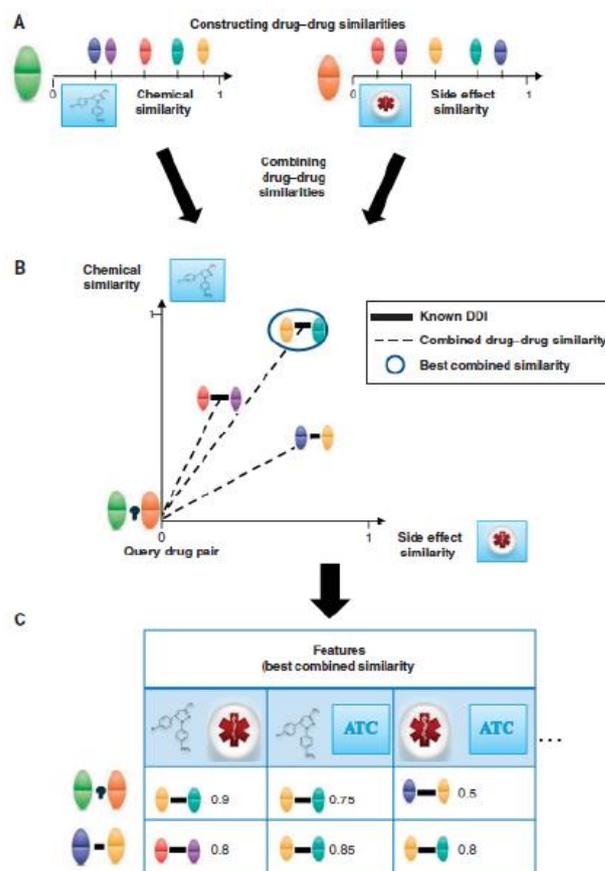


Fig. 2 Steps of INDI Algorithm [12]

III. METHODOLOGY

The methodology of this study consists of seven major steps: data acquisition, pre-processing, graph construction, identify similar drug pairs, predicting new DDIs, construct resulting matrix and output of final network.

A. Data Acquisition

This research is mainly based on three types of datasets:

FDA approved drug-target association data, FDA approved known DDIs data (2011 and 2014 snapshots of DrugBank database) and P-Score values.

The DDI data (both 2011 and 2014 snapshots) were downloaded from the DrugBank¹ database. FDA approved drug-target data and P-score values were obtained from the research "Systematic Prediction of Pharmacodynamic Drug-Drug Interactions through Protein-Protein Interaction Network" which was done by Huang et al., 2013 [3]. This dataset is published online at their personal website².

B. Pre-processing

The FDA approved drug target association data, FDA approved known DDIs (2011) dataset and P-score values were downloaded in text (.txt) file format. FDA approved known DDIs (2014) dataset was downloaded in XML file format. Mainly, the data pre-processing consists of four main phrases,

¹ <http://www.drugbank.ca>

² <http://www.picb.ac.cn/hanlab/DDI>

- *Retrieve intersection of DDIs from all available datasets.* There is a huge amount of DDIs available in the downloaded datasets: FDA approved drug target association data, FDA approved known DDIs (2011) dataset and P-score values. There were instances where a DDI is available in a selected dataset, but unavailable in the other dataset/s. In such instances, the model will produce erroneous outcomes due to the unavailability of required figures. Thus, to avoid such mistakes we have retrieved only the intersect DDIs from the available datasets.
- *Retrieve appropriate data columns from an existing dataset.* Downloaded datasets consist of several additional data fields which were not required for the proposed model. Such fields were simply ignored at the preprocessing stage to make the proposed model more optimize.
- *Eliminate the duplicate records.* There were some instances where the records were duplicated. In such instances, additional records were eliminated by retaining only distinct records. Main objective of the elimination process is to avoid any unnecessary calculations.
- *Process the XML to retrieve latest known DDIs (2014 snapshot of Drugbank database).* Downloaded XML consists of a vast amount of experimentally discovered DDIs up to the October 2014. Thus, only required set of interactions were filtered out at the pre-processing stage to construct the latest known DDIs network. The intension of this approach is to make the evaluation process more optimize.

C. Graph Construction

The drug network was constructed based on the pre-processed FDA approved known DDIs (2011) data. To construct the graph, known DDI data were represented using an adjacency matrix. To construct the DDI network, the above mentioned dataset with the set of known DDIs was transformed into a binary matrix M1 (with 352 rows and 352 columns). Where a matrix cell value of 1 represents a known interaction between a pair of drug and a value of 0 represent no interaction exists between selected pair of drug. Values on the diagonal of the matrix were set to 0, because the interaction of a drug by itself was omitted.

D. Identify Similar Drug Pairs

As a prerequisite for the prediction of DDIs, similar drug pairs were identified by using two distinct values:

Target Similarity Score: The target similarity score was calculated based on the available FDA approved drug target association data. This score was generated using the widely applied Jaccard coefficient (JC). The JC between the selected pair of drug was calculated as follows,

First, the number of common target protein count (intersection) was obtained for the selected drug pairs. Then, the number of all possible target protein count was calculated without considering duplicates (union). Then, the retrieved intersection target protein count was divided by the union of the target protein count to calculate the JC value. Formula is as follows,

$$J(A, B) = |A \cap B| / |A \cup B| \quad (1)$$

This calculation guarantees to give an answer between zero and one. Since the generated scores were always in between zero and one it is easy to manipulate the scores and to store them in an adjacency matrix for further calculations. The first instance of the similarity matrix M2 was constructed to capture the TC measure of similarity between pairs of drugs (the matrix cell value denotes the TC between pairs of drugs).

P-score: P-score values were directly taken from the research which was done by Huang et al., 2013 [3]. Since the obtained P-score values were spread in a vast range, it has been normalized using the tool Matlab in to a range of zero to one. The second instance of similarity matrix M2 was constructed by using normalized P-score values.

E. Predicting New DDIs

After the preparation of required inputs for the prediction model (established DDIs matrix, target similarity score matrix or P-score matrix) prediction process for novel DDIs is a three-step process in terms of technical perspective: Multiply matrixes (M1, M2), Retained the maximum value for each entry and Symmetric transformation considering the high value for each pair.

1.) Multiply matrix M1 (established DDIs matrix on 2011 dataset) and M2 (target similarity score matrix or P-score matrix) [M3 = M1xM2 (refer step 3a in Fig. 3)].

2.) Retained the maximum value for each entry (refer step 3b in Fig. 3).

The same interaction can be generated at different times based on similarities obtained from different pairs, in such occurrences only the maximum value was retained for each entry, so only the predicted interaction with the highest TC value was considered.

3.) Symmetric transformation considering the high value for each pair (refer step 3c in Fig. 3).

The symmetric transformation was carried out to obtain the final M3 matrix, considering the highest value for each pair of drugs (note that the matrix represented in 3b of Fig. 3 is not a symmetric matrix). In the example shown in Fig. 3, interaction 1-2 and 2-3 from M1 were retrieved in M3 with a TC>0.75. Interaction 1-4 was retrieved by the model with a low score (TC=0.3). The model was also predicted the new interaction 3-4 (TC=0.9).

F. Construct Resulting Matrix

Resulting matrix was constructed using the results which were obtained through target similarity matrix and P-score matrix (corresponding 3c matrices which was generated by the model). In the resulting matrix, cell value denotes the maximum TC which was generated as the result (step 3c) by either target similarity matrix or P-score matrix.

G. Output of Final Network

Final output networks were constructed by considering TC values which were indicated in step 3c as a threshold value. Where the threshold in a range of 0 – 0.5 was considered as a low strength interaction, interaction in the range of 0.5 – 0.75 was considered as a moderate strength interaction and finally interaction in the range of 0.75 – 1.0 was considered as a high strength interaction (refer Table 1 for the newly obtained interaction (X) strength ranges).

TABLE I
THRESHOLD VALUES TO MEASURE INTERACTION STRENGTH

Interaction Strength	Range
High	$0.75 \leq X \leq 1.0$
Moderate	$0.5 \leq X < 0.75$
Low	$0 < X < 0.5$

Based on the above threshold values, final output networks were created by considering two different approaches:

- The model was forced to preserve the previously known DDIs.

This approach uses the technique to gain the final output results without sacrificing the previously known DDIs. In this approach, the mechanism is followed to set the interaction to one in the final output matrix, if that interaction exists in the previously known DDIs dataset in 2011, but not retrieved as a predicted interaction. This approach was performed based on the assumption that there is a very less probability that the currently existing interaction may expire in the future.

- The model was not forced to preserve the previously known DDIs.

Without forced to preserve existing DDIs, the approach was performed by assuming that there is a probability that the currently existing interaction may expire in the future. In this approach, the model was freed to sacrifice previously known interactions. Thus, the final output network was solely constructed based on the results of target similarity matrix, P-score value matrix and the resulting matrix based on the threshold values.

Three different final output networks were constructed for each of these two-distinct approaches independently. These networks were generated based on the results of target similarity score matrix, P-score value matrix and the resulting matrix. Therefore, altogether six different final output networks were generated because of this process.

similarity score / P-score. In step 1, interaction matrix M1 was created. Where, the interactions in DrugBank were represented as ‘1’. In step 2, the similarity matrices M2 were created based on the Tanimoto coefficient values and P-score values for two independent instances. In step 3, matrix multiplication (M1xM2) was performed. Maximum value for each entry was retained. The final matrix M3, was formed using the technique symmetry based transformation. In all the matrices values in the diagonals were set to zero, because the interaction of a drug with itself was not considered.

In the final matrix (3c) red colour figures denote interactions, which were previously known and also obtained through the prediction model. Green and blue colour figures represent newly predicted drug interactions. Blue colour figures indicate interactions with higher confidence and green colour interactions can be considered as lower confidence interactions.

H. Evaluation

The accuracy of the model was evaluated by comparing the predicted results from the proposed model for various threshold values against the DrugBank database 2014 snapshot. The overall performance was summarized using the measures: accuracy, precision, recall and F-measure. The receiver operating characteristic (ROC) curve has been generated for more accurate interpretation of model performance. Also, the obtained results from these two approaches (force to preserve existing and without force to preserve existing DDIs) were compared with each other for detail exploration using the above discussed measures.

IV. RESULTS

Approach 01: Results Obtained Through the Model Forced to Preserve the Existing DDIs.

This section explores the results which were obtained by target similarity matrix, P-score matrix and resulting matrix using both evaluation measurement matrices and ROC curves when the model is forced to preserve the existing DDIs. These results were generated by changing the threshold values accordingly. Different threshold values generate different prediction outcomes with different confidence strengths.

A. Evaluation Measurement Matrices

In terms of target similarity matrix (refer Table II A), the accuracy increases along with the threshold value which means the prediction result of a selected DDI always tends to give better accuracy, when the model was executed with a higher threshold value. When considering the precision and recall figures, it shows that the model gives a very higher recall, but comparatively lower precision. The reason for a higher recall was that the model gives a lower number of false negatives. The reason for having such lower precision was due to the higher false positives. Since this is a prediction model having such false positive might be acceptable, because it may be an indication of novel DDIs, that might not experimentally have discovered yet in the current context. Also, both precision and recall figures get increased along with the threshold value. It’s an indication that predicting DDI with a higher confidence level always tends to give better results. F-measure figures also get increased along with the threshold value and this

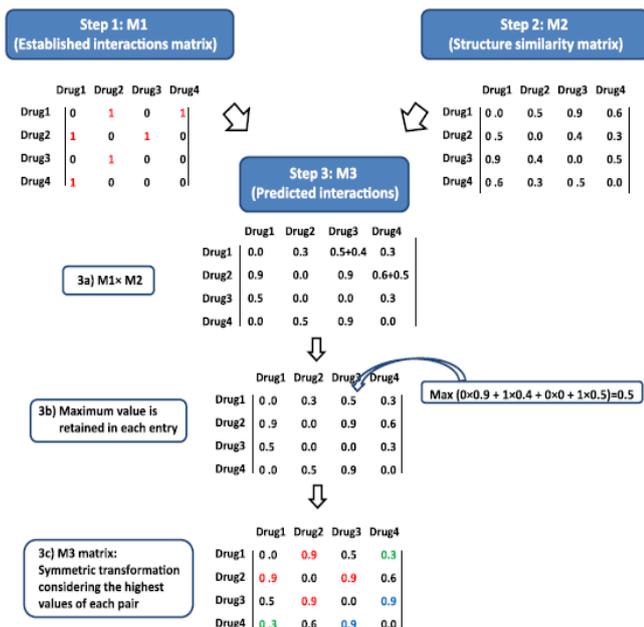


Fig. 3 Steps to predicting DDIs [6]

Figure 1: Process of predicting DDIs is a combined result of the DrugBank interaction database 2011 snapshot and target

measurement indicates the overall impact of precision and recall figures of proposed prediction model.

According to the results of P-score matrix (refer Table II B), the accuracy was slightly increased when the threshold gets increased from 0.25 to 0.5, but it has no impact on the accuracy when threshold gets increased from 0.5 to 0.75. It indicates that the accuracy gets saturated when the threshold reaches up to its maximum. The recall was constant throughout the process. This was due to fixed true positives and false negatives for all threshold values considered. The precision gets increased along with the threshold value and get saturated when the threshold value reaches up to 0.5. F-measure figures were also being in an acceptable level. This explains the scenario both precision and recall values were acceptable.

When considering the output of the resulting matrix (refer Table II C), the accuracy has a direct proportional relationship with the threshold value. Recall slightly decreased and precision figures increased when the threshold value increased. Recall figures were higher compared to the precision. The reason for such higher recalls was due to the lower false negatives. When considering the precision figures, they were still lower even though the threshold reached 0.75. The reason for such lower precision was due to the higher false positives. Having such false positives might acceptable, because it may consist of novel DDIs, that might not be experimentally discovered yet. F-measure figures were also get increased along with the threshold value. It clearly indicates that the model tends to give higher recall and precision values when the threshold values get increased.

TABLE III
EVALUATION MEASUREMENTS FOR ALL MATRICES – APPROACH 01

Force to Preserve Existing DDIs			
A) Target Similarity Matrix			
Evaluation Measurements %	T = 0.25	T = 0.50	T = 0.75
Accuracy	83.00	91.26	96.40
Precision	14.70	25.12	45.00
Recall	99.02	98.91	98.74
F-Measure	25.61	40.04	61.62
B) P-Score Matrix			
Accuracy	98.51	99.89	99.89
Precision	66.80	97.57	97.57
Recall	98.63	98.63	98.63
F-Measure	79.66	98.10	98.10
C) Resulting Matrix			
Accuracy	82.50	91.26	96.40
Precision	14.34	25.12	45.00
Recall	99.02	98.91	98.74
F-Measure	25.06	40.07	61.82

B. ROC Curves

All the coordinates to draw ROC curves for all approaches for the approach of forced to preserve the existing DDIs, were generated by varying the threshold value of the proposed model in a range of 0.1 to 1.00 with a step size of 0.1.

According to the ROC curve (refer Fig. 4 - Blue colour curve), plotted ROC was very much closer to the upper left corner which gives AUC of 0.9792. It clearly indicates that the proposed prediction model gives more valuable results.

With respect to the ROC curve for P-score matrix (refer Fig. 4 - Green colour curve), it has given a very higher true positive rate even from the very beginning (soon after false positive rate is greater than zero). The reason for such behaviour might be due to the very low number of false negatives and the very higher number of true negatives. The obtained curve was well ahead the diagonal line and very much closer to the 1.00 by giving an AUC of 0.9918.

When considering the ROC curve which was plotted for resulting matrix (refer Fig. 4 - Red colour curve), resulting ROC curve gives AUC of 0.9791 which was very much closer to one. It's a good indication of better results of this ROC as well.

When comparing the ROC figures obtained for each matrix type: target similarity, P-score and resulting matrices gives AUC figures as 0.9792, 0.9918 and 0.9791 respectively. This means P-score matrix generates the best results in terms of AUC. The target similarity matrix and resulting matrix gave almost similar AUC figures. Thus, there is no considerable impact on integrating the results of the two types of matrices to have a better prediction result, because it has not given a higher AUC value in the resulting matrix as expected for the force to preserve existing DDIs approach.

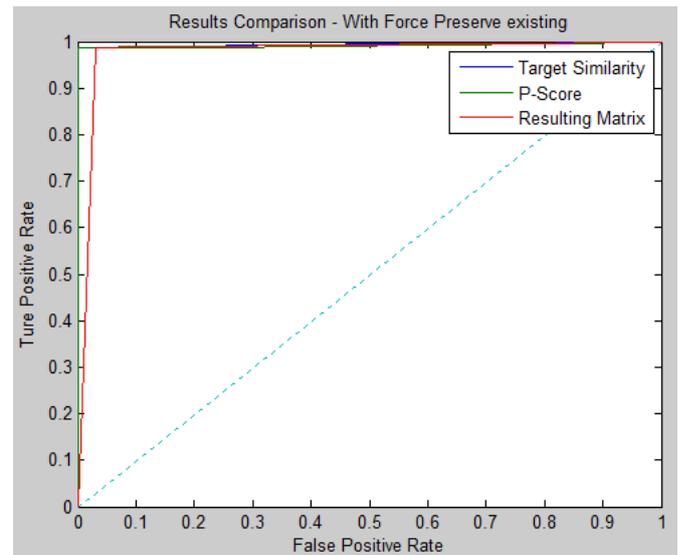


Fig. 4 ROC Curves for all matrices - Approach 01

Approach 02: Results Obtained Through the Model Without Preserving the Existing DDIs.

This section explores the results for the same set of matrix types: target similarity matrix, P-score matrix and resulting matrix, using a different approach named without force to preserve the existing DDIs. As same as the previous approach, this approach also addresses the evaluation methodology using two fundamental evaluation techniques named evaluation measurement matrices and ROC curves.

A. EVALUATION MEASUREMENT MATRICES

In terms of target similarity matrix (refer Table III A), the accuracy increased along with the threshold value. It's an indication that the prediction result of a selected DDI always

tends to give better accuracy, when the model was executed with a higher threshold value. In terms of precision and recall figures it clearly indicates that the proposed model tends to generate higher precision when threshold value increased, but recall figures decreased when threshold values increased. The reason for such behaviour was due to the reduction of false positives when the threshold value increased. The proposed model generates lower recall figures due to higher false negatives when the threshold value increased. Even though the threshold value was significantly higher, precision figures were not improved up to the expected level (precision was 29.37 even the threshold reached 0.75). The reason for such behaviour was, due to comparatively higher false positives even though the threshold reached a higher value. In terms of a prediction model, having such false positive might be acceptable because it might denote the novel DDIs that might not have experimentally discovered yet. Eventually, the F-measure increased along with the threshold value, but could not fully satisfy with the increased amount, because F-measure was 36.88 even when the threshold is 0.75. The reason for such behaviour was due to the slower growth of precision and decrement of recall when threshold increased.

According to the results of P-score matrix (refer Table III B), the accuracy was slightly increased when threshold increased from 0.25 to 0.5. But there was a slight reduction in accuracy when threshold increased from 0.5 to 0.75. It implies that the accuracy reaches its peak value and then decreases continuously when the threshold increases further in a range of 0.5 to 0.75. When considering the precision and recall figures, precision increases rapidly and reaches its maximum, while recall decreases rapidly. Ultimately, recall ended up with a very low figure (0.16) when threshold value further increased. The recall figures rapidly reduced due to the sudden decrease in true positives and sudden increase in false negatives. Precision values also show a rapid increase due to the fact of rapid reduction in false positive values. F-measure figures also rapidly decrease when the threshold increases. This behaviour was due to the rapid reduction of recall and precision figures. Thus, we can be satisfied with the results obtained in terms of accuracy and precision, but not with the recall and F-measure figures.

When analysing the results obtained from resulting matrix (refer Table III C), the accuracy has a direct proportional relationship with the threshold value. Recall was decreased and precision increased when the threshold value gets increased. Even though the recall figures get decreased, the figures, which were still higher, were being compared to precision values. The reason for such behaviour was due to the fact of sudden increment of false negatives when the threshold gets increased. The decrement of the false positives was the reason behind the increment of precision figures. F-measure figures were increased from 23.74 to 36.98 when the threshold get increased in a range of 0.25 to 0.75 because of corresponding precision and recall figures.

B. ROC Curves

As shown in the ROC curve (refer Fig. 3 - Blue colour curve), it covers AUC of 0.9030. It's a massive AUC while taking entire portion into the action, because there was only 0.097 (1.00 – 0.9030) area yet to be covered.

With respect to the ROC curve for P-score matrix (refer Fig. 5 - Green colour curve), it has the lowest AUC of 0.7708. Also, this ROC curve does not seem to be smooth like other

TABLE IIIII
EVALUATION MEASUREMENTS FOR ALL MATRICES – APPROACH 02

Without Force to Preserve Existing DDIs			
A) Target Similarity Matrix			
Evaluation Measurements %	T = 0.25	T = 0.50	T = 0.75
Accuracy	82.73	90.57	94.99
Precision	13.45	20.20	29.37
Recall	89.18	74.37	49.56
F-Measure	23.38	31.77	36.88
B) P-Score Matrix			
Accuracy	96.77	97.25	97.05
Precision	44.48	99.23	100.00
Recall	37.65	7.05	0.16
F-Measure	40.78	13.16	0.33
C) Resulting Matrix			
Accuracy	82.34	90.59	94.99
Precision	13.61	20.38	29.44
Recall	93.06	75.19	49.73
F-Measure	23.74	32.07	36.98

ROC curves. The reason for such behaviour was that it does not have an intermediate false positive rate coordinate in the range of 0.2079 to 1.00. Also, it does not have a true intermediate positive rate coordinate in the range of 0.6656 to 1.00 for the selected threshold values. This ROC was still acceptable, because still it is above the diagonal line and 0.2292 (1.00 – 0.7708) of AUC is yet to be covered.

The ROC curve which has plotted for resulting matrix (refer Fig. 5 - Red colour curve), gives AUC of 0.9215 which is almost closer to the one (0.0785 of AOC is yet to be covered). It's a good indication that this model gives more accurate results.

When comparing the ROC figures obtained for each matrix type: target similarity, P-score and resulting matrices gives AUC figures as 0.9030, 0.7708 and 0.9215 respectively. The P-score matrix has the lowest AUC among the available three types of matrices. The resulting matrix gives the best results in terms of AUC. Thus, the integration of target similarity matrix and the P-score matrix give more accurate results rather than considering them individually in this approach. Therefore, it is a good indication that the combination of more similar DDIs knowledge can be used to improve the accuracy of prediction results in the approach of without force to preserve existing DDIs.

Compare the Results Obtained through Approaches.

This section compares the results given by the mentioned two fundamental approaches used in this study: Force to preserve existing DDIs and without force to preserve existing DDIs. For each of these two approaches, we have discussed the effect of using three techniques: target similarity matrix, P-score matrix and resulting matrix using evaluation measurement matrices and ROC curves.

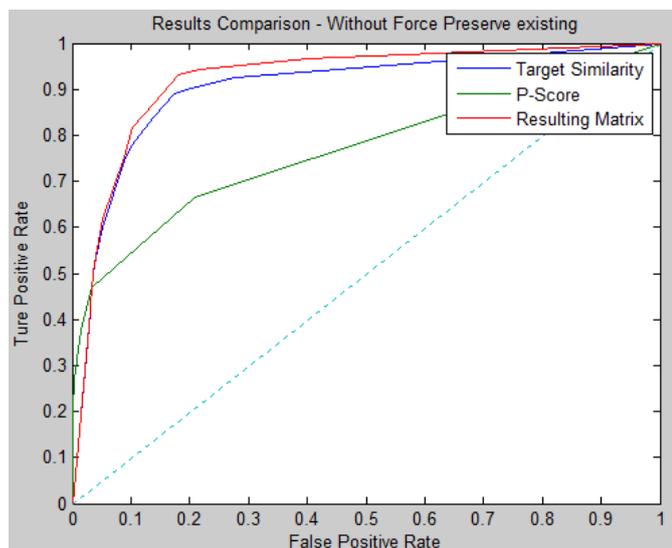


Fig. 5 ROC Curves for all matrices - Approach 02

A. Evaluation Measurement Matrices

According to the summarization of both approaches (refer Table II & III), the Forced to preserve existing approach performs better compared to the results which was given by the target similarity matrix in terms of accuracy, precision, recall and F-measure. In terms of accuracy, the approach 01 does not have a considerable amount of improvement, But when it comes to precision and recall approach 01 performs better especially when the threshold values were high.

When considering about the evaluation measurements given by P-score matrices for both approaches, 'Force to preserve existing DDIs' approach performs better compared to 'Without force to preserve approach' for all the evaluation measurements except the precision. There is no higher variation between accuracy figures, but there is a considerable amount of variation when it comes to recall and F-measure figures especially when the threshold values are higher. In terms of precision values, approach 02 gives better results compared to approach 01 when the threshold value gets increased.

According to the results obtained through resulting matrix, 'Forced to preserve existing DDIs' approach performs better for accuracy, precision, recall and F-measure compared to 'Without force to preserve existing DDIs' approach. Thus, in overall performance, approach 01 is better compared to second approach with respect to the results given by evaluation measurement matrices.

B. ROC Curves

When comparing the results given by 'Force to preserve existing DDIs' approach, the ROC curve drawn for P-score covers more area, where target similarity and resulting matrices cover almost similar amount of area (refer Fig. 4). When considering the exact AUC figures for 'Force to preserve existing DDIs' approach (refer Table IV), it indicates that the target similarity matrix has an AUC of 0.9792, P-score matrix has an AUC of 0.9918 and resulting matrix has an AUC of 0.9791. Thus, for the approach 01 P-score matrix technique performs better.

When comparing the results given by 'Without force to preserve existing DDIs' approach, it gives higher AUC for

the resulting matrix (refer Fig. 5) compared to target similarity and P-score matrix results. When considering the exact AUC figures for 'Without force to preserve existing DDIs' approach (refer Table IV), it indicates that the target similarity matrix has an AUC of 0.9030, P-score matrix has AUC of 0.7708 and resulting matrix has AUC of 0.9215. This implies that the combination of multiple results tends to give better results with the 'Without force to preserve existing DDIs' approach. Thus, the combination of multiple techniques gives more accurate results according to this approach.

TABLE IV
AUC FIGURES FOR APPROACH 1 AND APPROACH 2

		Matrix Type		
		Target Similarity Matrix	P-score Matrix	Resulting Matrix
AUC	Approach 1	0.9792	0.9918	0.9791
	Approach 2	0.9030	0.7708	0.9215

V. DISCUSSION

Different types of models have been published recently for predicting DDIs. These models are mainly based on two fundamental techniques named: knowledge based and similarity based. Often knowledge based approaches are focused on the post market stage. The similarity based approaches are focused on the pre-market stage and the initial part of the post-market stage is in the drug development process. However, both techniques suffer from several limitations such as the necessity to distinguish drug classes and inability to handle novel drugs for which limited reports exist [12].

We propose a large-scale method based on identifying target and side effect similarity to predict multiple types of DDIs. The model described in this article can exploit experimental knowledge to identify the possible causes for an interaction. The proposed model potentiates a visible pattern in the DrugBank database (similar drugs have similar interactions) by detecting drugs similar to the drugs implicated in the interactions described previously. Therefore, one limitation of this study is that the performance of the model depends on the comprehensiveness of the information in the original interaction database.

The proposed model only determines whether a given pair of drug is going to interact or not. It will determine the category of an interaction (whether the resulted interaction is harmful or harmless). Also, the proposed model does not consider the method of administration of a drug (e.g., Taken orally, injected or applied externally, such as an ointment or ophthalmic drops). According to the biology, the outcome can be varied depending on the method of administration. The effect changes due to the dose changes were not addressed in this research. Thus, such kind of facts can be further addressed as a future work of this study.

Although the similarity model provides valuable information associated with the initial interactions, a more reliable and complex system could be implemented through the integration of: structural similarity measures and knowledge in pharmacological databases containing

information about possible molecular structural similarity [6] and interaction profile fingerprints (IPFs) [5]. The proposed method could also be combined with other methodologies using different types of information, such as the Food and Drug Administration's Adverse Event Reporting System [2][11][13], which was created to provide post-marketing drug safety information, or the use of clinical data in electronic health records [4].

Several DDIs highlighted in our methodology was not recognized in the earlier stage and consequently it has considered as false positives in our evaluation. However, there is a possibility that some of these drugs do interact with each other, but not yet been identified. Therefore, it is possible to have a lower false positive rate than estimated.

VI. CONCLUSION

The results presented in this study demonstrate the usefulness of the proposed network based on drug-drug interaction methodology as a promising approach. The method described in this article is very simple, efficient, but biologically sound. This study addresses the problem of predicting novel DDIs using two different approaches. Under each of these approaches, we have analyzed the results using three different matrices. When concluding the prediction results, 'Force to preserve existing DDIs' approach performs better compared to 'Without force to preserving' approach. However, there were 3 700 experimentally discovered DDIs in the DrugBank database 2011 snapshot and in 2014 amount has been reduced up to 3 660. Thus, there is a reduction of forty interactions (3700 – 3660) along with the time. This does not indicate the fact that there were no new interactions discovered within those three years. It is an indication that there is a probability where the previously interact drug pairs were removed from the interaction drug pair list (ex: - let's assume that 10 new interactions were discovered within the time period and 14 of previously known interactions were not going to consider as a successful interaction. Furthermore, in such cases, the ultimate output looks as there is a reduction of DDIs).

In all approaches that we have considered, the accuracy gets increased along with the threshold value which means that there is a higher possibility of generating better prediction results when a higher threshold value was input to the proposed model. When considering about the obtained results through 'Without force to preserving existing DDIs', resulting matrix gives highest AUC, when compared to target similarity and P-score matrices (refer Fig. 3). In terms of exact figures, resulting matrix gives an AUC of 0.9215, but target similarity matrix and P-score matrix gives an AUC of 0.9030 and 0.7708 respectively. It shows that the integration of different similarity measures tends to give more accurate results rather than using them individually when it comes to without force to preserve existing (known) DDIs approach.

ACKNOWLEDGMENT

I would like to acknowledge Dr. Jialiang Huang who is a member of Dana-Farber Cancer Institute, Harvard Medical School and Boston Children's Hospital for providing me with some of the required data for my research. My thanks also go to Dr. Santiago Vilar who is the postdoctoral research scientist at the Columbia University for the support and technical guidance provided.

I also greatly appreciate the joyous nature, comments and encouragements provided by my friends and seniors, which have been a great strength to me. Finally, my thanks go to my loving parents whose assistance was tremendously helpful in completing my undergraduate studies, help throughout my whole life and especially this research, to the best of my ability.

REFERENCES

- [1] A. Cami, S. Manzi, A. Arnold, and B. Y. Reis, "Pharmacointeraction Network Models Predict Unknown Drug-Drug Interactions," *PLoS Comput. Biol.*, vol. 8, no. 4, 2013.
- [2] N. P. Tatonetti, G. H. Fernald, and R. B. Altman, "A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports," *Jamia*, 2011.
- [3] J. Huang, C. Niu, C. D. Green, L. Yang, H. Mei, and J.-D. J. Han, "Systematic prediction of pharmacodynamic drug-drug interactions through protein-protein interaction network," *PLoS Comput. Biol.*, vol. 9, no. 3, p. e1002998, Jan. 2013.
- [4] J. Pathak, R. Kiefer, and C. Chute, "Using Linked Data for Mining Drug-Drug Interactions in Electronic Health Records," *Stud. Heal. Technol. . . .*, pp. 682–686, 2013.
- [5] S. Vilar, E. Uriarte, L. Santana, N. P. Tatonetti, and C. Friedman, "Detection of Drug-Drug Interactions by Modeling Interaction Profile Fingerprints," *PLoS Comput. Biol.*, vol. 8, no. 3, 2013.
- [6] S. Vilar, R. Harpaz, E. Uriarte, L. Santana, R. Rabadan, and C. Friedman, "Drug - drug interaction through molecular structure similarity analysis," *Jamia*, 2012.
- [7] R. L. D. and S. P. M. Arthur J. Atkinson, Darrell R. Abernethy, Charles E. Daniels, *Principles of Clinical Pharmacology*, Second Edi. 2012.
- [8] van Puijenbroek EP, Egberts ACG, Heerdink ER, Leufkens HGM (2010), "Detecting drug-drug interactions using a database for spontaneous adverse drug reactions: an example with diuretics and non-steroidal anti-inflammatory drugs." *Pub Med.*, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11214785>.
- [9] Almenoff JS, DuMouchel W, Kindman LA, Yang XH, Fram D (2003), "Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting," *Sci. Transl. Med.*
- [10] Noren GN, Sundberg R, Bate A, Edwards IR (2008), "A statistical methodology for drug-drug interaction surveillance," *Pub Med.*, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/18344185>.
- [11] G. K. et al. NP Tatonetti, JC Denny, SN Murphy, GH Fernald, "Detecting Drug Interactions From Adverse-Event Reports," *PMC*, vol. 90, no. 1, pp. 133–142, 2011.
- [12] A. Gottlieb, G. Y. Stein, Y. Oron, E. Ruppim, and R. Sharan, "INDI : a computational framework for inferring drug interactions and their associated recommendations," *Mol. Syst. Biol.*, vol. 8, no. 592, pp. 1–12, 2012.
- [13] N. P. Tatonetti, P. P. Ye, R. Daneshjou, and R. B. Altman, "Data-driven prediction of drug effects and interactions," *Sci. Transl. Med.*, vol. 4, no. 125, p. 125ra31, Mar. 2012.
- [14] A. C. et Al., "Predicting Adverse Drug Events Using Pharmacological Network Models Aurel Cami et al .," vol. 127, 2011.
- [15] T. Arikuma, S. Yoshikawa, R. Azuma, K. Watanabe, K. Matsumura, and A. Konagaya, "Drug interaction prediction using ontology-driven hypothetical assertion framework for pathway generation followed by numerical simulation," *BMC Bioinformatics*, vol. 9 Suppl 6, p. S11, Jan. 2008.
- [16] "Jaccard index," [Online] 1999, http://en.wikipedia.org/wiki/Jaccard_index (Accessed: 20 August 2014).
- [17] M. Campillos, M. Kuhn, A.-C. Gavin, L. J. Jensen, and P. Bork, "Drug target identification using side-effect similarity," *Science*, vol. 321, no. 5886, pp. 263–6, Jul. 2008.
- [18] J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J.-P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. a Armstrong, S. J. Haggarty, P. a Clemons, R. Wei, S. a Carr, E. S. Lander, and T. R. Golub, "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease," *Science*, vol. 313, no. 5795, pp. 1929–35, Sep. 2006.
- [19] S. I. Berger and R. Iyengar, "Network analyses in systems pharmacology," *Bioinformatics*, vol. 25, no. 19, pp. 2466–72, Oct. 2009.

- [20] R. Guimerà and M. Sales-Pardo, "A network inference method for large-scale unsupervised identification of novel drug-drug interactions.," *PLoS Comput. Biol.*, vol. 9, no. 12, p. e1003374, Dec. 2013.
- [21] R. Guimerà and M. Sales-pardo, "Missing and spurious interactions and the reconstruction of complex networks," *PNAS*, vol. 106, no. 52, 2009.
- [22] Newman MEJ (2011), "Communities, modules and large-scale structure in networks.," *Nat Phys* 8: 25–31.
- [23] Eckert H, Bajorath J (2006), "Design and evaluation of a novel class-directed 2D fingerprint to search for structurally diverse active compounds.," *J Chem Inf Model* 46: 2515–2526.
- [24] X.-M. Zhao, M. Iskar, G. Zeller, M. Kuhn, V. van Noort, and P. Bork, "Prediction of drug combinations by integrating molecular and pharmacological data.," *PLoS Comput. Biol.*, vol. 7, no. 12, p. e1002323, Dec. 2011.
- [25] Martin YC, Kofron JL, Traphagen LM. "Do structurally similar molecules have similar biological activity?" *J Med Chem* 2002;45:4350e8.
- [26] Steffen A, Kogej T, Tyrchan C, et al. "Comparison of molecular fingerprint methods on the basis of biological profile data.," *J Chem Inf Model* 2009;49:338e47.
- [27] "Jaccard index," [Online] 1999, http://en.wikipedia.org/wiki/Jaccard_index (Accessed: 20 August 2017)
- [28] "Bayesian inference," [Online] 2010, http://en.wikipedia.org/wiki/Bayesian_inference (Accessed: 20 August 2017).