

# DISEASE AND MEDICATION NETWORKS: AN INSIGHT INTO DISEASE-DRUG INTERACTIONS

Completed Paper

**Dipanwita Dasgupta**

Department of Computer Science &  
Engineering  
Interdisciplinary Center for Network  
Science and Applications (iCeNSA)  
University of Notre Dame  
Notre Dame, Indiana, USA  
ddasgupt@nd.edu

**Nitesh V. Chawla**

Department of Computer Science &  
Engineering  
Interdisciplinary Center for Network  
Science and Applications (iCeNSA)  
University of Notre Dame  
Notre Dame, Indiana, USA  
nchawla@nd.edu

## Abstract

*The Electronic Health Records (EHRs) present a unique opportunity to longitudinally analyze and model disease progressions and interactions at a population level. EHRs can be used to construct various disease-disease and drug-drug networks for studying their interactions. Additionally, one can also construct rich disease-drug networks, which indicate the prescriptions that are assigned to treat those diseases. This analysis can illuminate not only the side-effects, but also potential candidates for disease targets. In this paper, we employ a network-based approach for studying disease-drug interactions. We construct an integrated bipartite network where the nodes are drugs and diagnoses. We found that the integrated network provides more information on these interactions, compared to the component or parent networks.*

## Introduction

The Electronic Health Records (EHRs) offer a compelling resource to longitudinally study the progression and co-occurrence of diseases at a patient level, as well as a population health level. Analytics performed on the EHRs can provide insight into the disease co-morbidities as well as cures recommended for the diseases (prescriptions or procedures). We leverage the richness of the EHRs to discover disease-disease interactions, as well as disease-drug (prescription) interactions. Our work can not only be used for tracking potential side effects but also, discovering new drug targets. During the drug development process, the efficacy of the drug is determined. In addition to the efficacy testing, tests are conducted to report the interaction with other drugs or proteins (Adams & Brantner, 2006; DiMasi, 2001). Clinical trials are conducted to test the efficacy and effectiveness of a drug on humans. The efficacy of the drug is an indicator of how well the drug performs in the trial, whereas efficiency indicates the performance of the

drug, even outside the trial. Moreover, leveraging these disease-disease and disease-drug interactions over time can, also, provide useful information towards personalized healthcare, as the disease and drug information can be potentially demonstrative of the socio-economic indicators of health (Marmot et al., 1991).

The medical history of an individual gives the list of reported health problems, confirmed principal diagnoses, diagnoses arising from side effects of drugs, allergies to drugs, or from co-morbid diagnoses. Looking at the drugs prescribed and diseases or conditions such as allergy diagnosed, one can infer the efficacy and efficiency of a drug in a real setting. We approach this work by constructing disease-disease networks to get a richer representation of the data.

The EHRs have been used extensively to build phenotypic networks of diseases (Hidalgo et al., 2009; Steinhäuser & Chawla, 2009). Network based approaches are used to study the genetic associations of diseases, pharmacological and pharmacokinetics of drugs, their interactions (Goh et al., 2007). Drug interactions have primarily been studied at the genetic or the protein level. However, studies show that phenotypic network of diseases is denser when compared to the genetic network of diseases (Davis & Chawla, 2011), and can be used to identify potential targets for genetic testing.

Similarly, we posit that as a drug targets diseases, it might be useful to leverage the richer phenotypic network as a guideline for finding possible side effects. This does not, however, eliminate the need for experimental verification. Nonetheless, it can be used to narrow down the number of options for which experiments have to be performed. This reduction also lowers the overall cost associated with drug development, as well as the market price of the drug. In this paper, we propose a network-based perspective to the problem. We derive drug-drug and disease-disease networks from the EHR data based on co-occurrence information. We, also, develop an integrated view of the drug-drug and disease-disease network to illuminate potential side-effects.

## Related Work

Network-based tools for analyzing protein-protein interactions, gene co-expression networks, and metabolic networks, in humans have been used for over a long period of time (Stelzl et al., 2005). Some of these tools use gene-expression data for building the network. The nodes in the networks are genes/diagnoses. These nodes are connected if the correlation (calculated using a distance metric) between the gene expression profiles of the nodes are above a certain threshold. These tools are aimed at analyzing the interactions between component inside human body or model organisms. Many of these tools use data for model organisms like yeast and transfer the knowledge gained from these studies to humans (Uetz et al., 2000).

There are multiple ways of modeling the disease-disease network. These techniques differ in the construction of the links or edges between the nodes. One such approach analyzes the disease-gene interactions as a bipartite network to identify the genes that are responsible for certain diagnoses (Goh et al., 2007). Projections from the bipartite network were made to form human disease network and human gene network. These projection networks were analyzed to make predictions for target genes of diseases. The co-morbidity of a disease depends on both phenotype and genotype (Schadt, 2009). After the successful completion of the human genome project, most network-based tools used some form of genotypic data for modeling disease network. Recent advances in disease modeling include building the phenotypic disease network where diseases (ICD-9-CM codes) were nodes. Edges indicated the co-morbidity of diseases, defined by correlation metric (Hidalgo et al., 2009; Steinhäuser & Chawla, 2009).

Networks-based approaches have, also, been extensively used for drug related studies. One approach is similar to that mentioned in (Goh et al., 2007). The only difference is that in this case, nodes represent drug and target proteins (Hidalgo et al., 2009). The analysis of this network revealed that most of the drugs were targeted towards a single protein. However, the trend is now changing towards multi-target

drugs, where a single drug is used for treating multiple diseases. The concept of multi-target drugs was, also, formulated using networks and is based on the fact that every drug was connected at least one other drug (Lu, Pan, Hu, & Wang, 2012). The networks, in this case, were drug-drug and target-target (protein-protein). All the drug-drug or drug-target networks are based on the genomic data. Neighbors of a node in the network are used to model drug interactions and find drug side-effect similarities (Brouwers, Iskar, Zeller, van Noort, & Bork, 2011; Ye, Liu, & Wei, 2014).

## Data

Our database comprises 543,571 de-identified disease/diagnosis records of 26,843 patients, spanning from the year 1999 to 2012. Each patient record in the Diagnosis database is represented by an encounter ID, patient ID and diagnosis codes represented by *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* ("International Classification of Diseases, 9th Revision, Clinical Modification(ICD-9-CM)", 2007). ICD-9-CM is the official system of assigning codes to diagnoses and procedures associated with hospital utilization in the United States, and are published by World Health Organization. The ICD-9-CM codes are 3-5 digits and have a hierarchical structure; wherein each 5-digit code is a subset of a 4-digit code, which is, further, a subset of a 3-digit code. For example, the code 29011 corresponds to Presenile dementia with delirium, which can be collapsed to 2901, a generic presenile dementia diagnosis, and further to 290, a family of dementia diagnoses. The number of diagnoses per patient record ranges from 1 to 15, with a median of 1. The highest occurring diagnoses in our database are 724(Back pain), 401(hypertension), 272(hyperlipidemia).

Another database comprises 315,437 medication records of 11,230 patients. The medications included both prescription drugs and procedures. There are 10,964 patients, which have medication records containing prescription drugs. Each record has a patient ID and an encounter ID. We only considered those records which contained prescription drugs. The number of medications per patient record ranges from 1 to 38, with a median of 2. The highest occurring medications in our database are acetaminophen-hydrocodone, acetaminophen-oxycodone, albuterol, gabapentin.

Table 1 shows a sample patient diagnosis records. Each line in the table represents a visit. Table 2 shows a sample patient medication history. Each line in the table represents a visit.

As these records came from a conglomeration of different databases, not every diagnosis had a corresponding prescription and vice versa. Based on the encounter IDs present in both databases, there are diagnosis/drug records of 4,629 patients. We only used these patients in our analysis. The diagnoses for each patient were collapsed to the leading 3-digit codes because the leading 3-digits represent the general diagnosis, whereas the rest of the digits provide the specific details. Some of the diagnosis codes were missing but had textual information describing the diagnosis. We matched the text with ICD-9-CM, for most of the cases. We ignored those entries for which the ICD-9-CM codes were missing, and the textual information could not be used to find the codes. Obviously, some inherent noise remains in the dataset, due to missing diagnoses, incorrect matching, wrongly entered information. Finally, we created a dataset with 23,365 diagnosis-drug pairs, which include 905 unique drugs and 623 unique diagnoses (3-digit ICD-9-CM codes).

<b>Table 1.A sample diagnosis history of a patient</b>		
Patient ID	Encounter ID	Diagnosis Code(ICD-9-CM)
10134	1246	40110
10134	1246	272
10134	1458	72401

<b>Table 2.A sample medication history of a patient</b>		
Patient ID	Encounter ID	Medication//Drug
10134	1246	acetaminophen
10134	1246	amoxicillin
10134	1458	aspirin

## **Results and Discussions**

### ***Network Construction***

We built a diagnosis phenotypic network (DD) from the filtered Disease/Diagnosis data with diagnoses as nodes. Figure 1 shows the network. An edge in this network represents the statistically significant co-occurrence of diagnoses in a patient's diagnosis record. We determined statistical significance using a one-tail two-proportion z-test with 95 percent confidence (Davis & Chawla, 2011). From the diagnosis' perspective, statistical significance means that the joint probability of diagnosis-pairs (end-nodes of an edge) is statistically significant greater than the product of the individual probabilities of the two diagnoses. The probability of a diagnosis reflects the prevalence of the diagnosis in the entire population (Diagnosis database). We included those diagnoses for building the network, which had more than 2 occurrences in the entire population. We built an undirected network as it is unclear from the data which diagnoses led to the other. The network is weighted where the weights denote the number of times that a diagnosis pair occurs in the filtered dataset. DD has unique 127 nodes (diagnoses) and 157 edges (diagnosis co-morbidity).

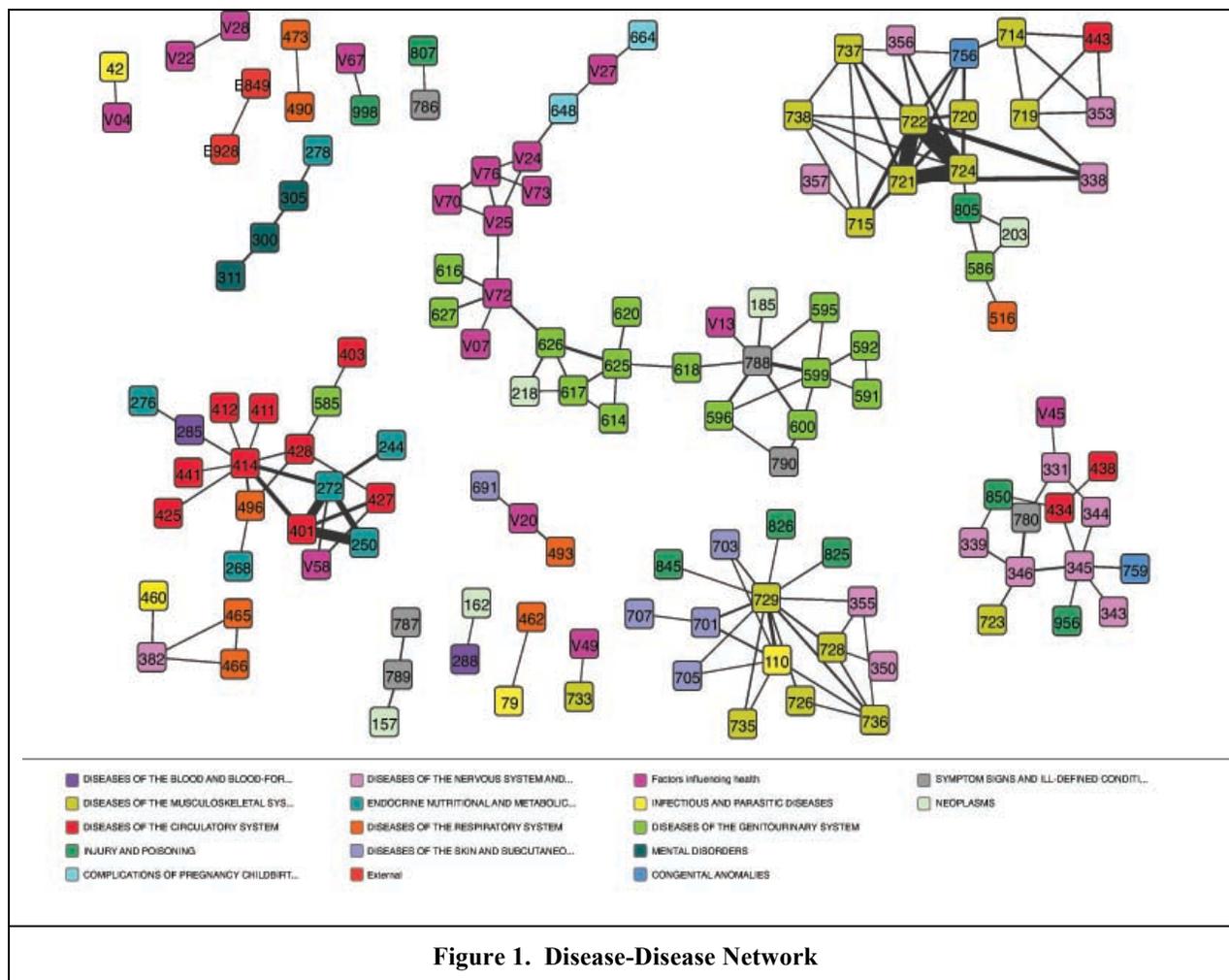
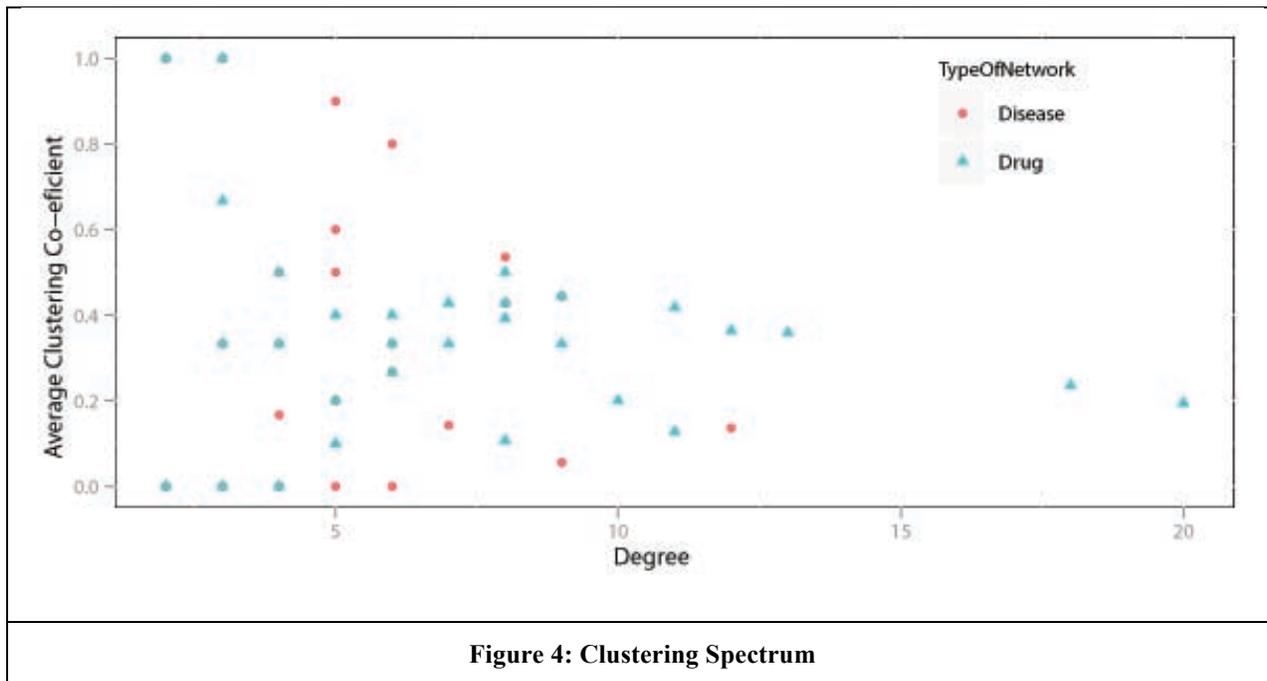
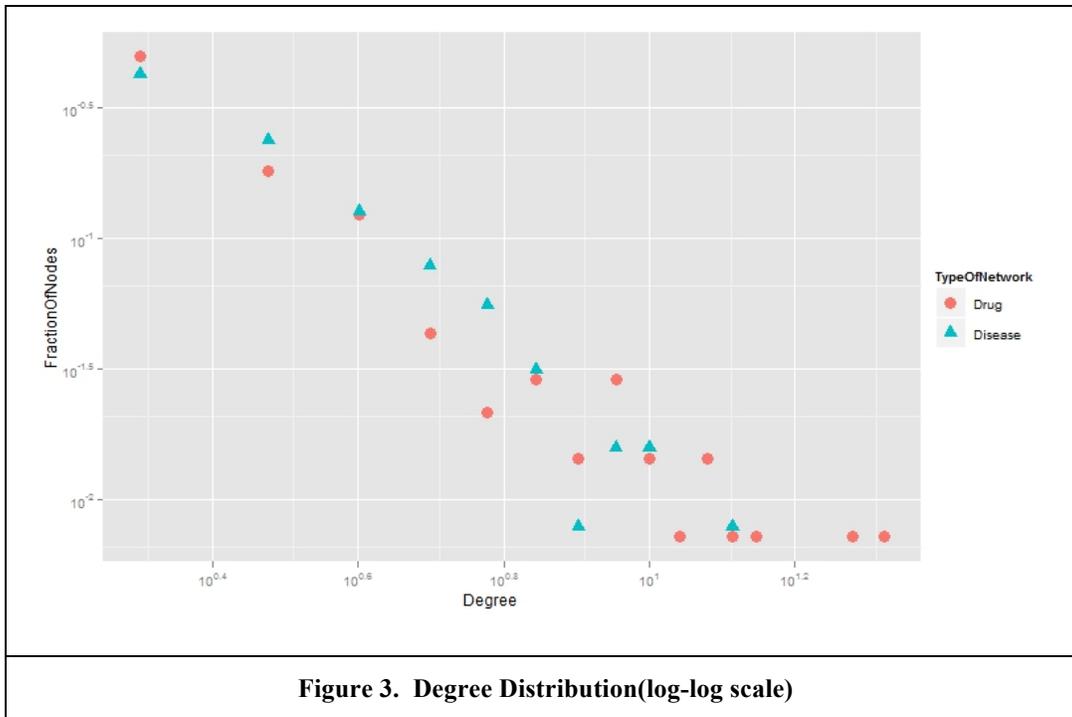


Figure 1. Disease-Disease Network

We also constructed a drug(or medication) network (MM) where the nodes are drugs or medications prescribed. To distinguish from DD (Disease-Disease) network, we will use the nomenclature of medication-medication networks here (MM). However, we will use medication and drug interchangeably in the paper. Figure 2 shows the MM network. Two drugs share a link (edge) if the drug-pair occurs a statistically significant number of times in the entire population. Statistical significance is calculated in a similar fashion as in DD by one-tail two-proportion z-test with 95 percent confidence. In MM, statistical significance indicates that the joint occurrence of two drugs is significantly greater than that of individual occurrence of each drug (product of marginal probabilities). We omitted those drug-pairs that had less than 3 occurrences in the filtered dataset. MM has 139 unique nodes (drugs) and 196 unique edges. MM is, also, an undirected and weighted network where the weights represent the number of times that a drug pair occurs in the filtered dataset (discussed in Section **Data**).





The networks were visualized using Cytoscape (Smoot, et al., 2011). We applied a weighted force-directed layout to get the connected components. The weight used by the layout represents the number of occurrences of each link (edge or node-pair) in the network. The thickness of each edge in the network represents the weight, relative to the maximum weight in the network. We used Cytoscape as a proof of concept. Advanced graph or network algorithms can be used to find the connected components or the clusters.

## Disease-Disease Network

There are 18 connected components in the diagnosis-diagnosis network (DD). One of the connected components suggests a strong connection between hypertension (ICD-9-CM 401) and hyperlipidemia (ICD-9-CM 272) i.e. there is high co-morbidity between them. It is also known that hypertension increases the risk of hyperlipidemia (high cholesterol) (Ames, 1991). Hyperlipidemia is strongly associated with acquired hypothyroidism (ICD-9-CM 244) because the treatment of acquired hypothyroidism reduces the levels of cholesterol (O'Brien et al. 1993). Further, the same connected component shows a strong association between hypertension and diabetes (ICD-9-CM 250). Studies have shown that hypertension is one of the major risk factors for diabetes (Grossman & Messerli, 2008; Kaplan, 2002). Moreover, according to the 2011 National Diabetes Fact Sheet, published by Centers for Disease Control and Prevention (CDC), 67 percent of the adults, aged 20 years or older with diabetes, had hypertension (for Disease Control, Prevention (CDC), Centers for Disease Control, Prevention (CDC), et al., 2011). The other neighbors of hypertension are coronary heart disease (ICD-9-CM 414) and cardiac dysrhythmias (ICD-9-CM 427). According to the Framingham Study, hypertension was identified as one of the primary causes for coronary heart disease (Castelli, 1984). Most of the diagnosis, related to heart disease, is present in this component of the network. Even though Chronic kidney disease (CKD)(ICD-9-CM 585) belongs to a separate category of ICD codes, compared to that of majority of the diagnoses present in this component, CKD is closely associated with heart failure(ICD-9-CM 428) and, in fact, can independently predict the risk of mortality and hospitalization for patients diagnosed with heart failure (Go et al., 2006).

The strongest link in the diagnosis-diagnosis network is between “Other and unspecified disorders of back” (ICD-9-CM724) and “Intervertebral disc disorders” (ICD-9-CM 722). The strongest link indicates the maximum number of occurrences of the diagnosis-pair in the entire dataset (maximum weight of an edge in the network). The most commonly occurring diagnosis in the filtered dataset is Backache (ICD-9-CM 724.5). This disease code (ICD-9-CM 724.5) can be collapsed to less-specific ICD-9-CM code, 724 (“Other and unspecified disorders of back”). Backache (Low back pain) is, indeed, a commonly occurring problem. Studies have shown that the common anatomical structures that are sites for back pain include bones, intervertebral discs, and ligaments (Hoy, Brooks, Blyth, & Buchbinder, 2010; Atlas & Deyo, 2001). “Intervertebral disc disorders” (ICD-9-CM 722) also shares an edge with “Pain, not elsewhere classified” (ICD-9-CM 338), even though they are not classified in the same category.

A connected component of the diagnosis-diagnosis network contains gynecological diagnoses and their associated procedures. “Special investigations and examinations” (ICD-9-CM V72) is one of the centers for this connected component. It has a specific sub-category “Gynecological examination” (ICD-9-CM V72.3). This connected component, also, contains diagnoses related to urinary system. Some of these diagnoses included “Symptoms involving urinary system” (ICD-9-CM 788), and “Other disorders of urethra and urinary tract” (ICD-9-CM 599).

## Drug-Drug Network

The drug-drug (MM) network has 17 components. A connected component corresponds to a cluster containing anti-hypertensive drugs, anti-hyperlipidemic drugs, and anti-diabetic drugs. Some of the anti-hypertensive drugs in the component are lisinopril, hydrochlorothiazide, amlodipine, metoprolol, and carvedilol. Some of the anti-hyperlipidemic drugs include simvastatin, and atorvastatin. As stated in subsection **Disease-Disease Network**, there is a strong association between hypertension, hyperlipidemia and diabetes. For example, simvastatin that treats high cholesterol (hyperlipidemia) has a large number of anti-hypertensive drugs (like amlodipine, and lisinopril) as neighbors in MM. Furthermore, the simvastatin node coexists with levothyroxine that is used in the treatment of hypothyroidism. However, according to the drug usage for levothyroxine, there is a high risk associated with levothyroxine and hyperlipidemia, hypertension (Drugs, 2011). Ideally, precautionary measures should be taken before prescribing levothyroxine for treatment of

thyroidism to patients with hyperlipidimia, and hypertension.

Similar to the diagnosis-diagnosis network, the drug-drug network, also, has a connected component containing pain medications. Some of the prevalent pain medications in the component are acetaminophen-hydrocodone, gabapentin, oxycodone, and pregabalin. The highest occurring link in the network is between acetaminophen-hydrocodone and gabapentin. This increases the risk of drug abuse as acetaminophen-hydrocodone is a Schedule II painkiller. According to the Controlled Substance Act (Title II of the Comprehensive Drug Abuse Prevention and Control Act of 1970) enacted by the Congress in the United States, the most commonly drugs used for abuse are Schedule II drugs: opioids and oxycodone (Administration, of Justice, & of America, 1978). Schedule II drugs have high potential for abuse, though accepted for medical use in treatment in the United States. According to the National Survey on Drug Use and Health (NSDUH) in 2011, the commonly used prescription drugs for abuse are pain relievers, tranquilizers, stimulants, and sedatives (Administration, 2012).

Another component or cluster in the network suggests a strong association between albuterol and azithromycin. Albuterol is prescribed for bronchitis or asthma. On the other hand, in spite of being anti-biotic, azithromycin is prescribed for bronchitis; though the correct dosage of azithromycin is not known (Evans et al., 2002). This validates the link between azithromycin and albuterol in the network. The component also shows a strong connection between albuterol and prednisone. Prednisone, used to treat allergies or any other inflammations, is prescribed in conjunction with albuterol for better treatment of asthma (Self et al., 1990).

Further, there is a cluster containing drugs related to depression, anxiety, mania disorders. For example, citalopram is used for the treatment of depression (Hyttel, 1982). Clonazepam is prescribed for seizures, panic disorder and anxiety. Similarly, quetiapine is used for treatment of manic-depressive disorders like schizophrenia and bipolar disorder. It is used in conjunction with other similar medications.

### **Network Integration**

The DD network shows how different diagnoses or diseases interact with each other in a patient. On the other hand, the MM network provides insight on how different drugs interact with each other. Drugs can be thought of representatives of diagnoses as they are, basically, the protein molecules aimed to treat the deformities caused in the protein structure of a gene by a diagnosis. Not at all possible side-effects of a drug can be analyzed experimentally. The integration of the phenotypic network of diagnosis like the DD network and the phenotypic network of drugs like the MM network can provide more insights on the side-effects of drugs and co-morbidities of different diagnoses.

To integrate the two networks, we built a bipartite network as shown in Fig 5. In a bipartite graph, there are only two types of nodes and no edges exist between nodes of same type. Edges are only allowed between nodes of different types. In our integrated network, nodes are diagnosis and prescription drugs. An edge represent joint occurrence of a diagnosis-drug pair in the filtered dataset, significantly more than that by chance. Significance is calculated in a similar fashion as that for DD or MM. In this case, statistical significance is, also, calculated using one tail two proportion z-test with 95 percent confidence. From the network's perspective, statistical significance means that the joint probability of drug-diagnosis pair (end-nodes of an edge) is statistically significant greater than the product of the individual probabilities of the two end nodes: drug and the diagnosis (marginal probability of drug/diagnosis). The probability of a diagnosis/drug reflects the prevalence of the diagnosis/drug in the entire population (filtered dataset). We only consider those drug-diagnosis pairs that have more than 2 occurrences in the filtered dataset. The figure 5 shows the visualized network and its connected components.

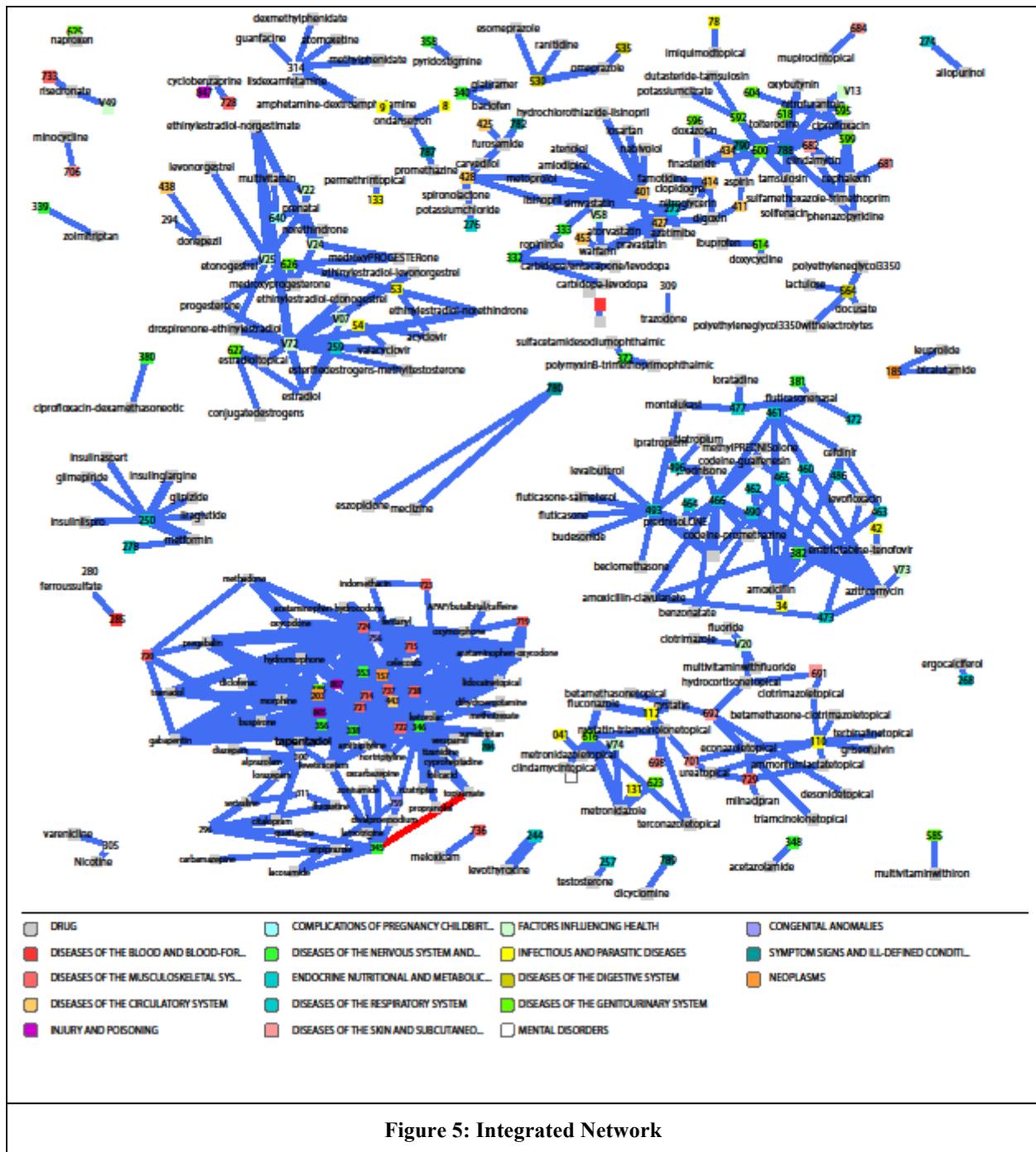


Figure 5: Integrated Network

Like the diagnosis-diagnosis network or the drug-drug network, the integrated network shows connected components or clusters centered on pain related diagnoses or drugs, gynecological diagnoses or drugs, or, diabetes or its associated drugs.

The major topological difference between the integrated network and the individual networks is that the integrated network has a gigantic component that includes all the pain-related diagnoses and drugs. The component containing epilepsy related drugs or diagnoses in the diagnosis-diagnosis

network or the drug-drug network, has been merged with the pain related ones in the integrated network. This suggests that drugs prescribed for treatment of epilepsy like levetiracetam can be used in the treatment of pain related diagnoses like neuropathic pain. Traditionally, researchers believed that levetiracetam or any other anti-convulsant drug has significant effects on pain (Guay, 2003). Recent clinical trials have proved it otherwise i.e. levetiracetam has no relevant effects on pain related diagnoses (Holbechl, Ottol, Bachl, Jensenl, & Sindrupl, 2011; Finnerup et al., 2009). The co-occurrence of levetiracetam and a pain related diagnosis in the filtered dataset (suggested by the edge between the two in the integrated network) indicates that the anti-convulsant drug was prescribed to patient complaining of pain, which should not have been ideally prescribed.

Bipolar disorder (ICD-9-CM 296) is one of the nodes in the above mentioned component. It is absent in the diagnosis-diagnosis network. It also shares a drug, quetiapine with epilepsy (ICD-9-CM 345). Quetiapine, an anti-psychotic drug, is generally used in the treatment for schizophrenia and bipolar disorder (Small et al., 1997). However, research studies state that the quetiapine has no effect on the patient, who is also taking anti-convulsant drugs. This statement is yet to be clinically proven (Nickl-Jockschat, Paulzen, Schneider, & Grözinger, 2009).

Additionally, “Anxiety, dissociative and somatoform disorders” (ICD-9-CM 300) is connected to drugs that are used in the treatment of anxiety, panic disorder, insomnia. For example, it is connected to alprazolam, diazepam and other benzodiazepines.

The network has a connected component with Diabetes mellitus (ICD-9-CM 250). Its neighbors or the drugs that share an edge with diabetes are prescribed for its treatment. According to Wikipedia, one of the connected drugs, metformin is the first-line drug of choice for the treatment of type 2 diabetes (Wikipedia, 2011). Metformin has a link with “Overweight, obesity and other hyperalimentation” (ICD-9-CM 278). Metformin is, in fact, used to control sugar levels with diet and exercise. Research has shown that obesity is one of the leading causes for diabetes (Hjartäker, Langseth, & Weiderpass, 2008; Barnett & Kumar, 2009). Unlike the diagnosis-diagnosis network, diabetes is not associated with the component containing hypertension (ICD-9-CM 401) and hyperlipidemia (ICD-9-CM 272). However, these two diagnoses are part of the same component, along with their prescribed drugs (like lisinopril for hypertension and simvastation for hyperlipidemia) and other co-morbid diagnosis (like ischemic heart disease (ICD-9-CM 414), heart failure (ICD-9-CM 427)).

Another connected component in the network is azithromycin and amoxicillin. These drugs are antibiotics, generally, provided to avoid secondary infection. Some of the nodes with amoxicillin form links or has an edge with are otitis media (ICD-9-CM 382), acute pharyngitis (ICD-9-CM 462), “Streptococcal sore throat and scarlet fever” (ICD-9-CM-034). According to the report (Cooper et al., 2001), the most common line of treatment for pharyngitis is antibiotics. Similarly, azithromycin is the most commonly prescribed for treatment of Acute bronchitis and bronchiolitis (ICD-9-CM 490), which can be seen from the network, though the correct dosage of azithromycin is not known (Evans et al., 2002). This component also contains asthma (ICD-9-CM 493). Some of the commonly prescribed drugs for asthma are inhaled corticosteroids (like fluticasone, budesonide, and beclomethasone), leukotriene modifiers (like montelukast), and combination inhalers (Fluticasone and salmeterol). The above mentioned drugs are all neighbors of asthma in the network. Some of drugs used to treat some co-morbid diseases with asthma, like obstructive pulmonary disease (COPD ICD-9-CM 496) are antibiotics.

The hormone related drugs, diagnoses and procedures form a single connected component in this integrated network. “Special investigations and examinations” (ICD-9-CM V72) is at the heart of this component for reasons similar to those in the drug-drug network. Some of commonly prescribed hormonal drugs are progesterone, and estradiol. Herpes zoster (ICD-9-CM 053) and Herpes simplex (ICD-9-CM 054), also, are nodes in this component. The drugs, associated with

these two nodes, are valacyclovir and acyclovir, which are commonly prescribed drugs for these diagnoses (Beutner, 1995).

A small component consists of “Functional digestive disorders, not elsewhere classified” (ICD-9-CM 564) and its associated drugs. The associated drugs like docusate and lactulose are prescribed for treatment of constipation (ICD-9-CM 564.0). ICD-9-CM 564.0 is a specific code for the generic category ICD-9-CM 564. Another smaller component consists of drugs like varenicline and nicotine. These drugs help to quit smoking or help to reduce tobacco dependence (Tobacco use disorder ICD-9-CM 305.1). ICD-9-CM 305.1 can be collapsed to ICD-9-CM 305 (“Nondependent abuse of drugs”).

The most common diagnosis is backache (ICD-9-CM 724.5). The backache node forms a large number of connections with pregabalin, morphine, acetaminophen-oxycodone, tapentadol, celecoxib, cyclobenzaprine, acetaminophen-hydrocodone, oxycodone, lidocaine topical. The backache node shares the maximum number of edges with or maximum number of co-occurrences in the filtered dataset with acetaminophen-hydrocodone and acetaminophen-oxycodone. These edges, also, represent some of the strongest connections in the network.

## Conclusions

In the light of our findings in the above sections, we conclude that a unified approach is necessary for analyzing the prescription drug data. The integrated network of drugs and diagnoses provides a better picture of their interactions. However, studying drugs in isolation can provide a different perspective on how two drugs can interact. Advanced network algorithms can provide better clusters, which will help to build a more compact and valid network. The previously undiscovered interactions can be found using link prediction algorithms on the networks. These algorithms can help to predict the unknown interactions. These predictions could provide a direction for performing potential experiments. The networks discussed in this paper cannot discard experiments for proving the effectiveness of drugs.

## Acknowledgements

We acknowledge financial support from grant #FA9550-12-1-0405 from the U.S Air Force Office of Scientific Research (AFOSR) and the Defense Advanced Research Projects Agency (DARPA).

## References

- Adams, C. P., & Brantner, V. V. (2006). Estimating the cost of new drug development: is it really \$802 million?. *Health Affairs*, 25 (2), 420-428.
- Admin, D. E., of Justice, U. D., & of America, U. S. (1978). Physician's Manual-An Informational Outline of the Controlled Substances Act of 1970.
- Administration, S. A. M. H. S. (2012, April). Results from the 2011 National Survey on Drug use and Health: Summary of National Findings. NSDUH Series, H-44.
- Ames, R. (1991). Hyperlipidemia in hypertension: causes and prevention. *American Heart Journal*, 122 (4), 1219-1224.
- Atlas, S. J., & Deyo, R. A. (2001). Evaluating and managing acute low back pain in the primary care setting. *Journal of general internal medicine*, 16 (2), 120-131.
- Barnett, T., & Kumar, S. (2009). *Obesity and diabetes* (Vol. 34). John Wiley & Sons.
- Beutner, K. R. (1995). Valacyclovir: a review of its antiviral activity, pharmacokinetic properties, and clinical efficacy. *Antiviral research*, 28 (4), 281-290.
- Brouwers, L., Iskar, M., Zeller, G., van Noort, V., & Bork, P. (2011). Network neighbors of drug targets contribute to drug side-effect similarity. *PLoS One*, 6 (7), e22187.
- Castelli, W. (1984). Epidemiology of coronary heart disease: the Framingham study. *The American*

*Journal of medicine*, 76 (2), 4-12.

- Cooper, R., Hoffman, J., Bartlett, J., Besser, R., Gonzales, R., Hickner, J., . . . others (2001). Principles of appropriate antibiotic use for acute pharyngitis in adults: background. *Annals of internal medicine*, 134 (6), 509-517.
- Davis, D. A., & Chawla, N. V. (2011). Exploring and exploiting disease interactions from multi-relational gene and phenotype networks. *PloS one*, 6 (7), e22670.
- DiMasi, J. A. (2001). New drug development in the United States from 1963 to 1999. *Clinical pharmacology and therapeutics*, 69 (5), 286-296.
- Drugs. (2011, December). *Levothyroxine*. [online]. (<http://www.drugs.com/levothyroxine.html>)
- Evans, A., Husain, S., Durairaj, L., Sadowski, L., Charles-Damte, M., & Wang, Y. (2002). Azithromycin for acute bronchitis: a randomised, double-blind, controlled trial. *The Lancet*, 359, 1648-1654.
- Finnerup, N. B., Grydehøj, J., Bing, J., Johannesen, I. L., Biering-Sørensen, F., Sindrup, S. H., & Jensen, T. S. (2009). Levetiracetam in spinal cord injury pain: a randomized controlled trial. *Spinal cord*, 47(12), 861-867.
- For Disease Control, C., Prevention (CDC), Centers for Disease Control, Prevention (CDC), et al. (2011). National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2011*.
- Go, A. S., Yang, J., Ackerson, L. M., Lepper, K., Robbins, S., Massie, B. M., & Shlipak, M. G. (2006). Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure the anemia in chronic heart failure: outcomes and resource utilization (anchor) study. *Circulation*, 113 (23), 2713-2723.
- Goh, K.-I., Cusick, M. E., Valle, D., Childs, B., Vidal, M., & Barabási, A.-L. (2007). The human disease network. *Proceedings of the National Academy of Sciences*, 104 (21), 8685-8690.
- Grossman, E., & Messerli, F. (2008). Hypertension and diabetes.
- Guay, D. R. (2003). Oxcarbazepine, topiramate, zonisamide, and levetiracetam: potential use in neuropathic pain. *The American journal of geriatric pharmacotherapy*, 1 (1), 18-37.
- Hidalgo, C. A., Blumm, N., Barabási, A.-L., & Christakis, N. A. (2009). A dynamic network approach for the study of human phenotypes. *PLoS computational biology*, 5 (4), e1000353.
- Hjartäker, A., Langseth, H., & Weiderpass, E. (2008). Obesity and diabetes epidemics. In *Innovative endocrinology of cancer* (pp. 72-93). Springer.
- Holbechl, J. V., Ottol, M., Bachl, F. W., Jensenl, T. S., & Sindrupl, S. H. (2011). The anticonvulsant levetiracetam for the treatment of pain in polyneuropathy: A randomized, placebo-controlled, cross-over trial. *European Journal of Pain*, 15 (6), 608-614.
- Hoy, D., Brooks, P., Blyth, F., & Buchbinder, R. (2010). The epidemiology of low back pain. *Best Practice & Research Clinical Rheumatology*, 24 (6), 769-781.
- Hyttel, J. (1982). Citalopram-pharmacological profile of a specific serotonin uptake inhibitor with antidepressant activity. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 6 (3), 277-295.
- International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). (2007). NC for Health Statistics. Retrieved from <http://www.cdc.gov/nchs/about/otheract/icd9/abticd9.htm>
- Kaplan, N. (2002). Hypertension and diabetes. *Journal of human hypertension*, 16 (3).
- Lu, J.-J., Pan, W., Hu, Y.-J., & Wang, Y.-T. (2012). Multi-target drugs: the trend of drug research and development. *PLoS One*, 7 (6), e40262.
- Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., . . . Smith, G. D. (1991). Health inequalities among British civil servants: the Whitehall II study. *The Lancet*, 337 (8754), 1387-1393.
- Nickl-Jockschat, T., Paulzen, M., Schneider, F., & Grözinger, M. (2009). Drug interaction can lead to undetectable serum concentrations of quetiapine in the presence of carbamazepine. *Clinical neuropharmacology*, 32 (1), 55.
- O'Brien, T., Dinneen, S. F., O'Brien, P. C., & Palumbo, P. J. (1993). Hyperlipidemia in patients with primary and secondary hypothyroidism. In *Mayo Clinic Proceedings* (Vol. 68, pp.860-866).
- Schadt, E. E. (2009). Molecular networks as sensors and drivers of common human diseases. *Nature*, 461 (7261), 218-223.

Self, T., Abou-Shala, N., Burns, R., Stewart, C., Ellis, R., Tsiu, S., & Kellermann, A. (1990). Inhaled albuterol and oral prednisone therapy in hospitalized adult asthmatics. Does aminophylline add any benefit? *Chest Journal*, 98 (6), 1317-1321.

Small, J. G., Hirsch, S. R., Arvanitis, L. A., Miller, B. G., Link, C. G., Group, S. S., et al. (1997). Quetiapine in patients with schizophrenia: a high- and low-dose double-blind comparison with placebo. *Archives of General Psychiatry*, 54 (6), 549.

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27 (3), 431-432.

Steinhaeuser, K., & Chawla, N. V. (2009). A network-based approach to understanding and predicting diseases. In *Social Computing and Behavioral Modeling* (pp. 1-8). Springer.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., . . . others (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, 122 (6), 957-968.

Tufte, E. R., & Graves-Morris, P. (1983). *The visual display of quantitative information* (Vol. 2). Graphics press Cheshire, CT.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., . . . others (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature*, 403 (6770), 623-627.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393 (6684), 440-442.

Wikipedia. (2011, December). *Metformin*. [online]. (<http://en.wikipedia.org/wiki/Metformin>)

Ye, H., Liu, Q., & Wei, J. (2014). Construction of Drug Network Based on Side Effects and Its Application for Drug Repositioning. *PLoS One*, 9 (2), e87864.