Analysis of Chronic Disease Related Inpatient Visits in Arizona Hospitals

Making Action Possible in Southern Arizona (MAP Dashboard)

White Paper #11
March 12, 2019

Prepared by:

Sudha Ram, PhD
Anheuser-Busch Professor of MIS, Entrepreneurship and Innovation
Director, INSITE: Center for Business Intelligence and Analytics
UA Eller College of Management

Karthik Srinivasan
MIS Doctoral Student, UA Eller College of Management

Sreya Chagarlamudi
MIS Undergraduate Student, UA Eller College of Management

Author contact information:

Sudha Ram: ram@eller.arizona.edu

Karthik Srinivasan: karthiks@email.arizona.edu

Sreya Chagarlamudi: schagarlamudi@email.arizona.edu

Executive Summary

Chronic diseases are health conditions that last three or more months and cannot be prevented by vaccines or cured by medication. They are the leading causes of death and disability in the United States. In the U.S., at least 25 percent of adults have at least two chronic conditions accounting for 86 percent of the $2.7 trillion annual healthcare expenditures related to chronic disease patients (Centers for Medicare and Medicaid Services 2016). It is important for healthcare policy makers to monitor and control the prevalence of chronic diseases.

Electronic health records (EHRs) have the potential to offer near real-time, inexpensive, standardized health data about multiple health conditions. Meaningful use objectives in the Medicare and Medicaid EHR incentive programs promote public health departments' use of EHR data for establishing registries, conducting syndromic surveillance, and monitoring reportable conditions. Despite increased adoption of EHR technologies in clinical practice and hospitals, only recently, have they been used for knowledge discovery through big data analysis techniques. In this study, we conduct an exploratory analysis of chronic diseases reported in EHRs of inpatient visits in Arizona hospitals for the time period 2012-2016. We analyze the spatial-temporal patterns in visits count of patients diagnosed with chronic diseases in these four years for 11 different regions in Arizona identified by their 3-digit zip codes.

This study is split into two phases. In the first phase, we investigate the incidence of individual chronic diseases and spatial-temporal patterns in related visit counts and average visit costs using summary statistics, simple visualizations and interactive dashboards. The 11 Arizona regions identified by the 3-digit zip codes map to the topographic boundaries of 11 major counties – Apache, Cochise, Coconino, Gila, Maricopa, Mohave, Navajo, Pima, Pinal, Yavapai and Yuma. We analyzed the prevalence of each chronic disease in different regions of Arizona and across time from 2012-2016. All 11 regions share the same set of top ten prevalent chronic disease/conditions – hypertension, disorders of lipoid metabolism, diabetes, nondependent abuse of drugs, cardiac dysrhythmias, chronic ischemic heart disease, acquired hypothyroidism, obesity, chronic renal failure and neurotic disorders; but with a difference in the relative ranking of these diseases. Diabetes is the most prevalent health condition in Apache County, whereas hypertension is the most prevalent condition observed in inpatient visits in other regions. Patients admitted with chronic diseases such as hypertension, disorders of lipoid metabolism and diabetes are higher in number in the second half of the year as compared to the first. The number of inpatient visits with neurotic-related disorder diagnoses is increasing at a rate of 18 percent every six months during 2012-2015. Circulatory system related diseases/conditions have the highest prevalence as admitting diagnoses in visits across Arizona, accounting for 19 to 38 percent of the proportion of total visits in the 11 Arizona regions. As 54 percent of chronic disease patients fall within the 50 to 79 years age group, residents in Arizona should be more health conscious during this time period of their life. The average visit cost for patients in the 0-9 year age group has increased, from $68,211 in the first half of 2012 to $100,689 in the second half of 2016, or by 48 percent.

In the second phase of the study, we investigate the phenomenon of simultaneous occurrence of two or more chronic diseases/conditions in the inpatient visits. We propose a network of diseases using the inpatient visits EHR data, where pairs of diseases are connected to each other if they simultaneously occur in one or more inpatient visits. We carry out a stratified analysis of EHRs of inpatient visits across 11 Arizona regions and 12 six-month periods, to understand the simultaneous occurrence of chronic diseases in inpatient visits across sub-populations and time periods. There exists a co-occurrence relationship among 24 percent of all possible pairs of diseases in the disease network, indicating that the disease network is fairly dense. The densities of networks corresponding

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

to inpatient visits EHRs in Mohave and Navajo are 14 and six percent higher, respectively, than density of the disease network created with all inpatient visits EHR data in Arizona. The top five diseases with the highest potential to co-occur with other diseases/conditions in an inpatient visit are chronic renal failure, chronic ischemic heart disease, disorders of lipoid metabolism, diabetes, and hypertensive renal disease. This list is common across the 11 Arizona regions, but the relative ranking of diseases is different for some regions. Disorders of lipoid metabolism is a disease/condition that has maximum potential to simultaneously occur with other diseases in hospital inpatient visits in Navajo, whereas chronic ischemic heart disease has highest potential to co-occur with other diseases/conditions in hospital inpatient visits in Pima. We vary the minimum value of edge weights of the network, to discover a repeating structural pattern in resultant networks. In this way, we analyze the disease network for identifying patterns of simultaneous occurrences of multiple chronic diseases in the inpatient visits in Arizona hospitals and present data-driven discoveries.

Our study explores chronic disease related visits in EHRs of Arizona hospitals. We hope that the findings presented in this paper can be helpful for county health departments, health care facilities and health policy makers for planning and informed decision making for chronic disease management in Arizona.

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Table of contents

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

1. Introduction

The health and social well-being of a region is closely tied to performance in all facets of human development. Improving healthcare quality and affordability is among the top priorities of the U.S. government. To improve healthcare, it is important to enumerate, understand and evaluate persistent problems. One such problem is chronic disease incidence and prevalence. In 2010, chronic diseases were responsible for 7 out of 10 deaths among Arizonans each year (ADHS Public Health Prevention 2018). Chronic diseases management is important for ensuring health and social wellbeing in Arizona.

A chronic disease is one lasting three months or more, based on the definition of the U.S. National Center for Health Statistics. Chronic diseases cannot be prevented by vaccines or cured by medication. In the U.S., at least 25 percent of adults have at least two chronic conditions with 86 percent of $2.7 trillion in annual healthcare expenditures related to chronic disease patients (Centers for Medicare and Medicaid Services 2016). In this paper, we conduct a retrospective analysis of chronic diseases reported in patient visits in Arizona hospitals. The inpatient visits electronic health records (EHR) is a repository of information about inpatients in all public hospitals across Arizona. The Arizona Department of Health Services (ADHS) maintains state inpatient records among other EHR that contain useful information about diseases and health conditions. We use inpatient visits EHR, curated and provided by ADHS, as public use files for the 2012-2016 time period as the primary dataset in our study.

The objective of our study is to explore the spatial-temporal patterns of one or more chronic diseases in the inpatient visits of Arizona hospitals. We conduct our study in the following two phases. In the first phase, we discuss the prevalence of individual chronic diseases and its relationship with different patient and spatial-temporal factors. We present distribution of visit counts and average visit costs across gender, age groups and regions through summary statistics, simple visualizations, and interactive dashboards. In the second phase of the study, we investigate the phenomenon of simultaneous occurrence of two or more chronic diseases/conditions in patient visits. For this purpose, we create a disease co-occurrence network (DCN) using the patient visits EHR, where diseases/conditions are represented as nodes and edges between each pair of nodes indicate evidence of co-occurrence between them. Through the exploratory findings from phases one and two, we effectively summarize the landscape of prevalence of chronic diseases observed in visit EHR of Arizona hospitals.

In the first phase of this study, of the 3,745,820 patient visit records present in Arizona EHR data during the 2012-2016 time period, we consider 915,355 visits which have a chronic disease as the admitting diagnosis. The dataset is divided further into visits reported in one of the 11 Arizona regions identified by 3-digit zip codes, as well as for six-month time periods (first half of 2012, second half of 2012, first half of 2013, and so on). The 3-digit zip codes approximately map to topographic boundaries of the 11 counties (Apache, Cochise, Coconino, Gila, Maricopa, Mohave, Navajo, Pima, Pinal, Yavapai and Yuma); hence, we refer to these regions using the county names in this study. Other fields in the EHR include the following information: (a) Patient related information such as gender, age group and marital status, (b) Up to 25 diagnoses and 25 procedure codes, and (c) Discharge related information such as cost of visit, payer type and discharge status. We examine whether the distribution of patient visits varies across values of gender, age group, payer type, time periods and regions. We observe that the patient visits count does not vary significantly over time, but there is a temporal pattern in the prevalence of multiple chronic diseases

such as hypertension, disorders of lipid metabolism and diabetes. Patient visits are more prevalent with these diseases in the second half of the year as compared to the first half. The number of patient visits with neurotic disorders related diagnoses are increasing at a rate of 18 percent every six months during the period 2012-2015. We further investigate the prevalence and average cost per visit for diseases grouped using the body systems classification system (circulatory system, mental disorders, musculoskeletal system, endocrine and immunity disorders, respiratory system, neoplasms, nervous system and sense organs, digestive system, genitourinary system, factors influencing contact with health services, blood and blood-forming organs related, congenital anomalies, skin and subcutaneous tissue, complications of pregnancy and childbirth, injury and poisoning, ill-defined conditions and infectious and parasitic diseases). Prevalence and average visit costs vary for diseases across all body systems.

In the second phase of this study, we create a disease co-occurrence network (DCN) using the inpatient visits EHR data of Arizona during period 2012-2015, that contains information about sets of all diseases/conditions and procedures performed in each visit. In each visit, international statistical classification of diseases (ICD-9) diagnosis codes represent the diseases/conditions. We define the nodes of the DCN as the 3-digit ICD-9 codes (the first three significant digits from the left of the 5-digit ICD-9 codes). The edges between each pair of nodes indicate the extent of co-occurrence between them, and various measures have been proposed in the literature to denote the strength or weight of these edges in the DCN. However, since existing approaches to assign DCN edge weights have limitations such as bias towards rare diseases and intractable ranges (Roque et al. 2011), we propose a new measure for edge weights in the network called the co-occurrence coefficient (CC). CC indicates the extent to which pairs of diseases/conditions co-occur across patient visits, varies in the range of zero to one and does not suffer from the limitations of previous measures. Data from 2016 was not considered for creating the DCN as diseases in 2016 were recorded using the ICD-10 diagnosis codes which have a different ontological classification from ICD-9 codes. The DCN can be constructed for ICD-10 diagnosis codes similar to ICD-9 without loss of generality.

In addition to creating a DCN for all inpatient visits, we create DCNs for subsets of the data corresponding to 11 Arizona regions and 12 six-month time periods, to understand the change in structure of the network across sub-populations and time periods. There exists a co-occurrence relationship among 24 percent of all possible diagnoses' pairs in the network, indicating that network is fairly dense. For comparison, the average Facebook user's friends list has a 12 percent network density (Hampton et al. 2012). The density of networks for some regions is higher than others, indicating that patient visits in those regions have a higher number of average diseases diagnosed per visit. The weighted degree of a disease node indicates the potential of the disease to co-occur with other diseases/conditions in a patient visit (Barrat et al. 2004). The list of top five diseases with highest weighted degrees in DCNs corresponding to EHRs across 11 Arizona regions is similar, but the relative ranking of these diseases is different. For example, *disorders of lipid metabolism* diagnosis has the highest weighted degree in Navajo County whereas chronic ischemic heart disease has the highest potential to co-occur with other diseases/conditions in hospital patient visits in Pima County. To investigate sensitivity of the network structure to our proposed edge weight measure, we vary the minimum value of edge weights of the DCN to discover a repeating structural pattern in resultant networks. We observe that disorders of lipid metabolism such as high cholesterol or hyperlipidemia (high amount of fat particles in blood) are linked to diabetes, chronic ischemic heart disease, and hypertension such that occurrence of disorders of lipid metabolism suggests a high possibility of any one or multiple of the other three health conditions. Each of these

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

chronic diseases are known to be caused by unhealthy lifestyles, thus underscoring the importance of diet and regular exercise for an individual's health and wellbeing.

The paper is organized as follows In Section 2, we describe and summarize the Arizona state inpatient visits EHR data. In Section 3, we conduct some exploratory analysis using an online dashboard developed in the first phase of the study. We provide some background and motivation for phase two of this study, i.e., exploring the phenomenon of co-occurring chronic diseases in patient visits in Section 4. In Section 5, we construct a disease co-occurrence network using Arizona state inpatient visits EHR and explore this network. Discussion and conclusions about the study are provided in Section 6.

2. EHR of inpatient visits in Arizona hospitals

Population-wide health information is directly available from healthcare providers as electronic health records (EHR). It contains information such as diagnoses, procedures, medications, biomarkers, and administrative information of patient visits. Such data is available at a patient visit level that can be queried across institutions in different regions for different timeframes. Health conditions related to cardiovascular problems, pulmonary diseases, hypertension, diabetes, ulcer, obesity, depression are recorded in inpatient visits EHR through the international statistical classification of diseases (ICD Version 9/10) coding standard. The inpatient visit electronic health records dataset (EHR) of Arizona contains visit information about patients who are transferred from outpatient and emergency departments for treatment. Data is anonymized to protect patient identity and each row contains visit-related information.

Each patient visit record can have up to 25 diagnosis codes, i.e., 25 co-occurring diseases and up to 25 procedure codes. Until the end of 2015, the ninth version of the international statistical classification of diseases (ICD Version 9 or ICD-9) coding standards were used to track the diagnosis codes. These were subsequently replaced by ICD Version 10 (ICD-10) in 2016. This data is available as public use files (PUF) and can be requested from azdhs.gov. A detailed data dictionary is available online which provides additional information about the individual fields present in the data.[1]

2.1 Data processing

The Arizona inpatient visits EHR dataset has 3,745,820 total records for the period 2012-2016. We filtered chronic disease patient visits using chronic condition indicators (CCI) for ICD-9 and ICD-10 codes provided by H-CUP, a federal-state-industry partnership sponsored by the Agency of Health Research and Quality.[2] The final dataset used in our analysis has 915,355 records corresponding to chronic disease related patient visits in Arizona. Apart from diagnosis codes, the visit related fields considered in this study are listed and summarized in Table 1. We performed the following data pre-processing on the variables: (a) Removing chronic diseases that have less than 20 related patient visits, (b) Filtering patient visits with 3-digit zip codes - 850, 851, 852, 853, 855, 856, 857, 859, 860, 863, 864, 865, (c) Grouping labels of payer type, ethnicity, and discharge status into 10 or fewer meaningful categories each, grouping smaller labels into a generic 'other' category, (d) Handling missing or abnormal values using mean value imputation, (e) Appending information about diagnoses such as body systems (circulatory system, mental disorders, etc.) and clinical classification software group codes and disease-related description from H-CUP to ICD-9/ICD-10 admitting

---

[1] ADHS Hospital discharge data dictionary - http://www.azdhs.gov/documents/preparedness/public-health-statistics/hospital-discharge-data/data-release/PUF_data_dictionary.pdf

[2] Chronic condition indicator files - https://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

diagnosis codes of patient visit, and (f) Linking patient visit 3-digit zip code information with the 11 Arizona regions and median income of those regions. [3,4,5,6]

Of the 14,568 5-digit ICD-9 diagnosis codes, 4,584 diagnosis codes are identified as chronic diseases/conditions. Of the 78,838 7-digit ICD-10 diagnoses, 11,346 diagnosis codes are identified as chronic diseases/conditions. We indicate existence of a chronic condition by flagging a 3-digit diagnosis code, when more than 90 percent of the corresponding set of expanded 5-digit diagnosis codes are identified as chronic conditions. Even though ICD-10 and ICD-9 are mapped to each other via one to one mapping lookup tables available online, we noticed a significant difference between distribution of diagnosis codes used for the same set of diseases/conditions and hence we present analysis for chronic diseases indicated by ICD-9 codes in this study.[7] In a future study, it is possible to extend the analysis for ICD-10 codes.

We use the 3-digit ICD-9 codes of chronic conditions/diseases for both phases of our study, as inference about 3-digit diseases/conditions are generalizable and easier to interpret than its more specific expansions (Slee 1978).

2.2 Summary of chronic diseases related inpatient visits in Arizona hospitals

Summary statistics for inpatient visits EHR of hospitals in Arizona are shown in Table 1(a) (patient visits) and Table 1(b) (counties). The average cost of patient visits is highest in Mohave ($80,340) and lowest in Coconino ($47,219). Of the total visits to inpatient wards, 66 percent are from an emergency ward. White Americans constitute around 85 percent of all patients. The highest proportion of patients are from the 60-69 years age group. Of all patients, 65 percent are discharged home while 33 percent are transferred to other institutions such as short-term general hospitals for inpatient care, skilled nursing facilities, organized home health service organizations, or federal health care facilities. Maricopa has the largest proportion of patient visits (41 percent), followed by Yuma (21 percent), and then Pima (15 percent).

---

[3] CCS codes and disease description related files - https://www.hcup-us.ahrq.gov/tools_software.jsp

[4] 3-digit zip code to location mapping tool - http://maps.huge.info/zip3.htm

[5] Population estimates files - https://population.az.gov/sites/default/files/documents/files/pop-estimates2016-04pla.pdf

[6] Factfinder tool to extract median income -
https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t#acsST

[7] CMS ICD-9 to ICD-10 and ICD-10 to ICD-9 code mapping - http://www.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html

MAP Dashboard White Paper　　　　　　www.mapazdashboard.arizona.edu

Table 1(a): Summary statistics of the Arizona inpatient visits EHR dataset (patient visits)

Variable		% of total patient visits
Age-group	0-9 years	2.57
	10-19 years	5.74
	20-29 years	7.13
	30-39 years	8.14
	40-49 years	10.70
	50-59 years	16.14
	60-69 years	20.05
	70-79 years	17.76
	80-89 years	9.83
	90+ years	1.95
Ethnicity	White	85.00
	African American	4.81
	Native American	3.57
	Asian	1.22
	Hispanic	4.26
	Others	1.13
Gender	Male	49.89
	Female	50.11
Priority of visit	Elective	33.41
	Emergency	43.89
	Trauma	0.25
	Urgent	22.40
	Unknown	0.05
Marital status	Divorced	9.75
	Married	42.34
	Separated	0.98
	Single	35.77
	Widowed	10.28
	Unknown	0.87
Discharge status	Discharged home	64.66
	Expired	1.43
	Discontinued care	1.09
	Transferred to another institution	32.83
Payer type	AHCCCS Medicaid	18.97
	Commercial	4.24
	HMO	14.82
	Medicare	34.10
	PPO	7.72
	Self Pay	5.18
	TRICARE	1.58
	Workers Compensation	0.31
	Medicare Advantage	11.15
	Other	1.93

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Table 1(b): Summary statistics of the Arizona inpatient visits EHR dataset (counties)

Variable		% of total patient visits	Average population	Median income ($)
County	*Apache*	0.36	72,195	32,127
	Cochise	4.90	129,314	44,873
	Coconino	1.91	139,000	50,953
	Gila	1.69	53,623	39,751
	Maricopa	41.41	4,053,850	54,550
	Mohave	3.16	204,106	38,743
	Navajo	1.20	108,660	36,480
	Pima	15.09	1,003,750	45,848
	Pinal	6.00	400,615	50,748
	Yavapai	3.70	217,127	45,684
	Yuma	20.57	208,466	41,127

2.3 Top ten prevalent chronic conditions in Arizona

The top 10 prevalent chronic diseases/conditions in Arizona are hypertension, disorders of lipoid metabolism, diabetes, nondependent abuse of drugs, cardiac dysrhythmias, other forms of chronic ischemic heart disease, acquired hypothyroidism, obesity and other hyper alimentation, chronic renal failure, and neurotic disorders.[8] Lists of top 10 prevalent chronic conditions in each of the 11 Arizona regions are provided in Appendix A. We observe that a majority of the 11 regions share a similar list of top 10 diseases, but with a slight difference in the ranking of diseases or introduction of diseases not in this list. For example, the most prevalent disease in Apache is diabetes followed by hypertension; osteoarthrosis and depressive disorder are present in the top 10 list for Cochise County. Figure 1 shows a bar chart of the top 10 prevalent chronic diseases/conditions in Arizona, arranged from left to right in decreasing order of their prevalence in 2012-2016 (i.e., number of patient visits recorded with the disease during 2012-2016 period).

---

[8] Note: *Other forms of chronic ischemic disease* is description for a billable ICD code corresponding to *chronic ischemic disease*. In this white paper, we use the terms *chronic ischemic disease* and *Other forms of chronic ischemic disease* interchangeably.

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu
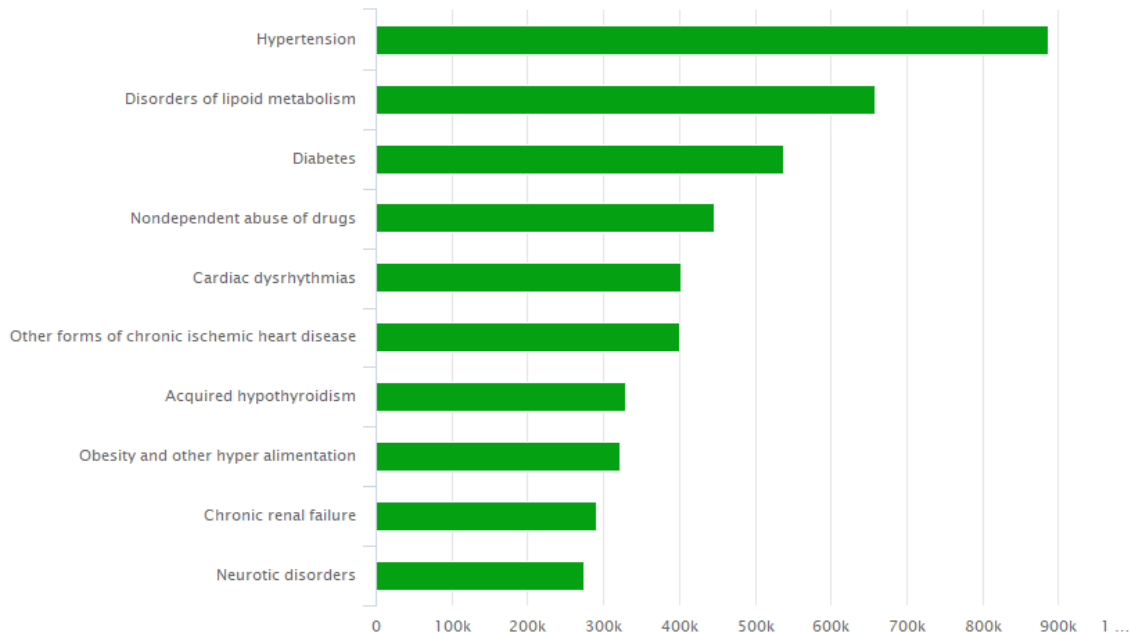
Figure 1: Prevalence of top ten chronic diseases in Arizona during 2012-2016 period

Next, we analyze the order of chronic disease prevalence from time period 2012-01 (first half of 2012) to 2015-01 (first half of 2015) for Arizona inpatient visits data. We haven't considered 2015-02 as the numbers are affected by transition from ICD-9 to ICD-10 across U.S. hospitals. We can clearly see a seasonal pattern in Figure 2, where the number of inpatient visits related to hypertension, lipoid metabolism, diabetes, chronic ischemic heart disease, cardiac dysrhythmias, and chronic renal failure are more prevalent in the second half of the year as compared to the first half. The prevalence of neurotic disorders has a visible positive trend, where it increases from 34,441 related visits in 2012-01 to 40,793 related visits in 2015-01; a 18 percent increase in number of cases.
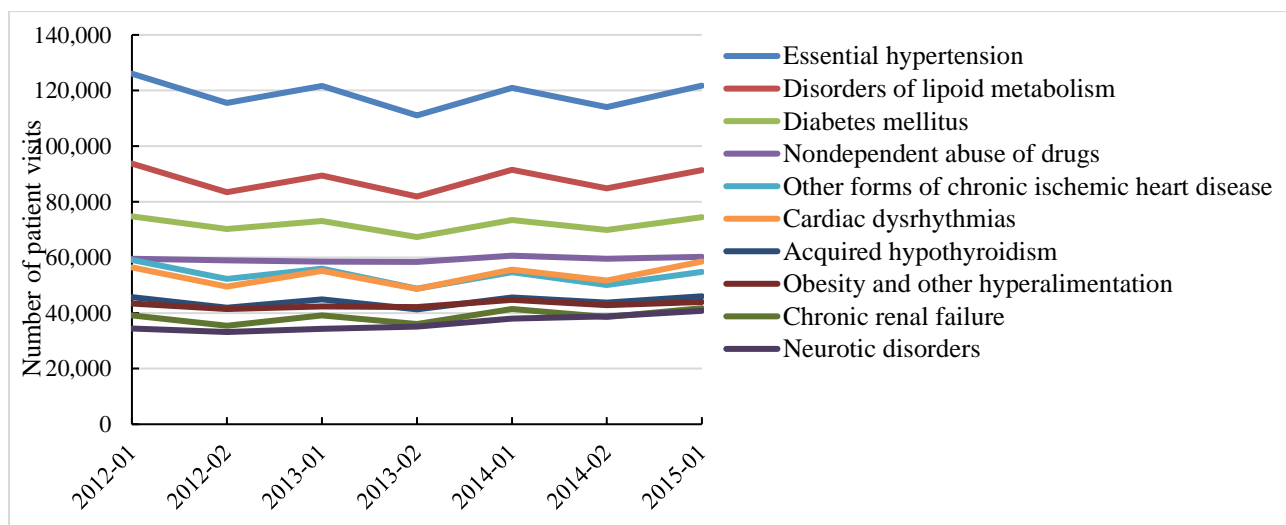
Figure 2: Prevalence of top ten chronic diseases as a function of time

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Further investigation into these exploratory findings is warranted using 2016 and 2017 inpatient visit records when they become available.

The *body system* field is a group variable used to classify the ICD-9 and ICD-10 diagnoses into 17 labels. The 4,584 chronic diseases/conditions can be classified into 17 categories, corresponding to 17 body systems (e.g., circulatory system, mental disorders, musculoskeletal system, etc.). Table 2 shows the proportion of total visits in a region having an admitting diagnosis classified into a particular body system. Of the total visits in Arizona, 26 percent have an admitting diagnosis code related to the circulatory system. The percentage of visits related to the circulatory system is lowest in Coconino County (19 percent) and highest in Mohave County (38 percent). Infectious and parasitic diseases are least prevalent among all body systems. Hospitals in Coconino County have a larger proportion of inpatient visits with mental disorders than other diseases/conditions.

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Table 2: Distribution of visits with admitting diagnoses across body systems

Body system	Apache	Cochise	Coconino	Gila	Maricopa	Mohave	Navajo	Pima	Pinal	Yavapai	Yuma	Arizona
Circulatory system	23.96	30.36	19.07	33.46	23.94	38.00	24.20	24.20	27.82	28.02	27.14	25.85
Mental disorders	18.53	13.76	29.09	14.44	26.28	5.60	25.69	25.51	17.67	19.04	20.28	22.70
Musculoskeletal system	17.39	19.96	16.61	18.03	14.65	17.91	20.93	16.17	17.03	20.64	16.32	16.13
Endocrine and immunity disorders	10.39	6.76	6.29	6.09	5.93	6.38	5.14	6.76	6.70	5.08	6.42	6.24
Respiratory system	3.81	5.98	4.63	5.98	5.90	11.49	4.03	5.71	6.89	5.68	6.35	6.14
Neoplasms	4.41	5.94	5.42	5.44	5.57	6.36	5.24	5.14	5.37	6.69	5.67	5.58
Nervous system and sense organs	6.73	4.39	5.55	4.24	4.28	3.04	3.29	4.54	4.22	4.04	4.77	4.39
Digestive system	3.78	3.47	3.03	4.03	3.69	3.90	3.06	3.30	3.80	2.83	3.71	3.59
Genitourinary system	1.77	2.78	2.50	2.26	2.28	2.23	1.61	2.18	3.29	2.72	2.46	2.39
Factors influencing contact with health services	1.89	1.77	1.51	1.60	1.96	1.34	2.05	1.61	1.97	1.81	1.91	1.85
Blood and blood-forming organs related	1.53	1.38	1.14	0.85	1.84	1.38	1.07	1.38	1.53	0.96	1.78	1.63
Congenital anomalies	2.61	0.99	1.84	0.96	1.16	0.55	1.40	0.94	1.08	0.77	1.04	1.07
Skin and subcutaneous tissue	1.68	0.89	0.95	1.29	0.89	0.95	0.98	0.84	0.99	0.74	0.83	0.88
Complications of pregnancy and childbirth	0.54	0.81	1.46	0.54	0.77	0.15	0.50	0.99	0.66	0.26	0.46	0.70
Injury and poisoning	0.72	0.35	0.53	0.52	0.42	0.34	0.53	0.32	0.53	0.27	0.40	0.40
Ill-defined conditions	0.15	0.34	0.29	0.23	0.31	0.25	0.16	0.34	0.35	0.36	0.35	0.32
Infectious and parasitic disease	0.09	0.06	0.10	0.06	0.14	0.14	0.12	0.08	0.09	0.07	0.10	0.11

3. Analyzing inpatient visits and average visit costs using interactive dashboards

The interactive dashboard created using Arizona state inpatient EHR data can be accessed at http://u.arizona.edu/~schagarlamudi/AZ_Map/ as well as at https://www.insiteua.org/research/chronic-disease-trends-arizona-regions. It has been created using Tableau 10.2 (academic edition). Readers are urged to utilize the interactive dashboards to compare the distributions of visits count and average visit costs across different age groups, body systems, time periods, regions, gender, and ethnicity.

3.1 Dashboard 1: Inpatient visits count of populations

We explore the distribution of visit count across sections of the overall inpatient population in Arizona in the dashboard shown in Figure 3. The bottom left graph shows the map of Arizona with

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

regions colored based on average median income. Clicking or selecting regions, will filter the other visualizations for those regions. The other two graphs in this dashboard show the distribution of inpatient visits count across the 17 body system categories (circulatory system, digestive systems, musculoskeletal system, neoplasms, etc.) and patient age groups (0-9 years, 10-19 years, 20-29 years, etc.). The visualizations show that a majority of chronic patient visits are related to circulatory and musculoskeletal system disorders. For a majority of disorders, females were over half of the affected population. Of all the chronic disease patients, 54 percent are in the 50-79 years age group. Given that approximately 30 percent of the total population in Arizona falls within the 50-79 years age group we can infer that this is the most vulnerable age group for residents in Arizona. The average visit cost for patients in 0-9 years age group has increased from $68,211 in first half of 2012 to $100,689 in second half of 2016, or by 48 percent. The underlying reason for this increase in care of 0-9 years age group patients during 2012 and 2016 needs to be investigated further.
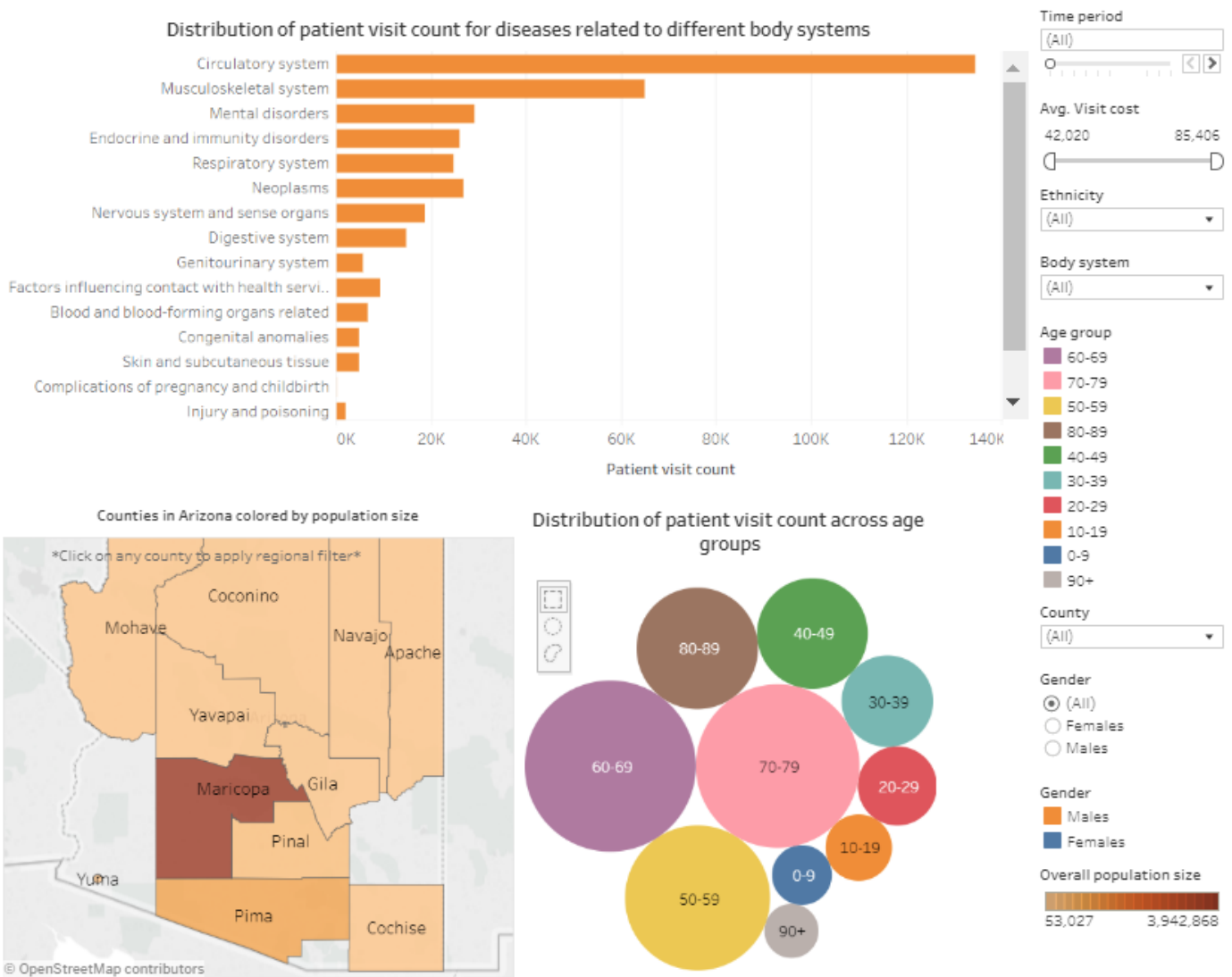

Figure 3: Dashboard 1 - Distribution of inpatient visits count for diseases related to different body systems

3.2 Dashboard 2: Average visit costs of population

This dashboard shown in Figure 4 presents the variability in average patient costs of visits across diseases classified under different body systems. We further include information on payer type using color codes to compare visit costs across different payer types such as Medicaid, health maintenance organizations (HMO), preferred provider organization (PPO), Medicare, etc.

Visits with congenital anomalies related diagnoses have the highest costs followed by neoplasm and circulatory system related visits. The proportion of average cost expenditure by Arizona Health Care Cost Containment System (AHCCCS) Medicaid is higher than other payers for congenital anomalies related visits, whereas PPOs spend more on an average for parasitic and infectious disease related visits. Commercial indemnity plans have higher average spending on visits related to injury/poisoning and skin problems related visits. These patterns could be due to different patient segments (differentiated by age groups, income-levels, education levels, occupations, etc.) having access to different insurance options and hence different payer types. We urge the reader to use the filters and click on the interactive (linked) visualizations to explore average visit cost patterns across different values of filter variables.
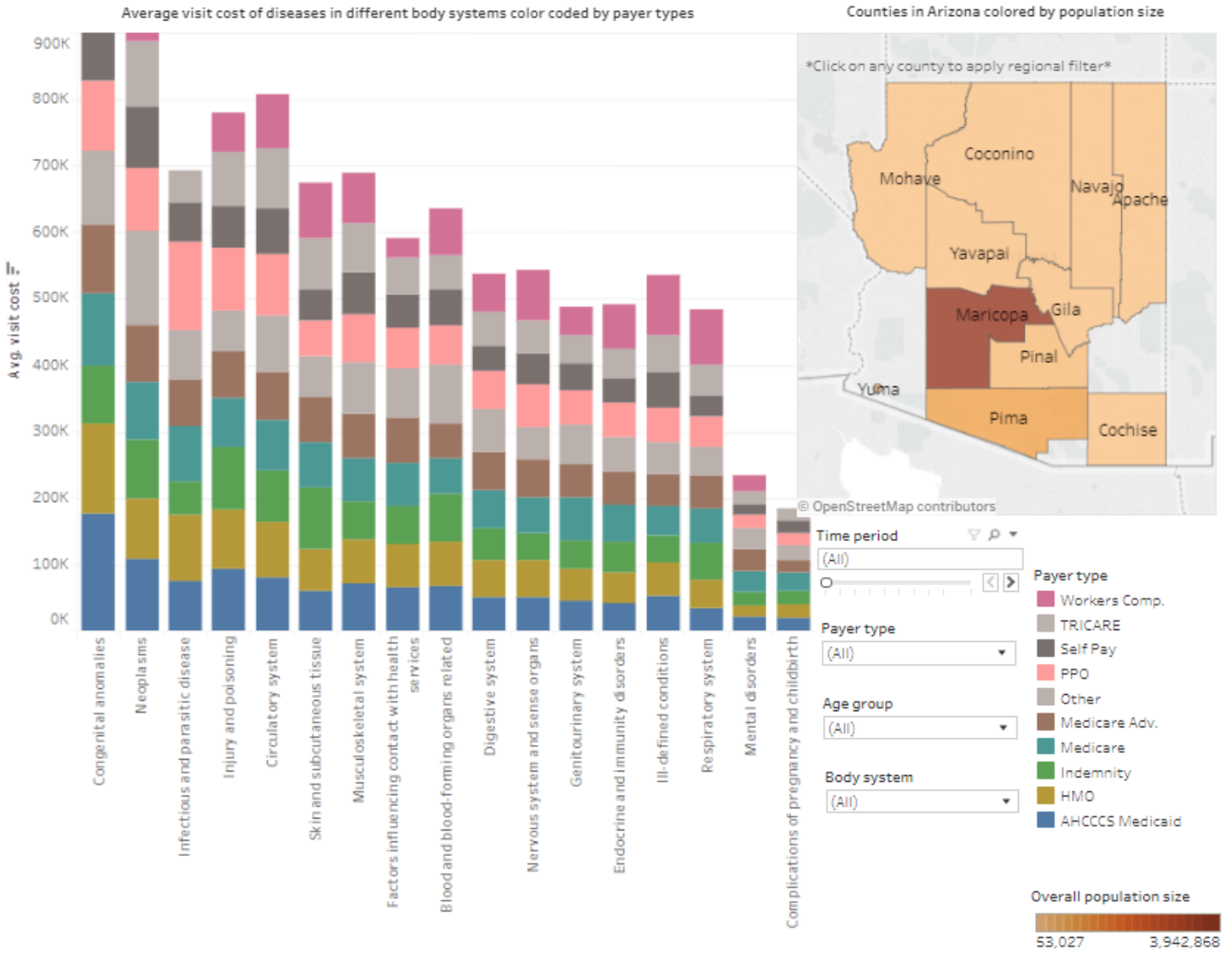
MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Figure 4: Dashboard 2 - Distribution of average visit costs for diseases related to different body systems across payer types

3.3 Dashboard 3: Visits count and costs across population

In dashboard 3 shown in Figure 5, we explore patterns in visits count and average visit costs simultaneously across two variables - age groups and discharge status. Discharge statuses have been grouped into four categories: discharged home, transferred to another institution, expired, and left against medical advice. The average cost of a visit where the patient has expired is 1.5 times higher than visit costs in other discharge status categories at any given time period. The distribution of patient visits count across age groups is consistent over time, with a majority of patients belonging to the 60-69 years and 70-79 years age groups. The average cost for patient visits in the 0-9 years age group is consistently increasing with time, from $68,211 in first half of 2012 to $100,689 in the second half of 2016. There is a data input issue for the second half of 2016 for discharge status and hence this particular set of observations can be ignored. The rate of increase

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

in average cost and total inpatient visits has steadily increased from 2012-01 (i.e., first half of 2012) to 2016-02 (i.e., second half of 2016).
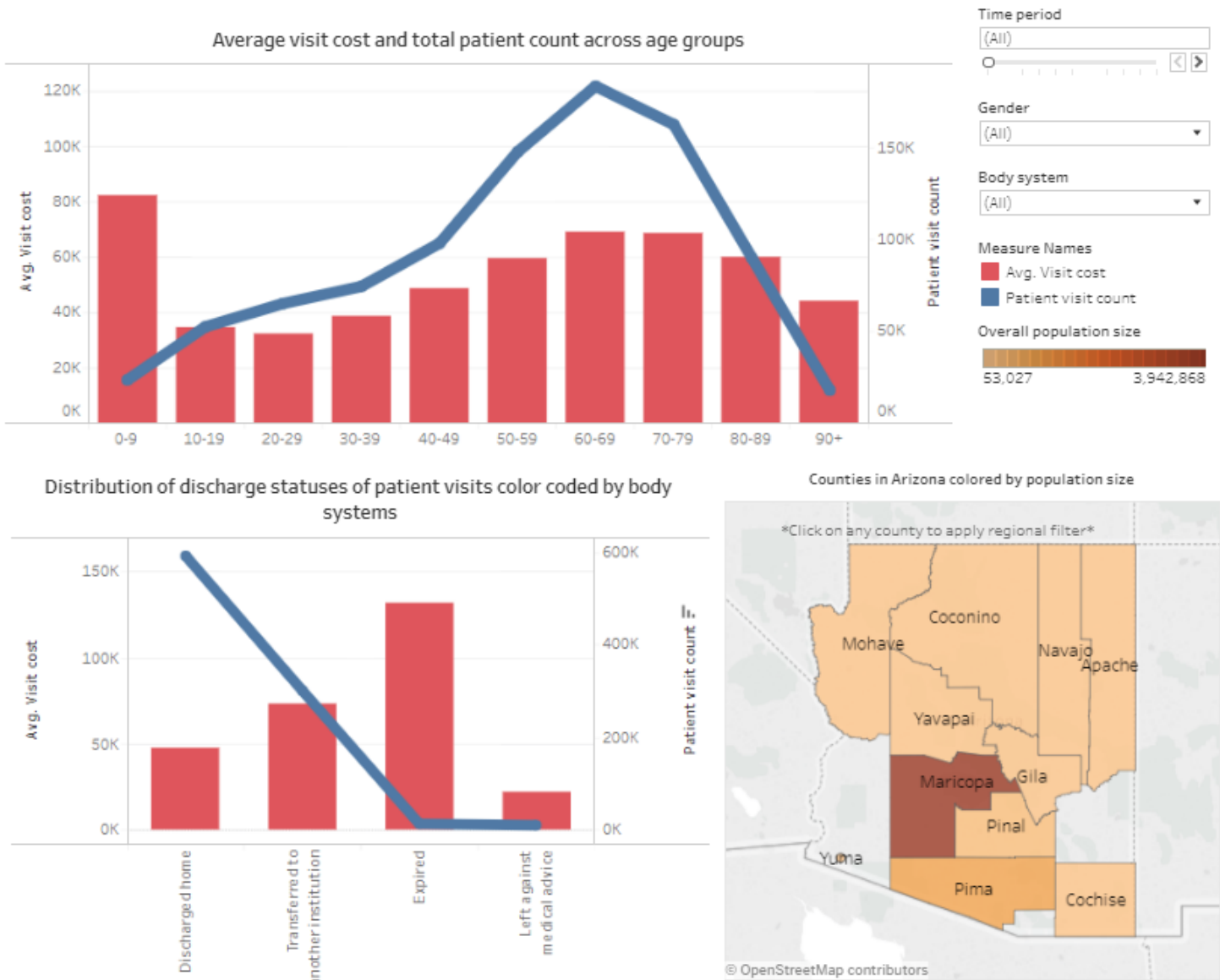
Figure 5: Dashboard 3 - Average visit costs and total visits count across age groups and discharge statuses

4. Analyzing disease co-occurrence

In phase two of this research, we conduct a detailed exploration of individual chronic conditions to investigate simultaneous occurrence of diseases in the inpatient visits across Arizona hospitals. Several studies have analyzed the inter-relationship between two or more diseases (Mannino et al. 2008; Ogle et al. 2000; Schellevis et al. 1993) and shown that simultaneous occurrence increases morbidity. Comorbidity indices have been proposed (Charlson et al. 1987; Elixhauser et al. 1998) to account for this common clinical phenomenon. Recently, network science has been proposed to

18

model the co-occurrence relationship between pairs of diagnoses/conditions using EHR data (Divo et al. 2015; Hidalgo et al. 2009; Srinivasan et al. 2018). Disease co-occurrence networks are created using diagnosis codes as nodes connected by edges indicating the strength of co-occurrence.

Disease co-occurrence networks (DCN) are constructed based on repeated evidence of two or more diseases. The nodes of these networks are ICD9/10 diagnosis codes and the edges represent co-occurrence relationship between pair-wise disease diagnoses. Figure 6 shows a disease co-occurrence network for a hypothetical example of three related conditions—diabetes, hypertension, and obesity. Even though there is evidence of co-occurrence between each disease pair in this network, the strength of the co-occurrence relationship is not straightforward. Multiple measures have been proposed in the past to capture the co-occurrence between disease pairs using edges weights defined in different ways as shown in Table 3.
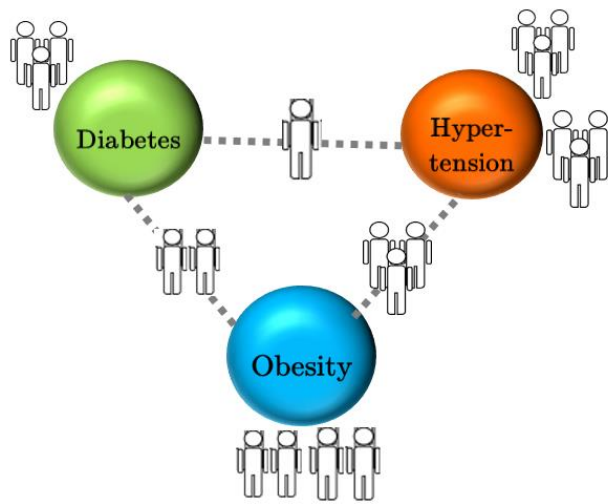
Figure 6: An illustrative representation of a 3-node disease co-occurrence network. The pairwise co-occurrence coefficients between diabetes, hypertension and obesity are derived using number of co-occurrences and prevalence of each disease.

Existing approaches to assign DCN edge weights have limitations such as bias towards rare diseases and intractable ranges (Roque et al. 2011). We propose a new measure for edge weights called co-occurrence correlation (CC) that is suitable for analyzing disease co-occurrence networks:

$$CC_{xy} = \frac{\sqrt{2}C_{xy}}{\sqrt{P_x^2 + P_y^2}}$$

Where, C_{xy} is the co-occurrence of diseases x and y across inpatient visits, P_x and P_y are prevalence of diseases x and y respectively.

C_{xy} is symmetric, reflexive, and constrained to the range zero to one, and is therefore semi-metric (Shi et al. 2012). It does not suffer the limitations of previously proposed measures. A two-sided t-test is used to retain significant edges in the network with the following test statistic:

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

$$t_{n-1} = \sqrt{\frac{(n-2)CC_{xy}^2}{(1-CC_{xy}^2)}} \; , n = \max(P_x, P_y)$$

Table 3: Edge weights for disease co-occurrence networks

No	Formula	Study	Application
1	$RR_{xy} = NC_{xy} / P_x P_y$	(Hidalgo et al. 2009)	Propose the DCN
2	$\phi_{xy} = \frac{(C_{xy}N - P_x P_y)}{\sqrt{P_x P_y (N - P_x)(N - P_y)}}$	(Hidalgo et al. 2009)	Propose the DCN
3	$SC_{xy} = \frac{C_{xy}}{(P_x + P_y)}$	(Steinhaeuser and Chawla 2009)	Compute patient similarity
4	$\phi'_{xy} = \ln_2 \left(\frac{C_{xy}+1}{(P_x P_y / N)+1} \right)$	(Roque et al. 2011)	Stratify patient cohorts
5	$RR_{xy} = NC_{xy} / P_x P_y$	(Klimek et al. 2015)	Identify diabetes comorbidity risks
6	C_{xy}	(Liu et al. 2016)	Comorbidities analysis of hypertension
N is the size of the total number of observations in the dataset, C_{xy} is the co-occurrence of diseases x and y across inpatient visits and $\{P_x, P_y\}$ are prevalence of diseases $\{x,y\}$.			

5. Disease co-occurrence networks

In this section, we briefly discuss how disease co-occurrence networks are extracted from the Arizona inpatient visits EHR data, followed by an exploratory analysis of components of this network.

5.1 ICD diagnosis codes

International statistical classification of diseases (ICD) are systematically versioned lists of diagnoses and procedure codes maintained by World Health Organization (WHO), consisting of an ontology of codes classifying diseases and related symptoms, causes and conditions (Centers for Disease Control and Prevention and National Center for Health Statistics 2013). Since 1900, they have undergone 11 revisions, with the latest version ICD-11 released on 18 June 2018. ICD codes are short alphanumeric codes (three to seven digits), with the first three digits from the left indicating disease categories/groups, and every subsequent digit added to capture more and more details of a diagnosis. For example, the 5-digit ICD9 code 53081 corresponds to *Gastroesophageal reflux*, where 3-digit codes 520-529 correspond to *Diseases of digestive system*, 3-digit code 530 corresponds to *Diseases of esophagus*, and 5308 maps to the category of *Other specified disorders of esophagus* consisting of *Gastroesophageal reflux* (53081) and *Barrett's esophagitis* (53085). More information

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

about ICD coding standards and details about ontology of each version can be found at the WHO webpage for international disease classification standards.[9]

We use the 3-digit ICD-9 codes of chronic conditions/diseases for both phases of our study, as inference about 3-digit diseases/conditions are generalized and easier to interpret than its more specific expansions (i.e. 5-digit ICD-9 codes or 7-digit ICD-10 codes). Even though ICD-10 to ICD-9 one to one mapping lookup tables have been created and made available online, we noticed a significant difference between distributions of diagnosis code used for the same set of diseases/conditions[10]. Hence, we fit the disease co-occurrence network (DCN) using 2012-15 inpatient EHR data that contains records with ICD-9 diagnosis codes corresponding to all diseases/conditions diagnosed in each visit.

5.2 Constructing DCNs with EHR datasets

There are 4584 5-digit ICD-9 diagnosis codes classified as chronic diseases/conditions, which are congregated into 290 3-digit ICD-9 diagnosis codes indicating chronic diseases/conditions. We assume that a 3-digit ICD-9 is a chronic disease/condition if at least 90 percent of its corresponding expanded 5-digit ICD-9 codes are chronic diseases/conditions. This ensures that we only consider diseases identified by 3-digit diagnoses that are known to be chronic in all forms rather than having one of their variants to be chronic. For example, viral hepatitis has ICD-9 code '070' and 24 corresponding expanded codes. Five of the 24 expanded diagnosis codes are indicated as chronic (07022 - chronic viral hepatitis B with hepatic coma without hepatitis delta, 07023 - chronic viral hepatitis B with hepatic coma with hepatitis delta, 07032 - chronic viral hepatitis B without mention of hepatic coma without mention of hepatitis delta, 07044 - chronic hepatitis C with hepatic coma, and 07054 - chronic hepatitis C without mention of hepatic coma), but viral hepatitis by itself is not a chronic condition. Hence, our filtering criteria are used to derive 290 3-digit ICD-9 diagnoses as being chronic conditions from existing chronic condition indicator mapping for 5-digit ICD-9 codes.[11] In addition to the above filter, we impose an exclusion filter of minimum 25th percentile of inpatient visit count per chronic disease, so that the network is not biased by the presence of rare chronic diseases/conditions. We develop the DCN as a weighted undirected network, with edge weights equal to the co-occurrence coefficient (CC) described in the previous section. As a final pre-processing step, we remove nodes in the network that are isolated or not linked to any other node. We repeat this process for all chronic disease related inpatient EHR data in Arizona hospitals, as well as datasets corresponding to inpatient EHR data of hospitals, in each of the 11 regions. We also create DCNs for Arizona inpatient EHR data for each six-month time period – 2012-01, 2012-02, 2013-01, 2013-02, 2014-01, 2014-02, 2015-01, and 2015-02 to explore possible temporal patterns in the DCNs (YYYY-01 indicates first six months of a year and YYYY-02 indicates the last six months of the year). A list of chronic conditions/diseases, 3-digit ICD-9 codes, related body systems, and respective exclusion indicators (i.e., less prevalent, isolated nodes) for DCNs with Arizona inpatient EHR data is provided in Appendix B.

---

[9] WHO ICD Code instruction manuals: http://www.who.int/classifications/icd/en/

[10] CMS ICD-9 to ICD-10 and ICD-10 to ICD-9 code mapping - http://www.nber.org/data/icd9-icd-10-cm-and-pcs-crosswalk-general-equivalence-mapping.html

[11] Chronic condition indicator files - https://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

5.3 Properties of DCNs across regions and time periods

The number of nodes, number of edges, and density of DCNs created with inpatient visits EHR datasets of Arizona, 11 Arizona regions, and time periods are shown in Table 4. In Figure 7, we can see that the disease co-occurrence network for Arizona has nodes that represent diseases/conditions from different body systems related to each other (with a color legend in Figure 8). This happens as a set of all diagnoses in a given inpatient visit often have individual diagnosis from multiple body systems.

Table 4: DCNs for 11 Arizona regions and 8 Time periods

Dataset		No. of nodes	No. of edges	Density
Overall Arizona (all time periods)		209	5113	0.2352
County (all time periods)	Apache	63	431	0.2207
	Cochise	122	1545	0.2093
	Coconino	109	1160	0.1971
	Gila	96	925	0.2029
	Maricopa	190	3641	0.2028
	Mohave	110	1614	0.2692
	Navajo	82	826	0.2487
	Pima	154	2536	0.2153
	Pinal	127	1694	0.2117
	Yavapai	118	1484	0.2150
	Yuma	167	2991	0.2158
Time periods (for Arizona)	2012-01	161	2412	0.1873
	2012-02	151	2296	0.2027
	2013-01	156	2379	0.1968
	2013-02	155	2433	0.2039
	2014-01	155	2567	0.2151
	2014-02	159	2574	0.2049
	2015-01	159	2697	0.2147
	2015-02*	138	2033	0.2151
2015-02 has fewer number of nodes due to nationwide coding transition from ICD-9 to ICD-10 diagnosis codes				

The density of a network is the ratio of number of actual edges to the number of possible edges between all the nodes in network. The density of the overall network created with all inpatient visits EHR data of Arizona is 0.2352; however, the density in each individual time period is lower than this by at least 10 percent. This may be attributed to the fact that the overall network is able to capture weaker evidence of co-occurrence as size of dataset increases. The density of networks corresponding to inpatient visits EHRs in Mohave and Navajo are 0.2692 and 0.2487, i.e., 14 percent and six percent higher, respectively, than the density of the overall DCN created with all inpatient visits EHR data of Arizona. Therefore, this disease co-occurrence phenomenon is relatively more visible in these two regions as compared to the overall DCN, despite their population being lower than most other regions.

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu
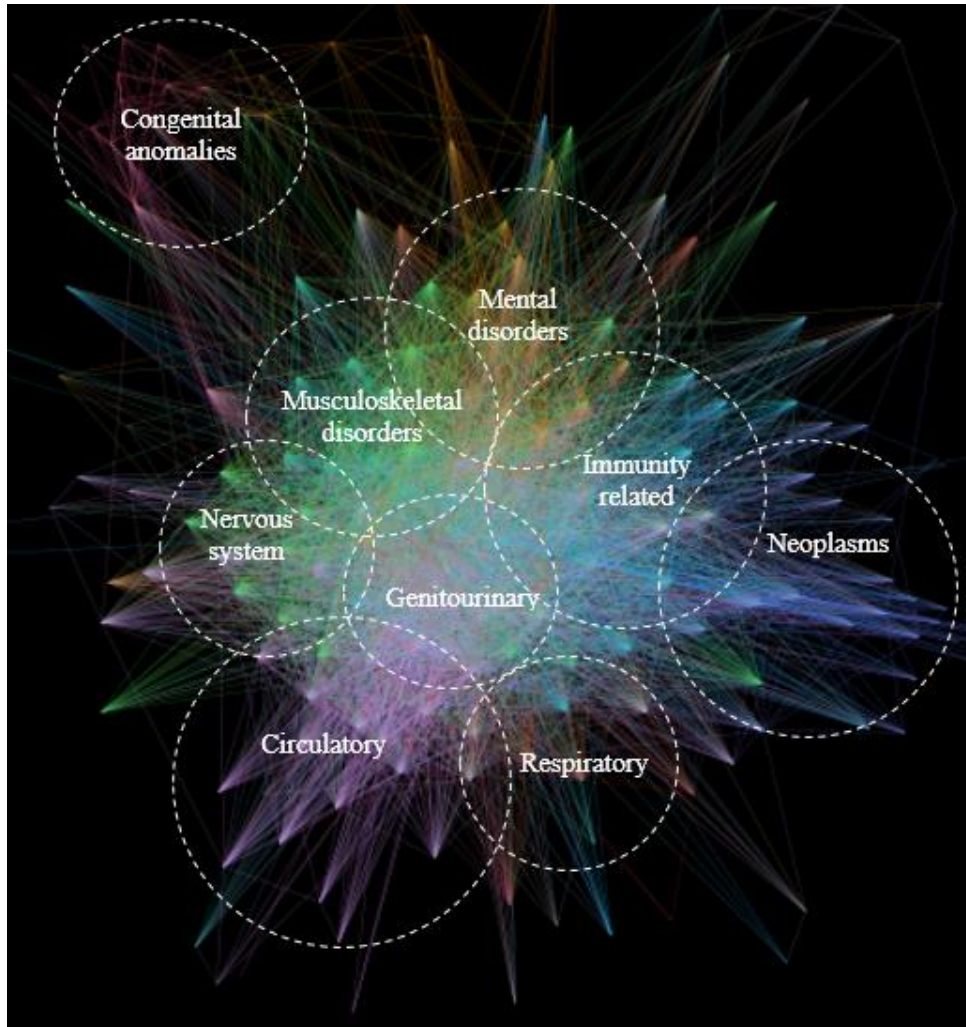
Figure 7: Disease co-occurrence network created with inpatient visits EHRs for Arizona; nodes and connecting edges are colored based on body system classification of the 3-digit ICD-9 diagnosis codes

Color	Body system	Proportion of nodes in network (%)
	Circulatory system	17.73
	Neoplasms	15.76
	Immunity related	11.82
	Nervous system	11.34
	Mental disorders	11.33
	Congenital anomalies	8.37
	Musculoskeletal disorders	5.92
	Genitourinary	5.91
	Respiratory	4.93
	Others (Digestive, blood, infectious, skin-related)	6.89

Figure 8: Color legend for Figure 7

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

The DCN for the entire Arizona dataset is shown in Figure 7. We can see that diseases/conditions in one body system commonly co-occur with diseases/conditions belonging to other body systems. The colors in the disease network indicate diseases/conditions belonging to different body systems as listed in Figure 8. DCNs for each body system (circulatory system, musculoskeletal disorders, nervous system, mental disorders, etc.) are shown in the Appendix C.

The weighted degree of a node is the sum of weights of all edges connected to the given node. It is also called strength of the node (Barrat et al. 2004). The top five diagnoses ranked based on highest weighted degree for each county is provided in Appendix D. It can be noted that the list of top five diagnoses is similar across regions, except that the order of these diseases for each county may be different. We compared the top five diagnoses for the networks across time periods, but we did not see any variation in the list as compared to the overall DCN based on the entire Arizona inpatient visits EHR dataset.

The edge weights of the disease co-occurrence network vary from zero to one, and their values indicate the extent to which pairs of diseases/conditions co-occur across inpatient visits. For a disease pair, a value of edge weight closer to zero suggests that the relationship between a pair of diseases/conditions is rare. Setting a threshold value of edge weight filters out co-occurrence relationships between disease pairs that are rare. We therefore vary the threshold value for the edge weight and examine changes in the resulting network structure. Table 5 shows the number of nodes, number of edges and density of networks created by varying the threshold value for edge weights.

Table 5: Threshold for edge weights and resultant network structural properties

Threshold for edge weight	Number of nodes	Number of edges	Density
0	203	4851	0.24
0.01	197	3712	0.19
0.02	180	2149	0.13
0.03	158	1352	0.11
0.04	142	915	0.09
0.05	128	685	0.08
0.1	83	246	0.07
0.2	34	51	0.09
0.3	21	23	0.11
0.4	6	4	0.27

We now analyze the networks created after varying the threshold value of edge weights. The highest edge weight in the network is 0.8959, corresponding to edge between nodes chronic renal failure and hypertensive renal disease, followed by 0.5191, weight of edge between hypertension and disorders of lipoid metabolism. Disorders of lipoid metabolism is linked to two other diagnoses, chronic ischemic heart disease and diabetes with edge weights 0.4472 and 0.4392 respectively. We can thus create a simple network with six nodes and four edges by setting threshold for edge weights as 0.4 as shown in Figure 9. We see that the network has two disjoint components: one consisting of disorders of lipoid metabolism linked to diabetes, chronic ischemic heart disease and hypertension and the second being a dyad of two diseases/conditions, chronic renal failure and hypertensive renal disease. All these diseases are in the top 10 prevalent diseases in Arizona (see appendix). One interesting aspect is the structure of the four-node subgraph where disorders of lipoid metabolism is focal and has other highly prevalent linked to each other through it indirectly. This indicates that pairwise occurrence of disorders of lipoid metabolism (such as obesity) with each of the other

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

disease/condition is high, but all three of these other conditions may not occur simultaneously with lipoid metabolism disorders during one single inpatient visit.

The 21-node network with threshold for edge weights as 0.3 (Figure 10) connects the renal failure and renal disease with the other four nodes as well introduces several independent components into the network. A color legend for Figure 10 through Figure 18 is shown in Figure 19. Diseases belonging to congenital disorders, neoplasms as well as mental disorders appear to co-occur with diseases that fall under the same body system category. However, there are some dyads that have diseases from different body systems such as chronic liver disease - alcohol dependence, hemiplegia - occlusion of cerebral arteries, organic psychotic conditions - cerebral degenerations.
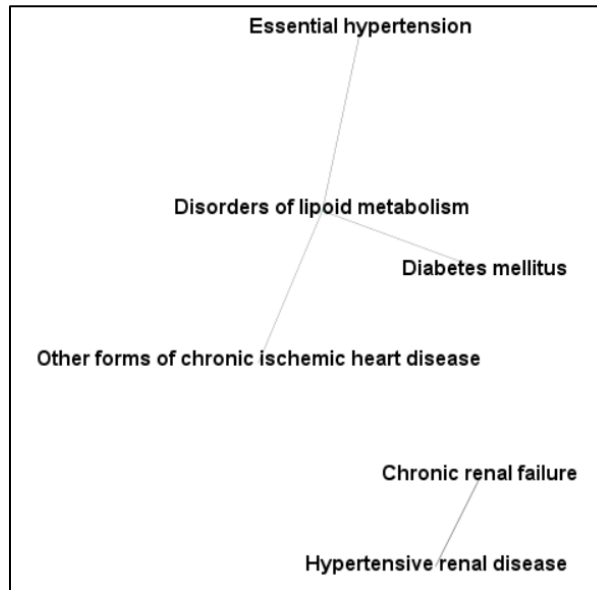
Figure 9: Six node network with threshold for edge weight as 0.4
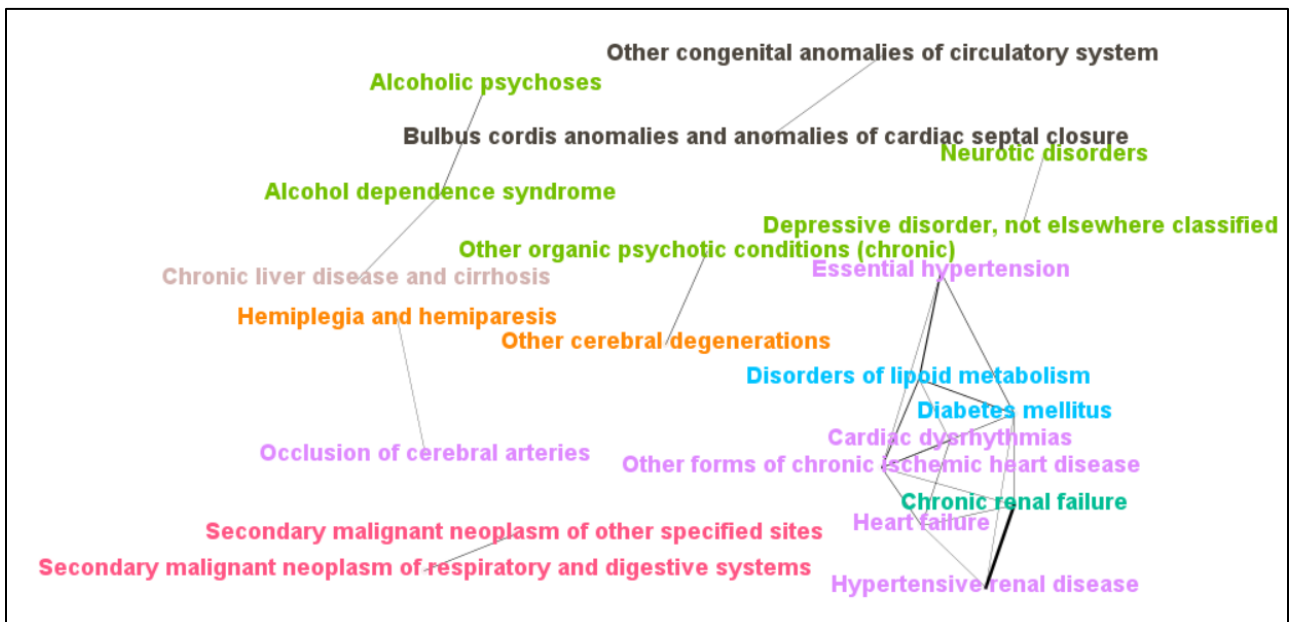
Figure 10: 21 node network with threshold for edge weight as 0.3

When the edge weight threshold value is reduced further to 0.2, the resulting network has 34 nodes as shown in Figure 11. Further reducing the threshold value to 0.1 results in a 83 node network (see Figure 12), with nine groups of nodes containing diseases including congenital anomalies, neoplasm and mental disorders.
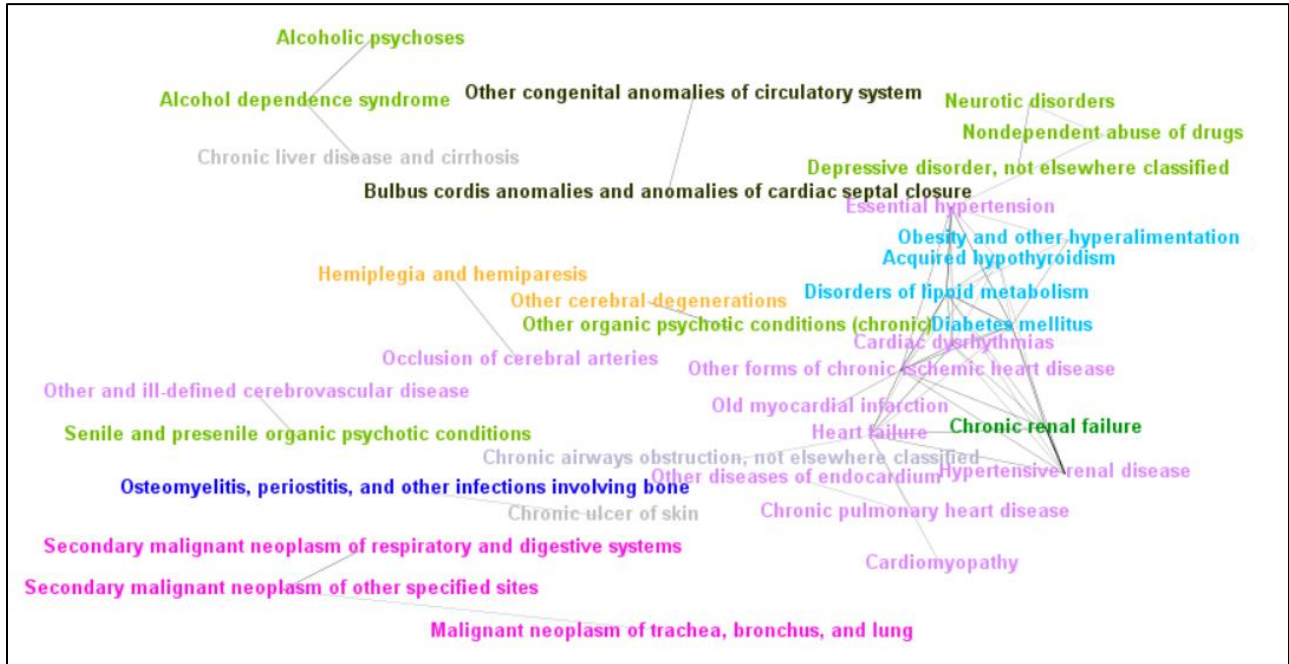
Figure 11: 34 node network with threshold for edge weight as 0.2
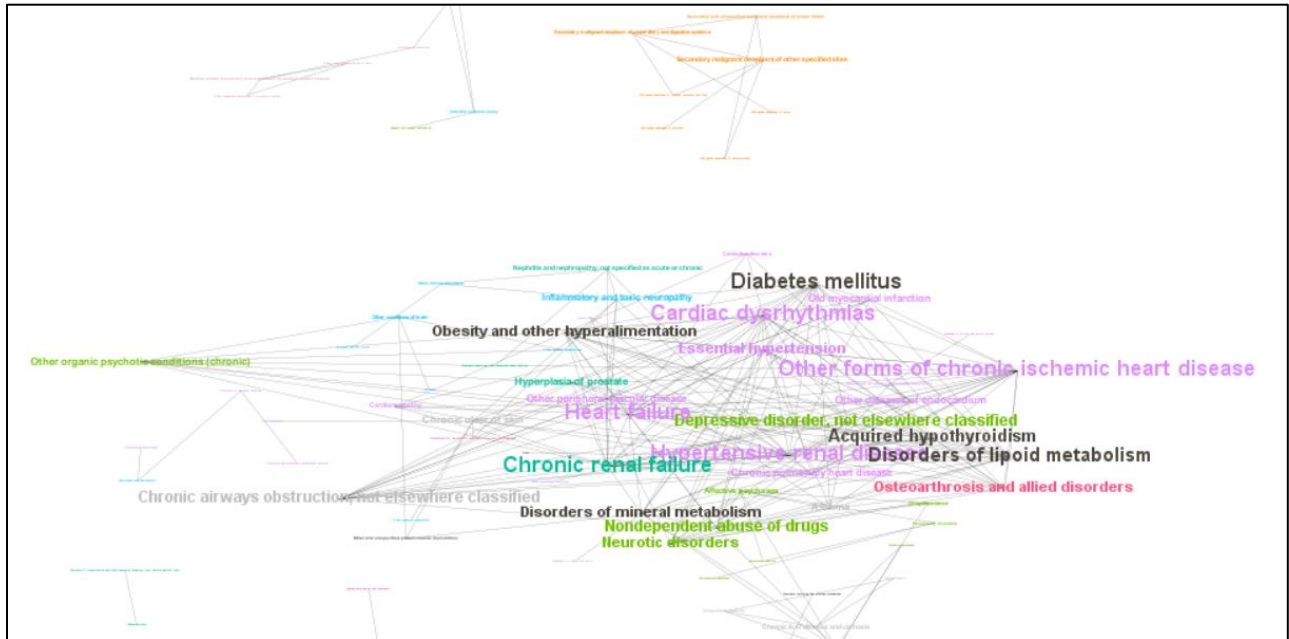
Figure 12: 83 node network with threshold for edge weight as 0.1

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

An ego network consists of a focal node ("ego") and nodes directly connected to it, with corresponding edges between all these nodes. Ego networks can be used to zoom into one single node and understand its neighborhood network consisting of nodes it is directly connected to. We examine the ego network for each of the six diseases in Figure 9 because these are the diseases that co-occur most frequently. These ego networks shown in Figure 13 through Figure 18 have considerable overlap. This is interesting because it indicates that in Arizona, these diseases are frequently co-occurring with a common set of other conditions – refer to Appendix E for this list.
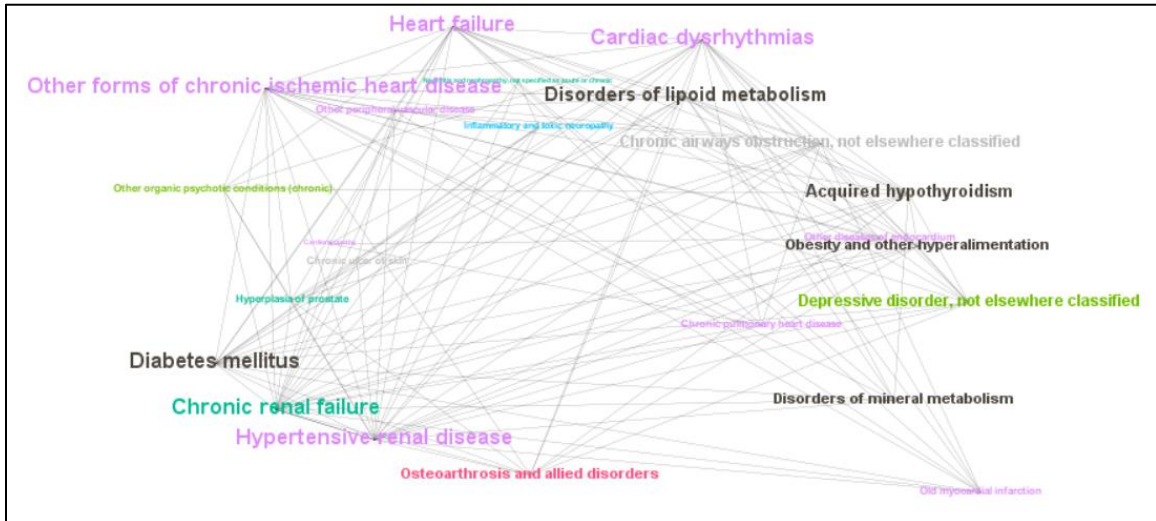
Figure 13: Chronic renal failure and hypertensive renal disease co-occurs with 22 diagnoses in network with threshold for edge weight as 0.1
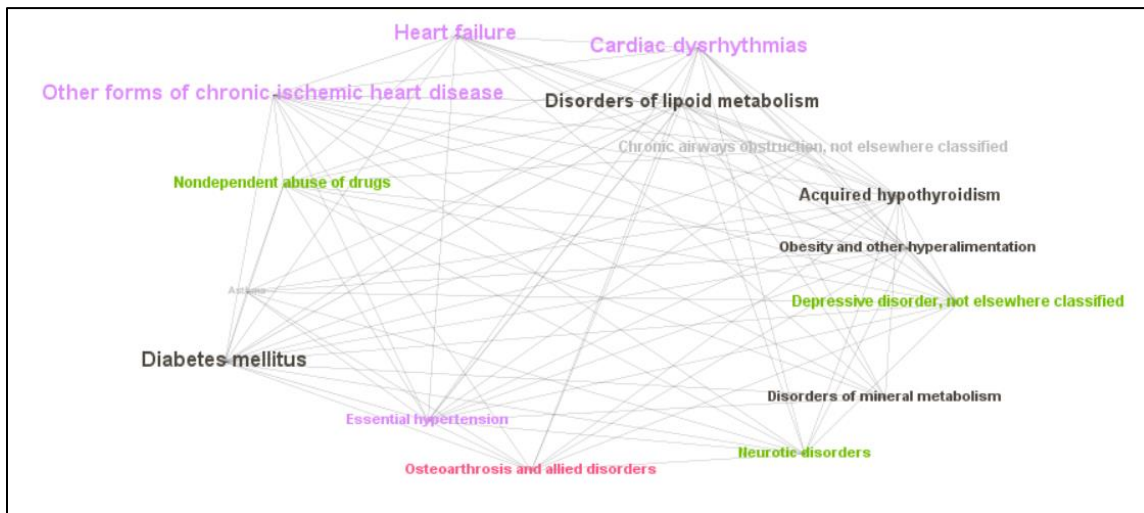
Figure 14: Hypertension co-occurs with 15 diagnoses in network with threshold for edge weight as 0.1

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu
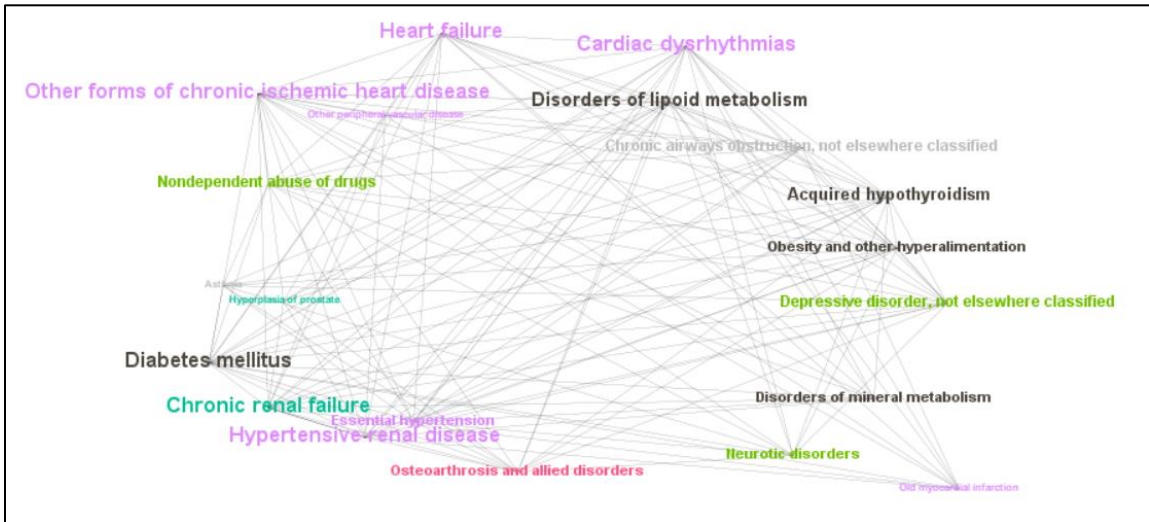
Figure 15: Disorders of lipoid metabolism co-occurs with 20 diagnoses in network with threshold for edge weight as 0.1
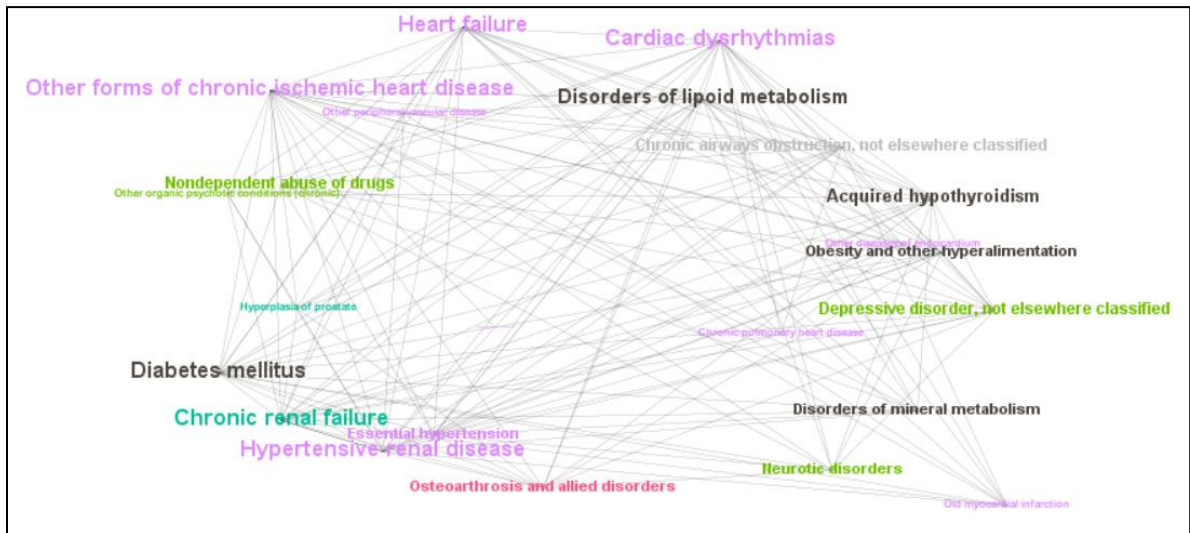
Figure 16: Chronic ischemic heart disease co-occurs with 23 diagnoses in network with threshold for edge weight as 0.1

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu
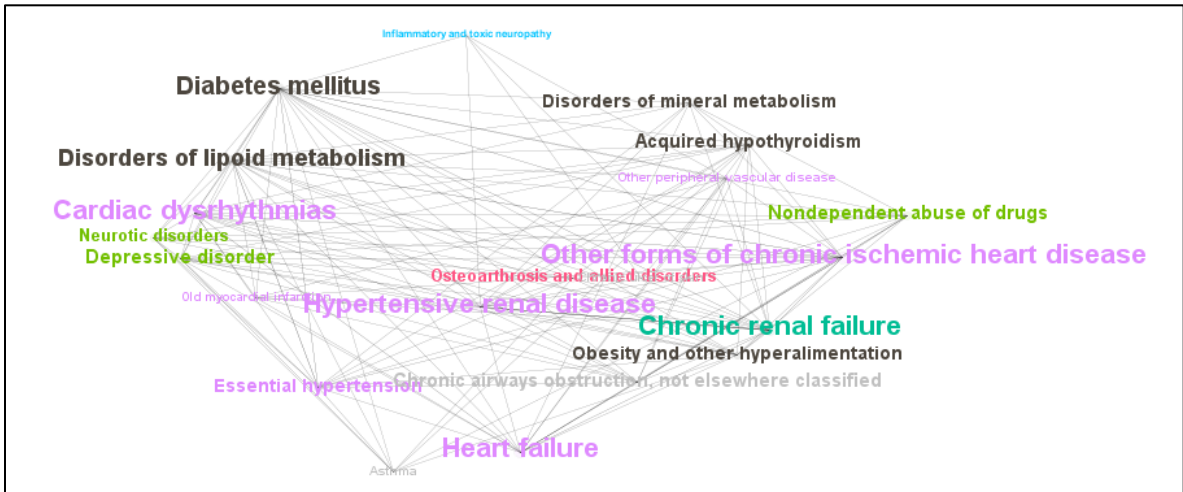
Figure 17: Diabetes co-occurs with 20 diagnoses in network with threshold for edge weight as 0.1
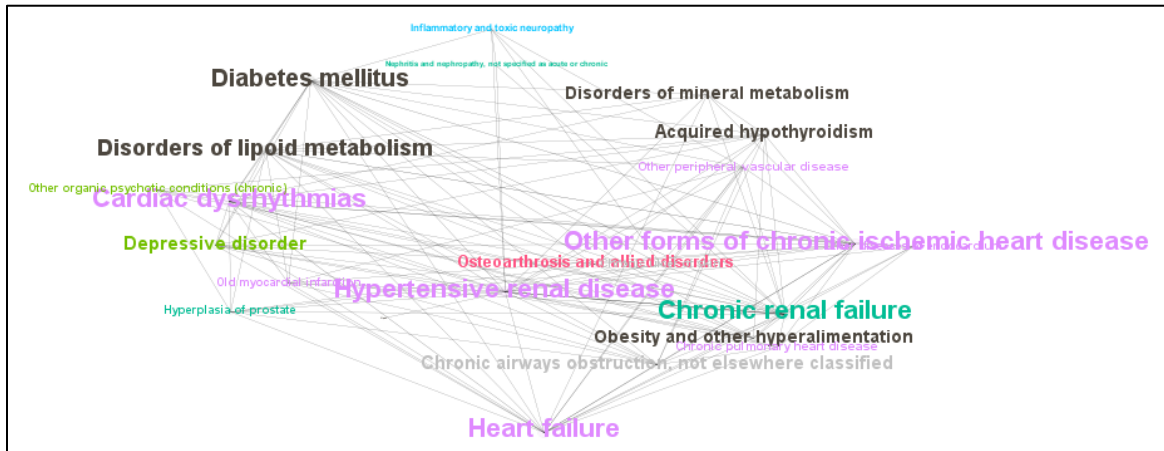
Figure 18: Hypertensive renal disease co-occurs with 22 diagnoses in network with threshold for edge weight as 0.1

Color	Body system
	Circulatory system
	Mental disorders
	Immunity related
	Congenital anomalies
	Nervous system
	Neoplasms
	Genitourinary
	Musculoskeletal
	Digestive system and other disorders

Figure 19: Color legend for figures - Figure 10 through Figure 18

We list the top three co-occurring diagnoses for the top six co-occurrent diseases in Table 6. The list of co-occurring diagnoses for each disease can be verified by medical professionals and used as a lookup of common co-occurring diseases/conditions corresponding to diseases already diagnosed.

29

The disease co-occurrence network is a decision support tool for practitioners useful while treating patients admitted with two or more diseases/conditions.

Table 6: Top three co-occurring diagnoses for the top six co-occurrent diseases

Disease	Co-occurring diagnosis 1	Co-occurring diagnosis 2	Co-occurring diagnosis 3
chronic renal failure	hypertensive renal disease	diabetes	heart failure
hypertensive renal disease	chronic renal failure	diabetes	heart failure
diabetes	hypertension	chronic ischemic heart disease	chronic renal failure
hypertension	disorders of lipoid metabolism	diabetes	chronic ischemic heart disease
disorders of lipoid metabolism	hypertension	chronic ischemic heart disease	diabetes
chronic ischemic heart disease	disorders of lipoid metabolism	heart failure	cardiac dysrhythmias

6. Discussion and conclusion

In the era of big data, healthcare information system technology and machine learning techniques can be used on existing data for knowledge discovery and explanatory statistical modeling, as well as for making accurate predictions. Data analytics can be used to reduce costs, improve personal care, and alleviate existing operational problems in chronic disease management.

We explored the spatial-temporal patterns in occurrence of one or more chronic diseases recorded in the inpatient visits EHRs of hospitals in Arizona in our study. The study was conducted in two phases. In the first phase, we investigated the incidence of individual chronic diseases and spatial-temporal patterns in related visit counts and average visit costs using summary statistics, simple visualizations, and interactive dashboards. In the second phase, we created networks using inpatient visits EHR data to study the co-occurrence of multiple chronic diseases.

In the first phase, we analyzed 915,355 inpatient visits with chronic diseases as admitting diagnoses in hospitals across 11 Arizona regions during the time period 2012-16. Each record in the EHR had information related to the visit (visit cost, discharge status, payer type), patient (age, gender, marital status, ethnicity), and procedures and diagnoses (diagnosis codes, body system category of diagnoses). We summarized the inpatient visits data and provided simple visualizations, as well as created an interactive dashboard enabling users to explore the data and make inferences.

In the second phase, we discussed the phenomenon of simultaneous co-occurrence of two or more diseases in an inpatient visit and presented the disease co-occurrence network (DCN) as a useful artifact for studying this phenomenon. The DCN was created using the inpatient visits EHR, where diseases/conditions represented as nodes and links between each pair of nodes indicate the evidence of co-occurrence between them. We proposed a novel metric that can be used as edge weight in the DCN. We summarized and visualized the networks created with the Arizona inpatient visit EHR dataset, as well as subset data corresponding to each county and time period.

Our study has a few limitations. First, we considered the inpatient visit EHR collected and processed by ADHS from all public licensed hospitals in Arizona without linking history of patients admitted in each visit. The patient history is an important factor affecting disease co-occurrence as well as spatial-temporal prevalence and incidence patterns. For example, it would be interesting to see the revisit pattern of patients across counties over different periods and how disease/conditions progress over time for these patients. Secondly, we have considered time periods of six months, as that is the least time grain available for analysis in the inpatient visit EHR provided by ADHS. Quarterly or monthly analysis of prevalence, incidence and co-occurrence of chronic diseases across counties can yield better insights into spatial-temporal patterns. Third, we consider the ICD-9 diagnosis codes for developing the disease co-occurrence network and making interpretations, but these codes are subject to manual errors by medical coders and misdiagnosis by physicians. Grouping the 5-digit ICD-9 codes to 3-digit codes for chronic diseases helps overcome classification errors within same sub-groups (i.e., having same 3-digital codes), thus partially resolving ICD-9 miscoding errors. Lastly, since our analysis is at a population level, it can aid health policy and decision-making, but cannot derive individual level insights for coping with chronic diseases. Future analysis can build upon our current methods for exploring inpatient visits EHR for chronic disease combined with patient history, higher temporal grain and patient-centric models thus overcoming the above limitations.

Our study has several implications. We observed that diabetes is the most prevalent health condition in Apache, whereas hypertension is the most prevalent condition observed in inpatient visits in other regions. Chronic diseases such as hypertension, disorders of lipid metabolism, and diabetes are more prevalent inpatient visits made during the second half of the year as compared to the first. The number of inpatient visits with neurotic disorders related diagnoses are increasing at a rate of 18 percent every six months during the period 2012-2015. Circulatory system related diseases/conditions have the highest prevalence as admitting diagnoses in visits across Arizona, accounting for 19 to 38 percent of total visits in the 11 Arizona regions. Of all patients suffering from chronic conditions, 54 percent are in the 50-79 age group, suggesting that this is the most vulnerable time period for residents in Arizona. The average cost for inpatient visits in the 0-9 years age group has increased systematically, from \$68,211 in the first half of 2012 to \$100,689 in the second half of 2016, a 48 percent increase in four years. The top five diseases with highest potential to co-occur with other diseases/conditions in an inpatient visit in Arizona are chronic renal failure, chronic ischemic heart disease, disorders of lipid metabolism, diabetes, and hypertensive renal disease. Chronic diseases are observed to co-occur more often in the inpatient visits in hospitals in Mohave and Navajo than in hospital patient visits in other regions. The underlying causes for higher chronic disease co-occurrence in these two counties need to be investigated from a clinical (treatment procedures, medications, distinctive patient socio-demographic distributions, etc.) as well as administrative (medical coding in hospitals in these counties, insurance payer types and claims, information systems technology, etc.) perspectives.

Disorders of lipid metabolism is a disease/condition that has maximum potential to simultaneously occur with other diseases in hospital inpatient visits in Navajo, whereas chronic ischemic heart disease has highest potential to co-occur with other diseases/conditions in hospital inpatient visits in Pima. Disorders of lipid metabolism such as high cholesterol or hyperlipidemia (high amount of fat particles in blood) are linked to diabetes, chronic ischemic heart disease, and hypertension such that occurrence of disorders of lipid metabolism suggests a high possibility of any one or multiple of the other three health conditions. The top 10 prevalent and co-occurring chronic diseases identified in this study are each known to be caused by unhealthy and stressful

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

lifestyles. Our findings thus underscore the importance of diet and regular exercise in an individual's health and wellbeing.

The DCN is a tool that can be used to model the disease co-occurrence phenomenon and used in several healthcare applications. For problems such as forecasting disease progression, reducing health costs, predicting time to readmission, and predicting rate of mortality, signals from the DCN can be used to account for the disease co-occurrence phenomenon in machine learning problems.

In the future, we plan to extend our current work for 2016 and 2017 EHR data that uses ICD-10 coding standard for recording diagnoses and procedures related to an inpatient visit. We plan to extend application of DCN in modeling disease progression and predicting time to readmission for different patient demographics.

Our study explores chronic disease related inpatient visits in Arizona hospitals using a data-centric exploratory approach. We hope that the findings presented in this paper can be helpful for policy makers for planning and facilitating informed decision making for chronic disease management in Arizona.

7. References

ADHS Public Health Prevention. 2018. "Tobacco & Chronic Disease," (available at https://www.azdhs.gov/prevention/tobacco-chronic-disease/index.php).

Barrat, A., Barthélemy, M., and Vespignani, A. 2004. "Weighted evolving networks: Coupling topology and weight dynamics," *Physical Review Letters*, (92:22), pp. 228701–1.

Centers for Disease Control and Prevention, and National Center for Health Statistics. 2013. "ICD-9-CM - International Classification of Diseases, Ninth Revision, Clinical Modification," *Classification of Diseases, Functioning, and Disability*, (2008), pp. 1–2.

Centers for Medicare and Medicaid Services. 2016. "National Health Expenditure Accounts," (available at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html; retrieved October 23, 2017).

Charlson, M. E., Pompei, P., Ales, K. L., and MacKenzie, C. R. 1987. "A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation," *Journal of Chronic Diseases*, (40:5), pp. 373–383.

Divo, M. J., Casanova, C., Marin, J. M., Pinto-Plata, V. M., De-Torres, J. P., Zulueta, J. J., Cabrera, C., Zagaceta, J., Sanchez-Salcedo, P., Berto, J., Davila, R. B., Alcaide, A. B., Cote, C., and Celli, B. R. 2015. "COPD comorbidities network," *European Respiratory Journal*, (46:3), pp. 640–650.

Elixhauser, A., Steiner, C., Harris, D. R., and Coffey, R. M. 1998. "Comorbidity measures for use with administrative data," *Medical care*, (36:1), pp. 8–27.

Hampton, K. N., Goulet, L. S., Marlow, C., and Rainie, L. 2012. "Why most Facebook users get more than they give," *Pew Research Center*.

Hidalgo, C. A., Blumm, N., Barabási, A. L., and Christakis, N. A. 2009. "A Dynamic Network Approach for the Study of Human Phenotypes," *PLoS Computational Biology*, (5:4).

Klimek, P., Kautzky-Willer, A., Chmiel, A., Schiller-Frühwirth, I., and Thurner, S. 2015. "Quantification of Diabetes Comorbidity Risks across Life Using Nation-Wide Big Claims Data," *PLoS Computational Biology*, (11:4), pp. 1–16.

Liu, J., Ma, J., Wang, J., Zeng, D. D., Song, H., Wang, L., and Cao, Z. 2016. "Comorbidity analysis according to sex and age in hypertension patients in China," *International Journal of Medical*

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Sciences, (13:2), pp. 99–107.

Mannino, D. M., Thorn, D., Swensen, A., and Holguin, F. 2008. "Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD," *European Respiratory Journal*, (32:4), pp. 962–969.

Ogle, K. S., Swanson, G. M., Woods, N., and Azzouz, F. 2000. "Cancer and comorbidity: Redefining chronic diseases," *Cancer*, (88:3), pp. 653–663.

Roque, F. S., Jensen, P. B., Schmock, H., Dalgaard, M., Andreatta, M., Hansen, T., Søeby, K., Bredkjær, S., Juul, A., Werge, T., Jensen, L. J., and Brunak, S. 2011. "Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts," *PLoS Computational Biology*, (7:8), p. e1002141.

Schellevis, F. G., van der Velden, J., van de Lisdonk, E., van Eijk, J. T. M., and van Weel, C. 1993. "Comorbidity of chronic diseases in general practice," *Journal of Clinical Epidemiology*, (46:5), pp. 469–473.

Shi, C., Kong, X., Yu, P. S., Xie, S., and Wu, B. 2012. "Relevance Search in Heterogeneous Networks," in *Proceedings of the 15th International Conference on Extending Database Technology*, pp. 180–191.

Slee, V. N. 1978. "The International Classification of Diseases: ninth revision (ICD-9)," *Annals of Internal Medicine*, (88:3), pp. 424–426.

Srinivasan, K., Currim, F., and Ram, S. 2018. "Predicting High Cost Patients at Point of Admission using Network Science," *IEEE Journal of Biomedical and Health Informatics*, (22:6), pp. 1970–1977.

Steinhaeuser, K., and Chawla, N. V. 2009. *Social Computing and Behavioral Modeling*.

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

8. Appendix

Appendix A - Top 10 prevalent diseases in Arizona and 11 regions

A.1 Arizona

ICD code	Prevalence	Disease/Condition
401	887,014	Hypertension
272	658,100	Disorders of lipoid metabolism
250	538,254	Diabetes
305	446,268	Nondependent abuse of drugs
427	401,372	Cardiac dysrhythmias
414	399,950	Other forms of chronic ischemic heart disease
244	328,957	Acquired hypothyroidism
278	322,282	Obesity and other hyper alimentation
585	291,362	Chronic renal failure
300	274,895	Neurotic disorders

A.2 Apache

ICD code	Prevalence	Disease/Condition
250	3,084	Diabetes
401	2,976	Hypertension
272	1,810	Disorders of lipoid metabolism
305	1,629	Nondependent abuse of drugs
278	1,142	Obesity and another hyper alimentation
414	1,092	Other forms of chronic ischemic heart disease
303	1,037	Alcohol dependence syndrome
585	1,012	Chronic renal failure
244	929	Acquired hypothyroidism
427	861	Cardiac dysrhythmias

A.3 Cochise

ICD code	Prevalence	Disease/Condition
401	40,711	Hypertension
272	29,290	Disorders of lipoid metabolism
250	24,043	Diabetes
414	19,006	Other forms of chronic ischemic heart disease
427	18,316	Cardiac dysrhythmias
305	15,296	Nondependent abuse of drugs
244	14,797	Acquired hypothyroidism
585	13,069	Chronic renal failure
715	11,947	Osteoarthrosis and allied disorders
311	11,754	Depressive disorder, not elsewhere classified

A.4 Coconino

ICD code	Prevalence	Disease/Condition
401	14,012	Hypertension
272	10,341	Disorders of lipoid metabolism
250	10,301	Diabetes
305	8,433	Nondependent abuse of drugs
278	7,112	Obesity and other hyper alimentation
244	6,139	Acquired hypothyroidism
427	5,549	Cardiac dysrhythmias
414	5,149	Other forms of chronic ischemic heart disease
311	5,124	Depressive disorder, not elsewhere classified
585	4,791	Chronic renal failure

A.5 Gila

ICD code	Prevalence	Disease/Condition
401	16,597	Hypertension
250	10,713	Diabetes
272	10,044	Disorders of lipoid metabolism
305	8,414	Nondependent abuse of drugs
414	7,674	Other forms of chronic ischemic heart disease
427	6,452	Cardiac dysrhythmias
244	5,892	Acquired hypothyroidism
585	5,260	Chronic renal failure
428	5,244	Heart failure
278	4,651	Obesity and other hyper alimentation

A.6 Maricopa

ICD code	Prevalence	Disease/Condition
401	332,578	Hypertension
272	243,800	Disorders of lipoid metabolism
250	204,808	Diabetes
305	184,440	Nondependent abuse of drugs
427	148,039	Cardiac dysrhythmias
414	142,490	Other forms of chronic ischemic heart disease
278	127,210	Obesity and other hyper alimentation
244	124,749	Acquired hypothyroidism
585	111,015	Chronic renal failure
300	106,990	Neurotic disorders

A.7 Mohave

ICD code	Prevalence	Disease/Condition
401	36,290	Hypertension
272	26,858	Disorders of lipoid metabolism
250	22,219	Diabetes
305	22,054	Nondependent abuse of drugs
414	21,698	Other forms of chronic ischemic heart disease
427	18,184	Cardiac dysrhythmias
428	14,234	Heart failure
585	13,385	Chronic renal failure
278	13,282	Obesity and other hyper alimentation
244	12,956	Acquired hypothyroidism

A.8 Navajo

ICD code	Prevalence	Disease/Condition
401	9,922	Hypertension
272	6,730	Disorders of lipoid metabolism
250	5,937	Diabetes
305	5,177	Nondependent abuse of drugs
414	4,732	Other forms of chronic ischemic heart disease
278	4,506	Obesity and other hyper alimentation
427	3,857	Cardiac dysrhythmias
244	3,249	Acquired hypothyroidism
311	2,882	Depressive disorder, not elsewhere classified
715	2,778	Osteoarthrosis and allied disorders

A.9 Pima

ICD code	Prevalence	Disease/Condition
401	116,529	Hypertension
272	83,101	Disorders of lipoid metabolism
250	70,620	Diabetes
305	62,316	Nondependent abuse of drugs
427	52,678	Cardiac dysrhythmias
414	49,429	Other forms of chronic ischemic heart disease
244	44,512	Acquired hypothyroidism
311	41,095	Depressive disorder, not elsewhere classified
278	40,640	Obesity and other hyper alimentation
300	40,533	Neurotic disorders

A.10 Pinal

ICD code	Prevalence	Disease/Condition
401	55,570	Hypertension
272	41,867	Disorders of lipoid metabolism
250	38,283	Diabetes
305	27,298	Nondependent abuse of drugs
414	25,021	Other forms of chronic ischemic heart disease
427	23,030	Cardiac dysrhythmias
278	21,647	Obesity and other hyper alimentation
244	18,737	Acquired hypothyroidism
585	17,955	Chronic renal failure
403	15,550	Hypertensive renal disease

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

A.11 Yavapai

ICD code	Prevalence	Disease/Condition
401	32,854	Hypertension
272	22,173	Disorders of lipoid metabolism
427	15,620	Cardiac dysrhythmias
250	15,456	Diabetes
414	14,661	Other forms of chronic ischemic heart disease
244	14,054	Acquired hypothyroidism
305	13,456	Nondependent abuse of drugs
278	9,431	Obesity and other hyper alimentation
311	8,731	Depressive disorder, not elsewhere classified
300	8,674	Neurotic disorders

A.12 Yuma

ICD code	Prevalence	Disease/Condition
401	188,267	Hypertension
272	151,382	Disorders of lipoid metabolism
250	113,317	Diabetes
414	90,528	Other forms of chronic ischemic heart disease
427	88,289	Cardiac dysrhythmias
305	79,760	Nondependent abuse of drugs
278	70,711	Obesity and other hyper alimentation
244	69,820	Acquired hypothyroidism
585	63,958	Chronic renal failure
300	58,693	Neurotic disorders

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Appendix B - List of 290 3-digit ICD-9 diagnosis codes identified as chronic conditions/diseases

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
042	Human immunodeficiency virus [HIV] disease	Infectious and parasitic disease		
046	Slow virus infection of central nervous system	Infectious and parasitic disease	Y	
076	Trachoma	Infectious and parasitic disease	Y	
135	Sarcoidosis	Infectious and parasitic disease		
137	Late effects of tuberculosis	Infectious and parasitic disease	Y	
138	Late effects of acute poliomyelitis	Infectious and parasitic disease		
139	Late effects of other infectious and parasitic diseases	Infectious and parasitic disease		Y
140	Malignant neoplasm of lip	Neoplasms	Y	
141	Malignant neoplasm of tongue	Neoplasms		
142	Malignant neoplasm of major salivary glands	Neoplasms	Y	
143	Malignant neoplasm of gum	Neoplasms	Y	
144	Malignant neoplasm of floor of mouth	Neoplasms	Y	
145	Malignant neoplasm of other and unspecified parts of mouth	Neoplasms	Y	
146	Malignant neoplasm of oropharynx	Neoplasms		
147	Malignant neoplasm of nasopharynx	Neoplasms	Y	
148	Malignant neoplasm of hypopharynx	Neoplasms	Y	
149	Malignant neoplasm of other and ill-defined sites within the lip, oral cavity, and pharynx	Neoplasms		Y
150	Malignant neoplasm of esophagus	Neoplasms		
151	Malignant neoplasm of stomach	Neoplasms		
152	Malignant neoplasm of small intestine, including duodenum	Neoplasms	Y	
153	Malignant neoplasm of colon	Neoplasms		
154	Malignant neoplasm of rectum, rectosigmoid junction, and anus	Neoplasms		
155	Malignant neoplasm of liver and intrahepatic bile ducts	Neoplasms		
156	Malignant neoplasm of gallbladder and extrahepatic bile ducts	Neoplasms		
157	Malignant neoplasm of pancreas	Neoplasms		
158	Malignant neoplasm of retroperitoneum and peritoneum	Neoplasms		Y
159	Malignant neoplasm of other and ill-defined sites within the digestive organs and peritoneum	Neoplasms	Y	

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
160	Malignant neoplasm of nasal cavities, middle ear, and accessory sinuses	Neoplasms	Y	
161	Malignant neoplasm of larynx	Neoplasms		Y
162	Malignant neoplasm of trachea, bronchus, and lung	Neoplasms		
163	Malignant neoplasm of pleura	Neoplasms	Y	
164	Malignant neoplasm of thymus, heart, and mediastinum	Neoplasms	Y	
165	Malignant neoplasm of other and ill-defined sites within the respiratory system and intrathoracic organs	Neoplasms	Y	
170	Malignant neoplasm of bone and articular cartilage	Neoplasms		
171	Malignant neoplasm of connective and other soft tissue	Neoplasms		
172	Malignant melanoma of skin	Neoplasms		
174	Malignant neoplasm of female breast	Neoplasms		
175	Malignant neoplasm of male breast	Neoplasms	Y	
176	Kaposi's sarcoma	Neoplasms	Y	
179	Malignant neoplasm of uterus, part unspecified	Neoplasms		Y
180	Malignant neoplasm of cervix uteri	Neoplasms		
181	Malignant neoplasm of placenta	Neoplasms	Y	
182	Malignant neoplasm of body of uterus	Neoplasms		
183	Malignant neoplasm of ovary and other uterine adnexa	Neoplasms		
184	Malignant neoplasm of other and unspecified female genital organs	Neoplasms		Y
185	Malignant neoplasm of prostate	Neoplasms		
186	Malignant neoplasm of testis	Neoplasms		Y
187	Malignant neoplasm of penis and other male genital organs	Neoplasms	Y	
188	Malignant neoplasm of bladder	Neoplasms		
189	Malignant neoplasm of kidney and other and unspecified urinary organs	Neoplasms		
190	Malignant neoplasm of eye	Neoplasms	Y	
191	Malignant neoplasm of brain	Neoplasms		
192	Malignant neoplasm of other and unspecified parts of nervous system	Neoplasms	Y	
193	Malignant neoplasm of thyroid gland	Neoplasms		
194	Malignant neoplasm of other endocrine glands and related structures	Neoplasms		Y
195	Malignant neoplasm of other and ill-defined sites	Neoplasms		Y
196	Secondary and unspecified malignant neoplasm of lymph nodes	Neoplasms		
197	Secondary malignant neoplasm of respiratory and digestive systems	Neoplasms		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
198	Secondary malignant neoplasm of other specified sites	Neoplasms		
199	Malignant neoplasm without specification of site	Neoplasms		
200	Lymphosarcoma and reticulosarcoma	Neoplasms		
201	Hodgkin's disease	Neoplasms		Y
202	Other malignant neoplasm of lymphoid and histiocytic tissue	Neoplasms		
203	Multiple myeloma and immunoproliferative neoplasms	Neoplasms		
204	Lymphoid leukemia	Neoplasms		
205	Myeloid leukemia	Neoplasms		
206	Monocytic leukemia	Neoplasms	Y	
207	Other specified leukemia	Neoplasms	Y	
208	Leukemia of unspecified cell type	Neoplasms		Y
230	Carcinoma in situ of digestive organs	Neoplasms	Y	
231	Carcinoma in situ of respiratory system	Neoplasms	Y	
233	Carcinoma in situ of breast and genitourinary system	Neoplasms		
234	Carcinoma in situ of other and unspecified sites	Neoplasms	Y	
240	Simple and unspecified goiter	Endocrine, nutritional, and metabolic diseases and immunity disorders		Y
241	Nontoxic nodular goiter	Endocrine, nutritional, and metabolic diseases and immunity disorders		
242	Thyrotoxicosis with or without goiter	Endocrine, nutritional, and metabolic diseases and immunity disorders		
243	Congenital hypothyroidism	Endocrine, nutritional, and metabolic diseases and immunity disorders	Y	
244	Acquired hypothyroidism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
245	Thyroiditis	Endocrine, nutritional, and metabolic diseases and immunity disorders		
250	Diabetes	Endocrine, nutritional, and metabolic diseases and immunity disorders		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
251	Other disorders of pancreatic internal secretion	Endocrine, nutritional, and metabolic diseases and immunity disorders		
252	Disorders of parathyroid gland	Endocrine, nutritional, and metabolic diseases and immunity disorders		
253	Disorders of the pituitary gland and its hypothalamic control	Endocrine, nutritional, and metabolic diseases and immunity disorders		
255	Disorders of adrenal glands	Endocrine, nutritional, and metabolic diseases and immunity disorders		
256	Ovarian dysfunction	Endocrine, nutritional, and metabolic diseases and immunity disorders		
257	Testicular dysfunction	Endocrine, nutritional, and metabolic diseases and immunity disorders		
258	Polyglandular dysfunction and related disorders	Endocrine, nutritional, and metabolic diseases and immunity disorders	Y	
260	Kwashiorkor	Endocrine, nutritional, and metabolic diseases and immunity disorders	Y	
261	Nutritional marasmus	Endocrine, nutritional, and metabolic diseases and immunity disorders		
262	Other severe protein-calorie malnutrition	Endocrine, nutritional, and metabolic diseases and immunity disorders		
263	Other and unspecified protein-calorie malnutrition	Endocrine, nutritional, and metabolic diseases and immunity disorders		
268	Vitamin D deficiency	Endocrine, nutritional, and metabolic diseases and immunity disorders		
270	Disorders of amino-acid transport and metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
271	Disorders of carbohydrate transport and metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
272	Disorders of lipoid metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
273	Disorders of plasma protein metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
274	Gout	Endocrine, nutritional, and metabolic diseases and immunity disorders		
275	Disorders of mineral metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
277	Other and unspecified disorders of metabolism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
278	Obesity and other hyper alimentation	Endocrine, nutritional, and metabolic diseases and immunity disorders		
279	Disorders involving the immune mechanism	Endocrine, nutritional, and metabolic diseases and immunity disorders		
282	Hereditary hemolytic anemias	Diseases of blood and blood-forming organs		
283	Acquired hemolytic anemias	Diseases of blood and blood-forming organs		Y
284	Aplastic anemia	Diseases of blood and blood-forming organs		
286	Coagulation defects	Diseases of blood and blood-forming organs		
288	Diseases of white blood cells	Diseases of blood and blood-forming organs		
290	Senile and presenile organic psychotic conditions	Mental disorders		
291	Alcoholic psychoses	Mental disorders		
294	Other organic psychotic conditions (chronic)	Mental disorders		
295	Schizophrenic psychoses	Mental disorders		
296	Affective psychoses	Mental disorders		
297	Paranoid states	Mental disorders		
299	Psychoses with origin specific to childhood	Mental disorders		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
300	Neurotic disorders	Mental disorders		
301	Personality disorders	Mental disorders		
302	Sexual deviations and disorders	Mental disorders		Y
303	Alcohol dependence syndrome	Mental disorders		
304	Drug dependence	Mental disorders		
305	Nondependent abuse of drugs	Mental disorders		
306	Physiological malfunction arising from mental factors	Mental disorders		Y
307	Special symptoms or syndromes, not elsewhere classified	Mental disorders		
308	Acute reaction to stress	Mental disorders		Y
309	Adjustment reaction	Mental disorders		
310	Specific nonpsychotic mental disorders following organic brain damage	Mental disorders		
311	Depressive disorder, not elsewhere classified	Mental disorders		
312	Disturbance of conduct, not elsewhere classified	Mental disorders		
313	Disturbance of emotions specific to childhood and adolescence	Mental disorders		
314	Hyperkinetic syndrome of childhood	Mental disorders		
315	Specific delays in development	Mental disorders		
316	Psychic factors associated with diseases classified elsewhere	Mental disorders	Y	
317	Mild mental retardation	Mental disorders		
318	Other specified mental retardation	Mental disorders		
319	Unspecified mental retardation	Mental disorders		
330	Cerebral degenerations usually manifest in childhood	Diseases of the nervous system and sense organs		Y
331	Other cerebral degenerations	Diseases of the nervous system and sense organs		
332	Parkinson's disease	Diseases of the nervous system and sense organs		
333	Other extrapyramidal disease and abnormal movement disorders	Diseases of the nervous system and sense organs		
334	Spinocerebellar disease	Diseases of the nervous system and sense organs		Y
335	Anterior horn cell disease	Diseases of the nervous system and sense organs		Y
336	Other diseases of spinal cord	Diseases of the nervous system and sense organs		

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
337	Disorders of the autonomic nervous system	Diseases of the nervous system and sense organs		
340	Multiple sclerosis	Diseases of the nervous system and sense organs		
342	Hemiplegia and hemiparesis	Diseases of the nervous system and sense organs		
343	Infantile cerebral palsy	Diseases of the nervous system and sense organs		
344	Other paralytic syndromes	Diseases of the nervous system and sense organs		
345	Epilepsy	Diseases of the nervous system and sense organs		
346	Migraine	Diseases of the nervous system and sense organs		
347	Cataplexy and narcolepsy	Diseases of the nervous system and sense organs		Y
348	Other conditions of brain	Diseases of the nervous system and sense organs		
353	Nerve root and plexus disorders	Diseases of the nervous system and sense organs		Y
354	Mononeuritis of upper limb and mononeuritis multiplex	Diseases of the nervous system and sense organs		
355	Mononeuritis of lower limb	Diseases of the nervous system and sense organs		
356	Hereditary and idiopathic peripheral neuropathy	Diseases of the nervous system and sense organs		
357	Inflammatory and toxic neuropathy	Diseases of the nervous system and sense organs		
359	Muscular dystrophies and other myopathies	Diseases of the nervous system and sense organs		
360	Disorders of the globe	Diseases of the nervous system and sense organs		Y
362	Other retinal disorders	Diseases of the nervous system and sense organs		

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
363	Chorioretinal inflammations and scars and other disorders of choroid	Diseases of the nervous system and sense organs	Y	
365	Glaucoma	Diseases of the nervous system and sense organs		
366	Cataract	Diseases of the nervous system and sense organs		
369	Blindness and low vision	Diseases of the nervous system and sense organs		
377	Disorders of optic nerve and visual pathways	Diseases of the nervous system and sense organs		
389	Hearing loss	Diseases of the nervous system and sense organs		
393	Chronic rheumatic pericarditis	Diseases of the circulatory system	Y	
394	Diseases of mitral valve	Diseases of the circulatory system		
395	Diseases of aortic valve	Diseases of the circulatory system	Y	
396	Diseases of mitral and aortic valves	Diseases of the circulatory system		
397	Diseases of other endocardial structures	Diseases of the circulatory system		
398	Other rheumatic heart disease	Diseases of the circulatory system		Y
401	Hypertension	Diseases of the circulatory system		
402	Hypertensive heart disease	Diseases of the circulatory system		
403	Hypertensive renal disease	Diseases of the circulatory system		
404	Hypertensive heart and renal disease	Diseases of the circulatory system		
405	Secondary hypertension	Diseases of the circulatory system		
410	Acute myocardial infarction	Diseases of the circulatory system		
411	Other acute and subacute form of ischemic heart disease	Diseases of the circulatory system		
412	Old myocardial infarction	Diseases of the circulatory system		
413	Angina pectoris	Diseases of the circulatory system		
414	Other forms of chronic ischemic heart disease	Diseases of the circulatory system		
416	Chronic pulmonary heart disease	Diseases of the circulatory system		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
424	Other diseases of endocardium	Diseases of the circulatory system		
425	Cardiomyopathy	Diseases of the circulatory system		
426	Conduction disorders	Diseases of the circulatory system		
427	Cardiac dysrhythmias	Diseases of the circulatory system		
428	Heart failure	Diseases of the circulatory system		
429	Ill-defined descriptions and complications of heart disease	Diseases of the circulatory system		
430	Subarachnoid hemorrhage	Diseases of the circulatory system		
431	Intracerebral hemorrhage	Diseases of the circulatory system		
432	Other and unspecified intracranial hemorrhage	Diseases of the circulatory system		
433	Occlusion and stenosis of precerebral arteries	Diseases of the circulatory system		
434	Occlusion of cerebral arteries	Diseases of the circulatory system		
435	Transient cerebral ischemia	Diseases of the circulatory system		
436	Acute but ill-defined cerebrovascular disease	Diseases of the circulatory system		Y
437	Other and ill-defined cerebrovascular disease	Diseases of the circulatory system		
438	Late effects of cerebrovascular disease	Diseases of the circulatory system		
440	Atherosclerosis	Diseases of the circulatory system		
441	Aortic aneurysm and dissection	Diseases of the circulatory system		
442	Other aneurysm	Diseases of the circulatory system		
443	Other peripheral vascular disease	Diseases of the circulatory system		
444	Arterial embolism and thrombosis	Diseases of the circulatory system		
445	Atheroembolism	Diseases of the circulatory system	Y	
446	Polyarteritis nodosa and allied conditions	Diseases of the circulatory system		
447	Other disorders of arteries and arterioles	Diseases of the circulatory system		
457	Noninfective disorders of lymphatic channels	Diseases of the circulatory system		
472	Chronic pharyngitis and nasopharyngitis	Diseases of the respiratory system		Y
473	Chronic sinusitis	Diseases of the respiratory system		

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

| 3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
474	Chronic disease of tonsils and adenoids	Diseases of the respiratory system		Y
476	Chronic laryngitis and laryngotracheitis	Diseases of the respiratory system	Y	
477	Allergic rhinitis	Diseases of the respiratory system		
491	Chronic bronchitis	Diseases of the respiratory system		
492	Emphysema	Diseases of the respiratory system		
493	Asthma	Diseases of the respiratory system		
494	Bronchiectasis	Diseases of the respiratory system		
495	Extrinsic allergic alveolitis	Diseases of the respiratory system	Y	
496	Chronic airways obstruction, not elsewhere classified	Diseases of the respiratory system		
500	Coalworkers' pneumoconiosis	Diseases of the respiratory system	Y	
501	Asbestosis	Diseases of the respiratory system		Y
502	Pneumoconiosis due to other silica or silicates	Diseases of the respiratory system	Y	
503	Pneumoconiosis due to other inorganic dust	Diseases of the respiratory system	Y	
504	Pneumopathy due to inhalation of other dust	Diseases of the respiratory system	Y	
505	Pneumoconiosis, unspecified	Diseases of the respiratory system	Y	
514	Pulmonary congestion and hypostasis	Diseases of the respiratory system		
515	Postinflammatory pulmonary fibrosis	Diseases of the respiratory system		
516	Other alveolar and parietoalveolar pneumopathy	Diseases of the respiratory system		
555	Regional enteritis	Diseases of the digestive system		
556	Ulcerative colitis	Diseases of the digestive system		
562	Diverticula of intestine	Diseases of the digestive system		
571	Chronic liver disease and cirrhosis	Diseases of the digestive system		
573	Other disorders of liver	Diseases of the digestive system		
581	Nephrotic syndrome	Diseases of the genitourinary system		
582	Chronic glomerulonephritis	Diseases of the genitourinary system		
583	Nephritis and nephropathy, not specified as acute or chronic	Diseases of the genitourinary system		

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
585	Chronic renal failure	Diseases of the genitourinary system		
586	Renal failure, unspecified	Diseases of the genitourinary system		Y
587	Renal sclerosis, unspecified	Diseases of the genitourinary system		
588	Disorders resulting from impaired renal function	Diseases of the genitourinary system		
600	Hyperplasia of prostate	Diseases of the genitourinary system		
607	Disorders of penis	Diseases of the genitourinary system		
617	Endometriosis	Diseases of the genitourinary system		
618	Genital prolapse	Diseases of the genitourinary system		
619	Fistula involving female genital tract	Diseases of the genitourinary system		Y
626	Disorders of menstruation and other abnormal bleeding from female genital tract	Diseases of the genitourinary system		
627	Menopausal and postmenopausal disorders	Diseases of the genitourinary system		
628	Infertility, female	Diseases of the genitourinary system	Y	
694	Bullous dermatoses	Diseases of the skin and subcutaneous tissue		Y
696	Psoriasis and similar disorders	Diseases of the skin and subcutaneous tissue		
707	Chronic ulcer of skin	Diseases of the skin and subcutaneous tissue		
710	Diffuse diseases of connective tissue	Diseases of the musculoskeletal system		
712	Crystal arthropathies	Diseases of the musculoskeletal system		
713	Arthropathy associated with other disorders classified elsewhere	Diseases of the musculoskeletal system		
714	Rheumatoid arthritis and other inflammatory polyarthropathies	Diseases of the musculoskeletal system		
715	Osteoarthrosis and allied disorders	Diseases of the musculoskeletal system		
716	Other and unspecified arthropathies	Diseases of the musculoskeletal system		
717	Internal derangement of knee	Diseases of the musculoskeletal system		

| 3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
720	Ankylosing spondylitis and other inflammatory spondylopathies	Diseases of the musculoskeletal system		Y
721	Spondylosis and allied disorders	Diseases of the musculoskeletal system		
722	Intervertebral disc disorders	Diseases of the musculoskeletal system		
725	Polymyalgia rheumatica	Diseases of the musculoskeletal system		
730	Osteomyelitis, periostitis, and other infections involving bone	Diseases of the musculoskeletal system		
732	Osteochondropathies	Diseases of the musculoskeletal system		Y
737	Curvature of spine	Diseases of the musculoskeletal system		
740	Anencephalus and similar anomalies	Congenital anomalies	Y	
741	Spina bifida	Congenital anomalies		
742	Other congenital anomalies of nervous system	Congenital anomalies		
743	Congenital anomalies of eye	Congenital anomalies		Y
744	Congenital anomalies of ear, face, and neck	Congenital anomalies		Y
745	Bulbus cordis anomalies and anomalies of cardiac septal closure	Congenital anomalies		
746	Other congenital anomalies of heart	Congenital anomalies		
747	Other congenital anomalies of circulatory system	Congenital anomalies		
748	Congenital anomalies of respiratory system	Congenital anomalies		
749	Cleft palate and cleft lip	Congenital anomalies		
750	Other congenital anomalies of upper alimentary tract	Congenital anomalies		
751	Other congenital anomalies of digestive system	Congenital anomalies		
752	Congenital anomalies of genital organs	Congenital anomalies		
753	Congenital anomalies of urinary system	Congenital anomalies		
754	Certain congenital musculoskeletal deformities	Congenital anomalies		
755	Other congenital anomalies of limbs	Congenital anomalies		
756	Other congenital musculoskeletal anomalies	Congenital anomalies		
757	Congenital anomalies of the integument	Congenital anomalies		
758	Chromosomal anomalies	Congenital anomalies		
759	Other and unspecified congenital anomalies	Congenital anomalies		
797	Senility without mention of psychosis	Symptoms, signs, and ill-defined conditions		Y

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

3-digit ICD-9 code	Disease/condition	Body system	Less prevalent	Isolated node
885	Traumatic amputation of thumb (complete) (partial)	Injury and poisoning	Y	
886	Traumatic amputation of other finger(s) (complete) (partial)	Injury and poisoning		Y
887	Traumatic amputation of arm and hand (complete) (partial)	Injury and poisoning	Y	
895	Traumatic amputation of toe(s) (complete) (partial)	Injury and poisoning	Y	
897	Traumatic amputation of leg(s) (complete) (partial)	Injury and poisoning	Y	
952	Spinal cord injury without evidence of spinal bone injury	Injury and poisoning		Y
984	Toxic effect of lead and its compounds (including fumes)	Injury and poisoning	Y	
985	Toxic effect of other metals	Injury and poisoning	Y	

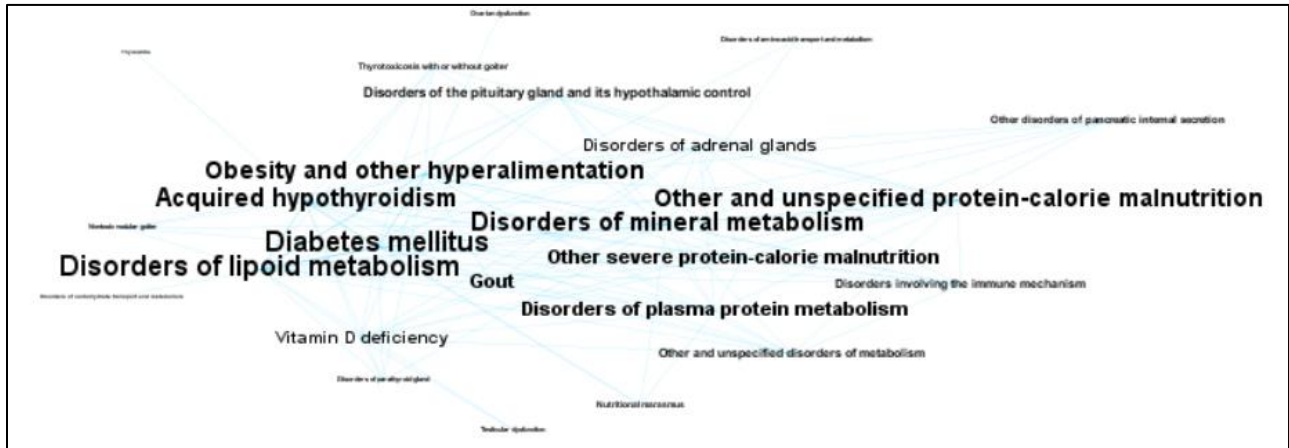# Appendix C   - DCN for five body systems

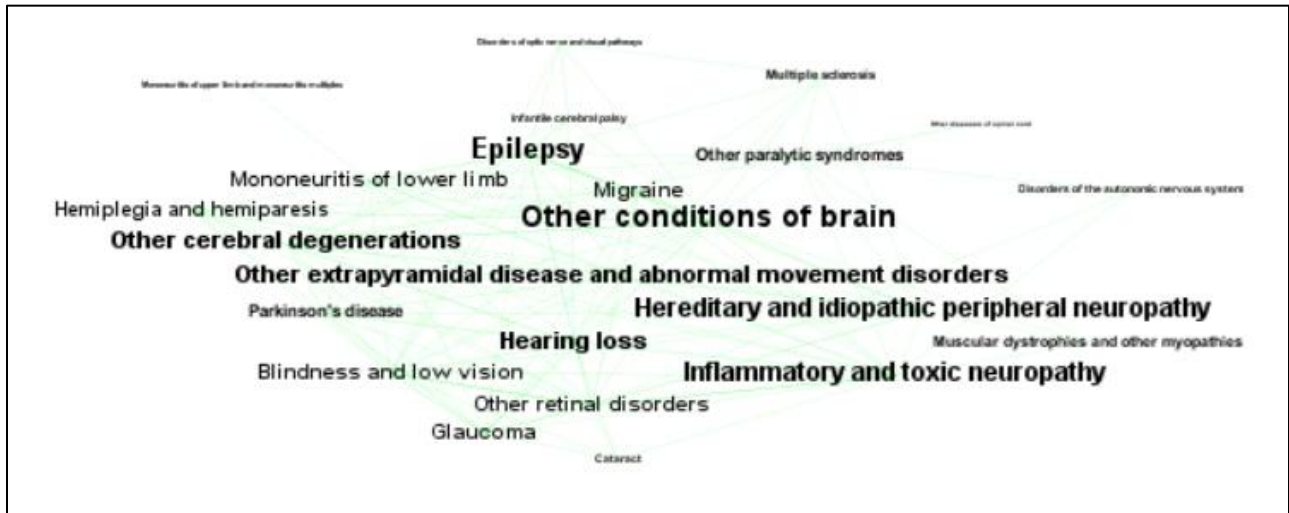## C.1 Circulatory system
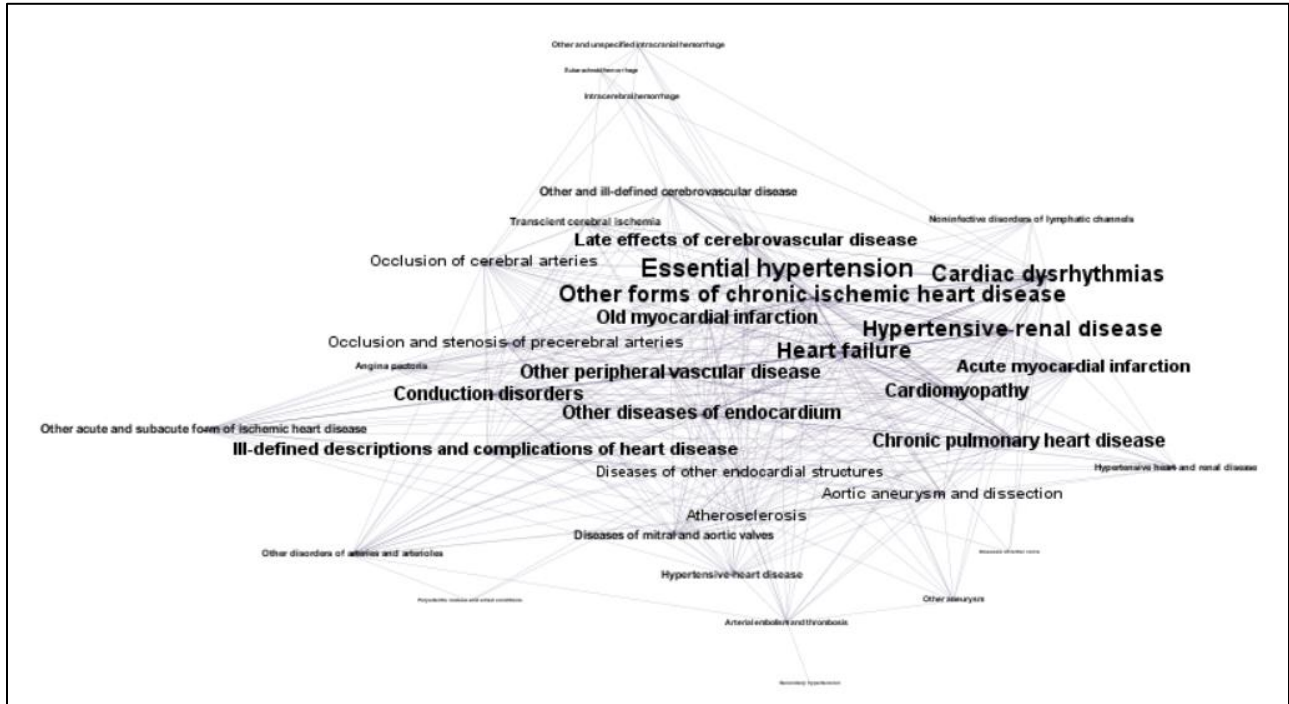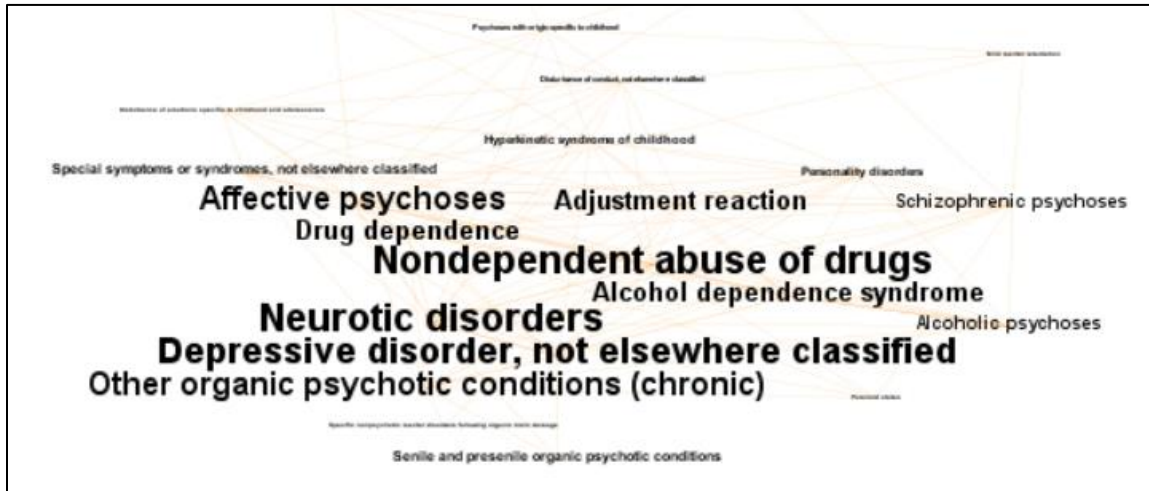

C.2 Neoplasms
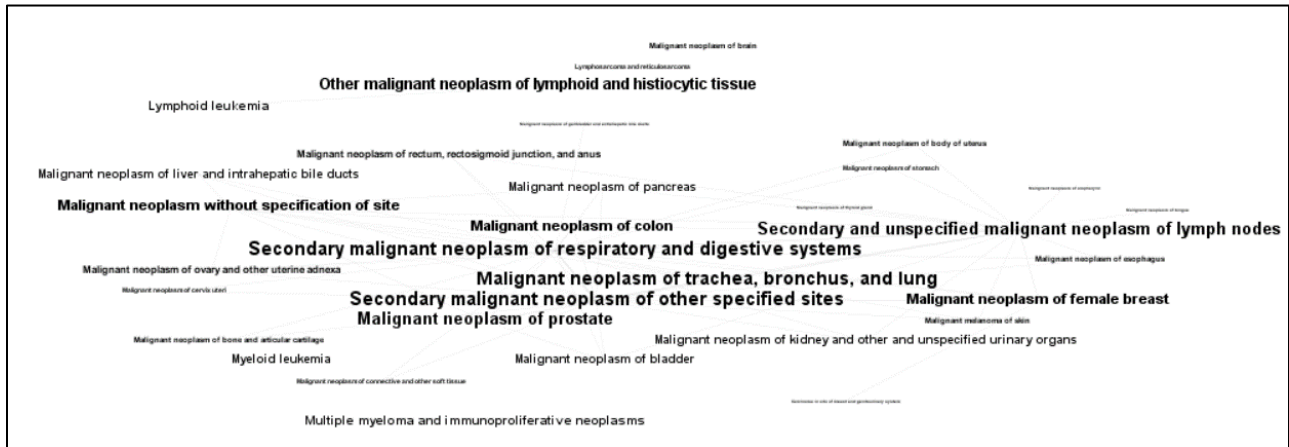
C.3 Immune system

C.4 Nervous system

C.5 Mental disorders

Appendix D - Top 5 nodes in terms of node strength for Arizona and 11 Arizona regions

Dataset	Diagnoses 1	Diagnoses 2	Diagnoses 3	Diagnoses 4	Diagnoses 5
Arizona	chronic renal failure	other forms of chronic ischemic heart disease	disorders of lipoid metabolism	diabetes	hypertensive renal disease
Apache	chronic renal failure	diabetes	disorders of lipoid metabolism	other forms of chronic ischemic heart disease	hypertensive renal disease
Cochise	other forms of chronic ischemic heart disease	chronic renal failure	disorders of lipoid metabolism	diabetes	heart failure
Coconino	chronic renal failure	disorders of lipoid metabolism	other forms of chronic ischemic heart disease	hypertensive renal disease	diabetes
Gila	other forms of chronic ischemic heart disease	chronic renal failure	disorders of lipoid metabolism	heart failure	diabetes
Maricopa	chronic renal failure	other forms of chronic ischemic heart disease	disorders of lipoid metabolism	diabetes	hypertensive renal disease
Mohave	other forms of chronic ischemic heart disease	heart failure	chronic renal failure	disorders of lipoid metabolism	diabetes
Navajo	disorders of lipoid metabolism	other forms of chronic ischemic heart disease	diabetes	chronic renal failure	heart failure
Pima	other forms of chronic ischemic heart disease	chronic renal failure	disorders of lipoid metabolism	heart failure	diabetes
Pinal	other forms of chronic ischemic heart disease	disorders of lipoid metabolism	diabetes	chronic renal failure	hypertensive renal disease

MAP Dashboard White Paper   www.mapazdashboard.arizona.edu

Yavapai	other forms of chronic ischemic heart disease	disorders of lipoid metabolism	chronic renal failure	heart failure	cardiac dysrhythmias
Yuma	other forms of chronic ischemic heart disease	chronic renal failure	disorders of lipoid metabolism	diabetes	hypertensive renal disease

Appendix E - List of 11 diseases/conditions that are present in ego networks of top 6 co-
     occurring diagnoses

ICD	Disease/condition
275	disorders of mineral metabolism
311	depressive disorder
278	obesity and other hyper alimentation
427	cardiac dysrhythmias
272	disorders of lipoid metabolism
250	diabetes
244	acquired hypothyroidism
414	chronic ischemic heart disease
428	heart failure
496	chronic airways obstruction, not elsewhere classified
715	osteoarthrosis and allied disorders