International Journal of Medical Informatics 108 (2017) 22–28

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Research Paper

Examining health disparities by gender: A multimorbidity network analysis of electronic medical record

Pankush Kalgotra[a,*], Ramesh Sharda[b], Julie M. Croff[c]

[a] Graduate School of Management, Clark University, 950 Main Street (Carlson Hall), Worcester, 01610 MA, USA
[b] Spears School of Business, Oklahoma State University, 102 Gunderson, Stillwater, 74078 OK, USA
[c] Oklahoma State University, 429 Willard Hall, Stillwater, 74078 OK, USA

ARTICLE INFO

Keywords:
Gender disparity
Multimorbidity
Network analysis

ABSTRACT

Problem: Multimorbidity health disparities have not been well examined by gender. Co-occurring diseases may be mutually deleterious, co-occurring independently, or co-occurring from a common antecedent. Diseases linked by a common antecedent may be caused by biological, behavioral, social, or environmental factors. This paper aims to address the co-occurrences of diseases using network analysis.

Methods: In this study, we identify these multi-morbidities from a large electronic medical record (EMR) containing diagnoses, symptoms and treatment data on more than 22.1 million patients. We create multimorbidity networks from males and females medical records and compare their structural properties.

Results: Our macro analysis at the organ-level indicates that females have a stronger multimorbidity network than males. For example, the female multimorbidity network includes six linkages to mental health, wherein the male multimorbidity network includes only two linkages to mental health. The strength of some disease associations between lipid metabolism and chronic heart disorders is stronger in males than females.

Conclusion: Our multimorbidity network analysis by gender identifies specific differences in disease diagnosis by gender, and presents questions for biological, behavioral, clinical, and policy research.

1. Introduction

Multiple ecological levels interact to influence disparities in health and health outcomes by gender. Health disparities observed between genders are caused by genetic, hormonal, physiological, behavioral, and sociocultural factors. Life expectancy at birth is notably longer for females at 81.4 years compared to males at 76.4 years [1]. During this longer lifetime, females are more likely to visit the hospital or health care provider, but less likely to die [2]. Notably this male-female health-survival paradox is explained by chronic diseases which are most prevalent by gender: females are more likely to experience pain, reproductive cancers, and depression, while males are more likely to experience cardiovascular disease and diabetes [3]. Additionally, when males and females are compared on the same chronic diseases, males may experience severe cases of chronic disease. Previous epidemiological studies of health disparities address individual diseases experienced by gender; however, most patients are diagnosed with multiple diseases. The goal of this paper is to explore disparities among males and females diagnosed with more than one disease, and present research and policy implications.

Two terms are often used to discuss the presence of more than one disease in a patient: comorbidity and multimorbidity. Comorbidity is a condition when an additional disease is diagnosed in presence of an index disease [4]. Multimorbidity is defined as the coexistence of multiple chronic diseases and conditions in a patient [5,6]. Throughout this manuscript we will use these terms interchangeably to denote co-occurrence of diseases, unless we need to specifically highlight the differences between comorbidity and multimorbidity. Previous studies on comorbidities have controlled for gender but rarely focused and reported differences in genders explicitly as pointed out by Short et al. [7]. Further examination of comorbidities by gender may be critically important for treatment of disease, and in identifying contraindications of common pharmaceuticals. The availability of large medical records affords the opportunity to study all possible disease relationships as observed in practice.

We adapt a network approach to model the multimorbidities [8]. Networks are formed from the interactions between the elements or nodes. Network analysis has been used in health and medical literature to understand the interaction of genes [9,10], molecular involvement in disease [11], drug trials [12], and historical epidemiological data on

* Corresponding author.
E-mail addresses: pkalgotra@clarku.edu (P. Kalgotra), ramesh.sharda@okstate.edu (R. Sharda), julie.croff@okstate.edu (J.M. Croff).

http://dx.doi.org/10.1016/j.ijmedinf.2017.09.014
Received 8 June 2017; Received in revised form 19 September 2017; Accepted 24 September 2017
1386-5056/ © 2017 Elsevier B.V. All rights reserved.

disease phenotypes [13,14]. Tai and Chiu [15] applied association rule mining to create comorbidity network in ADHD patients using clinical database. Similarly, Chmiel et al. [16] applied network approach to study the prevalence of different cluster of diseases over lifetime of genders. However, to the best of our knowledge, no one has applied this approach to study multimorbidity by gender in order to better understand health disparities.

In this paper, we develop and compare multimorbidity networks for males and females based on ICD-9 (International Classification of Diseases, Clinical Modification) codes of diagnoses. Our network comprises diseases connected based on the co-occurrences of diseases in 22.1 million patient records. The use of large dataset is another strength of our study. Knowing the relationships between diseases at the network level will enhance our understanding about disease associations at the patient population level.

2. Method and analysis

In this section, we begin by describing the data and explaining how we measure the multimorbidity in our context. Next, we present a method to develop a multimorbidity network. Then, we briefly describe the properties of the network that can explain the position of a disease in a web of other diseases, and help us understand differences between males and females.

2.1. Data description

We obtained data from the Oklahoma State University Center for Health Systems Innovation (CHSI), which houses HIPAA compliant patient data provided by Cerner Corporation, a major Electronic Medical Record (EMR) provider. The data warehouse contains an EMR on the visits of 58 million unique patients across 662 US hospitals (2000–2016). We used information about the demographics of the patients, hospitals and disease diagnoses coded by ICD-9 system.[1] We removed several hospital visits in which patients were either not diagnosed with a disease or were marked only for symptoms. After data preprocessing, we had approximately 22.1 million unique patients with the sufficient information to perform analysis.

We extracted medical records for males and females in two different datasets from this pseudo-population dataset for comparing comorbidities by gender. The datasets were further cleaned based on the detected anomalies in particular category. For example, there were a few patients who were coded as a male during one visit and a female or null in another. Although males can also have breast diseases biologically, we removed the male patients diagnosed with such diseases with a suspicion that these are erroneously coded (ICD9: 610–612).[2] We also removed males who were diagnosed with diseases such as inflammatory diseases of female pelvic organs (ICD9: 614–616)[3], and complications of pregnancy, childbirth, and the puerperium (ICD9: 630–679).[4] Similarly, we removed female patients diagnosed with diseases of male genital organs (ICD9: 600–608).[5] After cleaning the data, we had records of 12 million female patients and 9.9 million male patients. From the two samples, networks were created, one each for males and females.

[1] From the last quarter of 2016, the diagnoses in Cerner EMR are required to be coded in ICD-10 system. However, we did not consider the last quarter to maintain the consistency in our data analysis and considered only ICD-9 codes.

[2] There were 38,980 male patients with ICD-9 codes 610–612, which is 0.34% of the male database.

[3] 1594 patients.

[4] 20,009 patients.

[5] 8627 patients.

2.2. Measuring multimorbidity

In the past, comorbidity and multimorbidity were largely defined at the cross-sectional level [4,17]. The chronic diseases, which we would not expect to go away in one hospital visit, could be overestimated from the medical records because they are recorded multiple times in an EMR. However, we delineate multimorbidity considering the lifetime history of a patient rather than a single hospital visit. We measure multimorbidity as the presence of multiple diseases in the lifetime history of a patient. This measurement has two advantages over previous definitions. First, the EMR recording of a disease over multiple hospitals visits is only considered once. Considering the same disease as different across hospital visits can overestimate its presence and bias the analysis and conclusions. Second, our definition considers the impact of a disease in one visit on subsequent visits. Therefore, it incorporates a wider span of disease developments. However, there is a concern of taking into account the association between diseases diagnosed across hospital visits occurring after long period of time. Given the relatively short time span of the database (17 years), short average length between first and last hospital visit in the database (527 days), average number of hospital visits of a patient being 5.1 (all types of visits including inpatient, outpatient, etc.) and statistical analysis on millions of patients, we mitigate the concern of false positives.

2.3. Multimorbidity network

A multimorbidity network developed from patients contains a set of nodes connected through edges. In our network, nodes represent diseases. In an EMR, an ICD-9 code of a disease has three, four or five digits (xxx.xx). The first three digits represent the broader category of a disease. The fourth and fifth digits represent the sub-divisions of the disease. For example, the ICD-9 code for personality disorder is 301. At four-digit level (301.x), there are ten types of personality disorders and at five-digit level (301.xx), two other specific personality disorders are coded. We aggregated ICD-9-CM codes to three-digit level. Thus, variations of the same disease were considered as one node in the network. For example, multiple types of personality disorders mentioned above were aggregated into one node in our network.

An edge or connection between two diseases is created if these are comorbid. Since our focus is not to establish causality of a multimorbidity, we created a network with no direction in the relationships. For example, the comorbidity comprising congestive heart failure and rheumatic heart disease will be represented as an undirected edge between the two nodes representing the two diseases regardless of their causal relationship.

Historically, associations between diagnoses or comorbidities were modeled using a simple Pearson's correlation coefficient [18,13]. However, number of significant correlations is directly proportional to the number of observations used, and thus affected by the sample size. Power to detect rare comorbidities is low because of the rareness of events. Therefore, to establish the right measure to model a comorbidity, we use a cosine index known as Salton Cosine index [19]. SCI is immune to the total number of observations used [20] and measures the prevalence of a relationship between two diseases considering their individual prevalence. Salton Cosine Index, *SCI*, is calculated as in Eq. (1), where c_{ij} is the number of co-occurrences of diseases i and j, i_c is the prevalence of disease i and j_c is the prevalence of disease j. The cosine similarity has been used in the past to find phenotype overlaps [21,22]. We propose this as an appropriate measure for finding the strength of a comorbidity.

$$SCI_{ij} = \frac{(c_{ij})}{\sqrt{(i_c * j_c)}} \tag{1}$$

Statistical significance of SCI was determined by assessing the relationship between correlation and SCI, because this approach has been

suggested in the past to find the cut-off for SCI [23]. First, we determined the number of comorbidities significantly correlated in a network created using Pearson's correlation coefficient. Then, we related the number of comorbidities in the network created using Salton Cosine Index and found a cut-off where number of significantly correlated comorbidities are equal in both networks. In the network from entire database using Pearson's Correlation Coefficient, at $p < 0.01$, there were 14,463 significantly correlated comorbidities. Meanwhile, at the SCI cut-off of 0.04, the number of comorbidities were 14,463. Therefore, we used the cut-off of 0.04 for creating different networks for males and females. Then, the comparison between the networks was made using the network measures briefly described in the next section.

2.4. Network metrics

The structural properties of a network can be measured using several network metrics. These include degree, weighted degree, closeness and betweenness centrality [24]. In a multimorbidity network, the degree centrality of a disease (node) denotes the number of direct connections with other diseases. The weighted degree centrality of a disease considers the strength of the relationships with others and is calculated as a weighted sum of the strengths of the relationships. The closeness centrality of a disease determines an average number of steps it is away from other diseases in the network. A disease with higher closeness has a higher risk of being diagnosed with other diseases in less number of steps. Finally, the betweenness centrality of a disease describes its bridgeness. In other words, a disease with higher betweenness tends to be forming more bridges between other diseases.

3. Results and discussion

3.1. Comparison of male and female multimorbidity networks

The visualizations of female and male multimorbidity networks are presented in Fig. 1a and b respectively. In the visualization, the diseases are color coded based on the 17 categories/classes/organ systems described in the ICD-9 classification. These classes are also listed in Fig. 1. Size of a disease node represents an association with other disease(s), or its number of direct connections to other disease(s). The female multimorbidity network contains 300 diseases not connected to any other disease as compared to 265 diseases in the male multimorbidity network. In the female network, the diseases that are connected to at least one other disease in the network form three different sub-networks labelled as connected components. There is a primary cluster of diseases in the female network labelled as connected component-1 suggesting all diseases are associated to each other directly or indirectly. The two secondary clusters in the females were for burns (ICD9: 941–945, 948, 949) and a pair of appendicitis codes (ICD9: 540–541).

3.2. Network properties

The properties of each network are listed in Table 1. The number of nodes or diseases in two networks are different as some diseases are unique to each gender. There were 839 diseases reported in males and 899 unique diseases in females at three-digit ICD-9 codes, that is, 7% more unique disease diagnoses in females. In the male network, there were 12,498 comorbidities as compared to 14,810 in females. Recall the edge strength denotes the magnitude of comorbidity. Out of all the edges detected above, 10,607 were common between both sexes. A pair-wise comparison of these 10,607 edge strengths indicates stronger comorbidities among females (t value = 12.67, $p < 0.0001$).

Although females have stronger and more comorbidities overall, we found some disease associations to be stronger in males than females.[6] These include disorders of lipid metabolism – chronic ischemic heart disease; disorders of fluid electrolyte and acid base balance – acute kidney failure; benign neoplasm of parts of digestive system and hemorrhoids – diverticula of intestine; diabetes – chronic ischemic heart disease; anemias – hypertensive chronic kidney disease; and disorders of lipid metabolism –cardiac dysrhythmias. The average degree and weighted degree of the two networks were statistically different. Although the aggregated closeness and betweenness centralities of the two networks were not statistically different, we found several differences with respect to specific diseases in the two genders. For instance, acute upper respiratory infections and disorders of urethra & urinary tract form relatively more bridges between other diseases in females than males. On the other hand, the disorders of skin and subcutaneous tissue form a bridge between multiple other diseases more often in males than females.

3.3. Organ level network comparison

We aggregated the relationships depicted in the networks in Fig. 1a and b at the organ system level or class system categorized in the ICD-9 classification (See Table 2). We present two macro level networks at the class/organ system level in Fig. 2a and b for females and males, respectively. The diseases of different classes were aggregated at the class level by adding up their weights (Salton Cosine Index). We highlight the connections between diagnoses of different classes if their aggregated weight is more than ten. This cut-off of ten is to study the most prevalent relationships. However, one could select a lower cut-off to analyze the rare connections.

We present a unique way to visualize the relationships between disorders of different organ systems by creating an outline of a human body and mapping the categories of the diseases on it (See Fig. 2a and b). In the ICD-9 classification, some categories can be directly related to the organ system present on a specific position in human body such as circulatory system (class-8), mental disorders (class-5), digestive system (class-10), respiratory system (class-9), and genitourinary system (class-11). However, other classes such as 1–4, 6–7, and 12–18 cannot be related to a specific organ system as listed in Table 2. The classes directly related to an organ system are mapped at the positions of the particular organ system in the human body. The classes that are not related to a specific organ system are presented outside the human sketch. The size of a node denotes the number of connections to other nodes. The width of an edge between two classes represents the aggregated weight (aggregated Salton Cosine Index) or the strength between them. The same connections can be observed in Table 3 where a comparison is made between the two networks.

The Fig. 2a–b and Table 3 shows which organ systems diseases are diagnosed simultaneously more often in different genders. The female network is clearly denser than the male network with more connections. Notably, there are several multimorbidities present in the female network not present in male network at the selected cut-off. These are highlighted in Table 3 and notated by an F in each area of comorbidity. There is only one males-specific comorbidity as compared to eleven comorbidities noted as significant only among females. For example, mental disorders in males are associated with the disorders of circulatory and respiratory systems. However, female patients with mental disorders are at risk of diagnoses belonging to multiple other organ systems such as circulatory respiratory, digestive and musculoskeletal systems in addition to the injury, poisoning, endocrine, nutritional, metabolic, and immunity related disorders. Similarly, the disorders of genitourinary system are strongly associated with the disorders of respiratory system, musculoskeletal system and connective tissue in

[6] It has to be noted that we focus on the top comorbidities based on their strength and not the frequency. In addition, comorbidities are discussed if they belong to distinct classes or organ systems listed in Fig. 1 and Table 2.

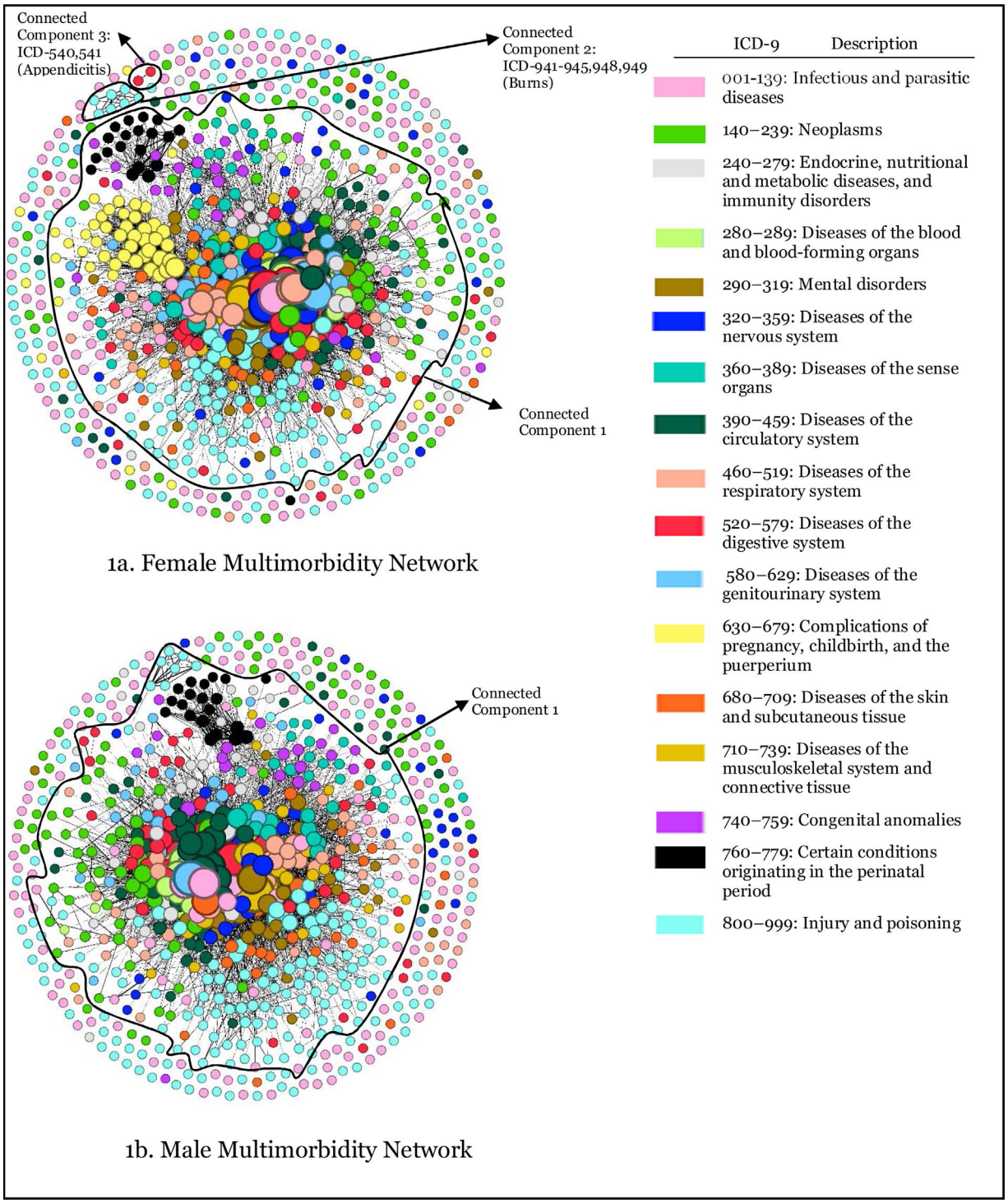

Fig. 1. a. Female Multimorbidity Network. b. Male Multimorbidity Network.

females than males. The disorders of musculoskeletal systems are also more strongly connected to other disorders in women than men. The musculoskeletal disorders such as osteoarthritis are known to be more prevalent in females [25] but other observed multimorbidity differences by gender need further research.

The above discussed relationships between diagnoses of different organ systems are more strongly connected in females than males. However, the comorbidity of endocrine, nutritional, metabolic diseases, and immunity disorders (3) with disorders of the blood and blood-forming organs (4) was only observed in males. Each connection needs further investigation so as to find the reasons for differences in genders. Recognition of these observed multimorbidities also may suggest greater precautions to be taken by patients themselves or the physicians to watch for related symptoms.

Our observed networks of comorbidities from the EMR data confirm the prevalence of higher comorbidities in females than males as supported by the previous research [26,27]. Notably, this study conforms previous work which identified a greater proportion of diagnosis of reproductive cancers and mental health diagnoses among females. However, contrary to previous research, we also note that the strength of some comorbidities are stronger in males than females.

Multimorbidity networks were different may be due to gender

Table 1
Gender multimorbidity networks properties.

	Female	Male	Pair-wise sample *t*-test (Female-Male)
No. of patients	12 M	9.9 M	N/A
Nodes (Diseases)	899	839	N/A
Edges (Comorbidities)	14,810	12,498	N/A
Avg. Degree (Degree of a disease is the number of diseases directly connected to it)	32.948	29.793	6.15, $p < 0.0001$
Avg. Wt. Degree (Degree calculated as a weighted sum of the strength of the comorbidities)	2.592	2.365	4.50, $p < 0.0001$
Avg. Betweenness (Number of times a disease is a bridge between pairs of diseases)	266.9	291.5	-1.78, $p = 0.07$
Avg. Closeness (Closeness centrality of a disease would represent how close a disease is to all the other diseases in the network)	0.293	0.285	0.80, $p = 0.42$
Graph Density	0.037	0.036	N/A

Table 2
ICD-9 code classification.

Class No.	Description	ICD-9 codes range	Mapped on organ system
1	Infectious and parasitic diseases	001–139	No
2	Neoplasms	140–239	No
3	Endocrine, nutritional and metabolic diseases, and immunity disorders	240–279	No
4	Diseases of the blood and blood-forming organs	280–289	No
5	Mental disorders	290–319	Yes
6	Diseases of the nervous system	320–359	No
7	Diseases of the sense organs	360–389	No
8	Diseases of the circulatory system	390–459	Yes
9	Diseases of the respiratory system	460–519	Yes
10	Diseases of the digestive system	520–579	Yes
11	Diseases of the genitourinary system	580–629	Yes
12	Complications of pregnancy, childbirth, and the puerperium	630–679	No
13	Diseases of the skin and subcutaneous tissue	680–709	No
14	Diseases of the musculoskeletal system and connective tissue	710–739	No
15	Congenital anomalies	740–759	No
16	Certain conditions originating in the perinatal period	760–779	No
17[a]	Symptoms, signs, and ill-defined conditions	780–799	N/A
18	Injury and poisoning	800–999	No

[a] Not considered in the analysis.

mental health multimorbidities in males and females is striking: perhaps physician implicit bias [29] and patient care seeking behaviors play a role in the diagnosis of mental health disorders by gender. Previous research suggests that social factors discourage men from seeking mental health care [28]. Therefore, the absence of strong multimorbidities with mental health among men was expected.

Notably, there was only one male dominated comorbidity in the network, disorders of blood and blood-forming collectively and disorders related to the endocrine, nutritional, metabolic, and immunity collectively, which were found more strongly connected in males than females. There could be a few potential explanations for this relationship: HIV infection or obesity. HIV infection is still the highest among men, and would be comorbid with an immunity diagnosis. Factors associated with obesity were strongly represented in the initial male multimorbidity network, many of these linked lipids, heart disease, diabetes, and digestive neoplasms; therefore, it is most likely that these diagnoses are linked to obesity, which has multiple antecedents addressed by public health.

Analysis of the network centralities (particularly weighted degree) suggested acute kidney failure, chronic kidney disease and chronic ischemic heart disease to be more strongly connected to other diseases in males than females. Moreover, diabetes mellitus emerged to be one of top diseases in males in terms of closeness (but the closeness number of diabetes in males was still smaller than females). Diabetes diagnosis is typically associated with overweight and obesity, and is often multimorbid with cardiovascular and other diagnoses related to overweight and obesity.

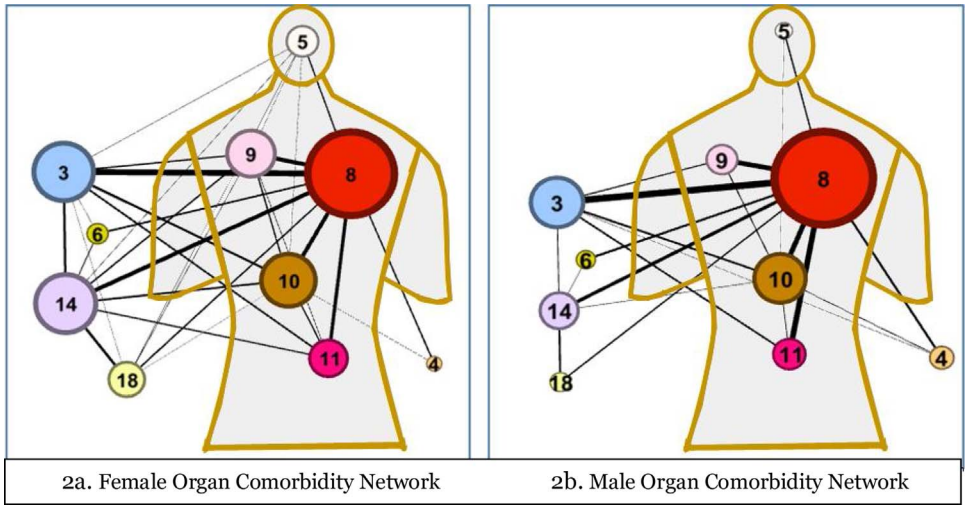

Fig. 2. a. Female Organ Comorbidity Network. b. Male Organ Comorbidity Network.

differences in care seeking behaviors among females because a greater frequency of care seeking behavior in females increases the risk of multiple disease diagnoses [28]. Moreover, the disparity between

4. Conclusions

Better understanding of multimorbidity networks may allow for

Table 3
Class associations in Female and Male Networks.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18
Infectious and parasitic [1]																	
Neoplasms [2]																	
Endocrine, nutritional, metabolic, and immunity disorders [3]																	
Blood and blood-forming organs [4]			M														
Mental disorders [5]			F														
Nervous system [6]			F														
Sense organs [7]																	
Circulatory system [8]			FM	FM	FM	FM											
Respiratory system [9]			FM		F			FM									
Digestive system [10]			FM	FM	FM			FM	FM								
Genitourinary system [11]			FM					FM	F	FM							
Pregnancy, childbirth, and the puerperium [12]																	
Skin and subcutaneous tissue [13]																	
Musculoskeletal system and connective tissue [14]			FM		F	FM		FM	F	FM	F						
Congenital anomalies [15]																	
Perinatal period [16]																	
Injury and poisoning [18]			F		F			FM	F	F				FM			

*Class 17 is symptoms and thus not included in the analysis

F Association only in Female Network

M Association only in Male Network

FM Association in both Networks

better screening and identification of diseases among patient populations, accounting for uniqueness for males and females in research measuring multimorbidity. These networks may improve health outcomes and reduce healthcare costs associated with hospital length of stay and readmission. The impact of comorbidity on the health outcomes has been studied in the past, but different network related properties have not been discussed in the public health literature. We shall establish the relationships between these concepts as a part of our future research.

Our study contributes both to the method and practice. With respect to the method contributions, we presented a novel approach to study multimorbidities at a population level. The network approach allowed us to study all the multimorbidities at once. Our paper is one of the first to apply a network approach to understand public health, particularly in the context of comorbidity/multimorbidity. The knowledge extracted from the large historical data can improve clinical decisions and outcomes as discussed by Tierney [30].

The analysis presented in this study has several practical implications. We mainly developed insights for health researchers. However, our study has implications for policy makers. In 1993, National Institute of Health (NIH) Revitalization Act was passed to encourage researchers to include women and minorities in clinical trials. Our analysis validates the disparities in diagnoses by genders, and thus we reinforce the need for considering the gender multimorbidities in clinical trials. In addition, education at every level should reinforce teaching of multimorbidity differences across population groups. We provide evidence to the gender disparities in public health through multimorbidity lens and support the global calls by Ovseiko et al. [31], Johnson et al. [32] and other thought leaders to recognize gender differences in health research.

This study has few limitations. First, our multimorbidities were based on the electronic health records and therefore, only the diagnoses recorded in specific hospitals were included. It is perhaps impossible to record lifetime history of a human in medical records. Hence, this limitation remains in all studies based on medical records. Second, we focused on the gender. However, we also recognize that the health disparities exist based on race and ethnicity [33]. Studying such disparities is part of our current research. Third, we only discussed simple network metrics such as degree, closeness and betweenness centrality. However, other complex measures such as clustering coefficient, cliques, clubs, eigenvector centrality, etc. can provide more information about the multimorbidities. Next, the differences were reported if the diseases were of different organ systems. However, the comorbidities related to the same organ system can also help enhance our understanding about the multimorbidities. We will explore these in future research. We also note that there were thousands of comorbidity differences in different population groups and we could not report them in this paper. Instead, we have attached Supplementary materials containing information on relationship of every disease with others. One can focus on one particular disease and find multimorbidities in our provided material.[7]

Notwithstanding these limitations, our study shows that big data and advanced analytics of large information can help gain new insights previously hard to discern [30]. We showed that advanced analytics methods such as network analysis can provide additional dimensions to understand the public health. Our study analyzed a dataset of millions of patients where diseases form a network and suggest that the structure of a network can have several implications. Moreover, there are several differences in different population groups in terms of multimorbidity network that should be considered while dealing with the comorbidities. Our study opens up an exciting and important area of research for policy makers, economists, social scientists and medical experts to treat different groups of population differently.

Author contributions

This study is a part of Pankush's Dissertation. He initiated the project under the supervision of his advisor, Dr. Sharda. Dr. Croff is a

[7] One can contact chsi.okstate.edu to request access to the dataset for research purposes.

domain expert and provided guidance related to the public health. This research is a result of collaboration between a public health researcher (Croff) and analytics researchers (Kalgotra and Sharda).

Conflicts of interest

None.

Summary points

What was already known before this study

- Health disparities in terms of diagnoses exist by gender
- Comorbidities impact health differently than independent occurrence of diseases

What this study added to our knowledge

- Gender-based health disparities can be better analyzed as multimorbidities through network analysis of large EMRs
- Females suffer more multimorbidities than males at organ system level
- More multimorbidities of musculoskeletal, genitourinary and mental disorders in females than males

Acknowledgments

This work was conducted with data from the Cerner Corporation's Health Facts database of electronic medical records provided by the Oklahoma State University Center for Health Systems Innovation (CHSI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Cerner Corporation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijmedinf.2017.09.014.

References

[1] National Center for Health Statistics, Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities, National Center for Health Statistics, Hyattsville, MD, 2016.
[2] A. Oksuzyan, K. Juel, J.W. Vaupel, K. Christensen, Men: good health and high mortality: sex differences in health and aging, Aging Clin. Exp. Res. 20 (2) (2008) 91.
[3] A.C. Case, C. Paxson, Sex differences in morbidity and mortality, Demography 42 (2) (2004) 189–214.
[4] A.R. Feinstein, The pre-therapeutic classification of co-morbidity in chronic disease, J. Chronic Dis. 23 (7) (1970) 455–468.
[5] M. van den Akker, F. Buntinx, J.A. Knottnerus, Comorbidity or multimorbidity: what's in a name? A review of literature, Eur. J. Gen. Pract. 2 (2) (1996) 65–70.
[6] M. van den Akker, F. Buntinx, J.F. Metsemakers, S. Roos, J.A. Knottnerus, Multimorbidity in general practice: prevalence, incidence, and determinants of co-occurring chronic and recurrent diseases, J. Clin. Epidemiol. 51 (5) (1998) 367–375.
[7] S.E. Short, Y.C. Yang, T.M. Jenkins, Sex, gender, genetics, and health, Am. J. Public Health 103 (S1) (2013) S93–S101.
[8] L. Euler, Leonhard Euler and the Königsberg bridges, Sci. Am. 189 (1953) 66–70.
[9] K.-I. Goh, M.E. Cusick, D. Valle, B. Childs, M. Vidal, A.-L. Barabasi, The human disease network, Proc. Natl. Acad. Sci. U. S. A. 104 (21) (2007) 8685–8690.
[10] F. Ferrazzi, P. Magni, L. Sacchi, A. Nuzzo, U. Petrovič, R. Bellazzi, Inferring gene regulatory networks by integrating static and dynamic data, Int. J. Med. Inform. 76 (2007) S462–S475.
[11] A.-L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nat. Rev. Genet. 12 (1) (2011) 56–68.
[12] B. Haslam, L. Perez-Breva, Learning disease relationships from clinical drug trials, J. Am. Med. Inform. Assoc. 24 (1) (2016) 13–23.
[13] C.A. Hidalgo, N. Blumm, A.-L. Barabási, N.A. Christakis, A dynamic network approach for the study of human phenotypes, PLoS Comput. Biol. 5 (4) (2009) e1000353.
[14] Y. Chen, R. Xu, Network analysis of human disease comorbidity patterns based on large-scale data mining, International Symposium on Bioinformatics Research and Applications, Springer International Publishing, 2014, pp. 243–254.
[15] Y.M. Tai, H.W. Chiu, Comorbidity study of ADHD: applying association rule mining (ARM) to National Health Insurance Database of Taiwan, Int. J. Med. Inform. 78 (12) (2009) e75–e83.
[16] A. Chmiel, P. Klimek, S. Thurner, Spreading of diseases through comorbidity networks across life and gender, New J. Phys. 16 (11) (2014) 115013.
[17] M. Jakovljevic, L. Ostojic, Comorbidity and multimorbidity in medicine today: challenges and opportunities for bringing separated branches of medicine closer to each other, Psychiatr. Danub. 25 (Suppl. 1) (2013) 18–28.
[18] M.J. Divo, C. Casanova, J.M. Marin, V.M. Pinto-Plata, J.P. de-Torres, J.J. Zulueta, et al., Chronic obstructive pulmonary disease comorbidities network, Eur. Respir. J. (2015) ERJ-01716-02014.
[19] G. Salton, M.J. McGill, Introduction to Modern Information Retrieval New York, McGraw-Hill, Inc., NY, 1986.
[20] P. Ahlgren, B. Jarneving, R. Rousseau, Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient, J. Am. Soc. Inf. Sci. Technol. 54 (6) (2003) 550–560.
[21] Y. Chen, X. Zhang, G.Q. Zhang, R. Xu, Comparative analysis of a novel disease phenotype network based on clinical manifestations, J. Biomed. Inform. 53 (2015) 113–120.
[22] K. Lage, E.O. Karlberg, Z.M. Størling, P.I. Olason, A.G. Pedersen, O. Rigina, et al., A human phenome-interactome network of protein complexes implicated in genetic disorders, Nat. Biotechnol. 25 (3) (2007) 309–316.
[23] L. Egghe, L. Leydesdorff, The relation between Pearson's correlation coefficient R and Salton's cosine measure, J. Am. Soc. Inf. Sci. Technol. 60 (5) (2009) 1027–1036.
[24] L.C. Freeman, Centrality in social networks conceptual clarification, Soc. Networks 1 (3) (1979) 215–239.
[25] A.D. Woolf, B. Pfleger, Burden of major musculoskeletal conditions, Bull. World Health Organ. 81 (9) (2003) 646–656.
[26] D.G. Blazer, S. Moody-Ayers, J. Craft-Morgan, B. Burchett, Depression in diabetes and obesity: racial/ethnic/gender issues in older adults, J. Psychosom. Res. 53 (4) (2002) 913–916.
[27] A. Moller-Leimkuhler, Gender differences in cardiovascular disease and comorbid depression, Dialogues Clin. Neurosci. 9 (1) (2007) 71.
[28] P. Corrigan, How stigma interferes with mental health care, Am. Psychol. 59 (7) (2004) 614.
[29] E.N. Chapman, A. Kaatz, M. Carnes, Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities, J. Gen. Intern. Med. 28 (11) (2013) 1504–1510.
[30] W.M. Tierney, Improving clinical decisions and outcomes with information: a review, Int. J. Med. Inform. 62 (1) (2001) 1–9.
[31] P.V. Ovseiko, T. Greenhalgh, P. Adam, J. Grant, S. Hinrichs-Krapels, K.E. Graham, et al., A global call for action to include gender in research impact assessment, Health Res. Policy Syst. 14 (1) (2016) 50.
[32] P. Johnson, T. Fitzgerald, A. Salganicoff, S.F. Wood, J.M. Goldstein, Sex-specific Medical Research: Why Women's Health Can't Wait. A Report of the Mary Horrigan Connors Center for Women's Health & Gender Biology at Brigham and Women's Hospital, Brigham and Women's Hospital, 2014.
[33] M.J. Fine, S.A. Ibrahim, S.B. Thomas, The role of race and genetics in health disparities research, Am. J. Public Health 95 (12) (2005) 2125–2128.