

Geo-Distinctive Comorbidity Networks of Pediatric Asthma

Eun Kyong SHIN ^{a,1} and Arash SHABAN-NEJAD ^a

^a *University of Tennessee Health Science Center – Oak Ridge National Laboratory- (UTHSC-ORNL), Center for Biomedical Informatics, Department of Pediatrics, Memphis, TN, USA*

Abstract. Most pediatric asthma cases occur in complex interdependencies, exhibiting complex manifestation of multiple symptoms. Studying asthma comorbidities can help to better understand the etiology pathway of the disease. Albeit such relations of co-expressed symptoms and their interactions have been highlighted recently, empirical investigation has not been rigorously applied to pediatric asthma cases. In this study, we use computational network modeling and analysis to reveal the links and associations between commonly co-observed diseases/conditions with asthma among children in Memphis, Tennessee. We present a novel method for geo-parsed comorbidity network analysis to show the distinctive patterns of comorbidity networks in urban and suburban areas in Memphis.

Keywords. Comorbidity networks, Pediatric asthma, Health geography, Health disparities, Population health surveillance

1. Introduction

Studying comorbidity networks can help us to detect and reveal hidden dependencies between different health conditions and therefore understand the etiology pathways of diseases [1-5]. Network-based comorbidity studies can generate more realistic pathways identifying overlapping symptoms [5], the relations between co-expressed symptoms and their interactions. However despite the importance of network studies, empirical investigation has not been rigorously applied to pediatric asthma cases [6]. Network analytics can reveal the links and associations between commonly co-observed diseases with asthma, which allows us to systematically explore the complexity of pediatric asthma comorbidity, and design effective preventive and therapeutic interventions.

Asthma is the most common chronic childhood disease, affecting 300 million people worldwide [7]. Asthma prevalence exhibits great geographical variances across different countries and disproportionate distribution within a country [8]. While in some countries less than 2.5% of the population have asthma, in other countries more than 10% of the population bearing the burden of the disease [9]. According to the U.S. Centers for Disease Control and Prevention (CDC), 13.5% of the U.S. population have experienced asthma and the number is growing [10]. Furthermore, the financial burden

¹ Corresponding Author, Center for Biomedical Informatics, Department of Pediatrics, 50 N Dunlap Street, Memphis, TN 38103, USA; E-mail: eshin3@uthsc.edu.

on patients with asthma in the U.S. is estimated approximately \$3,300 per patient for each year [10]. Asthma is more prevalent in children than adults [11]. Early experience of asthma has a wide range of disadvantages. In general, asthma patients are significantly limited in their physical activities and children with asthma are more likely to have more school absences [12, 13]. Most pediatric asthma cases involve a variety of complications, exhibiting complex manifestation of multiple symptoms [14]. In this paper, we first empirically reveal the comorbidity patterns in pediatric asthma and then examine geographical discrepancy between different neighborhood types. Our novel method based on geo-parsed comorbidity network analysis enables us to identify distinctive patterns of pediatric asthma comorbidity networks in urban and suburban areas in Memphis, TN.

2. Data and Methods

We use pediatric asthma encounter records, consist of 4,914 patients who visited a children's hospital in Memphis with asthma from Jan 1st, 2016 to Dec 31st, 2016. Demographically, our dataset consists of pediatric patients from 0 to 21 years old (mean is 7.29) and is composed mainly African American (86.86%) and 62.15% of male. We collected all encounter data with full diagnosis codes for each patient along with their residential zip codes. The medical records are in the ICD-10-CM format where full-length codes, commonly in 5 or 6 digits, are indicating symptom specifications in detail. To avoid the redundancy and over-representation of diseases and simplify the code interpretation, we reduce the ICD codes to 3-digits, which indicates the main disease category. For example, whereas at the 3-digits level, J45 is the code for asthma, at the 6-digits level, there are 18 different codes for asthma. Furthermore, this variation is not consistent across different diseases. In some cases we see over 30 different codes for one disease. By examining medical diagnosis of diseases of a patient, the relational structure of comorbidities can be reconstructed. Using a two-mode network analysis, we transform the matrix of patients by disease to disease by disease networks [15, 16]. We then use R for matrix transformation and Gephi (<https://gephi.org/>) for network visualization. Each node (circle in the network) represents a disease (one 3-digit code) and the line (link) between the two nodes portrays the co-appearance of the two diseases in a patient. By doing so, we generate multiple ego networks of asthma and then we merge them together. We count the first time visit exclusively within the observation window to avoid multi-count of one individual patient. We first generate a pediatric asthma comorbidity network to discover the scope of general comorbidity patterns in pediatric asthma. Next, we split the patients into geographically differentiated units using patients' residential zip codes. Finally, by producing comorbidity networks of the patients who are living in a same zip code area, we explore geo-distinctive patterns of pediatric asthma comorbidity.

3. Results and Findings

We extracted a pediatric asthma comorbidity network from the 4,914 individual medical records (Figure 1). The network has 680 nodes (diseases) and 8664 links (co-diagnoses). In other words, pediatric asthma is highly heterogeneous in its expressed comorbidity patterns and linked to 679 other diseases. This complexity may account for

high prevalence and persistence of asthma and its symptoms. The density of the network is 0.038 which is calculated by the existing number of the ties divided by all theoretically possible ties with a given number of the **nodes**. Overall the density is low due to **a large number of peripheral diseases**. Also, **network density** is not evenly distributed among all nodes. 418 diseases in the network have more than 10 degrees, which means they are connected to 10 other diseases.

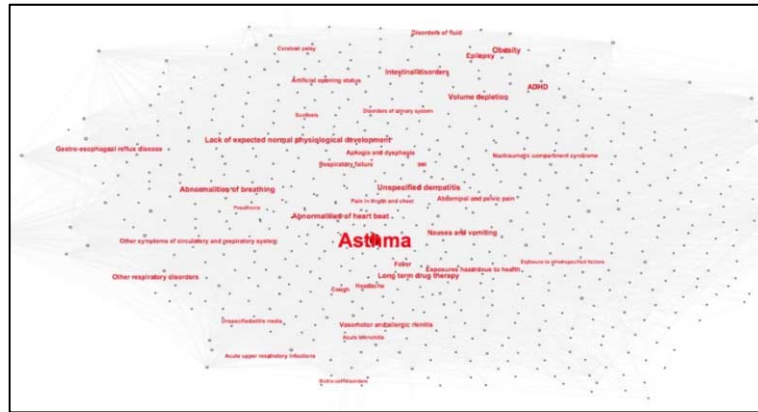


Figure 1. Pediatric Asthma Comorbidity Network

Obesity, the second central condition in the pediatric asthma network, is connected to 211 other diseases/conditions in the network and it is most closely linked to pediatric asthma. Dermatitis is the third largest health condition in the network. Interestingly, lack of normal physiological development and attention-deficit hyperactivity disorders (ADHD) are highly linked symptoms/conditions as well. Intestinal disorders are also commonly observed with pediatric asthma. Some diseases are naturally expectable to co-exist with asthma, i.e. respiratory disorders, and some are not, i.e. lack of physiological development, ADHD, scoliosis, exposures hazardous to health, and sickle-cell disease. Although some of these diseases may not be necessarily linked physiologically, these are pediatric asthma comorbidity patterns we empirically observed.

In continue, we explore if there is any geographical differential pattern of asthma comorbidity networks. We then parsed out patients living in different zip code areas to 29 different groups. Pediatric patients are disproportionately distributed across different areas in Memphis. What is more interesting is the fact that the presence of more patients does not necessarily mean more comorbidity symptoms. Some area exhibit more diverse symptoms than other places independent of the actual patient numbers. In addition, the densities of the networks vary. Furthermore, we found out that the total number of links are not coherently associated either with the number of diseases nor the total number of patients living in the area. Therefore, we can safely conclude that pediatric asthma comorbidity networks are not identical across geographical boundaries. We compare two zip code areas to demonstrate geo-distinctive comorbidity networks. Fixed with the same number of patients ($N=54$) living in a zip code area, the two areas are selected based on their drastic differences in residential conditions – one suburban area and one downtown area. Note that having the same number of patients residing in these areas does not necessarily mean that these two have any equivalent asthma prevalence in general. The suburban area has 9.40% of the residents and the downtown area has 48.50% of the residents under the poverty line, according to the 2010 U.S.

distinctive health problem in the U.S. This linkage between asthma and obesity exacerbates each other [6, 17, 18].

Therefore, understanding asthma comorbidities are essential to design better asthma surveillance systems and deliver effective interventions. The geo-parsed comorbidity network method can improve disease monitoring and the development of locally customized health interventions. Further research with a larger, more vibrant sample size is needed to widely generalize the findings of the current study. The sample population used for this study is composed predominantly African American. The relationships between asthma and other diseases may differ in other ethnic groups. As discussed earlier, we found that pediatric comorbidity network is not identical across different geographical boundaries. This study utilizes the medical records of one hospital in one city. If we can compare another pediatric comorbidity network in a place where asthma prevalence is low, we may be able to better understand some the driving forces of the pediatric asthma in Memphis with a higher degree of accuracy and confidence.

References

- [1] A.L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, *Nature reviews. Genetics* 12 (2011) 56.
- [2] D.S. Lee, J. Park, K. Kay, N. Christakis, Z. Oltvai, A.L. Barabási, The implications of human metabolic network topology for disease comorbidity, *Proc. of the National Academy of Sci.* 105 (2008) 9880-9885.
- [3] C.A. Hidalgo, N. Blumm, A.L. Barabási, N.A. Christakis, A dynamic network approach for the study of human phenotypes, *PLoS computational biology* 5 (2009) e1000353.
- [4] M. Oti, M.A. Huynen, H.G. Brunner, Phenome connections, *Trends in genetics* 24 (2008) 103-106.
- [5] A.O.J. Cramer, L.J. Waldorp, H.L.J.v.d. Maas, D. Borsboom, Comorbidity: A network perspective, *Behavioral and Brain Science* 33 (2010) 137-193.
- [6] E.P. de Groot, E.J. Duiverman, P.L. Brand, Comorbidities of asthma during childhood: possibly important, yet poorly studied, *Eur Respiratory Soc.* 2010.
- [7] M. Masoli, D. Fabian, S. Holt, R. Beasley, The global burden of asthma: executive summary of the GINA Dissemination Committee report, *Allergy* 59 (2004) 469-478.
- [8] D.R. Gold, R. Wright, Population disparities in asthma, *Annu. Rev. Public Health* 26 (2005) 89-113.
- [9] S.S. Braman, The global burden of asthma, *Chest Journal* 130 (2006) 4S-12S.
- [10] N.B. Johnson, L.D. Hayes, K. Brown, E.C. Hoo, K.A. Ethier, C.f.D. Control, Prevention, CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors—United States, 2005–2013, *MMWR Surveill Summ* 63 (2014) 3-27.
- [11] O.J. Akinbami, J.E. Moorman, C. Bailey, H.S. Zahran, M. King, et al., Trends in asthma prevalence, health care use, and mortality in the United States, 2001-2010, *NCHS Data Brief*. 94 (2012) 1-8.
- [12] D. Cypcar, R.F. Lemanske Jr, Asthma and exercise, *Clinics in chest medicine* 15 (1994) 351-368.
- [13] W.R. Taylor, P.W. Newacheck, Impact of childhood asthma on health, *Pediatrics* 90 (1992) 657-662.
- [14] A.M. Fitzpatrick, W.G. Teague, D.A. Meyers, et al. Heterogeneity of severe asthma in childhood: confirmation by cluster analysis of children in the NIH/National Heart, Lung, and Blood Institute Severe Asthma Research Program, *Journal of allergy and clinical immunology* 127 (2011) 382-389.
- [15] M. Latapy, C. Magnien, N. Del Vecchio, Basic notions for the analysis of large two-mode networks, *Social networks* 30 (2008) 31-48.
- [16] S.P. Borgatti, M.G. Everett, Network analysis of 2-mode data, *Social networks* 19 (1997) 243-269.
- [17] J. Gennuso, L.H. Epstein, R.A. Paluch, F. Cerny, The relationship between asthma and obesity in urban minority children and adolescents, *Archives of pediatrics & adolescent medicine* 152 (1998) 1197-1200.
- [18] H. Milgrom, L.M. Taussig, Keeping children with exercise-induced asthma active, *Pediatrics* 104 (1999) e38-e38.

Copyright of Studies in Health Technology & Informatics is the property of IOS Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.