An introduction to network analysis for studies of medication use

Mohsen Askar [a], Raphael Nozal Cañadas [b], Kristian Svendsen [a,*]

[a] *Department of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Norway*
[b] *Department of Informatics, Faculty of Science and Technology, UiT The Arctic University of Norway, Norway*

ABSTRACT

Background: Network Analysis (NA) is a method that has been used in various disciplines such as Social sciences and Ecology for decades. So far, NA has not been used extensively in studies of medication use. Only a handful of papers have used NA in Drug Prescription Networks (DPN). We provide an introduction to NA terminology alongside a guide to creating and extracting results from the medication networks.
Objective: To introduce the readers to NA as a tool to study medication use by demonstrating how to apply different NA measures on 3 generated medication networks.
Methods: We used the Norwegian Prescription Database (NorPD) to create a network that describes the co-medication in elderly persons in Norway on January 1, 2013. We used the Norwegian Electronic Prescription Support System (FEST) to create another network of severe drug-drug interactions (DDIs). Lastly, we created a network combining the two networks to show the actual use of drugs with severe DDIs. We used these networks to elucidate how to apply and interpret different network measures in medication networks.
Results: Interactive network graphs are made available online, Stata and R syntaxes are provided. Various useful network measures for medication networks were applied such as network topological features, modularity analysis and centrality measures. Edge lists data used to generate the networks are openly available for readers in an open data repository to explore and use.
Conclusion: We believe that NA can be a useful tool in medication use studies. We have provided information and hopefully inspiration for other researchers to use NA in their own projects. While network analyses are useful for exploring and discovering structures in medication use studies, it also has limitations. It can be challenging to interpret and it is not suitable for hypothesis testing.

Introduction

Studies in social pharmacy and pharmacoepidemiology often utilize highly complex data and require the use of sophisticated methods to discern important patterns. Data used for quantitative studies in social pharmacy and pharmacoepidemiology can be described as attribute data and relational data. Attribute data includes the characteristics of the studied objects (e.g. sex, age, medication use, sociodemographic information, etc.) while relational data contains the various relationships between subjects. The suitable way of studying attributes data is quantitative analyses, whereas, for relational data, Network Analysis (NA) is the appropriate approach.[1] The subjects studied in network analyses can take many different forms.

A network can be described as a graph that shows the interconnections between a set of actors. Each actor is represented by a *node* and each connection between these nodes is represented by an *edge*.[2] NA is a mathematical approach to study the relationships among nodes.[3] The mathematical background of NA are summarized elsewhere.[4,5]

Network Analysis has its roots in many research disciplines.[6] Network analysis is used, among others, in social studies,[7] ecological studies,[8] genetics[9] and systems pharmacology.[10]

As seen in Fig. 1, a network can be undirected (a and b) or directed (c and d). In a directed network, arrows show the direction of the relationship between nodes. In an undirected network, the relationship does not have a specific direction. The network edges can be weighted (b and d) or unweighted (a and c). In an unweighted network, the two nodes either have a relationship or not, while a weighed network considers the strength of the relationship.
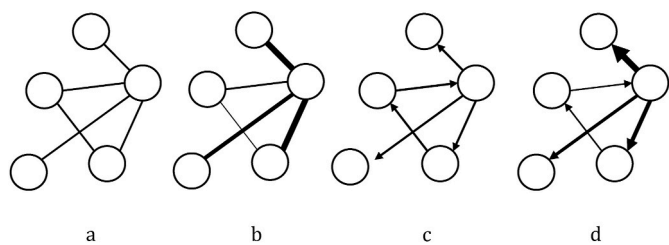

Fig. 1. Different types of networks, a) undirected and unweighted network, b) undirected and weighted network, c) directed and unweighted, and d) directed and weighted.

Use of Network Analysis in public health

Transmission networks have been used to examine the risk of disease transmission by investigating the relations between the infected people and healthy ones.[11–13] Another form of transmission is the transmission of information. NA has been used to visualize the dissemination of public health information to different organizations and consumers. Some network characteristics reveal the pattern and the main actors contributing the most to information spread. Simulated networks can be used to suggest how to accelerate information spread.[14] An example of this type of networks, the diffusion of information among physicians regarding a new drug. The study showed that more socially integrated physicians introduced the drug months before corresponding isolated physicians.[15] NA was also used to study how health workers' professional and personal behavior impact health services.[16,17]

Drug prescription network (DPN)

Pharmacoepidemiological studies of medications that are prescribed or dispensed is a relatively new application of NA. To our knowledge, Cavallo et al. were the first to study a drug prescription network in 2013. They used medications as the nodes and the number of patients being prescribed these medications as weighted edges. They aimed mainly at describing the topology of the co-prescription network to demonstrate which drug classes are most co-prescribed. They also compared the male/female networks and networks from different age strata and found that women in general were co-prescribed more drug classes.[18]

Bazzoni et al. were the first to use the term Drug Prescription Networks (DPN) in their paper published in 2015. They concluded that the DPNs are dense, highly clustered, modular and assortative. In this specific study , density reflected frequent co-prescribing. Modularity suggested that the network could be subdivided into clusters. The study also showed that it is possible to highlight spatial and temporal changes by comparing different networks.[19]

Network Analysis terminology

We organized the key measures that are useful in studies of medication use under 4 main categories: (1) Topology analysis (2) Modularity analysis (3) Network comparison (4) Bipartite networks.[20] (Fig. 2).

1. Topology analysis

Network topological features refer to a group of characteristics, which either describe the network as a whole (network-level) or define individual actors of the network (node-level). There are many topological measures and each of them gives information about a specific network attribute, which then may warrant further investigation.

a. *Global network description (network-level):* A group of measures that describe the network as a whole.
   - Number of nodes: the total number of drugs in the network. The network nodes can be grouped to show the number of drugs in each drug class. Different networks of different populations will have different distributions of drugs in the drug classes.
   - Density: the density of a network is the number of actual edges divided by the total number of edges that would exist if all the nodes in the network were connected. This potential number can be calculated by the formula below where *n* is the number of nodes:

$$\frac{n \times (n-1)}{2}$$

The network density can be useful in terms of comparison between different networks that describe the same type of drug-drug relation.

Assortativity: a network is assortative when the nodes that share a similar trait tend to connect. This trait can be many characteristics such as the nodes' degree. In this case, the assortativity means that nodes with a high number of edges tend to connect. Assortativity can be examined in terms of other common characteristics between the nodes as well. Assortativity coefficient is measured using Pearson correlation coefficient. Assortativity coefficient is scaled between −1 and 1, where 1 is most assortative.[21]

b. Node-level measures

Node-level measures describe the features of the different nodes across the network.
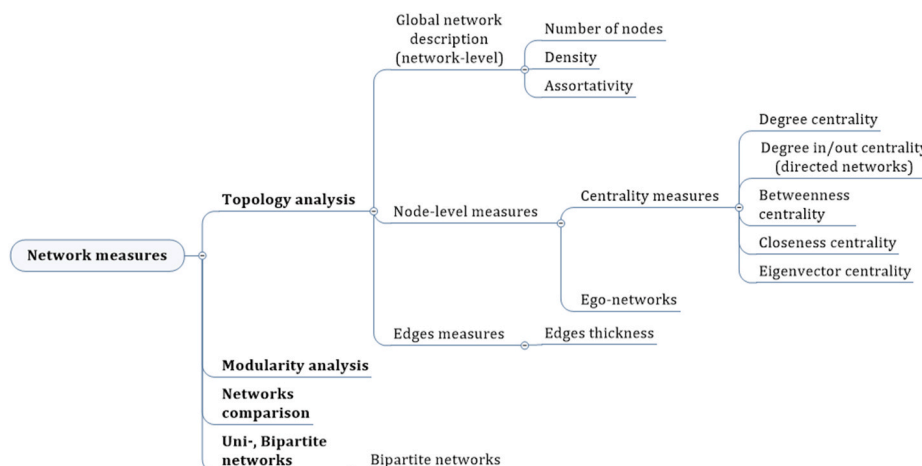

Fig. 2. Summarizing some of the Network Analysis measures that can be useful in the studies of medication use.

Centrality measures

Centrality measures indicate the importance of the network nodes by assigning a score to each of them. There are many different centrality measures and each of them can be used to describe a specific type of importance. By comparing the different centrality measures of a node, we can understand the different ways a node is influential to the network. This paper will discuss 4 of the most common types of centrality measures: degree, betweenness, closeness, and eigenvector centrality. The mathematical explanations of these measures are mentioned here.[22,23]

Degree centrality

Degree centrality is the number of edges that are connected to a node. A higher score indicates that the node is connected to many other nodes. Node A in Fig. 3 has a degree score of 4. In a directed network, the degree is split into In-degree, which is the number of edges that direct to a node and Out-degree, which is the number of edges that originate from the node. In- and out-degrees will therefore show the directions of relationships in a directed network. In Fig. 3, nodes C and D have an in-degree score of 3, while nodes A and G have an out-degree score of 3.

Betweenness centrality

The betweenness centrality of a node indicates how many times this node was used to connect two other nodes by the shortest possible path. Increasing the number of shortest paths will increase the betweenness centrality score.[22] In Fig. 3, node A has the highest betweenness centrality score of 1.5.

Closeness centrality

It is a measure of the average distance between the node and all other nodes in the network. Nodes with the highest closeness score have the shortest distances to all other network nodes. The nodes A, B and F have the highest closeness centrality score of 1.

Eigenvector centrality

It is a measure of the importance of a node in a network based on the node's connections with other vital nodes. Relative scores are given to all nodes in the network based on the concept that connections to high-scored nodes give 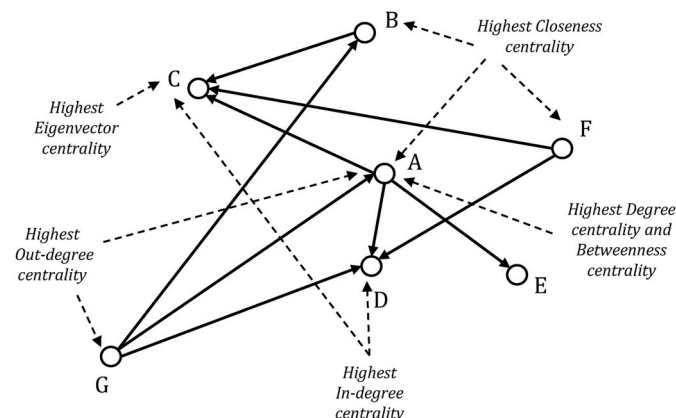a higher score to the node than equal connections to low-scored nodes. In other words, a high eigenvector score means that a node is connected to many nodes, which themselves are connected to important nodes in the network and have high scores of eigenvector centrality. This means that a node with a high eigenvector centrality score is not necessarily connected to the highest number of nodes in the network but is connected to the nodes with a high number of edges.[24] Node C in Fig. 3 has the highest eigenvector centrality score of 1. Assigning the centrality of each node in the network may lead us to visualize the network from a single specific important node perspective; this is called an *Ego-network* and it visualizes the part of the network that has the node of interest and the nodes that are directly connected to it.

c. Edge-level measures

Edge-thickness: in a weighted network, the edge-thickness represents a quantitative measure of the strength of the connection between two nodes. This representation is unique for NA and can be used to study many research questions. We will show an example where the number of users that co-medicated a pair of medications are used to represent the edge-thickness. In this context, thicker edges represent more frequently used pairs of medications.

2. Modularity analysis (Community detection)

One key feature of the network structure is its modularity. A module is a group of nodes that have many connections between each other and few(er) connections to the other nodes in the network.[25] There are many techniques of community detection including density-based, centrality-based, partition-based and hierarchical clustering techniques[20,26,27]

3. Network comparison

It is possible to compare two or more networks to show the changes over time (temporal), between different areas (spatial), or between different groups of patients. These comparisons can be done by comparing the characteristics of the networks to highlight the differences in numbers and influences of the nodes. Another way to compare different networks is to subtract or divide the values of the edges between two networks. This will create edges representing the differences between the networks. By comparing many networks, dynamic graphs can be created showing the topological changes from a network to the next. Nodes will appear, disappear or change their locations as the dynamic graph moves through the different networks.[28]

4. Bipartite networks

A network can be uni- or multipartite. We will only discuss uni- and bipartite networks. Unipartite networks have one set of nodes, while in bipartite networks the nodes belong to two disjointed sets (such as prescribers and patients). In a bipartite network, edges connect the nodes from different sets[29,30] (Fig. 4).

The aim of this paper is to introduce the readers to NA as a tool to

Fig. 3. Illustrating the different types of centrality. Node (A) represents the highest score of Degree and Betweenness centralities. The highest Eigenvector centrality score is assigned to the node (C). Nodes (A, B, F) have similar closeness Centrality. Nodes with the most in-degree edges are (C, D), while (A, G) have the most out-degree edges.
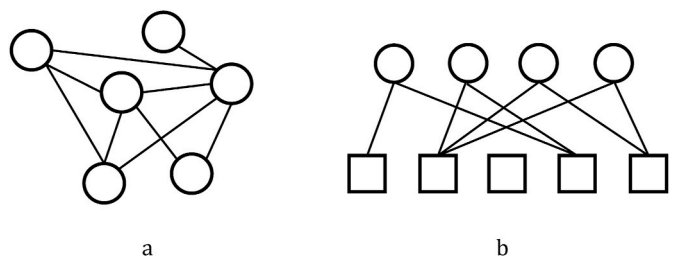
Fig. 4. (a) Unipartite network consisting of one type of nodes; (b) bipartite network consisting of two different types of nodes (circles and squares) in which the edges link between nodes from different types.

study medication use by demonstrating a practical real-life example of medication use in the elderly in Norway whenever it is possible, otherwise by giving an example from other studies.

Methods

We created a network of co-medication in elderly persons in Norway. We also created a network describing the severe drug-drug interactions (DDIs). Finally, we generated a network with the actual use of drugs with severe DDIs by combining the previous two networks.

Data sources

Co-medication network

The dataset used comes from the Norwegian Prescription Database (NorPD). It covers all dispensed prescriptions to elderly persons (≥65 years) in Norway between 2012 and 2014. The NorPD collects data from all pharmacies in Norway and covers all outpatient dispensing for the entire Norwegian population. Details on the NorPD are published elsewhere.[31] In total, the dataset included 765,383 patients, 344,285 men (45%) and 421,098 women (55%) with 75 years as mean patient age. Edges in this network represent the number of patients who combined pairs of medications. In order to define the co-medication, we created treatment episodes using the Proportion of Days Covered (PDC) approach.[32] We assumed that patients used one Defined Daily Dose (DDD)[33] per day and added 20% to each prescription duration to account for imperfect adherence. We also allowed a medication-free gap of 14 days before ending a treatment episode and starting another. This means that if the patient exceeds 14 days without the medication, the treatment episode for this medication ends and a new episode starts if the patient picks up a new prescription. Finally, co-medication was defined as the overlapping drug treatment episodes at the index date, January 1, 2013.

For each pair of nodes (drugs), we summed up the number of co-medication occurrences (i.e. number of patients combining these two drugs) to create a weighted and undirected network.

We excluded the medications that have no defined DDD such as the medications for topical use, vaccines, and ophthalmologicals. In total, we excluded 357 medications (217 local and 140 systematic drugs). The co-medication network is shown here: https://mohsenaskar.github.io/co-medication/network/. The network is searchable by substance name. Clicking on any node shows the ego-network of this node as well as some network measures.

Severe drug-drug interactions network

To create this network, we used a dataset derived from the Norwegian Electronic Prescription Support System (FEST). FEST is a national information service that provides common pharmaceutical data to the IT-systems that are involved in the drug prescribing process including systems used by physicians, hospitals and pharmacies.[34] Drug-drug interactions is a part of the FEST database. In FEST, the DDIs are divided into 3 categories; interactions that should be avoided (i.e. severe), interactions where precautions should be taken and interactions that do not require any action. Only severe DDIs were included in the study. There were 57,151 unique severe interactions. The edges in this network represent the presence of a severe interaction between the two nodes.

The network is undirected and unweighted. The severe DDIs network is shown here: https://mohsenaskar.github.io/DDI/network/

Combining co-medication and DDIs networks

Both DDI and co-medication network has drugs as nodes. When combining the two networks only edges that exists in both networks are included (only edges with any users combining the medications and where there is a severe DDI). The number of users for each edge from the co-medication network becomes the weight of the edges in the combined network.

This network is shown here: https://mohsenaskar.github.io/DDI-in-co-medication-network/network/

Preparing the data to create a network

The data from the NorPD contains attributable data including a patient identity number, sex, year of birth, and data about each individual dispensed drug. To create a network, this data needed to be reshaped. The first step was to create a file with only medications that were used on the index date. Secondly, the file was aggregated such that an edge list was created. The edge list contains 2 variables defining the pairs of drugs and one variable with the number of users co-medicating with each pair of drugs. This edge list can be used by various software as described below. The process of data preparation is summed up in Fig. 5 and the edge list is openly available at the UiT The Arctic University of Norway open data repository here: https://dataverse.no/dataset.xhtml?persistentId=doi:10.18710/1OUTYI.The Stata syntax for creating the edge list is supplied in supplementary 2.

Software to use for network analysis

There are many available tools to use for NA. We will focus on how to use the *nwcommands* suite of commands that can be downloaded into *Stata* and the *igraph* package in *R* as well as visualization in *Gephi*. Other packages like "*igraph*" or "*NetworkX*" for *Python* are popular as well. All these packages can be used for visualizing and computing different network measures with differences in their integrated features and performance.[35,36]

Using Stata (nwcommands, nwANND)

Using the edge list, *nwcommands* will create an adjacency matrix.[37] The adjacency matrix is a square matrix that contains the relationships between every pair of nodes in the network. The adjacency matrix can be saved as *Pajek* format that can be later imported and used by *Gephi*. In addition, *nwcommands* can display some network measures on both the network and node-level. *NwANND* is used for calculating the assortativity coefficient.[38] The syntax can be found in supplementary 2.
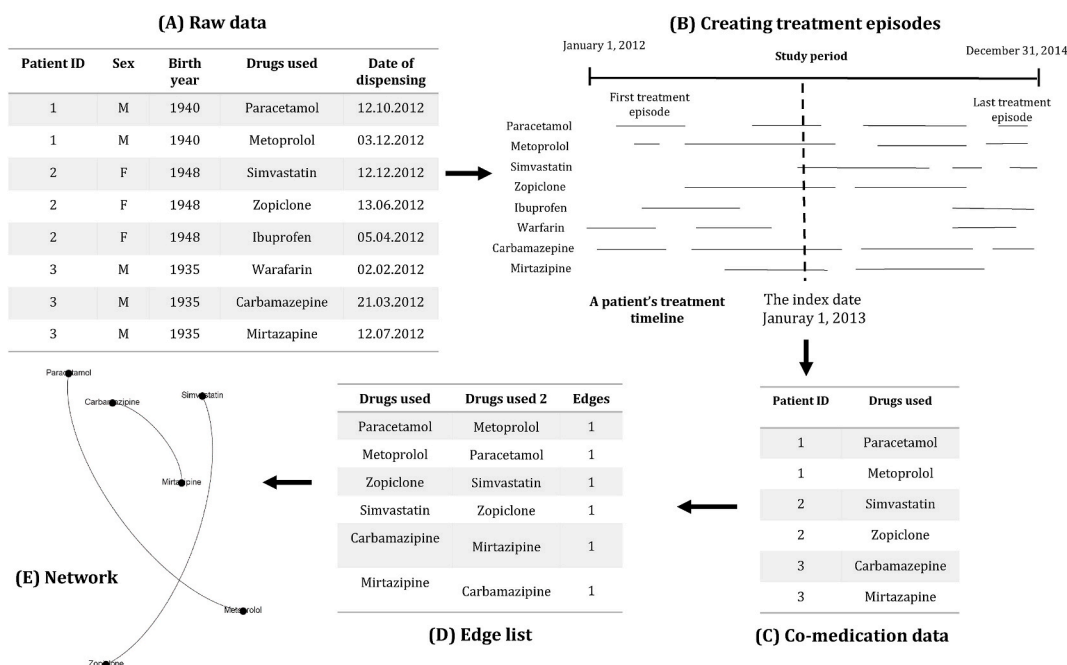
Using R (igraph)

Igraph (https://igraph.org/) is a library for creating and analyze graphs. It is widely used by network researchers to analyze graphs and networks. It is currently available for *C*, *C++*, *Python*, *R* and *Mathematica*.

One of the strengths of *igraph* is that it can be programmed with a high-level programming language and still be very efficient when handling large networks. In our *R* context, *igraph* integrates well with the visualization package (ggplot2) via the *ggraph* library.

Igraph uses an edge list and can link it with attribute data for each node as well. An example code for network visualization using *Igraph* and *ggraph*, is given in Supplementary 2.

Using gephi

Gephi (https://gephi.org/users/) is an open-source and free stand-alone software. The software can handle small to medium-sized networks (up to 150000 nodes). *Gephi* is user-friendly and requires no programming experience.[36] With many visualizing layouts and network measures,[39] *Gephi* can provide a good starting point for the drug network study.[39] After importing the adjacency matrix to *Gephi*, we can process the network by applying different visualizing layouts, adding filters and colors. The structure of most drug networks can be complex and the unprocessed form of the network is often uninformative. By using different attributes (e.g. sex, modularity, etc.) the network can become more easily interpretable.[28]

(A) Raw data

Patient ID	Sex	Birth year	Drugs used	Date of dispensing
1	M	1940	Paracetamol	12.10.2012
1	M	1940	Metoprolol	03.12.2012
2	F	1948	Simvastatin	12.12.2012
2	F	1948	Zopiclone	13.06.2012
2	F	1948	Ibuprofen	05.04.2012
3	M	1935	Warafarin	02.02.2012
3	M	1935	Carbamazepine	21.03.2012
3	M	1935	Mirtazipine	12.07.2012

(B) Creating treatment episodes

(E) Network

(D) Edge list

Drugs used	Drugs used 2	Edges
Paracetamol	Metoprolol	1
Metoprolol	Paracetamol	1
Zopiclone	Simvastatin	1
Simvastatin	Zopiclone	1
Carbamazipine	Mirtazipine	1
Mirtazipine	Carbamazipine	1

(C) Co-medication data

Patient ID	Drugs used
1	Paracetamol
1	Metoprolol
2	Simvastatin
2	Zopiclone
3	Carbamazepine
3	Mirtazapine

Fig. 5. Data preparation process. A) Raw clinical data in the long-form. B) Creating treatment episodes, a line indicates an episode. Gaps between lines indicate a medication-free period of more than 14 days. C) Data including only medications used at the index date D) Creating the edge list including 3 variables; two variables represent the nodes (Drug used, Drug used 2) and the third variable is the edge weight between the pair of nodes. E) Using the software to generate the network.

Results

We will present results from our networks using the same terms and order as in the introduction (Fig. 2). Table 1 provides the main topological features of the co-medication and the severe DDIs networks. The co-medication network is denser than the severe DDIs network, indicating that the drugs in the co-medication network are more connected. The assortativity coefficient shows that the co-medication network is non-assortative in terms of degree similarity, while the severe DDIs network is more assortative. Centrality measures in the co-medication network revealed that the same 5 drugs are the most central in all measures, while in the severe DDIs network; there is more variation in the top 5 drugs in each centrality measure. The results also showed that both networks are modular.

Fig. 6 shows that the majority of anatomical drug classes were assortative. This means that the drugs from the same anatomical group tend to be more co-prescribed. We also investigated the assortativity of the drugs on the pharmacological level (3rd level Anatomical Therapeutic Chemical classification) in supplementary 3.

Ego-networks as a measure can be seen by accessing the online networks we created and selecting individual nodes. The different network links can be found in the method section.

The top 10 edge weights for the severe DDIs in the co-medication network and co-medication only network are shown in Tables 2 and 3 respectively. We see in Table 2 that the number of patients using drugs causing severe DDIs are relatively low (less than 1000 users for all) while the most commonly co-medicated drugs seen in Table 3 is much higher with acetylsalicylic Acid (aspirin) and simvastatin having around 83000 users representing almost 11% of the population.

Modularity analysis

We found 4 modules in the co-medication network and 11 modules in the severe DDI network. For the co-medication network, there was one large community and 3 other smaller communities. Nervous and Respiratory system groups (N- and R-groups) drugs are just found in module 0, while Cardiac-, Alimentary-, Blood groups(C-, A-, B- groups)

are common groups between modules 1 and 2, but with considering the number of users in each module we can locate in which module these ATC groups represent the most importance. Drugs used for diabetes, (A10) group, present only in module 2. The complete tables of modules are listed in supplementary 1. For the severe DDIs network, the modules found are shown below in Fig. 7.

Discussion

A Network visualizes the relationships of a dataset in one graph. This unique ability of data representation is combined with many measures that are helpful for many research disciplines. A starting point for generating any network is to select the nodes and define the edges. A precise definition of the edges allows the researcher to extract the correct information. NA is a well-suited approach to study complex systems. Although the approach has been widely used in many fields of research, only a few studies studied the drug-relations in a network.[18,19]
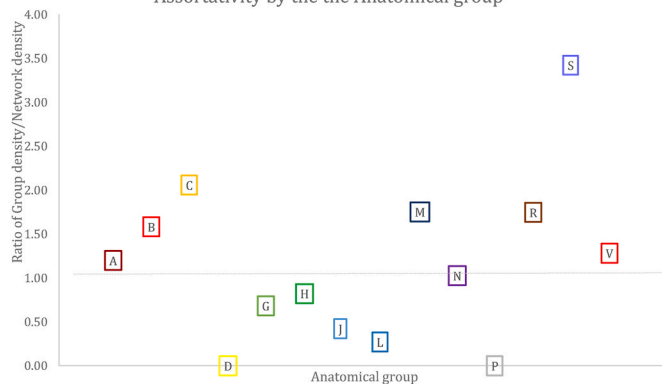
Our results show that many network outcomes can be useful in the studies of medication use. Moreover, some results are unique measures that only NA can perform such as edge measure and modules detection. Employing centrality measures in the drug study introduce an opportunity to observe the influence of the different drugs in the drug-network. Determining this influence can be useful for clinicians and decision-makers.

After generating a network, some topological features have to be reported first to get a general idea about the network content and its basic characteristics. Network-level measures such as assortativity and density reveal many clues for further investigation. Centrality measures show how influential each node is in the network. It is possible to have high centrality of one type and a low of another for the same node.[8] In order to study the importance of the nodes, it is necessary to use more than one measure of centrality. Recent studies suggest using centrality measures as an alternative approach for variables selection. Lutz et al. used the centrality measures to identify 4 additional variables contributing to the predicting of treatment dropout in patients with anxiety disorders.[40] Valenzuela et al. described a methodology based on degree and centrality measures to obtain the most representative variables for

Table 1

The topological measures of co-medication and severe DDIs networks.

Outcome	Co-medication network	DDIs network	Indicates
1. Topology analysis			
a) Network-level measures			
Number of nodes	762	1699	Number of drugs present in the network.
Number of edges	75052	57151	Number of connections between the network nodes
Density	0.26	0.04	The extent of connections between the network's nodes
Average degree	99	34	The average number of connections that each node has.
Assortativity co-efficient	−0.26	0.4	To what extent nodes with higher degree tends to correlate.
b) Node-level measures			
Centrality measures			
Nodes with the highest Degree centrality scores	Acetylsalicylic acid, Simvastatin, Zopiclone, Paracetamol, Metoprolol	Typhoid vaccine, Erythromycin, Prikkperikum, Clarithromycin, Moxifloxacin	Combining these centrality measures can be used to assign the importance of each drug to the network.
Nodes with the highest Betweenness centrality scores	Acetylsalicylic acid, Zopiclone, Simvastatin, Paracetamol, Metoprolol	Typhoid vaccine, Padeliporfin, Hyperici herba (St John's-wort), Tuberculosis vaccine, Ginkgo leaves	
Nodes with the highest Closeness centrality scores	Acetylsalicylic acid, Simvastatin, Zopiclone, Paracetamol, Metoprolol	Bromelains, Telbivudine, Peg interferon alfa-2a, Diazepam, Oxazepam	
Nodes with highest Eigenvector centrality scores	Acetylsalicylic acid, Simvastatin, Zopiclone, Metoprolol, Paracetamol	Typhoid vaccine, Erythromycin, Clarithromycin, Chloramphenicol, Moxifloxacin	
c) Edge-level measures			
Average path length	1.77	3.09	Average shortest path between two nodes.
Thickest edge weight	82948	1	For the weighted co-medication network the number reflects the highest number of patients co-medicating, This highlights clinically important combinations.
Edges range	1–82948	0–1	
2. Modularity			
Modularity	0.088	0,54	Indicates presence of modules in the network.
Number of modules (communities)	4	11	
Number of nodes in largest module	530 (module 0)	372 (module 4)	

Fig. 6. Assortativity of network nodes in terms of similarity by the anatomical group. Squares above 1 represent a drug group with a higher density than the general density of the network (0.26). The S (Sensory organs) anatomical group had the highest assortativity. D (Dermatologicals) and P (Antiparasitic products) groups had no edges because these drug classes were excluded from the study.

Table 2

The top 10 clinically relevant severe DDIs in the co-medication network.

	The severe DDI drug pair		No. of patients co-medicating
1	Codeine and paracetamol	Tramadol	855
2	Esomeprazole	Clopidogrel	823
3	Simvastatin	Carbamazepine	480
4	Metoprolol	Paroxetine	454
5	Metoprolol	Verapamil	380
6	Lansoprazole	Clopidogrel	308
7	Diclofenac	Ibuprofen	305
8	Diazepam	Oxazepam	300
9	Carbamazepine	Zopiclone	280
10	Omeprazole	Clopidogrel	277

Table 3

The top 10 combined drugs in the co-medication network.

	Most combined drugs		No. of patients co-medicated
1	Acetylsalicylic acid	Simvastatin	82948
2	Acetylsalicylic acid	Metoprolol	52577
3	Acetylsalicylic acid	Atorvastatin	42753
4	Metoprolol	Simvastatin	36792
5	Acetylsalicylic acid	Amlodipine	32628
6	Acetylsalicylic acid	Zopiclone	29173
7	Amlodipine	Simvastatin	22554
8	Acetylsalicylic acid	Ramipril	19660
9	Simvastatin	Zopiclone	18845
10	Metformin	Acetylsalicylic acid	18507

predicting successful aging.[41] These approaches are interesting and represent an alternative method to the other variable selection methods. Edge-level measures are the core of the networks and the principal for many network measures.

Modularity analysis exposes the network structure. This measure is believed to introduce special importance in the drug study. Bazzoni et al. found the DPN to be modular,[19] which is consistent with what we found in our networks. Further investigation is needed to assess the underlying patterns in the modules we found in our co-medication network. Modules can be interpreted as clusters of patients with similar diagnoses using the same medications. In our initial analysis of modularity, we identified 4 modules, further work could be done to identify smaller groups by detecting the sub-clusters inside each module. Modules in the DDI network could be connected to pharmacological data to see the importance of pharmacokinetic interactions through systems such as the cytochrome P450 system. We have not explored this but there is a great
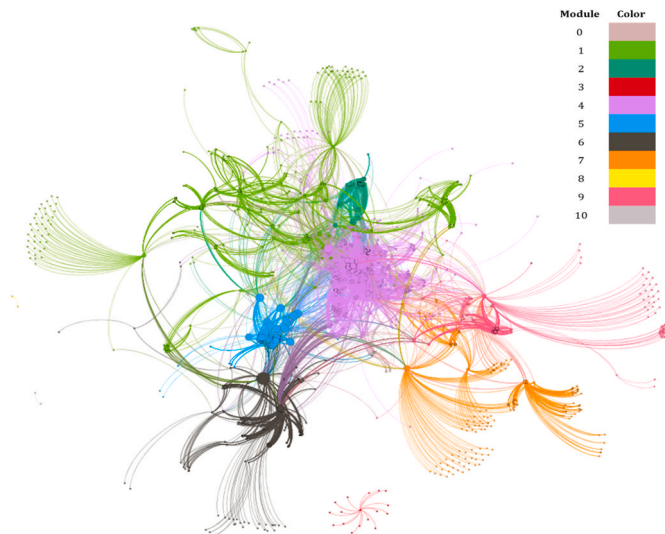
Fig. 7. The 11 modules that were detected in the severe DDIs network. Different colors indicate different modules. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 8. An example of a bipartite network representing a sub-graph of two types of nodes (i.e. prescribers and patients) linked by the number of Fentanyl® patches prescriptions. The bigger nodes indicate more number of connections. The thicker edges indicate a higher number of prescriptions. (redrawn from "Network analysis and visualization of opioid prescribing data" [43]).

potential in using modularity analyses in order to understand networks.

Comparing different networks can reveal the change in patterns over time, place and different populations. Networks that describe the relation between drug-use and morbidities for a patient or a group of similar patients over time may identify the development of co-morbidities and drug use.
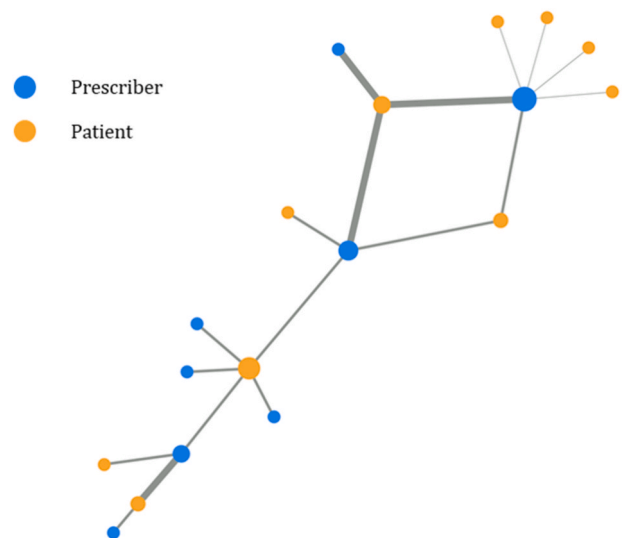
Bipartite networks provide a variety of possibilities to study many situations in which drugs are involved with other network actors such as physicians or diseases. Dasgupta & Chawla created a bipartite drug-disease network to study the interactions between drugs and co-morbidities.[42] Hu et al. studied the prescribing of some opioids by creating a bipartite network of patients and prescribers and using the network to analyze the relationship between patients and prescribers and detect "doctor shopping" and suspicious network nodes.[43] A redrawn example from this study is shown in Fig. 8.

Our study has some limitations. As we used the DDD to outline the treatment episodes, we excluded the medications that have no defined DDD. This reduced the represented co-medication in our networks to the actual co-medication at the index date.

NA also has some important limitations. As a tool, it can be used to explore data, to find unusual structures, group nodes together and find unusual individual nodes. However, it can be hard to interpret results from NA and it is only suited for hypothesis generation. It also cannot explore many sets of relationships between variables at the same time as well as determining causal relationships. For such research questions, other hypothesis testing methodologies will be more needed. However, in research focused on exploration, NA can be a valuable tool.

Conclusion

The main purpose of this paper was to demystify the NA as a method. We have explained the terminology of network analyses and showed, with examples, how network analyses can be used for hypothesis generation. The online links to our networks visualize the data much better than a static picture can and we hope that we have provided enough information, and inspiration, to explore how you can use NA on your own data. We are confident that the future will see many new applications of NA and interesting results for researchers in social pharmacy and pharmacoepidemiology.

CRediT authorship contribution statement

Mohsen Askar: Conceptualization, Methodology, Software, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Raphael Nozal Cañadas:** Software, Writing – review & editing. **Kristian Svendsen:** Conceptualization, Software, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no conflicts of interest related to this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sapharm.2021.06.021.

References

1. Scott J. *Social Network Analysis*. SAGE Publications Ltd; 2017. https://doi.org/10.4135/9781529716597.
2. Prell C. *Social Network Analysis: History, Theory & Methodology*. Sage; 2012.
3. Wasserman S, Faust K. *Social Network Analysis*. Cambridge University Press; 1994. https://doi.org/10.1017/CBO9780511815478.
4. Newman M. *Networks: An Introduction*. Oxford University Press, Inc.; 2010.
5. Newman MEJ. Mathematics of networks. In: *The New Palgrave Dictionary of Economics*. Palgrave Macmillan UK; 2018:8525–8533. https://doi.org/10.1057/978-1-349-95189-5_2565.
6. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. *Annu Rev Publ Health*. 2007;28:69–93. https://doi.org/10.1146/annurev.publhealth.28.021406.144132.

7. Freeman L. *The Development of Social Network Analysis, A Study in the Sociology of Science.* Vancouver, BC Canada: ΣP Empirical Press; 2004.

8. Delmas E, Besson M, Brice M-H, et al. Analysing ecological networks of species interactions. *Biol Rev.* 2019;94(1):16–36. https://doi.org/10.1111/brv.12433.

9. Ghazalpour A, Doss S, Zhang B, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet.* 2006;2(8):e130. https://doi.org/10.1371/journal.pgen.0020130.

10. Zhao S, Iyengar R. Systems pharmacology: network analysis to identify multiscale mechanisms of drug action. *Annu Rev Pharmacol Toxicol.* 2012;52(1):505–521. https://doi.org/10.1146/annurev-pharmtox-010611-134520.

11. Ken TDE, Matt JK. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc Natl Acad Sci U S A.* 2002;99(20):13330. https://doi.org/10.1073/pnas.202244299.

12. Keeling MJ, Eames KTD. Networks and epidemic models. *J R Soc Interface.* 2005;2 (4):295–307. https://doi.org/10.1098/rsif.2005.0051.

13. Grande KM, Stanley M, Redo C, Wergin A, Guilfoyle S, Gasiorowicz M. Social network diagramming as an applied tool for public health: lessons learned from an HCV cluster. *Am J Public Health.* 2015;105(8):1611–1616. https://doi.org/10.2105/AJPH.2014.302193.

14. Valente TW, Davis RL. Accelerating the diffusion of innovations using opinion leaders. *Ann Am Acad Polit Soc Sci.* 1999;566(1):55–67. https://doi.org/10.1177/0002716299566001005.

15. Coleman J, Katz E, Menzel H. The diffusion of an innovation among physicians. *Sociometry.* 1957;20(4):253–270. https://doi.org/10.2307/2785979.

16. Poghosyan L, Lucero RJ, Knutson AR, Friedberg M W, Poghosyan H. Social networks in health care teams: evidence from the United States. *J Health Organisat Manag.* 2016;30(7):1119–1139. https://doi.org/10.1108/JHOM-12-2015-0201.

17. Scott J, Tallia A, Crosson JC, et al. Social network analysis as an analytic tool for interaction patterns in primary care practices. *Ann Fam Med.* 2005;3(5):443. https://doi.org/10.1370/afm.344.

18. Cavallo P, Pagano S, Boccia G, De Caro F, De Santis M, Capunzo M. Network analysis of drug prescriptions. *Pharmacoepidemiol Drug Saf.* 2013;22(2):130–137. https://doi.org/10.1002/pds.3384.

19. Bazzoni G, Marengoni A, Tettamanti M, et al. The drug prescription network: a system-level view of drug co-prescription in community-dwelling elderly people. *Rejuvenation Res.* 2015;18(2):153–161. https://doi.org/10.1089/rej.2014.1628.

20. Kim E-Y. Knowledge-based bioinformatics: from analysis to interpretation, edited by gil alterovitz & marco Ramoni. *Healthc Inform Res.* 2010;16(4):312. https://doi.org/10.4258/hir.2010.16.4.312.

21. Barrenas F, Chavali S, Holme P, Mobini R, Benson M. Network properties of complex human disease genes identified through genome-wide association studies. In: Mailund T, ed. *PLoS One.* vol. 4. 2009, e8090. https://doi.org/10.1371/journal.pone.0008090, 11.

22. Freeman L, Freeman L. A set of measures of centrality based on betweenness. *Sociometry.* 1977;40(1):35–41. https://doi.org/10.2307/3033543.

23. Bonacich P. Power and centrality: a family of measures. *Am J Sociol.* 1987;92(5):1170–1182. https://doi.org/10.1086/228631.

24. Ruhnau B. Eigenvector-centrality — a node-centrality? *Soc Networks.* 2000;22(4):357–365. https://doi.org/10.1016/S0378-8733(00)00031-9.

25. Ji X, Machiraju R, Ritter A, Yen PY. Examining the distribution, modularity, and community structure in article networks for systematic reviews. *AMIA. Annu Symp proceedings AMIA Symp.* 2015;2015:1927–1936.

26. Zhang M, Lu LJ. Modules in networks, algorithms and methods. In: *Encyclopedia of Systems Biology.* New York: Springer; 2013:1447–1450. https://doi.org/10.1007/978-1-4419-9863-7_557.

27. Newman MEJ. Modularity and community structure in networks. *Proc Natl Acad Sci Unit States Am.* 2006;103(23):8577. https://doi.org/10.1073/pnas.0601602103.

28. Cherven K. *Mastering Gephi Network Visualization Produce Advanced Network Graphs in Gephi and Gain Valuable Insights into Your Network Datasets.* 2015. Published online.

29. Introduction to bipartite graphs. In: Asratian AS, Häggkvist R, Denley TMJ, eds. *Bipartite Graphs And Their Applications. Cambridge Tracts in Mathematics.* Cambridge University Press; 1998:7–22. https://doi.org/10.1017/CBO9780511984068.004.

30. Guillaume J-L, Latapy M. Bipartite structure of all complex networks. *Inf Process Lett.* 2004;90(5):215–221. https://doi.org/10.1016/j.ipl.2004.03.007.

31. Furu K. Establishment of the nationwide Norwegian Prescription Database (NorPD); new opportunities for research in pharmacoepidemiology in Norway. *Epidemiology.* 2008;18(2):129–136. https://doi.org/10.5324/nje.v18i2.23.

32. Raebel AM, Schmittdiel JJ, Karter LA, Konieczny FJ, Steiner FJ. Standardizing terminology and definitions of medication adherence and persistence in research employing electronic databases. *Med Care.* 2013;51(Suppl 8):S11–S21. https://doi.org/10.1097/MLR.0b013e31829b1d2a, 8 Suppl 3.

33. Üstün TB. International classification systems for health. In: *International Encyclopedia of Public Health.* Elsevier; 2017:304–311. https://doi.org/10.1016/B978-0-12-803678-5.00237-X.

34. Norwegian Medicines Agency. Fest implementation guidelines. https://lege middelverket.no/Documents/Andretemaer/FEST/Hvordan.bruke.FEST/Impleme ntation.guide.FEST.v3.0.pdf; 2019.

35. Pavlopoulos GA, Paez-Espino D, Kyrpides NC, Iliopoulos I. Empirical comparison of visualization tools for larger-scale network analysis. *Adv Bioinformatics.* 2017;2017:1–8. https://doi.org/10.1155/2017/1278932.

36. Akhtar N. Social network analysis tools. In: *Fourth International Conference on Communication Systems and Network Technologies. IEEE; 2014.* 2014:388–392. https://doi.org/10.1109/CSNT.2014.83.

37. Grund T. Network analysis using stata. https://nwcommands.wordpress.com/tutor ials-and-slides/; 2015.

38. Joyez C. NWANND: Stata module to compute ANND (average nearest neighbor degree) and related measures. *Stat Softw Components*; 2016. Published online December 18 https://ideas.repec.org/c/boc/bocode/s458261.html. Accessed October 29, 2020.

39. Bastian M, Heymann S, Jacomy M. *Gephi: An Open Source Software for Exploring and Manipulating Networks.* 2009. https://doi.org/10.13140/2.1.1341.1520.

40. Lutz W, Schwartz B, Hofmann SG, Fisher AJ, Husen K, Rubel JA. Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: a methodological proof-of-concept study. *Sci Rep.* 2018;8(1):7819. https://doi.org/10.1038/s41598-018-25953-0.

41. Valenzuela JFB, Monterola C, Tong VJC, Fülöp T, Ng TP, Larbi A. Degree and centrality-based approaches in network-based variable selection: insights from the Singapore Longitudinal Aging Study. In: Mariño IP, ed. *PLoS One.* vol. 14. 2019, e0219186. https://doi.org/10.1371/journal.pone.0219186, 7.

42. Dasgupta D, Chawla NV. *Disease and Medication Networks: An Insight into Disease-Drug Interactions.* 2014:1.

43. Hu X, Gallagher M, Loveday W, Dev A, Connor JP. Network analysis and visualisation of opioid prescribing data. *IEEE J Biomed Heal Informatics.* 2020;24(5):1447–1455. https://doi.org/10.1109/JBHI.2019.2939028.