2018

An Evaluation of Provider-Sharing Networks of Patients Using Prescription Opioids

Hilary A. Aroke
University of Rhode Island, aharoke@gmail.com

AN EVALUATION OF PROVIDER-SHARING NETWORKS OF PATIENTS USING PRESCRIPTION OPIOIDS

BY

HILARY A. AROKE

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

STATISTICS

UNIVERSITY OF RHODE ISLAND

2018

MASTER OF SCIENCE THESIS

OF

HILARY A. AROKE

APPROVED:

Thesis Committee:

Major Professor Natallia V. Katenka

Ashley L. Buchanan

Prabhani Kuruppumullage Don

Stephen S. Kogut

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2018

ABSTRACT

IMPORTANCE: Patients using prescription opioid are embedded in a network due provider-sharing and living in the same community. As a result, they may exert influence on each other's treatment preferences and share attitudes towards prescription opioid use and misuse.

OBJECTIVE: To determine patient characteristics associated with the observed pattern of shared prescribers in a network and identify influential patients in the network.

DESIGN, SETTING, AND PARTICIPANTS: We conducted a cross-sectional network-based study using the Rhode Island (RI) Prescription Drug Monitoring Program (PDMP) data for the 2015 calendar year. All patients who filled at least one opioid prescription at a retail pharmacy were eligible. The analysis was limited to patients who were on a stable opioid regimen and used only one source of payment, and filled only one type of opioid medication (oxycodone, hydrocodone or buprenorphine/naloxone) from ≥ 3 prescribers, and visited ≥ 3 pharmacies during the year. To minimize the influence of less relevant network connections, we excluded institutional providers and providers who issued opioid prescriptions to ≤ 6 patients. We applied social network analysis (SNA) methods to a sample of 372 patients connected to each other through provider-sharing. We used the exponential random graph model (ERGM) assuming conditional dyadic independence to examine the relationship between patient attributes and the likelihood of forming network ties. Homophily was defined as the tendency of patients to associate with others who have similar characteristics. Three centrality measures (degree, closeness, and betweenness)

were used to identify patients with potential influence in the opioid prescription network.

MAIN OUTCMES AND MEASURES: We provide a visual and descriptive characterization of the network, used centrality measures to identify influential patients, and ERGM to assess homophily and differential homophily.

RESULTS: The mean age of patients included in the analysis was 51 years; 53% were female; 57% took oxycodone, 34% took hydrocodone and 9% took buprenorphine/naloxone. On average, 53% of patients received less than 50 morphine milligram equivalents (MME) daily, and the mean (standard deviation [SD]) number of opioid prescriptions per patient was 14.4 (6.6). Sixty-four percent of patients had commercial insurance, 28% had Medicaid, 5% had Medicare, and almost 2.5% used cash payment only. All three centrality measures were in agreement on the identification of the most influential patient in the opioid prescription network but overall correlation between the measures was low. After controlling for the main effects in the ERGM model, homophily was associated with age group, method of payment, number and type of opioid prescription filled, mean daily MME, and number of providers seen.

CONSLUSIONS: Characteristics of patients in an opioid prescription network may influence which provider they choose and which patients they are connect to through provider sharing. Interventions targeted at influential patients in the network may have potential to influence social norms around the use and misuse of prescription opioids that may lead to reductions in prescription opioid-related overdose deaths.

ACKNOWLEDGMENTS

I would like to express my profound gratitude to several people whose assistance and encouragement made the completion of this master's degree possible. First and foremost, I would like to thank my co-Major Advisors, Drs. Natallia Katenka and Ashley Buchanan for their guidance, encouragements, and patience with me. You made me feel as if I was the only student you were mentoring and advising by paying close attention to details of the study concept, proposal, and analysis. Our weekly meetings provided an opportunity for you to know me as person, my weaknesses, my strengths, and above all, my determination to succeed. I learned a lot during those meetings from your insightful comments, and sometimes tough questions. You made the best effort to make this the best thesis I could hope for. Any remaining errors in the thesis are mine.

I thank Professor Liliana Gonzalez for encouraging me to add a Statistics degree to my studies. It was not an easy decision, but her support and that of the whole department is unforgettable. My mentor in the Pharmaceutical Science program, Professor Stephen Kogut, was kind enough to accept my request to pursue a Master of Science in Statistics. I have enjoyed every bit of the application of the knowledge acquired to my PhD program!

I consider myself lucky to have been taught by Drs. Gavino Puggioni and Steffen Ventz who are some of the greatest statisticians in the field. Special thanks to Dr. Prabhani Kuruppumullage Don for accepting to serve on my thesis committee. Your regular encouragement cannot be forgotten. I will be remised if I did not extend my

thanks to Beth Larimer and Lorraine Berube who kept the department running smoothly at all times.

Finally, I would also like to thank my family- my parents for providing the best foundation for my education, my siblings for their moral support, my kids (Cynthia, Kevin, and Eric) for enduring my partial absence from home during the past few years, and my wife Bridget for her selflessness and endless love. Edwin, accept my heartfelt gratitude for always being there for me. Last but not least, I thank Fr. Maurice Agbaw Ebai for his timely spiritual support and guidance.

This thesis is dedicated to GOD for HIS Amazing Grace!

PREFACE

This thesis is written in the manuscript format, and is comprised of a single manuscript, which applied network analysis to advance our understanding of an opioid prescription network in the state of Rhode Island. We propose approaches to identify and target influential patients for interventions to alter social norms around opioid misuse. Its focus is on the application of the methods and it is written for a non-statistical audience. A more statistical discussion of essential concepts is presented in the Technical Appendix.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

MANUSCRIPT

TITLE: An Evaluation of Provider-sharing Networks of Patients Using Prescription Opioids

Hilary Aroke, MD MPH[1,2], Natallia Katenka, PhD MS[2], Stephen Kogut, PhD MBA RPh[1], Ashley Buchanan, DrPH MS[1]

[1]Department of Pharmacy Practice, College of Pharmacy, University of Rhode Island, Kingston, Rhode Island, USA

[2]Department of Computer Science and Statistics, College of Arts & Sciences, University of Rhode Island, Kingston, Rhode Island, USA

Corresponding Author: Hilary Aroke, Department of Pharmacy Practice, College of Pharmacy, University of Rhode Island, 7 Greenhouse Road, Suite 265G, Kingston, RI 02881; e-mail: haroke@uri.edu

Funding: Unfunded.

Target Journal: JAMA Internal Medicine

Format: Long-form paper

Publication status: In preparation for publication

ABSTRACT

IMPORTANCE: Patients using prescription opioid are embedded in a network due to provider-sharing and living in the same community. As a result, they may exert influence on each other's treatment preferences and share attitudes towards prescription opioid use and misuse.

OBJECTIVE: To determine patient characteristics associated with the observed pattern of shared prescribers in a network and identify influential patients in the network.

DESIGN, SETTING, AND PARTICIPANTS: We conducted a cross-sectional network-based study using the Rhode Island (RI) Prescription Drug Monitoring Program (PDMP) data for the 2015 calendar year. All patients who filled at least one opioid prescription at a retail pharmacy were eligible. The analysis was limited to patients who were on a stable opioid regimen and used only one source of payment, and filled only one type of opioid medication (oxycodone, hydrocodone or buprenorphine/naloxone) from ≥ 3 prescribers, and visited ≥ 3 pharmacies during the year. To minimize the influence of less relevant network connections, we excluded institutional providers and providers who issued opioid prescriptions to ≤ 6 patients. We applied social network analysis (SNA) methods to a sample of 372 patients connected to each other through provider-sharing. We used the exponential random graph model (ERGM) assuming conditional dyadic independence to examine the relationship between patient attributes and the likelihood of forming network ties. Homophily was defined as the tendency of patients to associate with others who have similar characteristics. Three centrality measures (degree, closeness, and betweenness)

were used to identify patients with potential influence in the opioid prescription network.

MAIN OUTCMES AND MEASURES: We provide a visual and descriptive characterization of the network, used centrality measures to identify influential patients, and ERGM to assess homophily and differential homophily.

RESULTS: The mean age of patients included in the analysis was 51 years; 53% were female; 57% took oxycodone, 34% took hydrocodone and 9% took buprenorphine/naloxone. On average, 53% of patients received less than 50 morphine milligram equivalents (MME) daily, and the mean (standard deviation [SD]) number of opioid prescriptions per patient was 14.4 (6.6). Sixty-four percent of patients had commercial insurance, 28% had Medicaid, 5% had Medicare, and almost 2.5% used cash payment only. All three centrality measures were in agreement on the identification of the most influential patient in the opioid prescription network but overall correlation between the measures was low. After controlling for the main effects in the ERGM model, homophily was associated with age group, method of payment, number and type of opioid prescription filled, mean daily MME, and number of providers seen.

CONSLUSIONS: Characteristics of patients in an opioid prescription network may influence which provider they choose and which patients they are connect to through provider sharing. Interventions targeted at influential patients in the network may have potential to influence social norms around the use and misuse of prescription opioids that may lead to reductions in prescription opioid-related overdose deaths.

KEY POINTS

Questions: What patient characteristics explain the pattern of shared-provider connections among patients in an opioid prescription network and can we identify influential patients as potential targets for opioid misuse prevention interventions?

Findings: In this social network analysis of PDMP data, we found extensive homophily that was associated with age group, method of payment, number and type of opioid prescription filled, mean daily dose, and number of prescribers ordering opioid prescriptions. All three commonly used centrality measures identified the same individual as the most influential patient in the network.

Meaning: Some patients in an opioid prescription network occupy influential positions through a large number of shared providers or by virtue of their location on paths between other patients in the network. Patients with similar characteristics tend to share providers with each other. These findings suggest that interventions targeted at influential, well-connected patients in the network may alter social norms around prescription opioid use and misuse in a community.

INTRODUCTION

The United States is experiencing an unprecedented prescription opioid overdose crisis driven in part by few patients who possibly engage in doctor shopping which may be identified in this context as patients obtaining opioid prescriptions from multiple providers without the prescribers' knowledge of other opioid prescriptions.[1-3] Prescribers may be sought by patients using opioids because of their reputation around opioid prescribing patterns such as use of high daily dose, use of combination opioids, and frequent refills. Knowledge about individual prescriber clinical practices and preferences may be shared among patients during co-visitation or social encounters in the community. A recent study demonstrated that health care providers tended to share patients with providers who have similar patients in their practice.[4] This suggests that patients prescribed opioids in a single state could be conceptualized as a network of patients with connections through shared providers which we define in an opioid prescription network. We hypothesized that patients within an opioid prescription network may exert influence on each other's opioid prescription utilization, including opioid misuse as a result of living in the same community or sharing a common opioid prescriber in the network, thereby impacting their network member's opioid prescription utilization and social norms around opioid use and misuse.[5]

Limited data suggests that a few high-intensity prescribers play a central role in sustaining the prescription opioid epidemic.[6,7] The pattern of provider-sharing may help identify corresponding influential or central patients in a network, thereby providing a clearer picture of where doctor shopping for prescription opioids may be occurring. This understanding can inform the implementation of targeted interventions

designed to improve prescription opioid utilization, prevent misuse, and treat opioid use disorder among patients in a network. A network-based perspective has been used to study a wide range of relational processes involving the flow of information between network members connected to each other in a social network. This perspective provides a framework that can be used to understand the structure of a network and how it influences the behavior of individual members in the network.[8-10] Landon et al. recently used network-based methods to demonstrate that characteristics of patient-sharing networks and the position of providers in the network are associated with healthcare resource utilization and cost.[11] Another study used network analysis to show racial differences in referral patterns for total hip replacement between communities with low and high concentrations of back residents.[12] Similar studies have not been done using an opioid prescription network.

There is a dearth of knowledge about characteristics of patients possibly engaged in doctor shopping for opioid prescriptions and methods to identify prescription opioid doctor shopping behavior are limited. To the best of our knowledge, network analysis has not yet been used to study an opioid prescription network within any state. The purpose of this study was to explore and characterize a patient-based opioid prescription network using social network analysis (SNA) methods. Specifically, we described patterns of relationships between patients within an opioid network, identified patients who have an influential role in the network, and examined patient characteristics that may explain the observed pattern of provider-sharing relationships. We used the exponential random graph model (ERGM) assuming conditional dyadic independence to examine the influence of some

characteristics of individual patients in the network on their likelihood to form network connections through provider sharing.

BACKGROUND

Over the past three decades opioid prescribing has increased tremendously in the United States, with a corresponding rise in opioid misuse and opioid overdose-related deaths.[13,14] An important feature of this opioid epidemic is the association between increasing rates of opioid prescribing and opioid-related morbidity and mortality.[15-17] Among people who died of opioid overdoses, up to 66% used prescription opioid analgesics originally prescribed for someone else; with doctor shopping being an important means for acquiring these prescription opioids for misuse.[18-24] In one study designed to determine the prevalence of doctor shopping for different controlled substances, prescription opioid medications (12.8%) were the most frequently involved, followed by benzodiazepines (2.4%), and stimulants (1.4%).[25] A cross-sectional study of French patients on buprenorphine maintenance treatment identified opioid misuse as a significant problem.[3] Similar findings were reported in a review of buprenorphine misuse.[26] Doctor shopping for prescription opioid medications often precedes fatal overdose, and accounts for about 40% of opioid-related overdoses, and up to 30% of deaths.[18,27-30] According to data from the Rhode Island Department of Health, overdose deaths increased by more than 90% between 2011 and 2016. There were 426 overdose deaths, of which 32% were related to fentanyl products, about 70% were males, and 25% were in their 50s.[31]

In response to the epidemic of opioid misuse, state-run Prescription Drug Monitoring Programs (PDMPs) were implemented in 49 states. These electronic

databases collect information on controlled substance prescriptions including opioid prescriptions, regardless of the sources of payment. These databases have been used by prescribers and others to examine filling patterns consistent with prescription opioid shopping behavior and potential misuse.[32-34] The number of providers involved in the care of a patient is often considered to be one of the strongest predictors of potential opioid misuse because until recently providers often did not have ready access to complete and accurate medication history at time of opioid prescribing.[35-41] The use of the number of prescribers to fill controlled substances, referred to as multiple prescriber episodes (MPE), has frequently been used as an indicator of doctor shopping in PDMP databases. The absence of a universally accepted definition for a threshold has led to wide variations in national estimates of doctor shopping, and associated difficulties in making comparisons across different settings to evaluate the effectiveness of interventions to prevent or reduce prescription opioid doctor shopping behavior.[18,32,34,42] We sought to evaluate the influence of MPE, number of opioid prescriptions, number of pharmacies used, method of payment, age group, and gender on the likelihood to form network connections through provider sharing.

Standard statistical approaches often assume independence of patients and/or providers and ignore contextual relationships between providers and patients, and among patients due geographic proximity, social influence, and local medical practice norms; thereby, limiting our ability to evaluate prescription opioid doctor shopping behavior. The goal of this study was to incorporate relational information using SNA. These findings will better inform future intervention policies designed to improve

social norms around prescription opioid use and prevent potential misuse among patients within a community of patients using prescription opioids.

METHODS

Data source: We conducted a cross-sectional network-based study using the Rhode Island (RI) PDMP data for the 2015 calendar year when the opioid crisis was a major statewide concern to patients, prescribers and public health regulators. The 2015 data contains records of schedule II to IV controlled substances dispensed by all retail pharmacies in the state. It includes de-identified unique patient, prescriber, and dispensing pharmacy information, and a limited number of variables such as age (in years) and sex of the patient, National Drug Code (NDC), product name, strength, formulation, and therapeutic class code of the drug plus number of days' supply, metric quantity dispensed, method of payment, and the date each prescription was filled. Daily morphine milligrams equivalents (MME) were estimated using standard conversion factors published by the Center for Disease Control and Prevention.

Cohort selection: A total of 2,058,816 controlled substance prescriptions were dispensed in RI in 2015 including opioid analgesics, benzodiazepines, psychostimulants, skeletal muscle relaxants, and sleep aids.[43] Unique NDC codes were used to identify 809,195 schedules II-IV opioid prescriptions filled at retail pharmacies by 222,513 patients (Figure 1). To minimize the impact of less clinically relevant network connections, we excluded institutional providers and prescribers who issued opioid prescriptions to ≤ 6 patients during the study year. An institutional provider was defined as any prescriber who had more than 2,400 patients on opioid prescriptions medications per year attributed to their Drug Enforcement

Administration (DEA) number. Institutional DEA numbers are used by medical students and residents to prescribe controlled substances under the supervision of a licensed attending physician. Patients were required to have used only one type of opioid medication and one source of payment method during the study period in order to capture patients who were more likely on a stable pain management regimen and to facilitate meaningful interpretation of the impact of these patient attributes. We limited our analysis to commonly used and misused types of opioid medications (i.e., oxycodone and hydrocodone), and buprenorphine/naloxone that is used to treat opioid use disorders. We hypothesized that patients on buprenorphine/naloxone for opioid use disorders would connect to one another more often than expected by chance. In addition, we restricted our analysis to patients who saw ≥ 3 prescribers for the same opioid prescription and filled their opioid prescriptions at ≥ 3 pharmacies within one year in order to capture patients with meaningful involvement in the patient-based opioid prescription network. Multiple visits allow the network to capture relationships between patients using opioid prescriptions. Several studies have used higher thresholds of 4 or 5 to flag doctor/pharmacy shopping behavior when applied to multiple types of opioid prescription per patient.[42,44] We evaluated the influence of age or age category, sex, source of payment, and type of opioid medication, number of opioid prescriptions, average daily MME, number of providers, and number of pharmacies on the likelihood of having a network connection defined by having one or more shared providers in an opioid prescription network.

Network-based framework

A network may be defined as a collection of points (i.e., vertices, nodes) and lines (i.e., edges, ties, links, connections) joining them. In a social network, these vertices represent people or groups of people and edges represent a kind of interaction between them. PDMP data links each patient who received at least one opioid prescription to one or more providers who ordered the opioid prescription(s). The receipt of one or more opioid prescriptions from a prescriber was used as a proxy for a relationship or interaction between a patient and a provider because state regulation requires a physician visit for a written opioid prescription. These prescription records were used to create an *edge list*, a two-column table, mapping patients to providers, each row representing an individual opioid prescription. The edge list was used to create a *bipartite* (two-mode) network where all pairs of patients and providers are joined by an edge. The bipartite network was represented as a provider-by-patient incidence matrix with cell entries indicating whether a provider wrote an opioid prescription to a particular patient (Figure 2). The rows of the matrix consisted of individual patients and columns identified providers. Pre-multiplying the bipartite incidence matrix by its transpose gave a symmetric *unipartite* (one-mode) adjacency matrix with either providers or patients only. The diagonal elements of the unipartite square matrix corresponded to the number of providers who wrote at least one opioid prescription to a given patient while the off-diagonal elements indicate the number of providers any two patients had in common. This analysis focused on the patient-based network where all nodes are represented by patients and connections (or edges) correspond to shared-provider relationships. To avoid the creation of loops and multi-edges,

diagonal elements were set to zero and off-diagonal elements with values greater than one were set to 1, respectively. The construction of a simple patient-based network graph is illustrated in Figure 3.

Network visualization: Network visualizations were selected to optimally place nodes in positions that visually convey important information in the network, such as the overall structure, location of influential patients in the network, and the presence of distinctive subgroups (or clusters within the network).[45,46] Some patient characteristics were incorporated into visualizations using different node colors. Graphical representation was used to examine degree, triangles, *dyad-wise shared partners* (DSP), and *edgewise shared partners* (ESP) distributions to explore the network in order to understand its structures. A DSP is a linked or unlinked dyad (i.e., patient pair) where both patients are linked to a third network member. ESP is a linked dyad in which both patients of the dyad are linked to a third network member. The distribution of ESP in a network was used to show how many dyads had one shared partner, two shared partners, and so on. Similarly, the distribution of DSP was used to show the number of dyads in the network with one shared partner, two shared partners, and so on. Node degree and triangles are defined below.

Network description: We evaluated the network with basic description of the network size, density, and number of components, diameter, clustering, centrality and modularity. The *network size* is defined as the number of nodes (i.e., patients) in the network and its *density* is defined as the proportion of observed connections in the network to the maximum number of possible connections in a randomly-generated network of the same size. A *path* is a series of steps required to go from patient A to

patient B. The shortest path is called the *geodesic* (distance) and the longest path is its *diameter*. A component is a subgroup of patients in the network such that there is a path connecting any two patients in the component directly or indirectly. A network is said to be *connected* if all pairs of nodes are connected directly or indirectly. When the largest connected component (LCC) is much bigger than the other components in a network it is called the *giant component*. The LCC was used to improve visualization, apply centrality measures, and develop ERGM models.

Global clustering measures the tendency of a network to form *closed triangles* (i.e., connections between three patients). A triangle closes when three patients share an opioid prescriber. *Transitivity* or clustering coefficient is the proportion of paths of length two that are closed. For each patient, it refers to the ratio of the total number of connections that exist among neighbors of the patient in the network to the total number of possible connections that could exist if they were completely connected.[47] This *local clustering coefficient* describes the extent to which network neighbors of a particular patient are directly connected to each other and may be interpreted as the probability that any two randomly selected neighbors of a particular patient in the network are connected to each other. Lower local clustering coefficients indicate fewer structural holes in the network and greater patient centrality.[47,48]

Network measures of *centrality* attempt to determine which patients are the most influential or central persons in a network.[47] The influence of an individual patient on others in a network through dissemination of information may influence social norms around opioid misuse and the sharing of opioid prescriptions in the network. In general, we expect patients with more connections to exert greater

influence on others in the network by sharing their attitudes towards opioid prescription use with a wider group of patients. We employed three commonly used centrality measures: degree, closeness, and betweenness centralities. Firstly, a patient's *degree centrality* is the number of other patients with direct connections to the patient or simply the number of shared providers. Secondly, *closeness centrality* measures how close a patient is to every other patient in the network, and reflects how fast information and influence of a particular patient can disseminate to other patients in the network. Formally, it is the inverse of the sum of all distances between patient i and all other patients in the network. Thirdly, *betweenness centrality* measures the extent to which a patient acts as a bridge between pairs of other patients in the network to facilitate the flow of information through the network. This implies that patients with larger betweenness centralities are more likely to have contacts with many other patients and may have greater influence regarding social norms around the use and potential misuse of prescription opioids in the network.[49,50] To estimate this measure, we used a commonly applied algorithm proposed by Freeman.[51] The most influential patient in the network was identified using each centrality measure and a subgraph corresponding to that individual and his or her immediate neighborhood was constructed.[52]

The LCC was used to calculate centrality measures that were standardized for comparison. Degree and closeness centralities were standardized dividing the estimate by their maximum possible values, $n-1$ and $1/(n-1)$, respectively, while the betweenness was normalized by dividing through the number of pairs of vertices not including the index vertex, $(n-1)(n-2)/2$. A chi-squared test was used to

compare the distribution of patient attributes across tertiles of the standardized centrality measures in the LCC. Assuming standardized centrality measures are independent and identically distributed across patients, three separate multivariable logistic regression models were fit to predict membership in the highest tertile of the standardized centrality measure. We used Pearson's correlation coefficient and 28 (10%) patients with the highest values for each standardized centrality measure to assess the level of agreement in identifying the most influential patients.

Modularity, a chance-corrected statistic, is defined as the proportion of connections that fall within observed groups based on patient characteristics minus the expected proportion if the connections were randomly distributed. The scores ranges from -5 to $+1$, and the closer the score value is to 1, the more the network exhibits clustering with respect to the grouping factor. Modularity scores were used to examine the influence of patient characteristics on patterns of connections in the opioid prescription network. Furthermore, because network visualization and modularity showed some evidence of clustering, mixing matrices and Pearson's correlation were examined. Mixing matrices were used to examine the number of connected dyads for each possible combination of levels of categorical patient characteristic. These exploratory analyses identified patient characteristics that were included in the ERGM.

Statistical network modeling: ERGMs were used to estimate the influence of covariates on the likelihood of ties in the opioid prescription network. This class of models formulates the probability of observing a set of network edges (and non-edges) as:

$$\mathrm{P}(\boldsymbol{Y} = y|\mathbf{X}) = \exp[\theta^T g(y, \mathbf{X})] / \mathbf{K}(\theta),$$

where, $\mathbf{Y}$ is the random set of relationships (edges and non-edges) in a network, y is the observed set of relationships, $\mathbf{X}$ is a matrix of attributes for the vertices in that network, $g(y, \mathbf{X})$ is a vector of the network statistics, θ is the vector of coefficients, and $\boldsymbol{K}(\theta)$ is a normalizing constant. Alternatively, the model states that the log odds for any given edge to exist conditional on the remaining network connections, and can be written as:

$$logit(Y_{ij} = 1) = \theta^{T} \delta[g(y, \boldsymbol{X})]_{ij},$$

where, Y_{ij} is an indicator for a connection between pairs of patients in **Y**, and $\delta[g(y, \boldsymbol{X})]_{ij}$ is the change in $g(y, \boldsymbol{X})$ value as y_{ij} is toggled from 0 to 1 (See Technical Appendix for details).

We first modeled a simple random graph (i.e., null model) which contained only an *edges* term to capture the network density.[53] A simulated network of the same size and density as the observed opioid prescription network was compared to the observed network in order to identify important differences between the two networks. The main effects and pairwise homophily interaction terms were added sequentially to the null model to represent attributes of patients in the network. *Homophily* was defined as the tendency for patients to connect with others like themselves. To examine the influence of node attributes on the likelihood of having a shared provider in the network, patient attributes were added to the model as main effects. We hypothesized that specific patient attributes, including number of opioid prescriptions, sex, age group, type of insurance coverage, type of opioid prescription, number of prescribers and pharmacies, explain the pattern of patient connections through provider sharing.

Homophily or assortative mixing is a tendency of patients to associate with similar patients, while disassortative mixing is the tendency to associate with dissimilar patients. Two types of dyadic interaction terms were added to the main effects model to assess assortative and disassortative mixing in the network leading to patterns of homophily or heterophily, respectively. First, we assessed the likelihood of provider sharing when both patients in a dyad had the same level of a categorical attribute. The number of opioid prescriptions was added as a continuous attribute. We hypothesized that two patients with a similar number of opioid prescriptions filled during the study year were more likely to form a network connection based on having a shared provider. Secondly, we assessed the likelihood of provider sharing when both patients in a dyad had different levels (i.e., dissimilar) of a categorical attribute such as type of opioid prescription (differential homophily).

We limited this analysis to ERGM models that assume dyadic independence of network connections.[54] This assumption specified that patients sharing a provider were dependent but independent if they had no provider in common. The null and main effects models with and without homophily and differential homophily terms were compared using Log L and related measures of deviance (-2LogL), the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).[55,56] All tests of statistical significance were two-sided and performed at the 0.05 significance level. Data manipulation was performed with SAS, version 9.4 (SAS Institute, Cary, NC) and network analysis was implemented with R statistical software, version 3.2.3 (R Core Team 2016). The study was approved by the Institutional Review Board at University of Rhode Island.

RESULTS

A total of 372 patients prescribed opioids by 746 providers during a one-year study period met the inclusion criteria for meaningful involvement in the opioid prescription network in RI. Table 1 presents a summary of the characteristics of all the patients compared to those in the LCC. The mean (SD) age of all patients in the sample was 51 (14) years with 50% aged 45-64 years, and 53% were female. More than 5,000 opioid prescriptions were filled, of which 57% were prescriptions for oxycodone, 34% were for hydrocodone and 9% were for buprenorphine/naloxone. The mean number of opioid prescriptions filled per patient was 14.4 (SD=6.6) with 53% receiving on average less than 50 MME daily. However, 25% of patients had on average more than 90 MME per day. Most patients paid for all their opioid prescriptions with commercial insurance and only 2.4% used cash payment exclusively. About half of the patients filled opioid prescriptions written from ≥4 prescribers while 31% filled their opioid prescriptions at ≥4 pharmacies during the one-year study period.

Network characteristics: The bipartite network had a total of 1,118 nodes (746 providers plus 372 patients) with 1,460 unique connections between them corresponding to unique patient-provider relationships resulting from one or more opioid prescriptions. There were more prescribers in the network than patients. The overall bipartite density was only 0.5%. The full opioid network contained 372 patients with 1,980 connections among them; 32 (8.6%) had no shared providers (isolates) [Figure 4]. The full network had 55 connected components including 32 isolates, and the LCC contained 74% of all the patients in the network including 95.66% of all connections; the second largest connected component consisted of only

2.4% (Figure 5). The full network and its LCC demonstrated apparent clustering with at least 4-5 large clusters. There was some evidence of clustering by age group, opioid type, average daily dose, and method of payment; however, no pattern by gender was apparent. The most obvious clustering is seen among patients who were on buprenorphine/naloxone. One possible explanation for this clustering is that under the Drug Addiction Treatment Act (DATA) of 2000, only certain qualified providers are authorized to prescribe buprenorphine/naloxone as medication-assisted treatment for people diagnosed with an opioid use disorder.

Characteristics of the full opioid network and its LCC are presented in Table 2. The LCC had an overall density of 5% compared with 3% for the whole network. The average number of shared providers was higher among those in the LCC (13.8) compared to the whole network (10.6). The average path length and longest path were the same (≈ 10) suggesting that the rate of flow of information diffusion in the LCC would be similar to that of the full network. However, the density around the most central patient was 65%, average number of shared providers was 25 (SD=9) and the average path length only 1.4 (Figure 6). Assortative mixing and the fraction of transitive triples (transitivity) were higher for the whole cohort. About 85% of patients who shared a provider were connected to other patients who also shared a provider with each of them. Seventy-five of patients were connected to one or more patients with at least one similar characteristic.

The number of shared providers was quite heterogeneous across patients (Figure 7). While there are many patients with few shared providers, there was a non-trivial number with many shared providers. In particular, there are 28 patients with 29

shared prescribers. This may correspond to providers in the same practice or on-call group. Given the nature of the decay in the degree distribution, a log-log scale was used to assess the results. The middle panel in Figure 7 shows a somewhat linear decay in the log-frequency as a function of the log-degree. A plot of the average neighbor degree versus vertex degree suggests that while there is a tendency for patients with many shared providers to connect to each other, those with fewer shared providers tend to connect with both patients having lower and higher number of shared providers (assortative degree network). This is illustrated by the high network density around the most influential patient (Figure 6).

Centrality measures: Overall, there was moderate correlation between degree and closeness ($r = 0.53$; $p < 0.001$), and between closeness and betweenness($r = 0.48$; $p < 0.001$) centralities. However, correlation between degree and betweenness centralities was low ($r = 0.19$; $p=0.002$). Among 56 patients with the highest standardized centrality values for any of the three measures, 14 (25%) were identified by degree and closeness centralities, 12 (21%) by closeness and betweenness centralities, 7 (12%) by degree and betweenness centralities, and only 5 (8.93%) by all three measures. However, all three measures identified the same patient as the most influential patient in the network; a 48-year-old female on Medicare taking Oxycodone who filled 19 prescriptions for an average daily MME > 90 and saw 5 different providers and visited 5 different pharmacies during one calendar year (data not shown).

Tertiles of standardized centrality measures estimated from the LCC are presented in Table 3. Age group, type of opioid used, average daily dose, and number

of opioid prescribers were associated with at least one standardized centrality measure while gender, method of payment, and number of pharmacies were not associated with any centrality measure. Based on multivariable logistic regression model, age group, type of opioid used, average daily dose, and number of opioid prescribers were associated with the highest tertile of at least one standardized centrality measure, after adjusting for other covariates in the model (Table 4). Patients aged 45-64 years were most likely to be classified as having the highest levels of standardized degree centrality tertile as compared to those ages 65 years and older. Furthermore, patients who took on average > 90 daily MME were 6.7 times more likely to have the highest standardized degree centrality tertile compared to those on < 50 MME per day. This suggests that patients on higher daily doses of opioids tend to have more shared providers. Based on standardized closeness and betweenness centralities, patients who had ≥ 4 providers were more likely to be classified in the highest tertile of their respective standardized centrality measures. As compared with patients on hydrocodone, patients on buprenorphine/naloxone were less likely to be in the highest standardized degree centrality tertile and more likely to be in the highest standardized betweenness centrality tertile. This suggests that patients with few connections may be crucial for the diffusion of information and prescription opioids in the network.

Network connectivity: Graphical examination of triangles, degree, DSP, and ESP were used to understand the network structure (Figure 8). The observed LCC of the network had many more completed triples than a randomly-generated network of the same size and density. Similarly, the LCC had many patients with few shared providers (low-degree nodes) and few patients with many shared providers (high-degree nodes)

compared with a random network of the same size and density. Edgewise and dyad-wise shared partner distributions also differed in the observed LCC and random networks with the observed LCC having more patients with multiple ESP and DSP compared with the random network, which indicated a large number of patients with one or two shared partners, and hardly any higher level multiples.

Mixing matrices is presented in Table 5. Provider sharing tended to be between two patients who are both female, one younger and the other middle-aged, both on commercial insurance, both on oxycodone, one on low-dose and the other on intermediate daily dose, or both with 4 or more opioid prescribers. For example, of 825 connected pairs of patients who took hydrocodone in opioid type mixing matrix, 659 (80%) are connected to a patient who took only hydrocodone. This suggested a higher likelihood of patients who took hydrocodone to share providers with other patients who also took hydrocodone (i.e., homophily of opioid prescribing). From Table 6 opioid type, number of providers, average daily MME and age group had the highest modularity score and may explain some of the clustering observed in the network. From the perspective of the network connections and patient attributes, the GC was a reasonable representation of the full opioid prescription network.

Results from ERGMs: Based on a null model with only the *edge* term to account for the number of connections in the network, the probability of a connection between any randomly selected two patients was 0.06 (i.e., density of the network). This baseline model was compared to models with more constraints. Although the null model provided a reasonable representation of the observed network density, it failed to represent other features of the network such as transitivity. A plot of 1,000 simulated

networks of the same size and density as the null model was used to assess how well the null model captured transitivity. The point on the x-axis in Figure 9 corresponds to the location of 12,514 triangles in the observed patient network, which was much higher than the number of triangles in any of the 1,000 simulated networks. This suggested that a more complex model with constraints on the number of triangle was needed to capture transitivity and other network characteristics in the observed opioid prescription network.

Model with main effects of patient attributes: Based the modularity score and clinical importance, we selected gender, age category, payment method, opioid type, number of prescriptions, average daily dose in MME, number of provider who wrote opioid prescriptions to the patient in one calendar year (categorical) and number of pharmacies visited during the year to fill opioid prescriptions (categorical) to include in the main effects model. The null hypothesis was that there was no association between each patient attribute and the likelihood of a patient having a connection through provider sharing, after controlling for all other attributes in the model. The results of the main effects model are summarized in Table 7. Positive coefficients (i.e., log odds of a connection) indicate a higher likelihood of sharing a provider with another patient in the network (compared to the reference level for categorical attributes) and negative coefficients indicate lower likelihood. The total number of opioid prescription filled was positively associated with an increased likelihood of having a shared provider. Male patients were less likely to have a shared provider in the network than female patients, although this difference was not statistically significant. Unlike patients on Medicare, patients who used Medicaid or cash

payments exclusively were less likely to have a shared provider with other patients in the opioid prescription network than patients who used commercial insurance but the difference was not statistically significant for the use of cash term. With only 9 patients who used cash exclusively, there may not have been enough power to detect any difference. Patients who took either hydrocodone or oxycodone were more likely to have a shared provider in the network. A higher average daily dose of opioids was associated with a greater the chance of having a shared provider in the network. Furthermore, patients who had ≥ 4 opioid prescribers in one year were more likely to have at least one shared provider in the opioid prescription network than those with fewer providers. However, patients who filled their opioid prescriptions at ≥ 4 pharmacies were less likely to have a shared provider in the network than those who used fewer pharmacies. These results are consistent with our results of network visualizations, mixing matrices, Pearson's correlation coefficients and modularity scores.

Model with main effects and homophily terms: We hypothesized that two patients with the same level of a categorical attribute, or similar number of opioid prescriptions filled during the study year, were equally likely to form a network connection based on having a shared provider. The results of the model with main effects and homophily terms are summarized in Table 8. Positive and significant parameter estimates for gender, age category, opioid type, average daily MME, and number of providers all indicated the presence of homophily effects for these patient attributes, after controlling for their main effects in the model. All homophily interaction terms were statistically significant at 0.05 level except for method of payment that was only

significant at 0.10 significance level. Use of ≥ 4 pharmacies had no homophilic effects. For the number of opioid prescriptions filled during the year, a negative coefficient indicated homophily because the absolute difference in size decreases as the sizes of both values in a dyad becomes more identical. The addition of homophilic terms did not alter the qualitative associations of the main attributes.

Model with main effects and differential homophily terms: The results of the model with differential homophily terms are presented in Table 9. Overall homophilic effects of gender were seen mainly among females; for age mainly among patients 65 years and older; for opioid type among all levels especially those on buprenorphine/naloxone subgroup; and for number of providers mainly among those who saw ≤ 4 providers in a year, after controlling for other variables in the model, respectively. Overall homophily effects of payment type were significant at the 0.10 level only among those on commercial insurance and there were no homophilic effects within other subgroups of payment method. Adjusted homophily effects of opioid dose and number of pharmacies did not achieve statistical significance. The addition of homophily terms did not alter the qualitative association of the main attributes, except for the subgroup of patients aged 65 and older. The model with main effects and differential homophily terms had a lower AIC than the baseline model and the model with main effects and homophilic terms.

DISCUSSION

This study suggests that patients in an opioid prescription network were highly connected. Our sample of 372 patients had 1,980 shared-provider connections and almost 75% of patients were connected to each other either directly or indirectly in

one giant component that consisted of 96% of all connections in the full network. The intensity of prescription opioid use and possible misuse is reflected in the number of prescribers in the network. There were twice as many providers as patients using prescription opioids. More than half the patients saw at least four providers although they appeared to be on a stable opioid regimen. A majority of patients were female and aged 45-64 years old. The underlying structure of the network was significantly different from that of a randomly-generated network of the same size and density. The random network was never designed as a model for observed networks because it ignores node attributes that may explain observed clustering. The distribution of the number of shared providers was bimodal with a second peak corresponding to 29 shared prescribers. This may represent a group of patients who belong to a large medical group with many providers who cater for all the patients, including pain management.

Our opioid prescription network demonstrated homophily by opioid type, opioid dose, age group, sources of payment and number of providers. The most obvious clustering on visualization was seen among patients on buprenorphine/naloxone. One possible explanation for this clustering is that under the Drug Addiction Treatment Act (DATA) of 2000, only certain qualified providers are authorized to prescribe buprenorphine/naloxone as medication-assisted treatment for people diagnosed with an opioid use disorder. Such patients are more likely to seek certified providers in the network. Overall, 75% of patients were connected to one or more patients with at least one similar characteristic. Furthermore, patients with many shared providers were more likely to connect to each other and such patients may have

patterns of use consistent with potential abuse. Although older patients with more medical conditions tend to be on opioid therapy, they had fewer shared providers compared to younger patients. This may be due older patients having a stable health insurance and an established primary care provider who meets their medical needs. Centrality measures suggested that relatively few patients were at the center of the opioid prescription network. Similar conclusions have been drawn about providers using standard statistical methods.[6,7] The level of connectedness was captured by degree and closeness centrality measures, which identified higher daily opioid dose to be associated with the number of shared providers (i.e., degree centrality) and closeness centrality, which is a measure of how quickly information emanating from one patient in the network could spread to other patients assuming each shared provider relationship offers ample opportunities to disseminate information and training on opioid misuse prevention.

If network connections represent the flow of information and influence, then a measure of how often a patient in the network acts as a bridge between other patients may provide a more useful measure. Betweenness centrality assumes there is flow of information in a network and attempts to capture the influence of each member over the spread of that information. However, in calculating this measure it is assumed that all patients in the network have the same probability of sharing information received and that the information spreads around via the shortest paths. This suggests that more information would pass through patients with larger betweenness centralities whose removal from the network could disrupt the network cohesion. In practice, patients may not spread information at the same rate and information may not spread through

the shortest paths.[47] Despite these limitations, betweenness centrality remains a very useful guide to the potential influence a network member over the flow of information and may serve as a useful way to identify patients for prevention and treatment interventions. It also has a wide range of values making it easier to distinguish between central and less central patient targets. Our analysis suggested that such intervention would seek patients with MPE or those on opioid use disorder treatment first. These findings support the use of multiple provider episodes as an indicator of potential opioid misuse.[44,57,58] Additional studies are needed to evaluate the practical advantages of using betweenness centrality with or without MPE. This can be easily implemented through sequential analysis of PDMP data which were instituted or strengthened primarily in response to the prescription opioid epidemic.

Our results suggest that patients on prescription opioid medications are not isolated from their social environment, but rather are connected to each other via provider sharing. Sharing an opioid prescriber may increase the probability of establishing a personal relationship with another patient on opioid therapy through a chance encounter in a physician office or the community because, unlike some other controlled substances, oxycodone, hydrocodone and buprenorphine/naloxone require an office visit for a written prescription. These relationships form a basis for constructing an opioid prescription network using comprehensive and reliable prescription information captured by the PDMP at the state-level irrespective of the payer. New state regulations and laws mandates quantity limits on opioid prescriptions. For patients requiring chronic opioid therapy, this may lead to biweekly or monthly office visits for opioid prescription renewals. Refills may be provided by a

partner in the same practice or a provider on-call. After office hours, the patient may visit the emergency room or a walk-in clinic for a short supply. Broadly defined, doctor shopping involves visiting many providers during an episode of illness, or to acquire controlled substances illicitly. In the context of acquiring prescription opioid medications for potential misuse, patients may engage in doctor shopping because of long waiting times for an appointment, inconvenient office hours, persistence of painful condition, provider attitude, or absence because of vacation.[59-61] Hence, doctor shopping behavior may be a reflection of fragmented care in a patient with persistent pain. We had no clinical diagnosis information, place of service, geographic location, or provider specialty to impugn any diagnosis or clinical condition for opioid prescriptions.

A recent social network analysis of physicians in the United States showed that providers tend to be connected to other providers with similar patients in a patient-sharing network.[5] Our study shows substantial clustering and assortative mixing driven by patient characteristics. The extent of homophily is similar to what has been observed in other social settings.[62,63] The cross-sectional nature of our study does not allow us infer the reasons for preferential connections of patients using prescription opioid to other patients with similar characteristics. One implication of homophily is that it could facilitate the spread of valuable health information through targeted training of influential patients in the network and reduce the diffusion of risky behavior around prescription opioid use and misuse.

The successful application of network science to a systematic problem such as the opioid crisis requires careful consideration of the choice of nodes and the connections between them to ensure clinical significance. This study demonstrates several advantages of using PDMP to identify connections within a network of patients on prescription opioid therapy. The sample for this analysis was derived from a statewide PDMP data which provided a unique opportunity to identify virtually all provider-sharing relationships among patients on prescription opioids regardless of the sources of payment. Because the data is collected for use in clinical decision-making at the point of patient care, it is assumed to be accurate and comprehensive. It provides data that can be used to identify patterns of opioid prescription filling that raise concerns for potential opioid misuse.

LIMITATIONS

These analyses are subject to a few limitations. First, the sharing of an opioid prescriber was used to infer an information sharing relationship between patients on prescription opioids. Patterns of provider sharing may be a reflection of fragmented care rather than doctor shopping of prescription opioids for illicit use. We cannot identify doctor shopping from prescription data because in this context doctor shopping often involves an illegal and covert activity. PDMP data do not reveal the reasons why patients obtained multiple prescriptions from so many different prescribers. Second, we excluded providers with very few patients or potential institutional providers. We have no evidence that all prescribers were licensed in the state for the whole calendar year and such providers may have fewer patients. Their removal may cause network fragmentation, fundamentally altering some of the

properties of the network. Third, we used a threshold of 2,400 for so-called institutional providers but there could have been other institutional prescribers with fewer patients similar to non-institutional providers. Furthermore, providers in single group practice are more likely to see the same patients creating a false pattern of doctor shopping behavior. Because our analysis did not account for the grouping of physician practices, there is potential for misclassification of the number of providers seen. Moreover, our analysis failed to capture evidence of geographic proximity, co-visitation to providers, or community social interactions. Fourth, our data was limited to a single calendar year for one state with very few variables making it difficult to characterize the nature of the relationships. We did not account for prescription opioid fills across the state-lines. Indeed, PDMP data may underestimate the prevalence of prescription opioid misuse because it focuses on prescribers and patients ignoring other sources of illicit prescription opioid use, including theft, illicit drug use, and unlicensed internet pharmacies. Fifth, the enforcement of regulations on electronic filing of controlled substance fillings, merging of prescriber practices, entry and exit of providers and pharmacies from the market may alter the patterns of relationships substantially. Sixth, the usefulness of closeness centrality in observed networks is often limited by the existence of several components because distance between two patients in different components is considered infinite. In addition, the range of closeness centrality values is relatively narrow, making it difficult to distinguish between an influential and a less influential patient in the network for the purpose of designing interventions. Seventh, the ERGM model of characteristics associated with network connections, homophily and differential homophily assumes conditional

independence of dyads, which is largely unrealistic in practice.[64] ERGM models of large networks perform poorly due time-consuming algorithms and non-converging Markov chains, and associated model degeneracy. Furthermore, nodes are assumed fixed and homogenous except for differences captured in available nodal attributes. The use of a subnetwork in which these assumptions hold limits generalizability of results because every network is unique due to inconsistency under sampling.

CONCLUSIONS

Patterns of provider sharing in a patient-based prescription opioid network, suggest that patients prescribed opioids may have extensive connections with other that could be leveraged to improve dissemination of health promotion or disruption of negative behaviors with the use and misuse of opioid prescriptions. The characteristics of these patients and the structure of the network uniting them influence their health choices. The analysis suggests that opioid prescribers could easily be sought by patients who doctor shop for prescription opioids. Interventions targeted at influential patients in the network may have potential to influence social norm around the use and misuse of prescription opioids eventually leading to reductions in opioid overdoses and opioid use disorders.

REFERENCES

1. Martyres RF, Clode D, Burns JM. Seeking drugs or seeking help? Escalating "doctor shopping" by young heroin users before fatal overdose. *Med J Aust.* 2004;180(5):211-214.

2. Klienschmidt R, Price J, Caught K. Doctor shopping. *Aust Fam Physician.* 1995;24(6):1037-1041.

3. Pradel V, Thirion X, Ronfle E, Masut A, Micallef J, Begaud B. Assessment of doctor-shopping for high dosage buprenorphine maintenance treatment in a French region: development of a new method for prescription database. *Pharmacoepidemiol Drug Saf.* 2004;13(7):473-481.

4. Landon BE, Onnela JP, Keating NL, et al. Using administrative data to identify naturally occurring networks of physicians. *Med Care.* 2013;51(8):715-721.

5. Landon BE, Keating NL, Barnett ML, et al. Variation in patient-sharing networks of physicians across the United States. *JAMA.* 2012;308(3):265-273.

6. McDonald DC, Carlson KE. Estimating the prevalence of opioid diversion by "doctor shoppers" in the United States. *PLoS One.* 2013;8(7):e69241.

7. Paulozzi LJ, Strickler GK, Kreiner PW, Koris CM, Centers for Disease C, Prevention. Controlled Substance Prescribing Patterns--Prescription Behavior Surveillance System, Eight States, 2013. *MMWR Surveill Summ.* 2015;64(9):1-14.

8. Ennett ST, Bauman KE. Peer group structure and adolescent cigarette smoking: a social network analysis. *J Health Soc Behav.* 1993;34(3):226-236.

9. Hall JA, Valente TW. Adolescent smoking networks: the effects of influence and selection on future smoking. *Addict Behav.* 2007;32(12):3054-3059.

10. Rothenberg RB, Sterk C, Toomey KE, et al. Using social network and ethnographic tools to evaluate syphilis transmission. *Sex Transm Dis.* 1998;25(3):154-160.

11. Landon BE, Keating NL, Onnela J, Zaslavsky AM, Christakis NA, O'Malley A. Patient-sharing networks of physicians and health care utilization and spending among medicare beneficiaries. *JAMA Intern Med.* 2018;178(1):66-73.

12. Ghomrawi HMK, Funk RJ, Parks ML, Owen-Smith J, Hollingsworth JM. Physician referral patterns and racial disparities in total hip replacement: A network analysis approach. *PLoS One.* 2018;13(2):e0193014.

13. Okie S. A flood of opioids, a rising tide of deaths. *N Engl J Med.* 2010;363(21):1981-1985.

14. Paulozzi LJ, Mack KA, Hockenberry JM. Vital signs: variation among States in prescribing of opioid pain relievers and benzodiazepines - United States, 2012. *MMWR Morb Mortal Wkly Rep.* 2014;63(26):563-568.

15. Jones CM, Mack KA, Paulozzi LJ. Pharmaceutical overdose deaths, United States, 2010. *JAMA.* 2013;309(7):657-659.

16. Paulozzi LJ. Prescription drug overdoses: a review. *J Safety Res.* 2012;43(4):283-289.

17. Manchikanti L, Helm S, 2nd, Fellows B, et al. Opioid epidemic in the United States. *Pain Physician.* 2012;15(3 Suppl):ES9-38.

18. Hall AJ, Logan JE, Toblin RL, et al. Patterns of abuse among unintentional pharmaceutical overdose fatalities. *JAMA.* 2008;300(22):2613-2620.

19. Kuehn BM. Opioid prescriptions soar: increase in legitimate use as well as abuse. *JAMA.* 2007;297(3):249-251.

20. Barrett K, Watson A. Physician perspectives on a pilot prescription monitoring program. *J Pain Palliat Care Pharmacother.* 2005;19(3):5-13.

21. Inciardi JA, Surratt HL, Kurtz SP, Cicero TJ. Mechanisms of prescription drug diversion among drug-involved club- and street-based populations. *Pain Med.* 2007;8(2):171-183.

22. Joranson DE, Gilson AM. Drug crime is a source of abused pain medications in the United States. *J Pain Symptom Manage.* 2005;30(4):299-301.

23. Hurwitz W. The challenge of prescription drug misuse: a review and commentary. *Pain Med.* 2005;6(2):152-161.

24. White AG, Birnbaum HG, Schiller M, Tang J, Katz NP. Analytic models to identify patients at risk for prescription opioid abuse. *Am J Manag Care.* 2009;15(12):897-906.

25. Wilsey BL, Fishman SM, Gilson AM, et al. Profiling multiple provider prescribing of opioids, benzodiazepines, stimulants, and anorectics. *Drug Alcohol Depend.* 2010;112(1-2):99-106.

26. Lofwall MR, Walsh SL. A Review of Buprenorphine Diversion and Misuse: The Current Evidence Base and Experiences from Around the World. *Journal of addiction medicine.* 2014;8(5):315-326.

27. Lanier WA, Johnson EM, Rolfs RT, Friedrichs MD, Grey TC. Risk factors for prescription opioid-related death, Utah, 2008-2009. *Pain Med.* 2012;13(12):1580-1589.

28. CDC grand rounds: prescription drug overdoses - a U.S. epidemic. *MMWR Morb Mortal Wkly Rep.* 2012;61(1):10-13.

29. Gwira Baumblatt JA, Wiedeman C, Dunn JR, Schaffner W, Paulozzi LJ, Jones TF. High-risk use by patients prescribed opioids for pain and its role in overdose deaths. *JAMA Intern Med.* 2014;174(5):796-801.

30. Peirce GL, Smith MJ, Abate MA, Halverson J. Doctor and pharmacy shopping for controlled substances. *Med Care.* 2012;50(6):494-500.

31. Overdose Death Data. Overdose Prevention and Intervention Task Force. 2017; http://preventoverdoseri.org/overdose-deaths/. Accessed April 10, 2018.

32. Wilsey BL, Fishman SM, Gilson AM, et al. An analysis of the number of multiple prescribers for opioids utilizing data from the California Prescription Monitoring Program. *Pharmacoepidemiol Drug Saf.* 2011;20(12):1262-1268.

33. Deyo RA, Irvine JM, Millet LM, et al. Measures such as interstate cooperation would improve the efficacy of programs to track controlled drug prescriptions. *Health Aff (Millwood).* 2013;32(3):603-613.

34. Katz N, Panas L, Kim M, et al. Usefulness of prescription monitoring programs for surveillance--analysis of Schedule II opioid prescription data in Massachusetts, 1996-2006. *Pharmacoepidemiol Drug Saf.* 2010;19(2):115-123.

35. Tamblyn RM, McLeod PJ, Abrahamowicz M, Laprise R. Do too many cooks spoil the broth? Multiple physician involvement in medical management of elderly patients and potentially inappropriate drug combinations. *CMAJ.* 1996;154(8):1177-1184.

36. Price D, Cooke J, Singleton S, Feely M. Doctors' unawareness of the drugs their patients are taking: a major cause of overprescribing? *Br Med J (Clin Res Ed).* 1986;292(6513):99-100.

37. Frank C, Godwin M, Verma S, et al. What drugs are our frail elderly patients taking? Do drugs they take or fail to take put them at increased risk of interactions and inappropriate medication use? *Can Fam Physician.* 2001;47:1198-1204.

38. Claoue C, Elkington AR. Informing the hospital of patients' drug regimens. *Br Med J (Clin Res Ed).* 1986;292(6513):101.

39. Bedell SE, Jabbour S, Goldberg R, et al. Discrepancies in the use of medications: their extent and predictors in an outpatient practice. *Arch Intern Med.* 2000;160(14):2129-2134.

40. Fisher ES, Staiger DO, Bynum JP, Gottlieb DJ. Creating accountable care organizations: the extended hospital medical staff. *Health Aff (Millwood).* 2007;26(1):w44-57.

41. Boockvar K, Fishman E, Kyriacou CK, Monias A, Gavi S, Cortes T. Adverse events due to discontinuations in drug use and dose changes in patients transferred between acute and long-term care facilities. *Arch Intern Med.* 2004;164(5):545-550.

42. Cepeda MS, Fife D, Chow W, Mastrogiovanni G, Henderson SC. Opioid shopping behavior: how often, how soon, which drugs, and what payment method. *J Clin Pharmacol.* 2013;53(1):112-117.

43. Aroke H, Buchanan A, Wen X, Ragosta P, Koziol J, Kogut S. Estimating the Direct Costs of Outpatient Opioid Prescriptions: A Retrospective Analysis of Data

from the Rhode Island Prescription Drug Monitoring Program. *J Manag Care Spec Pharm.* 2018;24(3):214-224.

44. Barnett ML, Olenski AR, Jena AB. Opioid-Prescribing Patterns of Emergency Physicians and Risk of Long-Term Use. *N Engl J Med.* 2017;376(7):663-673.

45. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Software: Practice and Experience.* 1991;21(11):1129-1164.

46. Kamada T, Kawai S. An algorithm for drawing general undirected graphs. *Information processing letters.* 1989;31(1):7-15.

47. MEJ N. *Networks: an introduction.* Oxford University Press, Oxford; 2010.

48. Burt RS. *Structural holes: The social structure of competition.* Harvard university press; 2009.

49. Keating NL, Ayanian JZ, Cleary PD, Marsden PV. Factors affecting influential discussions among physicians: a social network analysis of a primary care practice. *J Gen Intern Med.* 2007;22(6):794-798.

50. Wasserman S, Faust K. *Social network analysis: Methods and applications.* Vol 8: Cambridge university press; 1994.

51. Freeman LC. A Set of Measures of Centrality Based on Betweenness. *Sociometry.* 1977;40(1):35-41.

52. Kolaczyk ED, Csárdi G. *Statistical analysis of network data with R.* Vol 65: Springer; 2014.

53. Erdös P, Rényi A. On random graphs, I. *Publicationes Mathematicae (Debrecen).* 1959;6:290-297.

54. Holland PW, Leinhardt S. An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association.* 1981;76(373):33-50.

55. AKAIKE H. Information theory and an extension of the maximum likelihood principle. Paper presented at: Second International Symposium on Information Theory; 1973.

56. Schwarz G. Estimating the dimension of a model. *The annals of statistics.* 1978;6(2):461-464.

57. Cepeda MS, Fife D, Berwaerts J, Friedman A, Yuan Y, Mastrogiovanni G. Doctor shopping for medications used in the treatment of attention deficit hyperactivity disorder: shoppers often pay in cash and cross state lines. *Am J Drug Alcohol Abuse.* 2015;41(3):226-229.

58. Cochran G, Woo B, Lo-Ciganic WH, Gordon AJ, Donohue JM, Gellad WF. Defining Nonmedical Use of Prescription Opioids Within Health Care Claims: A Systematic Review. *Subst Abus.* 2015;36(2):192-202.

59. Yeung RY, Leung GM, McGhee SM, Johnston JM. Waiting time and doctor shopping in a mixed medical economy. *Health Econ.* 2004;13(11):1137-1144.

60. Feroni I, Peretti-Watel P, Paraponaris A, et al. French general practitioners' attitudes and prescription patterns toward buprenorphine maintenance treatment: does doctor shopping reflect buprenorphine misuse? *J Addict Dis.* 2005;24(3):7-22.

61. Kasteler J, Kane RL, Olsen DM, Thetford C. Issues underlying prevalence of "doctor-shopping" behavior. *J Health Soc Behav.* 1976;17(4):329-339.

62. McPherson M, Smith-Lovin L, Cook JM. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology.* 2001;27(1):415-444.

63. Apicella CL, Marlowe FW, Fowler JH, Christakis NA. Social networks and cooperation in hunter-gatherers. *Nature.* 2012;481(7382):497-501.

64. Kolaczyk E. *Statistical Analysis of Network Data: Methods and Models (Springer Series in Statistics).* Springer; 2009.

TABLES AND FIGURES

Table 1. Characteristics of Patients in an Opioid Prescription Network of Patients in the RI PDMP in 2015 (n=372)

Characteristic	**All (N=372)**	**LCC (N=274)**
	Mean (SD)	
Number of providers	3.92 (1.27)	3.99 (1.34)
Number of pharmacies	3.47 (0.82)	3.49 (0.85)
Number of opioid Rx	14.36 (6.56)	14.01 (5.52)
Total daily MME, mg	1,238 (1,958)	1262 (1976)
Age (years)	50.59 (13.84)	49.34 (13.45)
	N (%)	
Age group (years)		
21-44	126 (33.87)	103 (37.59)
45-64	189 (50.81)	134 (48.91)
65+	57 (15.32)	37 (13.50)
Gender		
Female	197 (52.96)	139 (50.73)
Male	175 (47.04)	135 (49.27)
Method of payment for opioids		
Commercial	239 (64.25)	178 (64.96)
Medicare	104 (27.96)	5 (1.82)
Medicaid	20 (5.38)	15 (5.47)
Cash	9 (2.42)	76 (27.74)
Number of providers		
3	187 (50.27)	134 (48.91)
4+	185 (49.73)	140 (51.09)
Number of pharmacies		
3	257 (69.09)	187 (68.25)
4+	115 (30.91)	87 (31.75)
Type of opioid medication used		
Buprenorphine/naloxone	33 (8.87)	32 (11.68)
Hydrocodone	126 (33.87)	93 (33.94)
Oxycodone	213 (57.26)	149 (54.38)
MME category, mg		
< 50	197 (52.96)	136 (49.64)
50-90	81 (21.77)	67 (24.45)
>90	94 (25.27)	71 (25.91)

Abbreviations: LCC= largest connected component; Rx=prescription(s)
SD=standard deviation; MME=morphine milligram equivalent

Table 2. Summary of Characteristics of a Network of Patients and a Network of Providers in the RI PDMP in 2015

Network characteristics	Patient-based Network		Provider-based Network	
Basic characteristics	All	LCC	All	LCC
Number of vertices	372	274	746	479
Number of edges	1,980	1,894	1,901	1,437
Density	0.0287	0.0506	0.0068	0.0126
Average degree (SD)	10.7 (11.7)	13.8 (12.0)	5.1 (4.1)	6.0 (4.7)
Average path length	4.18	4.19	5.05	5.10
Diameter	10 (13)	10 (13)	11	11
Global transitivity	0.853	0.852	0.498	0.460
Mean Local transitivity	0.731	0.718	0.839	0.789
Assortative coefficient	0.785	0.751	0.156	0.040
Number of components	55	1	55	1
Size of LCC	274	274	479	479

Abbreviations: LCC= largest connected component; SD=standard deviation.

Table 3. Distribution of standardized centrality measures by patient characteristics using the largest connected component of the opioid prescription network

	Quartiles of standardized centrality measures											
	Betweenness centrality				Degree centrality				Closeness centrality			
Characteristic	T 1 (n=91)	T 2 (n=92)	T 3 (n=91)	p-value	T 1 (n=93)	T 2 (n=89)	T 3 (n=92)	p-value	T 1 (n=92)	T 2 (n=91)	T 3 (n=91)	p-value
Age group (years)												
21-44	36	28	39	0.022	36	35	32	0.035	37	33	33	0.178
45-64	36	55	43		41	38	55		38	45	51	
65+	19	9	9		16	16	5		17	13	7	
Gender												
Female	45	51	43	0.518	51	45	43	0.545	45	45	49	0.765
Male	46	41	48		42	44	49		47	46	24	
Payment method												
Commercial	54	61	63	0.883	61	54	63	0.825	55	65	58	0.750
Medicare	29	24	23		25	27	24		29	22	25	
Medicaid	6	5	4		5	7	3		6	3	6	
Cash	2	2	1		2	1	2		2	1	2	
Number of providers												
3	58	42	34	0.001	50	43	41	0.453	53	46	35	0.032
4+	33	50	57		43	46	51		39	45	56	
Number of pharmacies												
3	62	69	56	0.148	64	61	62	0.976	64	61	62	0.934
4+	29	23	35		29	28	30		28	30	29	
Most commonly filled opioids												
Bup/naloxone	3	12	17	0.024	10	21	1	< 0.001	7	14	11	0.106
Hydrocodone	36	31	26		46	25	22		38	32	23	
Oxycodone	52	49	48		37	43	69		47	45	57	
MME category, mg												
< 50	49	43	44	0.784	66	41	29	< 0.001	53	49	34	0.057
50-90	20	26	21		12	19	36		18	19	30	
>90	22	23	26		15	29	27		21	23	27	

Abbreviations: MME=morphine milligram equivalent, Bup=Buprenorphine; T1-T3=Tertile 1, Tertile 2 & Tertile 3.

Table 4. Odds Ratios (95% CI) Associated with a Patient being Classified in the Upper Tertile of Standardized Centrality Measures

	Standardized betweenness centrality		Standardized degree centrality		Standardized closeness centrality	
Characteristic	**Unadjusted OR**	**Adjusted OR**	**Unadjusted OR**	**Adjusted OR**	**Unadjusted OR**	**Adjusted OR**
Number of opioid Rx	0.96 (0.92, 1.01)	NS	1.10 (1.047, 1.155	NS	1.01 (0.96 1.05)	NS
Age group (years)						
65+	Ref.	NS	Ref.	Ref.	Ref.	NS
21-44	1.90 (0.81, 4.43)		2.88 (1.03, 8.08)	4.27 (1.41, 12.86)	2.02 (0.80, 5.07)	
45-64	1.47 (0.64, 3.38)		4.46 (1.63, 12.15)	5.12 (1.77, 14.83)	2.63 (1.08, 6.44)	
Gender						
Female	Ref.	NS	Ref.	NS	Ref.	NS
Male	1.232 (0.744, 2.04)		1.27 (0.77, 2.10)		0.83 (0.50, 1.37)	
Payment method						
Cash	Ref.	NS	Ref	NS	Ref.	NS
Commercial	2.19 (0.24, 20.03)		0.82 (0.13, 5.05)		0.73 (0.12, 4.46)	
Medicare	1.46 (0.12, 17.23)		0.69 (0.11, 4.42)		0.74 (0.12, 4.69)	
Medicaid	1.74 (0.18, 16.39)		0.38 (0.04, 3.36)		1.00 (0.13, 7.89)	
# of providers						
3	Ref.	Ref.	Ref.	NS	Ref.	Ref.
4+	2.02 (1.21, 3.38)	2.62 (1.49, 4.59)	1.30 (0.79, 2.15)		1.89 (1.13, 3.15)	2.09 (1.21, 3.64)
# of pharmacies						
3	Ref.	NS	Ref.	NS	Ref.	NS
4+	1.60 (0.93, 2.68)		1.06 (0.62, 1.82)		1.01 (0.59, 1.73)	
Opioid type						
Hydrocodone	Ref.	Ref	Ref.	Ref.	Ref.	NS
Bup/naloxone	2.92 (1.28, 6.69)	4.18 (1.73, 10.07)	0.10 (0.01, 0.81)	0.02 (0.00, 0.16)	1.59 (0.67, 3.80)	
Oxycodone	1.23 (0.69, 2.16)	1.13 (0.63, 2.02)	2.78 (1.56, 4.95)	1.10 (0.52, 2.33)	1.89 (1.06, 3.35)	
MME category, mg						
< 50	Ref.	NS	Ref.	Ref.	Ref.	Ref.
50-90	0.96 (0.51, 1.79)		4.29 (2.28, 8.06)	4.18 (1.97, 8.84)	2.43 (1.31, 4.51)	2.32 (1.24, 4.34)
>90	1.21 (0.66, 2.21)		2.26 (1.21, 4.26)	6.73 (2.69, 16.86)	1.84 (0.99, 3.41)	2.31 (1.20, 4.44)

Abbreviations: CI=confidence interval; OR=odds ratio; Rx=prescription; MME=morphine milligram equivalent, Bup=Buprenorphine, NS=not statistically significant

Table 5. Mixing Matrices of Categorical Patient Attributes for an Opioid Prescription Network of Patients

	All, N=372					LCC, N=274				
Patient Attributes		**Mixing matrix**					**Mixing matrix**			
	N (%)	1	2	3	4[a]	N(%)	1	2	3	4[a]
Gender										
Female	197 (52.96)	496	949			139 (50.73)	479	904		
Male	175 (47.04)	949	535			135 (49.27)	904	511		
Age category										
<44	126 (33.87)	263	745	119		103 (37.59)	262	723	111	
45-64	189 (50.81)	745	635	192		134 (48.91)	723	606	170	
65+	57 (15.32)	119	192	26		37 (13.50)	111	170	22	
Use of cash										
Cash only	9 (2.42)	1	73			5 (1.82)	1	69		
Insurance only	33 (97.58)	73	1906			269 (98.18)	69	1824		
Payment method										
Commercial	239 (64.25)	887	95	791	52	178 (64.96)	855	92	664	49
Medicaid	20 (0.00)	95	3	60	2	15 (0.00)	92	2	57	2
Medicare	104 (27.96)	701	60	160	19	76 (27.74)	664	59	154	18
Cash	9 (2.42)	52	2	19	1	5 (1.82)	49	2	18	1
Type of opioid medication										
Bup/Naloxone	33 (0.00)	88	14	76		32 (0.00)	88	14	76	
Hydrocodone	126 (33.87)	14	158	687		93 (33.94)	14	152	659	
Oxycodone	213 (57.26)	76	687	957		149 (54.38)	76	659	905	
Daily MME										
<50	197 (52.96)	339	471	370		136 (49.64)	301	457	347	
50-90	81 (21.77)	491	231	336		67 (24.45)	457	231	329	
>90	94 (25.27)	370	336	233		71 (25.91)	347	329	229	
Number of prescribers										
3	187 (50.27)	526	718			134 (48.91)	512	693		
4+	185 (49.73)	718	736			140 (51.09)		689		
Number of pharmacies										
3	257 (69.09)	908	883			187 (68.25)	856	852		
4+	115 (30.91)	883	189			87 (31.75)	852	186		

Abbreviations: LCC= largest connected component; Bup=Buprenorphine; MME=morphine milligram equivalent,
[a]Number of categories for each attribute

Table 6. Clustering by Node Attribute using Modularity Score

Node attributes	Total	LCC
Gender	0.021	0.023
Age group (years)	-0.012	0.025
Age, years (continuous)	0.024	-0.013
Payment method	0.013	0.015
Number of providers (continuous)	0.080	0.078
Number of providers (categorical)	0.132	0.130
Number of pharmacies (continuous)	-0.009	-0.009
Number of pharmacies (categorical)	-0.012	-0.012
Type of opioid prescription used	0.080	0.082
Number of opioid prescriptions	0.018	0.018
Total daily dose, MME (mg)	-0.006	-0.007
MME category, mg	0.068	0.066

Abbreviations: LCC= largest connected component; MME=morphine milligram equivalent.

Table 7. Main Effects Model to Estimate the Log odds of a tie between two patients in the Opioid Prescription Network of Patients Using the LCC

Term	Estimate	SE	P-value
Edges	-5.661	0.174	< 0.001
Number of opioid prescriptions	0.009	0.003	0.004
Gender			
Female	Ref.		
Male	-0.002	0.035	0.953
Age category			
<44	Ref.		
45-64	0.064	0.038	0.091
65+	-0.620	0.074	< 0.001
Payment method			
Commercial	Ref.		
Medicaid	-0.468	0.086	< 0.001
Medicare	0.132	0.045	0.003
Cash	-0.201	0.127	0.115
Opioid type			
Buprenorphine/Naloxone	Ref.		
Hydrocodone	0.875	0.092	< 0.001
Oxycodone	0.983	0.080	< 0.001
Daily MME			
<50	Ref.		
50-90	0.533	0.048	< 0.001
>90	0.703	0.060	< 0.001
Number of prescribers			
3	Ref.		
4+	0.133	0.036	< 0.001
Number of pharmacies			
3	Ref.		
4+	-0.040	0.037	0.284

Abbreviations: SE=standard error; MME=morphine milligram equivalent; AIC=Akaike information criterion; BIC=Bayesian information criterion, LCC= largest connected component.

Table 8. Main Effects Model with Homophily Terms to Estimate the Log odds of a tie between two patients in the Opioid Prescription Network of Patients Using the LCC

Term	Estimate	SE	P-value
Edges	-6.283	0.191	< 0.001
Number of opioid prescriptions	0.031	0.004	< 0.001
Gender			
Female	Ref.		
Male	-0.011	0.034	0.746
Age category			
<44	Ref.		
45-64	0.064	0.039	0.845
65+	-0.637	0.076	< 0.001
Payment method			
Commercial	Ref.		
Medicaid	-0.431	0.094	< 0.001
Medicare	0.178	0.050	< 0.001
Cash	-0.116	0.134	0.386
Type of opioid prescription			
Buprenorphine/Naloxone	Ref.		
Hydrocodone	0.786	0.091	< 0.001
Oxycodone	0.776	0.081	< 0.001
Average daily MME, mg			
<50	Ref.		
50-90	0.476	0.048	< 0.001
>90	0.711	0.060	< 0.001
Number of prescribers			
3	Ref.		
4+	0.094	0.033	0.005
Number of pharmacies			
3	Ref.		
4+	-0.066	0.043	0.130
Homophily terms			
Number of opioid prescriptions	-0.060	0.006	< 0.001
Gender	0.106	0.048	0.027
Age category	0.107	0.053	0.016
Method of payment	0.107	0.060	0.074
Type of opioid prescription	0.456	0.059	< 0.001
Average daily MME, mg	0.169	0.052	0.001
Number of prescribers	0.579	0.050	< 0.001
Number of pharmacies	-0.062	0.056	0.265

Abbreviations: SE=standard error; MME=morphine milligram equivalent; AIC=Akaike information criterion; BIC=Bayesian information criterion.

Table 9. Main Effects Model with Differential Homophily Terms to Estimate the Log odds of a tie between two patients in the Opioid Prescription Network of Patients Using the LCC

	Main effects			Differential homophily		
Term	Estimate	SE	P-value	Estimate	SE	P-value
Edges	-9.485	0.412	< 0.001	NA	NA	NA
Number of opioid prescriptions	0.031	0.004	< 0.001	-0.060	0.006	< 0.001
Gender						
Female	Ref.	Ref.		0.212	0.097	0.028
Male	0.094	0.059	0.110	NA	NA	NA
Age category						
<44	Ref.	Ref.		-0.058	0.148	0.697
45-64	-0.071	0.126	0.576	0.152	0.139	0.273
65+	-0.744	0.096	< 0.001	0.505	0.264	0.056
Payment method						
Commercial	Ref.	Ref.		0.263	0.159	0.098
Medicaid	-0.304	0.136	0.025	-0.480	0.740	0.516
Medicare	0.349	0.151	0.021	-0.152	0.179	0.397
Cash	-0.031	0.164	0.852	0.774	1.097	0.480
Type of opioid prescription						
Buprenorphine/Naloxone	Ref.	Ref.		4.776	0.327	< 0.001
Hydrocodone	1.760	0.138	< 0.001	0.919	0.315	0.004
Oxycodone	2.661	0.276	< 0.001	-0.719	0.300	0.016
Average daily MME, mg						
<50	Ref.	Ref.		0.083	0.118	0.480
50-90	0.447	0.088	< 0.001	0.151	0.118	0.202
>90	0.725	0.088	< 0.001	0.028	0.130	0.831
Number of prescribers						
3	Ref.	Ref.		1.104	0.101	< 0.001
4+	0.642	0.058	< 0.001	NA	NA	NA
Number of pharmacies						
3	Ref.	Ref.		-0.122	0.112	0.273
4+	-0.127	0.086	0.139	NA	NA	NA

Abbreviations: NA=not applicable; SE=standard error; MME=morphine milligram equivalent; AIC=Akaike information criterion; BIC=Bayesian information criterion.

FIGURES

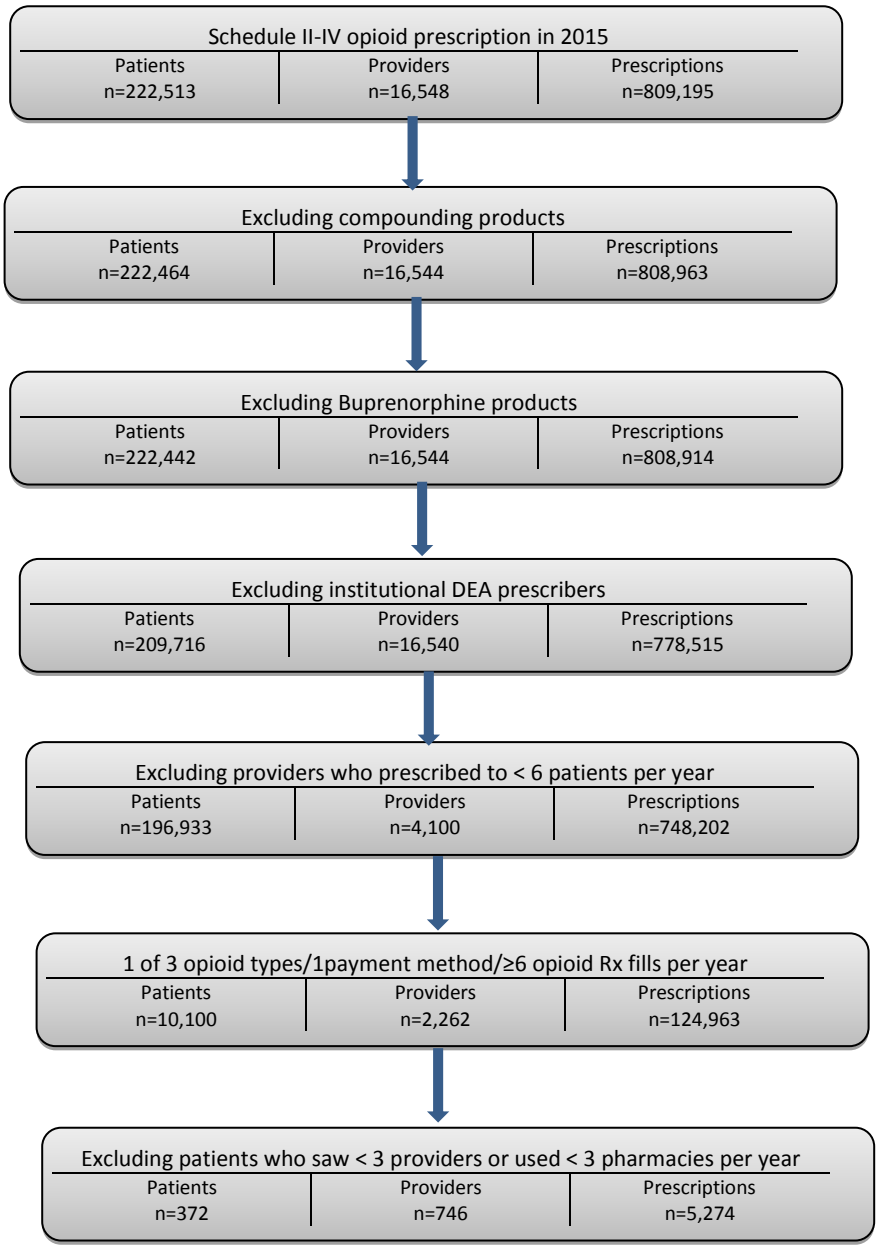

Figure 1. Study Sample Selection Flowchart showing number of patients, providers, and opioid prescriptions with exclusion criteria

A

		Providers		
		X	**Y**	**Z**
Patients	**1**	1	1	0
	2	1	0	0
	3	0	1	0
	4	0	1	1
	5	0	1	1
	6	0	0	1

$\mathbf{AA}^{T}$

B

	1	**2**	**3**	**4**	**5**	**6**
1	2	1	1	1	1	0
2	1	1	0	0	0	0
3	1	0	1	1	1	0
4	1	0	1	2	2	1
5	1	0	1	2	2	1
6	0	0	0	1	1	1

Figure 2: Matrix illustration of the construction of a simple network. Bipartite incidence matrix, **A** (left panel) shows the relation "prescribed opioid analgesic(s) to" and the one-mode projection adjacency matrix **B** (right panel) shows the provider-sharing relationships between patients. One-mode projection of the bipartite network is obtained by post-multiplying the matrix **A** by its transpose, $\mathbf{A}^{T}$.

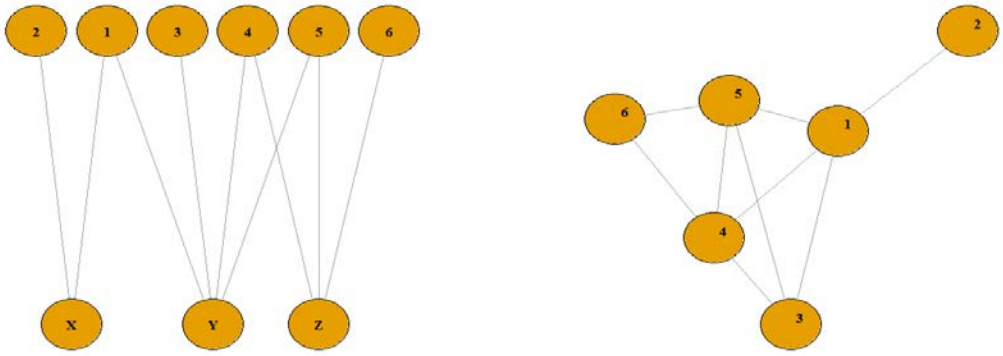

Figure 3. Schematic illustrating a projection from a two-mode to a one-mode network**.** Bipartite network graph of "prescribed opioid analgesic(s) to" relation for 3 providers (X, Y, Z) and 6 patients (1-6) [left panel] and a unipartite network graph of "shares a provider with" relation for 6 patients (right panel).

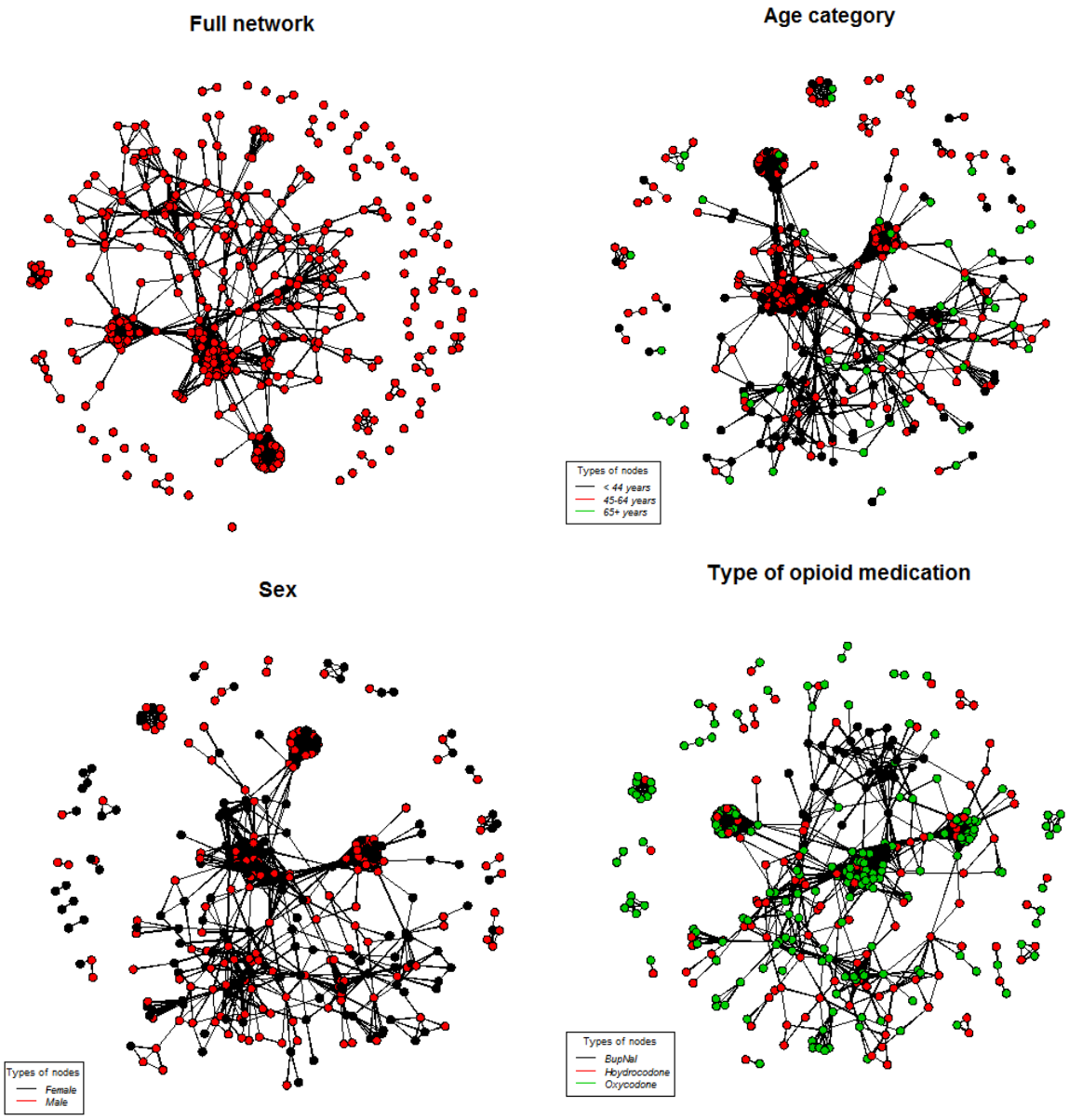

Figure 4. Four depictions of the full patient network: (a) Full network with few isolates and clusters; (b) clustering by age category; (c) clustering by sex; and (b) clustering by opioid type.

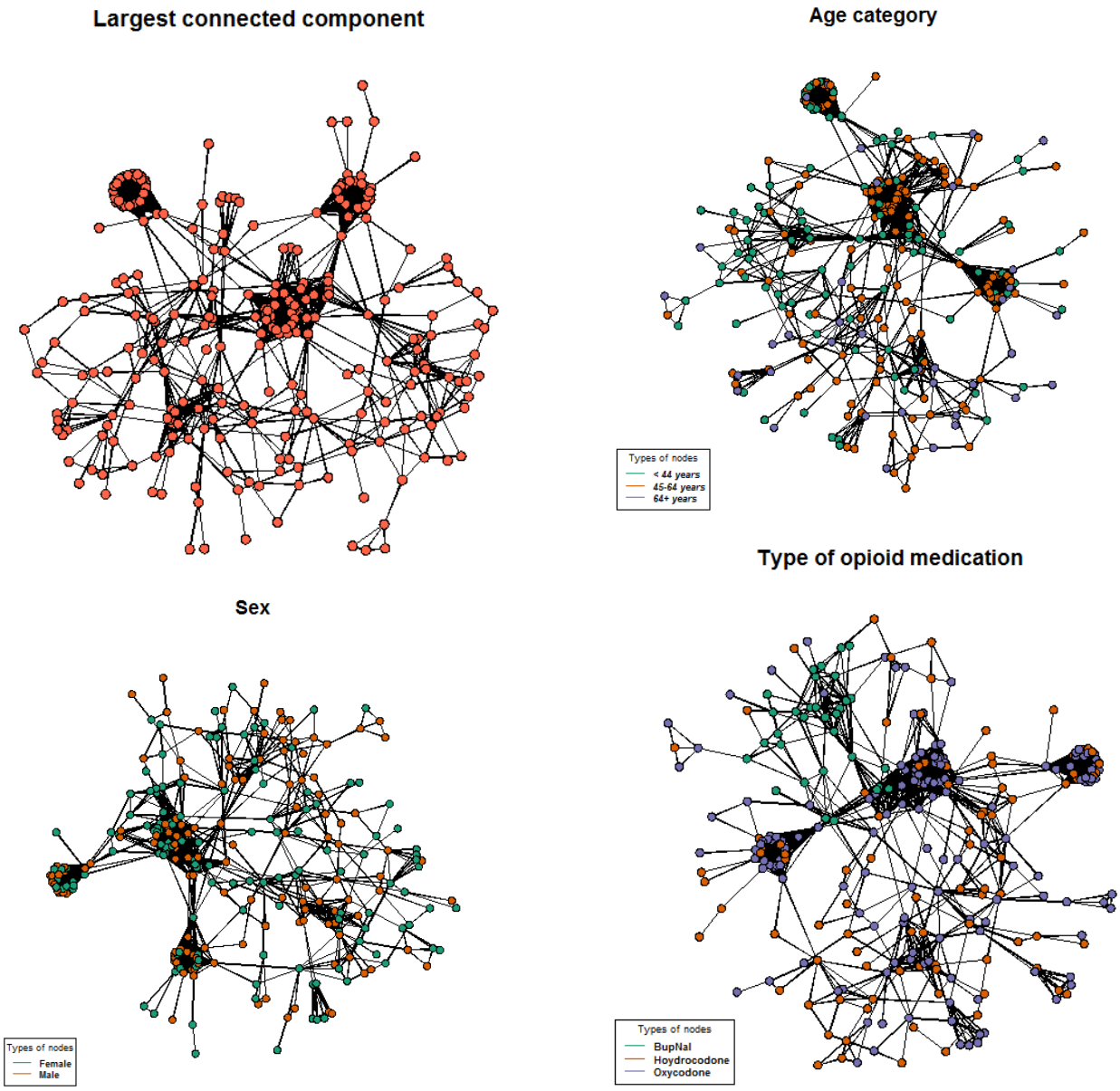

Figure 5. Four depictions of the largest connected component: (a) The largest connected component with few clusters; (b) clustering by age category; (c) clustering by sex; and (b) clustering by opioid type.

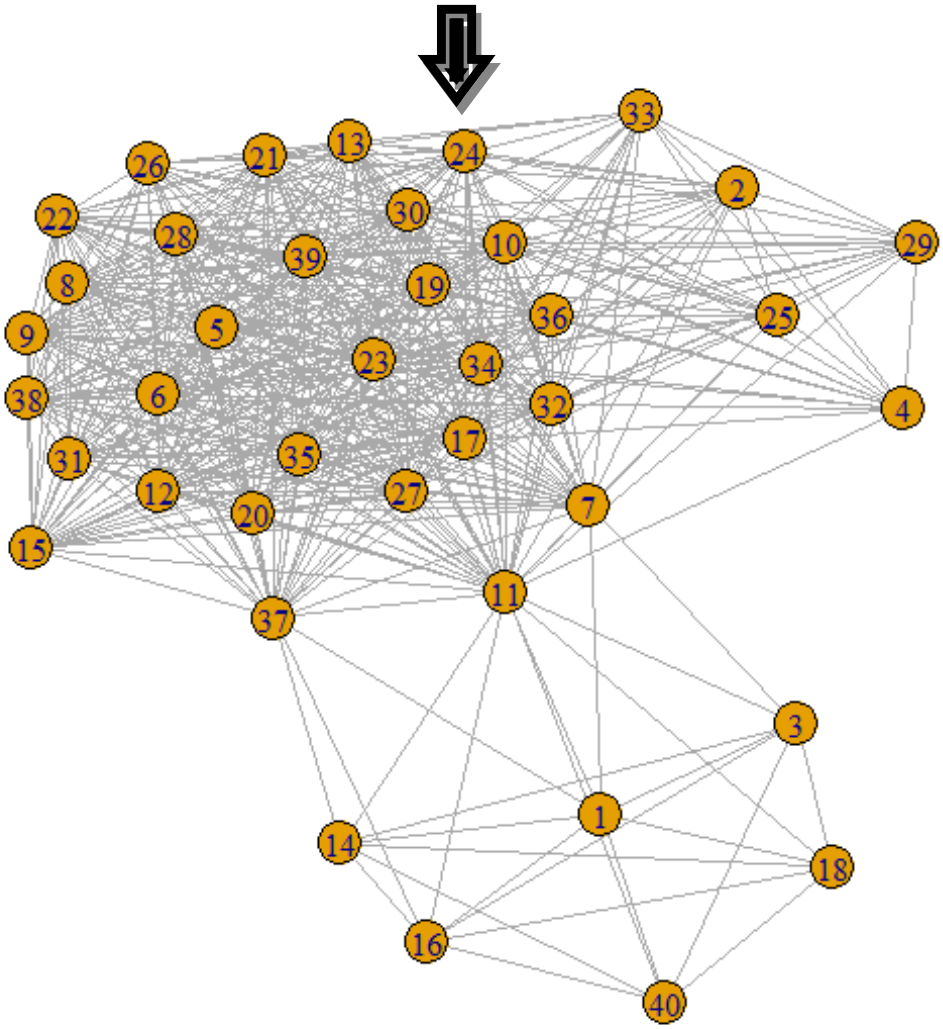

Figure 6: Ego-centric network around the most influential patient (# 24) identified by degree, closeness, and betweenness centrality measures. This ego-centric network is subgraph corresponding to the most influential patient with his or her immediate neighborhood.

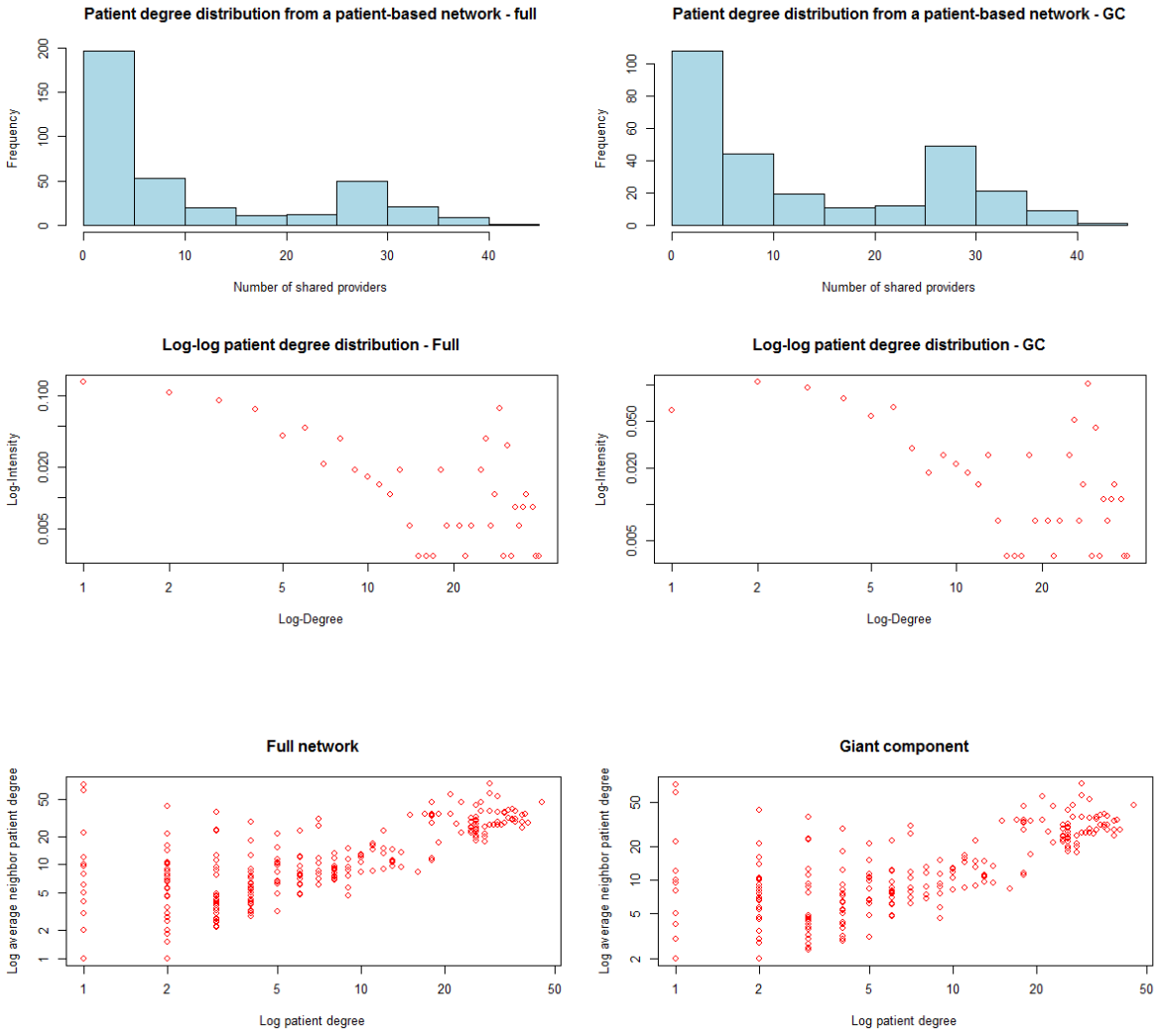

Figure 7. Degree distribution in original scale (top panels), log-log scale (middle panels), and the average neighbor degree versus vertex degree on a log-log scale (bottom panels) for complete network (left) and its largest connected component (right).

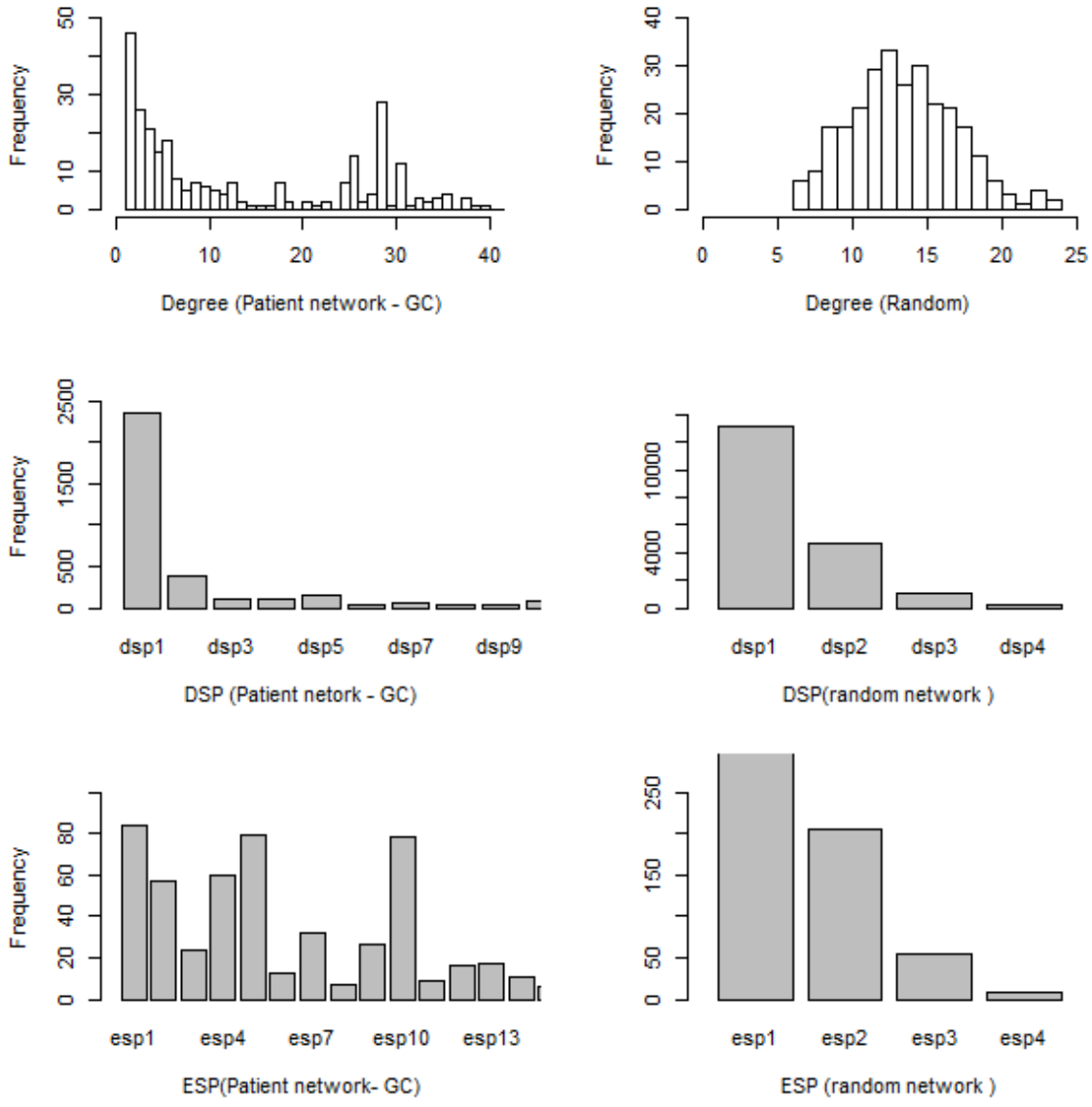

Figure 8: Plots of degree and shared partnerships (DSP and ESP) in the largest connected component of the network (left) and a randomly generated network of the same size and density (right). Abbreviations: DSP=dyad-wise shared partners, ESP= edgewise shared partners, GC=giant component.

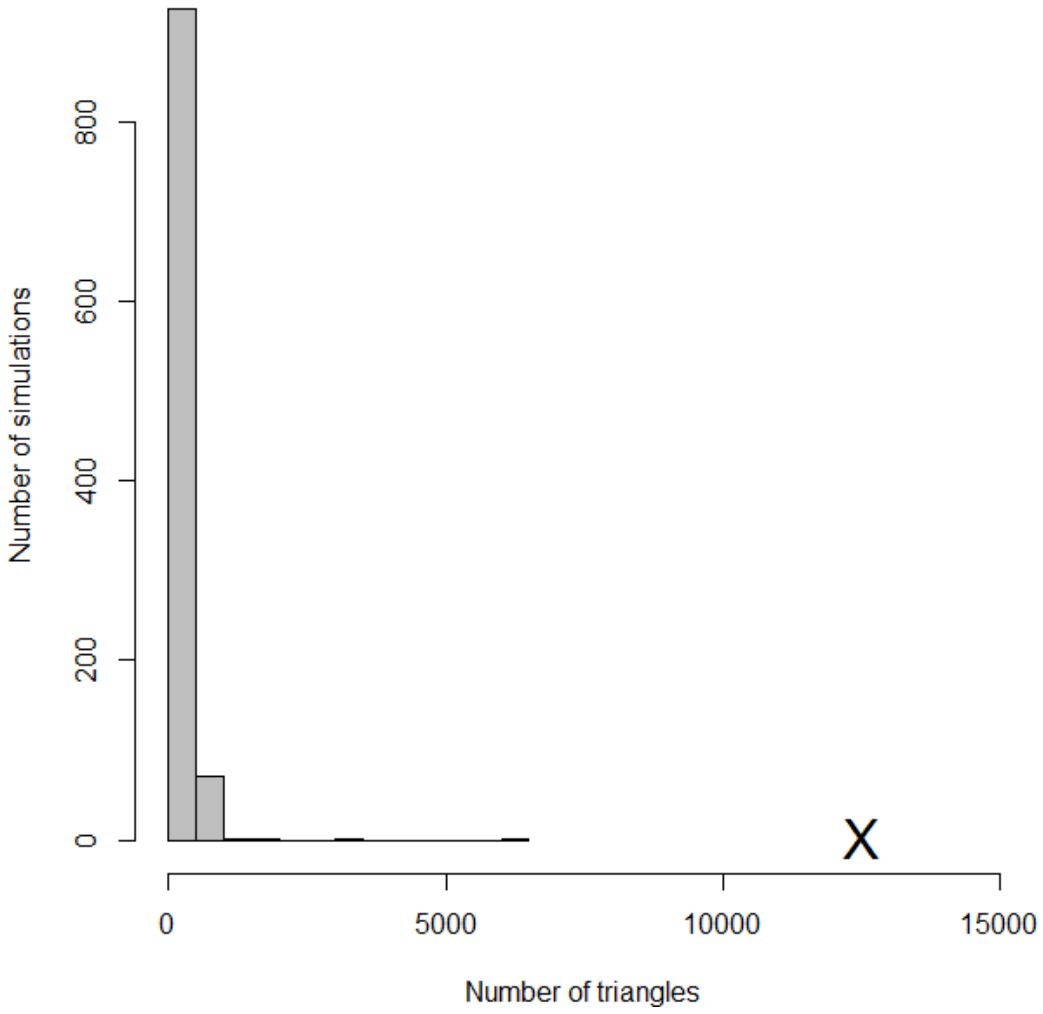

Figure 9: Number of triangles in 1000 networks simulated based on the null model of the largest connected component of the network.

FUTURE RESEARCH WORK

In the future we plan to incorporate structural properties of the observed opioid network as model covariates thereby allowing the observed network to be conditioned on observed degree distribution and level of transitivity. This will be implemented in R using a Bayesian approach which exhibits better convergence properties than non-Bayesian method used in this analysis. We also plan to analyze corresponding provider- and pharmacy-based networks where connections between providers or pharmacies represent patient-sharing. These analyses will also use centrality measures to identify the most influential providers and pharmacies and evaluate whether patterns of opioid prescribing or dispensing vary by communities in the network.

TECHNICAL APPENDIX

A. NETWORK CHARACTERIZATION AND MEASUREMENTS

This Technical Appendix provides statistical formulations of key terms used in this thesis. The notations and definitions are adapted from Kolaczyk.[1] In general, let uppercase letters denote random variables and lowercase denote realizations of those random variables. We assume the observed network is fixed and does not vary over time. We also assumed that we ascertained the full network sample.

Defining a Network: A *graph* $G = (V, E)$ is a mathematical structure consisting of a set V of *vertices* or nodes and a set E of *edges* or *links*, where elements of E are unordered pairs $\{u,v\}$ of distinct vertices $u,v \in V$ for an undirected graph. The number of vertices $N_v = |V|$ and the number of edges $N_e = |E|$ are called the *order* and *size* of the graph G, respectively. A graph $H = (V_H, E_H)$ is a *subgraph* of another graph $G = (V_G, E_G)$ if $V_H \subseteq V_G$ and $E_H \subseteq E_G$. An *induced subgraph* of G is a subgraph $G' = (V', E')$, where $V' \subseteq V$ is a pre-specified subset of vertices and $E' \subseteq E$ is the collection of edges to be found in G among that subset of vertices. A *simple graph* has no edges for which both ends connect to a single vertex (i.e., no *loops*) and no pairs of vertices with more than one edge between them (i.e., no *multi-edges*). Two vertices $u,v \in V$ are said to be *adjacent* if joined by an edge in E, and two edges $e_1,e_2 \in E$ are adjacent if joined by a common endpoint in V. A vertex $v \in V$ is *incident* on an edge $e \in E$ if v is an endpoint of e.

A network therefore consists of a set of nodes and the relationships (ties, links, edges, connections) among them. The relationship can be directed or undirected and dichotomous (present or absent). All pairs of nodes in a network are *dyads* while all sets

3 nodes form *triads*. These dyads and triads can be linked or unlinked. A group of k nodes taking a star format with a node at the center linked to all others in the group is called *k*-star. An undirected network has two types of dyads (null or present) and four types of triads defined by the number of connected edges (0, 1, 2 or 3) and may have a 3-star, 4-star, and 5-star formats. We also distinguish *edgewise shared partnership* (ESP) and *dyad-wise shared partnership* (DSP). A DSP is a linked or unlinked dyad where both members of the dyad are linked to a third network member. ESP is a subset of DSP with a linked dyad in which both members of the dyad also have a link to a third network member. The distribution of ESP in a network shows how many dyads have one shared partner, two shared partners, and so on. Similarly, the distribution of DSP shows the number of dyads in the network with one shared partner, two shared partners, and so on.

A *bipartite* network is a graph G = (V, E) such that the vertex set V may be partitioned into two disjoint sets, say V_1 and V_2, and each edge in E has one endpoint in V_1 and the other in V_2. Specifically, a graph $G_1 = (V_1, E_1)$ may be defined on the vertex set V_1 by assigning an edge to any pair of vertices that both have edges in E to at least one common vertex in V_2. Similarly, a graph G_2 may be defined on V_2.

Several notions are related to the concept of movement around a graph. A *walk* on a graph G, from v_0 to v_l, is an alternating sequence $\{v_0, e_1, v_1, e_2, \ldots, v_{l-1}, e_l, v_l\}$, where the endpoints of e_i are $\{v_{i-1}, v_i\}$. The *length* of this walk is said to be l. A *trail* is a walk without repeated edges and a *path* is a trail without repeated vertices. A vertex *v* in a graph G is said to be *reachable* from another vertex *u* if there exists a walk from *u* to *v*. The graph G is said to be *connected* if every vertex is reachable from every other. A *component* of a graph is a maximally connected subgraph. *Geodesic* (distance) is the

length of the shortest path(s) between the vertices (which we set equal to infinity if no such path exists). The *diameter* of the graph is the value of the longest distance in a graph.

Algebraic graph theory has several applications in social network analysis. The connectivity of a graph G may be captured and stored in an $N_v \times N_v$ binary, symmetric, adjacency matrix A with entries:

$$A_{ij} = \begin{cases} 1, & \text{if } \{i,j\} \in E, \\ 0, & \textit{otherwise}, \end{cases}$$

where A is non-zero for entries whose row-column indices correspond to vertices in G joined by an edge, and zero, for those that are not. The row sum $A_{i+} = \Sigma_j A_{ij}$ is equal to the degree d_i of vertex *i*. and by symmetry, $A_{i+} = A_{+i}$. The structure of a graph G may also be captured in an $N_v \times N_e$ binary, incidence matrix B with entries:

$$B_{ij} = \begin{cases} 1, & \text{if vertex } i \text{ is incident to edge } j, \\ 0, & \textit{otherwise}. \end{cases}$$

Suppose that G = (V, E) is a graph corresponding to an observed social network among individuals $i \in V$, with a social tie between individuals $i, j \in V$ indicated by an edge $\{i, j\} \in E$. Let $Y_{ij} = Y_{ji} = 1$ if $\{i, j\} \in E$, and zero if not. $Y = [Y_{ij}]$ is the adjacency matrix for G, and treated as a random matrix.

Modularity: The process of community detection can be approached as an optimization problem using computational algorithms developed for studying similar networks.[2,3] The algorithm detects subgroups within networks that are more inter-connected than would be expected by chance alone.[3-5] In our example, each provider was assigned to a single community, such that communities are comprised of distinct, non-overlapping groups of providers in the network. The null model adjusts for node degree so that patients with

high nodal degree are more likely to be connected than those with low nodal degree thereby maintaining the expected degree distribution of the network.[4] The optimization process involves the maximization of the quantity:

$$Q = \frac{1}{2m}\sum_i \sum_j \left[A_{ij} - \frac{k_i k_j}{2m}\right] \delta(s_i, s_j),$$

where A_{ij} is nonzero if and only if node i and j are connected by a tie, and its value quantifies the number of providers the two patients share; k_i is the degree of node i, m is the number of edges in the network (or their total weight in weighted networks), s_i is the community assignment of node i, and $\delta(s_i, s_j)$ is the Kronecker delta which is equal to 1 if the arguments are identical, otherwise it is zero. We used the greedy optimization method which has been shown to perform well for a variety of networks[6].

Centrality measures: A patient that is connected to many other patients in a network is in a prominent or influential position within the network. This simplest measure of centrality is based on the notion that a patient with more direct connections in the network is more influential than one with fewer or no connections at all. The degree d_v of a vertex *v*, in a network graph G = (V, E), is the number of edges in E incident upon *v*, that is, at distance one and mean degree is the average degree of all patients in the opioid network. Vertex degree is arguably the most widely used measure of vertex centrality. In our setting, the patients with higher degrees are more central because in many social settings people with more connections tend to be more influential. A patient's degree is the total number of other patients within the network who are connected to the patient through provider sharing. Degree centrality we can be standardized by dividing by the maximum possible value of |V|-1.

Given a network graph G, we define f_d to be the fraction of vertices $v \in V$ with degree $d_v = d$. The collection $\{f_d\}_{d\geq 0}$ is called the degree distribution of G. The degree distribution provides a summary of the connectivity in the graph.

Another notion of a 'central role in the network' is that a vertex be 'close' to many other vertices. The standard approach, introduced by Sabidussi, is to let the centrality vary inversely with a measure of the total distance of a vertex from all others.[7]

$$cl(v) = \frac{1}{\sum_{u\in V} \mathrm{dist}(v,u)},$$

where dist(v,u) is the geodesic distance between the vertices $u,v \in V$. Using this formulation, the more central a node is, then the lower its total distance to all other nodes. For comparison across networks and with other centrality measures, this measure is normalized or standardized to lie in the interval [0,1], through multiplication by a factor $N_v - 1$.

Betweenness centrality is based upon the perspective that importance relates to where a vertex is located with respect to the paths in the network graph. If we assume those paths as the routes by which communication and other exchanges takes place, vertices that sit on many paths are likely more critical to the communication and other dissemination processes. Betweenness centrality measures are aimed at summarizing the extent to which a vertex is located 'between' other pairs of vertices. This is defined as the proportion of times the provider lies on the shortest paths in the network. A patient is central in the network if he/she is located between many non-adjacent patients on their geodesics (shortest) paths. The most commonly used betweenness centrality, introduced by Freeman,[8] is defined as

$$c_B(v) = \sum_{s\neq t\neq v\in V} \frac{\sigma(s,t|v)}{\sigma(s,t)},$$

where σ (*s*, *t*|*v*) is the total number of shortest paths between *s* and *t* that pass through *v*, and σ (*s*, *t*) = Σ_vσ (*s*, *t*|v). In the event that shortest paths are unique, c_B(v) just counts the number of shortest paths going through *v*. This centrality measure can be normalized through division by a factor of $(N_v-1)(N_v-2)/2$.

Network cohesion: The definition of network cohesion depends on the context. Generally, network cohesion refers to the extent to which subsets of nodes are connected to each other to form triads, components, clusters, and communities.

Network density: For a graph G with no self-loops and no multiple edges, the density of a subgraph H = (V_H, E_H) is

$$\text{den}(H) = \frac{|E_H|}{|V_H|(|V_H|-1)/2}$$

The value of den(*H*) lies between zero and one.

Clustering coefficient and transitivity: A triangle is a complete subgraph of order three. A connected triple is a subgraph of three vertices connected by two edges. A measure of the frequency with which connected triples form closed triangles provides some indication of the extent to which edges are 'clustered' in the graph. Let $\tau_\Delta(v)$ denote the number of triangles in G into which $v \in V$ falls, and $\tau_3(v)$, the number of connected triples in G for which the two edges are both incident to *v*. The local clustering coefficient, den(H_v) can be re-expressed as $cl(v) = \tau_\Delta(v)/\tau_3(v)$ for those vertices v with $\tau_3(v) > 0$[9]. The corresponding clustering coefficient for G takes the form:

$$cl(G) = \frac{1}{V'}\sum_{v\in V'} \text{cl}(v),$$

where V' ⊆V is the set of vertices *v* with $d_v \geq 2$.

The clustering coefficient is the ratio of total the number of connections that exist among neighbors of the patient in the network to the total number of potential connections that could exist if they were completely connected. It is used to describe the extent to which network neighbors of a particular patient are directly connected to each other and interpreted as the probability that any two randomly selected neighbors of a particular patient in the network are connected to each other.

Assortativity and mixing: Assortative mixing is the selective linking among vertices, according to a certain characteristic(s), and measures that quantify the extent of assortative mixing in a given network have been referred to as assortativity coefficients. Suppose that each vertex in a graph G can be labeled according to one of M categories. Let f_{ij} be the fraction of edges in G that join a vertex in the i^{th} category with a vertex in the j^{th} category; denote the i^{th} marginal row and column sums of the resulting matrix f by f_{i+} and f_{+i}, respectively. We then define the assortativity coefficient r_a to be

$$r_a = \frac{\sum_i f_{ii} - \sum_i f_{i+} f_{+i}}{1 - \sum_i f_{i+} f_{+i}}.$$

The value r_a is equal to zero when the mixing in the graph is no different from that obtained through a random assignment of edges that preserves the marginal degree distribution. Similarly, it is equal to one when there is perfect assortative mixing (i.e., when edges only connect vertices of the same category). When the mixing is perfectly disassortative, the value takes its minimum value, that is, every edge in the graph connects vertices of two different categories.

B. The Exponential Random Graph Model (ERGM)

A discrete random vector **Z** is said to belong to an *exponential family* if its probability mass function may be expressed in the form

$$P_\theta(Z = z) = exp\{\theta^T \boldsymbol{g}(\boldsymbol{z}) - \psi(\theta)\},$$

where $\theta \in \mathbb{R}^p$ is a $p \times 1$ vector of parameters, $\mathbf{g}(\cdot)$ is a p-dimensional function of **z**, and $\psi(\theta)$ is a normalization term, ensuring that $P_\theta(\cdot)$ sums to one over its range. The class of discrete exponential families includes many familiar distributions, such as the binomial, geometric, and Poisson. In the case of continuous exponential families, where an analogous form of the equation holds for probability density functions, examples include the Gaussian and chi-square distributions.

Consider G = (V, E) as a random graph. Let $Y_{ij} = Y_{ji}$ be a binary random variable indicating the presence or absence of an edge e ∈ E between the two vertices *i* and *j* in V. The matrix $Y = [Y_{ij}]$ is thus the (random) adjacency matrix for G. Denote by $y = [y_{ij}]$ a particular realization of Y. An exponential random graph model is a model specified in exponential family form for the joint distribution of the elements in **Y**. An ERGM takes the form

$$P_\theta(\mathbf{Y} = \mathbf{y}) = \left(\frac{1}{\kappa}\right) \exp\{\textstyle\sum_H \theta_H g_H(\boldsymbol{y})\},$$

where the following conditions hold:

(i) each H is a configuration is a set of possible edges among a subset of the vertices in G;

(ii) $g_H(\mathbf{y}) = \prod_{y_{ij} \in H} y_{ij}$, and is either one if the configuration **H** occurs in **y**, or zero, otherwise;

(iii) a non-zero value for θ_H means that the $\mathbf{Y}_{ij}$ are dependent for all pairs of vertices $\{i, j\}$ in **H**, conditional upon the rest of the graph; and

(iv) $\boldsymbol{\kappa} = \kappa(\boldsymbol{\theta})$ is a normalization constant, where

$\boldsymbol{\kappa}(\boldsymbol{\theta}) = \sum_{y} exp\{\sum_{H} \theta_H g_H(\boldsymbol{y})\}$.

The summation in the previous equation is over all possible configurations **H**. Note that this model implies a certain (in)dependency structure among the elements in **Y**. Generally, such assumptions specify that the random variables $\{\mathbf{Y}_{ij}\}_{(i,j)\in A}$ are independent of $\{Y_{i'j'}\}_{(i'j')\in B}$, conditional on the values of $\{Y_{i'j'}\}_{(i'j')\in C}$, for given index sets A ,B, and C .

Bernoulli or Simple Random Graph: The simple random graph model randomly distributes ties or connections among network members based on the same specified probability.[10,11] Network density or the probability of a tie occurring is the proportion of observed ties out of all possible ties:

$$\text{den}(H) = \frac{|E_H|}{|V_H|(|V_H|-1)/2}$$

This calculation assumes that ties are independent and identically distributed and ignores the attributes of network members that may influence the probability of a tie. Suppose we specify that, for any given pair of vertices, the presence or absence of an edge between that pair is independent of the status of possible edges between any other pairs of vertices. That is, for each pair $\{i, j\}$, we assume that $\mathbf{Y}_{ij}$ is independent of $Y_{i',j'}$, for any $\{i', j'\} \neq \{i, j\}$. This assumption implies that $\theta_H = 0$ for all configurations **H** involving three or more vertices. In this case, the only relevant functions g_H are those of the form: $g_H(\mathbf{y}) = g_{ij}(\mathbf{y}) = y_{ij}$, and the ERGM reduces to

$$P_\theta(\mathbf{Y} = \mathbf{y}) = \left(\frac{1}{\kappa}\right) \exp\left\{\sum_{i,j} \theta_{ij} y_{ij}\right\},$$

This implies that each edge $\{i, j\}$ is present in the graph independently with probability

$$p_{ij} = \exp(\theta_{ij})/[1+\exp(\theta_{ij})] .$$

However, this entails a model with N_v^2 parameters, which is likely far too parameterized for many data sets.

In order to reduce the total number of parameters, it is common to impose an assumption of homogeneity across certain vertex pairs. For example, assuming homogeneity across all of G (i.e., $\theta_{ij} \equiv \theta$, for all $\{i, j\}$) yields

$$P_\theta(\mathbf{Y} = \mathbf{y}) = \left(\frac{1}{K}\right) \exp\{\theta L(y)\},$$

where $\boldsymbol{L}(\mathbf{y}) = \Sigma_{i,j}\, y_{ij} = N_e$ is the number of edges in the graph. In this case, the Bernoulli random graph model is recovered, with $p = \exp(\theta)/[1+\exp(\theta)]$.

Assumptions of complete independence among possible edges are largely untenable in practice. In general, Bernoulli-like random graphs lack the ability to reproduce many of the most basic structural characteristics observed in most real-world networks. However, the simple random graph model provides a baseline to compare with more complex models and assess improvements in model fit using simulation methods.

Markov Random Graphs: Frank and Strauss introduced the notion of Markov dependence for network graph models, which specifies that two possible edges are dependent whenever they share a vertex, conditional on all other possible edges, and independent if they do not.[12] That is, the presence or absence of $\{i, j\}$ in the graph will depend upon that of $\{i,k\}$, for a given $k \neq j$, even given information on the status of all other possible edges in the network. A random graph G arising under Markov dependence conditions is called a Markov graph. This model was extended by assuming a more general conditional dependence among ties in a network (i.e., two connections are conditionally dependent if the conditional probability that both connections exist in the

network does not equal the product of their marginal conditional probabilities, given all other network ties.) Under an assumption of homogeneity, Frank and Strauss showed that G is a Markov graph if and only if $P_\theta(\cdot)$ may be expressed as

$$P_\theta(\mathbf{Y} = \mathbf{y}) = \left(\frac{1}{K}\right) \exp\left\{\sum_{k=1}^{N_v-1} \theta_k S_k(y) + \theta_\tau T(y)\right\},$$

where $S_1(\mathbf{y}) = N_e$ is the number of edges, $S_k(\mathbf{y})$ is the number of k-stars, for $2 \leq k \leq N_v - 1$, and $\mathbf{T}(\mathbf{y})$ is the number of triangles. The statistics S_k in and $\mathbf{T}$, can be correlated. We see from the definitions of the statistics S_k and $\mathbf{T}$ that Markov dependence results are explicitly parameterized to account for some effects of transitivity.

In practice, it is common to include star counts S_k no higher than k=2 or at most k=3, by setting $\theta_4 = \ldots = \theta_{N_v-1} = 0$. This often leads to model degeneracy. Inclusion of a large number of higher order terms does not solve this problem. Partial conditional dependence assumption has been proposed to address issues of degeneracy. For example, Snijders et al proposed a solution by imposing a parametric constraint of the form $\theta_k \propto (-1)^k \lambda^{2-k}$ upon the star parameter, for all $k \geq 2$, for some λ larger than one.[13] This tactic combines all k-star statistics $S_k(\mathbf{y})$, for $k \geq 2$, into a single alternating k-star statistic of the form

$$\mathrm{AKS}_\lambda(y) = \sum_{k=2}^{N_v-1} (-1)^k \frac{S_k(y)}{\lambda^{k-2}},$$

and weighting that statistic by a single parameter $\boldsymbol{\theta}_{AKS}$ that takes into account the star effects of all orders simultaneously. The alternating signs allow the counts of k-stars of successively greater order to balance each other, rather than simply ballooning. We often assume that dependence between ties that do not share a node is due to the presence of other ties in the network.[14] To account for this partial conditional dependence, three non-

linear terms are often added to the model: geometrically weighted degree (GWD), geometrically weighted DSP (GWDSP), and geometrically weighted ESP (GWESP). The statistic AKS_{λ} (**y**) is a linear function of GWD count. The GWD term is designed to account for the decreasing degree distribution in observed networks while GWESP term is designed to account for clustering in observed networks. Finally, the GWDSP term accounts for the number of dyads with shared partners, often found within clusters in the network.

C. Constructing the Exponential Random Graph Models

Several packages are available for estimating network models. Our analysis was conducted in R-statnet, a suit of packages for building ERGMs in R. We first employed the null model which corresponds to the simple random graph model and can be written as:

$$logit\{P(Y_{ij} = 1 | n\ patients, Y_{ij}^{c})\} = \theta_{edges}\delta_{edges}.$$

The model was estimated by maximum likelihood estimation and served as a comparator for assessing model fit as more useful and complex models were constructed.

Adding attributes: We first considered whether the addition of node attributes influenced the likelihood of a tie in the network. These nodal attributes accounted for the characteristics of each individual network member. To examine the effects of these attributes on the likelihood of a tie, these attributes were added to the model as main effects. The null and alternate hypotheses are:

H_0: There is no association between node attribute and the likelihood of a patient to form ties.

H_a: There is association between node attribute and the likelihood of a patient to form ties.

In statnet, categorical and continuous main effects are added using *nodefactor* and *nodecov()*, respectively. *Nodefactor* main effect term adds multiple statistics to the model output, each corresponding to the number of times a node with the specified attribute is at one end of an edge. The δ corresponding to a categorical node attribute can be summarized as follow:

$$\delta_{category} = \begin{cases} 2 \text{ if both nodes } i \text{ and } j \text{ have the characteristic} \\ 1 \text{ if } i \text{ or } j \text{ has the characteristic} \\ 0 \text{ if neither } i \text{ nor } j \text{ has the characteristic} \end{cases}$$

The reference group which is omitted in the output can be changed using the base argument.

The *nodecov* main effect term adds one network statistic to the output that sums the attribute of interest for the two nodes in a dyad.

Interaction terms for nodal attributes to account for the attributes of both members of a dyad in the network. Homophily interaction terms were included in the model using *nodematch*. Differential homophily was requested by specifying *diff=TRUE* after the name of the attribute in a *nodematch* term. The homophily change statistics is defined as:

$$\delta_{hom} = \begin{cases} 1 \text{ if both nodes } i \text{ and } j \text{ have the same value for a categorical covariate} \\ 0 \text{ otherwise} \end{cases}$$

And the differential homophily change statistics is as:

$$\delta_{diff} = \begin{cases} 1 \; if \; both \; nodes \; i \; and \; j \; have \; the \; same \; value \\ 0 \; otherwise \end{cases}$$

A potential limitation is that models that include the interaction terms are dyadic independence models which assume that each dyad is independent of all other dyads in the model.

Model fit and diagnostic assessments: Model fit assessment involves a systematic examination of how well the model actually captures the observed network structures being modeled. We compared models using the statistical measures of log-likelihood and the related deviance (-2LogL), the Akaike information criterion (AIC), or the Bayesian information criterion (BIC). The log-likelihood is calculated by summing the difference between predicted probabilities of Y_{ij} and the observed value of Y_{ij}.

$$loglikelihood = \sum_{i=1}^{N} [Y_{ij} \ln(P(Y_{ij})) + (1 - Y_{ij}) \ln(1 - P(Y_{ij}))]$$

The deviance is a measure of lack of fit and a larger deviance indicates a greater the lack of fit. The deviance gets smaller as more parameters are added to the model. The AIC and BIC account for this by penalizing models with more parameters that do not improve the model fit.

$$AIC = Deviance + 2p,$$

$$BIC = Deviance + p*2\ln(N),$$

where p is the number of parameters and N is the network size.

Both values of the AIC and BIC were used to compare nested and non-nested models. These measures of model fit were developed for the analysis of data that are assumed to meet the independence of observation assumption. The null, main effect, and homophily models which assume dyadic independence were compared using deviance, AIC, and BIC. Models to account for non-uniform degree distribution and transitivity resulting from complex of dependence in observed social networks GWD, GWESP, and GWDSP were not evaluated in this analysis.

REFERENCES FOR TECHNICAL APPENDIX

1. Kolaczyk E. *Statistical Analysis of Network Data: Methods and Models (Springer Series in Statistics).* Springer; 2009.

2. Fortunato S. Community detection in graphs. *Physics Reports.* 2. 2010;486(3–5):75-174.

3. Pollack CE, Lemke KW, Roberts E, Weiner JP. Patient sharing and quality of care: measuring outcomes of care coordination using claims data. *Med Care.* 2015;53(4):317-323.

4. Girvan M, Newman MEJ. Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences of the United States of America.* 2002;99(12):7821-7826.

5. Newman ME. Fast algorithm for detecting community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2004;69(6 Pt 2):066133.

6. Blondel V, D., Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment.* 2008;2008(10):P10008.

7. Sabidussi G. The centrality of a graph. *Psychometrika.* 1966;31(4):581-603.

8. Freeman LC. A Set of Measures of Centrality Based on Betweenness. *Sociometry.* 1977;40(1):35-41.

9. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature.* 1998;393:440.

10. Erdös P, Rényi A. On random graphs, I. *Publicationes Mathematicae (Debrecen).* 1959;6:290-297.

11. Gilbert EN. Random Graphs. *The Annals of Mathematical Statistics.* 1959;30(4):1141-1144.

12. Frank O, Strauss D. Markov Graphs. *Journal of the American Statistical Association.* 1986;81(395):832-842.

13. Snijders TA, Pattison PE, Robins GL, Handcock MS. New specifications for exponential random graph models. *Sociological Methodology.* 2006;36(1):99-153.

14. Robins G, Pattison P, Kalish Y, Lusher D. An introduction to exponential random graph (p*) models for social networks. *Social Networks.* 2007;29(2):173-191.