

Detecting Anomalous Networks of Opioid Prescribers and Dispensers in Prescription Drug Data

Katie Rosman, Daniel B. Neill

Machine Learning for Good Laboratory, New York University
ksr351@nyu.edu, daniel.neill@nyu.edu

Abstract

The opioid overdose epidemic represents a serious public health crisis, with fatality rates rising considerably over the past several years. To help address the abuse of prescription opioids, state governments collect data on dispensed prescriptions, yet the use of these data is typically limited to manual searches. In this paper, we propose a novel graph-based framework for detecting anomalous opioid prescribing patterns in state Prescription Drug Monitoring Program (PDMP) data, which could aid governments in deterring opioid diversion and abuse. Specifically, we seek to identify connected networks of opioid prescribers and dispensers who engage in high-risk and possibly illicit activity. We develop and apply a novel extension of the Non-Parametric Heterogeneous Graph Scan (NPHGS) to two years of de-identified PDMP data from the state of Kansas, and find that NPHGS identifies subgraphs that are significantly more anomalous than those detected by other graph-based methods. NPHGS also reveals clusters of potentially illicit activity, which may assist law enforcement and regulatory agencies. Our paper is the first to demonstrate how prescription data can systematically identify anomalous opioid prescribers and dispensers, illustrating the efficacy of a network-based approach. Additionally, our technical extensions to NPHGS offer both improved flexibility and graph density reduction, enabling the framework to be replicated across jurisdictions and extended to other domains.

Introduction

Misuse and abuse of opioids represents a widespread public health crisis in the United States. Between 2019 and 2020, deaths from opioid overdoses increased by nearly 40%, and opioid overdoses now account for 75% of all fatal drug overdoses (National Institutes of Health 2022). While the majority of opioid overdoses involve illicit substances, such as fentanyl and heroin, prescription opioids contribute to this crisis in several ways. Like illicit opioids, prescription opioids contain highly addictive properties that can lead to abuse. They are often diverted from the healthcare system to be shared or sold illegally and consumed in high-risk volumes. In 2020, prescription opioids accounted for 24% of all fatal opioid-related overdoses (Centers for Disease Control and Prevention 2021). In addition, prescription opioids commonly serve as a gateway to illicit opioids; in a nationally

representative survey of heroin users, 80% of respondents reported using prescription opioids prior to heroin (Lankenau et al. 2012).

To better monitor the prescribing and dispensing of opioids, states have created electronic databases known as prescription drug monitoring programs (PDMPs), which collect data on dispensed prescriptions of controlled substances statewide. By providing prescribers and dispensers with full visibility into their patient’s prescription history, PDMPs are intended to help reduce or prevent “doctor shopping” and “pharmacy shopping,” behaviors in which patients request prescriptions from multiple prescribers or dispensaries simultaneously. As of 2022, all 50 states operate PDMPs, and most licensed clinicians are required to check patient records in the PDMP before prescribing or dispensing controlled substances (PDMP Training and Technical Assistance Center 2022).

Over the last decade, PDMP data has also been used to evaluate effects on clinical decision-making, identify public health trends, and help plan for community-level interventions (Substance Abuse and Mental Health Services Administration 2017; Schell et al. 2022). However, the potential of this rich data source for detecting high-risk and potentially illicit behavior among opioid prescribers and dispensers remains largely unexamined. Because clinicians are largely responsible for determining the recipients of legally prescribed opioids, down to the specific drug, duration, and dosage, and are afforded wide discretion while doing so, the individual behaviors of prescribers and dispensers can significantly affect patients’ exposure to and supply of opioids. In some cases, most prominently illustrated by Purdue Pharma’s targeted marketing of OxyContin to physicians and the subsequent lawsuits and settlement, unsafe prescribing can lead to or exacerbate incidents of opioid abuse and addiction. Furthermore, given their essentially unfettered access to highly addictive and in-demand substances, some clinicians have participated in illicit monetary schemes to divert prescription opioids to illegal channels, a process known as drug diversion. By increasing the supply of opioids sold on the “black market,” drug diversion further increases the risk of adverse opioid-related public health outcomes. For these reasons, we believe it is imperative for state governments to develop systems that proactively monitor opioid prescribers and dispensers for unsafe and potentially illicit prescribing

behaviors. With improved monitoring systems, states may be able to more efficiently detect the most pressing high-risk behavior by opioid prescribers and dispensers, which could ultimately reduce opioid addiction and abuse.

To the extent that this topic has been explored in the ML field, it has generally been represented as a point-based anomaly detection problem: researchers seek to identify pharmacies and patients with individually anomalous opioid prescriptions. However, this approach fails to consider several key dynamics of illicit behavior; namely that individuals are more likely to commit crime if they are surrounded by others who do so, that individuals often commit crime in pairs or groups, and that illicit actors may not appear suspicious unless they are linked together (Akoglu, Tong, and Koutra 2015). In the context of prescription opioids, both the prescriber (physician or nurse practitioner) and dispenser (pharmacist or pharmacy technician) must sign off in order for patients to access these substances, making this domain an ideal candidate for a network-based approach. From the perspective of state regulatory and law enforcement agencies, which operate with a strong inclination to take on cases with the greatest potential impact, a network-based approach is a more efficient and pragmatic use of public resources. Indeed, many governments have already demonstrated a motivation to investigate and prosecute networks of illicit opioid activity (New York State Attorney General 2018; Department of Justice 2019). Therefore, rather than detect individual actors, this paper seeks to detect connected networks of opioid prescribers and dispensers engaging in high-risk and potentially illicit behavior. We believe that this approach will not only yield more prescribers and dispensers of interest, but that it is a more effective and impactful strategy for public agencies seeking to identify and disrupt these networks.

For all of its potential impact, detecting connected networks of anomalous opioid prescribing poses many challenges. Substantial variation exists in the prescription data among patients, pharmacies, and prescribers; factors such as location, patient medical condition, medical specialty of the prescriber, and size and organizational structure of the pharmacy all contribute. Without additional data sources to provide more context, such as patient electronic medical records or a clinician's profile, it is often difficult to distinguish between medically necessary and unnecessary prescriptions. Additionally, the sheer volume of prescription data for any state presents a computational challenge, especially when transforming data into directed and attributed graphs.

To address these challenges, we extend the Non-Parametric Heterogeneous Graph Scan (NPHGS), an algorithm originally developed for event detection (Chen and Neill 2014). Consistent with Chen and Neill's approach, we model the prescription data as a heterogeneous graph. Each graph node represents either a prescriber or dispenser, contains ten distinct "red flag" attributes, and includes edges to nodes of the opposite type based on the volume of prescriptions between the two providers. We then generate empirical p-values to rank the anomalousness of each node and use a nonparametric scan statistic to identify the most anomalous connected subgraphs of opioid prescribers and dispensers in the data. However, we extend Chen and Neill's original work

in two crucial ways: 1) we integrate high-performing point-based anomaly detection algorithms into the process of ranking each node, and 2) we reduce the graph's density in order to find more meaningful patterns in the data. We subsequently describe the detected clusters, both in aggregate as well as through two in-depth case studies that visually and qualitatively illustrate the utility of this approach.

The main contributions of our paper are as follows:

- **Novel application of PDMP data to improve opioid-related public health outcomes.** To our knowledge, this is the first work to address opioid abuse and addiction by focusing on anomalous patterns among both opioid prescribers and dispensers in prescription drug data.
- **Novel representation and analysis of PDMP data as a heterogeneous graph.** To our knowledge, this is the first work that takes a network-based approach to addressing opioid abuse and addiction by transforming prescription drug data into a heterogeneous, directed and attributed graph, which enables the use of nonparametric scan statistics for scalable anomaly detection.
- **Flexible and modular integration of anomaly detection into NPHGS.** We extend Chen and Neill's original work by incorporating a flexible mechanism to generate the first stage of empirical p-values for nodes in the graph. Rather than use a single method to generate empirical p-values, we test a variety of distinct anomaly detection methods, and demonstrate improved detection performance as compared to the original NPHGS approach.
- **Generating sparser graphs with more meaningful connections.** We implement a threshold for directed edges among nodes in the graph, based on a parameter of the nonparametric scan. This reduces the density of the graph by eliminating weaker connections, enabling NPHGS to detect more meaningful relationships in the data.
- **Comprehensive evaluation of proposed techniques.** We compare our technique to three competing methods, including two well-known graph-based anomaly detection algorithms. Our results show that our proposed technique produces clusters that are significantly more anomalous than those generated by existing methods.
- **Development of case studies from select subgraphs.** We generate case studies that visually and qualitatively illustrate the utility of our approach, and serve as a model for how our findings could be effectively presented to and within state regulatory and law enforcement agencies.

Related Work

While anomaly detection methods have uncovered illicit behavior across a wide range of problem domains, including money laundering, identity theft, and cyberattacks, only one published study has applied these techniques to detecting anomalous patterns of opioid prescriptions. Hu et al. (2015) proposed a framework to identify anomalous behavior among 8,000 Australian patients who were dispensed fentanyl patches, and created three probabilistic models to rank patients' degree of anomalousness based on variables

including age, gender, number of prescribers, and days between prescriptions. An unpublished paper by Bertsimas, Fazel-Zarandi, and Ivanhoe (2020) proposed supervised, unsupervised, and semi-supervised methods for detecting suspicious patterns of oxycodone and hydrocodone shipments to pharmacies across the U.S. However, neither of these papers examine the problem from a network perspective.

Chen and Neill (2014) originally proposed the Non-Parametric Heterogeneous Graph Scan (NPHGS), a non-parametric graph-based algorithm which we extend and apply here, for detecting events on social media. They found that NPHGS could accurately forecast civil unrest events and achieve early detection of rare disease outbreaks, substantially outperforming competing methods in both cases, but did not consider applications to the opioid crisis.

Data

Our dataset consists of ~ 15 million prescriptions for controlled substances dispensed in the state of Kansas between 2014-2015. The data was provided by the Kansas Board of Pharmacy, which oversees K-TRACS, the state’s prescription drug monitoring program first established by law in 2008. K-TRACS became operational in 2011, and requires all Kansas-based pharmacies to report their dispensing of controlled substances (Kansas State Board of Pharmacy 2020).

Each record contains detailed information about the prescription itself, including the dates it was written and filled, the number of days’ supply, the brand and generic names of the drug as well as its National Drug Code (NDC) number, the form of the drug (e.g., tablet, vial), the quantity of units dispensed, the dosage (quantified in morphine milligram equivalent units, or MME, per day), and the method of payment (e.g., cash, insurance, Medicaid). The record additionally indicates if the drug was categorized as an opioid, benzodiazepine, muscle relaxant, or stimulant, if it had short- or long-acting effects, and if the prescription was new or a refill. Each record also includes anonymized identifiers for each patient, prescriber, and dispenser, which allowed us to aggregate prescriptions for a given individual or group of individuals. Additionally, zip code, city, and state are listed for each patient, prescriber, and dispenser in the data.

Methods

We developed a novel extension of NPHGS to discover subgraphs of opioid prescribers and dispensers that were anomalous along multiple dimensions, as described below.

Data Preprocessing

To more accurately capture temporal trends, we split the two years of data provided by K-TRACS into eight quarters. We excluded prescriptions for which neither the patient nor dispenser was located in Kansas. For patients with missing geographical information, we assumed that they resided in the state of Kansas. We used the centroid of each zip code to calculate the approximate distances between patients, providers, and dispensers, since this was the most precise location information contained in the data. Due to

Aggregation Level	Node Attribute
Patient	Average # of prescribers
Patient	Average # of dispensers
Patient	90-day MME patient average
Patient	90-day MME patient maximum
Patient	% patients w/ opioid + benzodiazepine
Patient	# of prescribers 200+ miles away
Prescription	% of prescriptions paid in cash
Prescription	% of opioid prescriptions paid in cash
Prescription	% of opioid prescriptions
Prescription	% of opioid prescriptions >90 MME/day

Table 1: Attributes of prescriber and dispenser nodes.

its antagonist properties, we did not consider buprenorphine as an opioid, but simply treated it as a non-opioid prescription. We also created 90-day MME trajectories for each patient. MME trajectories were calculated by aggregating all prescriptions for each patient across the 90-day period, then calculating the total MME—potentially including multiple, overlapping prescriptions—they were prescribed each day. These trajectories enabled us to create two new features for each patient: the maximum MME found in their trajectory within the 90 days, and the average MME they were prescribed per day.

Data Representation

We modeled the data from each quarter as a heterogeneous graph with both prescriber and dispenser nodes. As indicated in Table 1, each node contained a variety of prescription- and patient-level attributes, for a total of ten attributes per node. The attributes were partially based on the “red-flag” patterns of opioid misuse, abuse, and diversion that have been widely identified in public health literature. For example, paying for prescriptions with cash, visiting prescribers far from the patient’s home, and obtaining prescriptions from multiple prescribers and dispensers are all frequently cited suspicious behaviors (Cepeda et al. 2013). Two additional attributes, the 90-day MME maximum and average per patient, are grounded in the fact that higher opioid dosages are associated with more high-risk and illicit activity (Dowell, Haegerich, and Chou 2016). Other attributes were created because they represented a direct contradiction of established medical guidance. For instance, the Centers for Disease Control (CDC) strongly advises medical professionals against prescribing opioids above 90 MME or prescribing a patient opioids and benzodiazepines concurrently (Fuiden 2017).

Nodes were connected to each other through directed edges. Edges only connected two nodes of different types (e.g., one prescriber and one dispenser), and contained one attribute: the number of prescriptions shared by both nodes, relative to all prescriptions associated with each node individually. To be included as an edge, this proportion needed to reach a given threshold, which will be discussed below.

NPHGS Background

NPHGS models a heterogeneous graph as a “sensor network,” in which each entity senses its local “neighborhood” and reports an empirical p-value measuring its degree of anomalousness (Chen and Neill 2014). Mathematically, for every node in the graph, this entails a two-stage empirical calibration process. First, each attribute d for each node v is assigned an empirical p-value, defined as:

$$p_d(v) = \frac{1}{T} \sum_{t=1 \dots T} \mathbf{1}\{f_d(v^{(t)}) \geq f_d(v)\},$$

where $f_d(v)$ refers to the d -th component of the feature vector $f(v)$, and $\{v^{(t)}\}_{t=1 \dots T}$ is the reference set for node v . Here the reference set includes all other nodes of the same type (prescriber or dispenser). Thus $p_d(v)$ can be interpreted as the proportion of observations in the reference distribution that meet or exceed the value of attribute d for node v .

In the second stage, NPHGS identifies the minimum p-value for each node, and re-calibrates these values as follows:

$$p(v) = \frac{1}{T} \sum_{t=1 \dots T} \mathbf{1}\left\{ \min_d p_d(v^{(t)}) \leq \min_d p_d(v) \right\}.$$

For a given node v , $p(v)$ can be interpreted as the proportion of nodes in the reference distribution with lowest p-values at least as significant as its lowest p-value.

As Chen and Neill note, this two-stage empirical calibration process offers several benefits for heterogeneous graphs: it calibrates all node types so that each $p(v)$ is uniformly distributed on $[0,1]$ under the null hypothesis that node v and its reference set are exchangeable, allows for the consideration of multiple attributes for a single node, and accounts for correlations between the first-stage empirical values and the overall empirical p-value of the node through the second step of the calibration process.

Once the node-level p-values are obtained, NPHGS uses a nonparametric scan statistic to identify subgraphs with a surprisingly high concentration of low p-values. The general form of this nonparametric scan statistic is defined as:

$$F(S) = \max_{\alpha} F_{\alpha}(S) = \max_{\alpha} \phi(\alpha, N_{\alpha}(S), N(S)),$$

where $S \subseteq V$ refers to a subgraph of connected nodes, $N_{\alpha}(S)$ is the number of p-values in S that are significant at level α , and $N(S)$ is the total number of p-values in S . The score $F(S)$ is maximized over all connected subgraphs S using a greedy subgraph expansion search. Rather than using a fixed significance level, α is optimized between 0 and some constant $\alpha_{\max} < 1$, giving NPHGS high detection power for signals consisting of either a small number of extremely significant p-values or a larger number of moderately significant p-values. Chen and Neill use $\alpha_{\max} = 0.15$.

The function ϕ compares the observed number of p-values $N_{\alpha}(S)$ that are significant at level α to the expected number of p-values $\alpha N(S)$ that are significant at level α , given the null hypothesis that p-values are uniformly distributed over $[0,1]$. The nonparametric scan statistic used by NPHGS is the Berk-Jones statistic (Berk and Jones 1979), defined as

$$\phi_{BJ}(\alpha, N_{\alpha}, N) = N \times KL\left(\frac{N_{\alpha}}{N}, \alpha\right),$$

where KL refers to the Kullback-Liebler divergence between the expected and observed numbers of p-values less than α ,

$$KL(a, b) = a \log \frac{a}{b} + (1 - a) \log \frac{1 - a}{1 - b}.$$

The Berk-Jones statistic can be interpreted as the log-likelihood ratio statistic for testing if the empirical p-values in the subgraph follow a uniform distribution.

NPHGS Extensions

In this paper, we propose two novel extensions to the NPHGS framework, with the goal of improving the method’s effectiveness, flexibility, and computational efficiency. First, instead of setting a single value for α_{\max} , we maximize $F(S)$ over four different α_{\max} thresholds: 0.01, 0.05, 0.10, and 0.15. As in Chen and Neill (2014), for each α_{\max} value, the significance level α is determined by maximizing $F_{\alpha}(S)$ over all α values between 0 and α_{\max} . Note that a different graph is created for each distinct α_{\max} value; otherwise, this modification would be unnecessary, since it would be equivalent to running NPHGS for the single largest α_{\max} value. This extension allows for greater flexibility when searching for the most surprising concentration of low (significant) observed p-values among a set of connected nodes.

For a given $\alpha_{\max} \in \{0.01, 0.05, 0.1, 0.15\}$, we use the α_{\max} value as the minimum threshold proportion for creating a directed edge between nodes. For example, with α_{\max} set to 0.05, an edge from a prescriber to a dispenser would be included only if 5% or more of the prescriber’s total prescriptions went to that dispensary. Thus the larger the value of α_{\max} , the higher the proportion of prescriptions needed to surpass the minimum threshold, which results in sparser graphs. This extension reduces the density of the graph by eliminating weaker connections, so that NPHGS identifies more meaningful patterns in the data. In particular, Wang, Neill, and Chen (2022) show that, for highly connected graph structures and large values of the significance threshold α , the nonparametric scan may incorrectly find a large, high-scoring connected subgraph that includes almost all of the significant nodes in the graph even under the null hypothesis H_0 that all nodes’ empirical p-values are uniformly distributed on $[0,1]$. This reduces both detection power and precision unless the scan is recalibrated using a computationally intensive approach. However, applying recent results in percolation theory (Krivelevich 2016), we show in the Technical Appendix that using α_{\max} as the threshold for including an edge guarantees that the graph is sufficiently sparse so that this undesirable result does not occur. Additional theoretical and empirical justification for the use of α_{\max} is provided as well.

Our second novel extension introduces a flexible mechanism to generate the empirical p-values $p(v)$ for each node in the graph. Instead of relying upon a single, fixed method to compute the anomalousness of each node, we recognize that the most appropriate method may differ based on the unique characteristics of the dataset and problem domain. We therefore separate the method that evaluates each node’s degree

of anomalousness from the process of generating empirical p-values. In practice, this separation enables a user to rank nodes using different anomaly detection techniques, generating different respective sets of empirical p-values, and then choose an appropriate method.

In this paper, we test six different methods to generate empirical p-values for each node. The first method (“first min”) follows the original NPHGS framework, in which an empirical p-value is calculated for every attribute of the node, and the minimum attribute p-value is selected and re-calibrated. The second and third methods we test (“second min” and “third min”) follow the same two-stage calibration process, but with one modification: instead of selecting the node’s attribute with the lowest p-value in the first stage, the second- or third-lowest attribute p-value, respectively, is selected and re-calibrated. We hypothesize that using the second- or third-lowest p-value may yield a subgraph with nodes that are anomalous across multiple attributes, rather than just one. For the remaining three methods, we use a separate, point-based anomaly detection algorithm in the first stage to generate outlier scores for each node. Based on the outlier score $o(v)$, with lower scores reflecting more anomalous nodes, we generate empirical p-values for each node by re-calibrating the outlier scores:

$$p(v) = \frac{1}{T} \sum_{t=1 \dots T} \mathbf{1}\{o(v^{(t)}) \leq o(v)\}.$$

We selected three well-known anomaly detection algorithms to generate outlier scores in the first stage: Local Outlier Factor (LOF) (Breunig et al. 2000), One-Class Support Vector Machines (OCSVM) (Schölkopf et al. 1999), and Isolation Forests (IF) (Liu, Ting, and Zhou 2008). These were implemented using the open-source sklearn package.

For each of the six methods, we selected the top five prescribers and five dispensers with the lowest node-level p-values as seed nodes. Starting with these nodes, we implemented NPHGS over the four α_{\max} values for each quarter.

Experiments

To evaluate the six methods used to generate empirical p-values for NPHGS, and to compare NPHGS to competing methods in the literature, we implemented five-fold cross validation on the data. For each split, the training set (comprising 80% of the data) was used to implement NPHGS and competing methods, and to generate the resulting anomalous clusters. The testing set (comprising 20% of the data) was used to identify patients that were associated with the prescribers and dispensers from the detected clusters. The patients identified from all five testing sets were combined, and the resulting patient populations were compared for each of the six node-level p-value selection methods, the three competing methods described below, and the “baseline” population of all patients contained in the data for that quarter.

Competing Methods

To evaluate the performance of NPHGS relative to other methods, we applied three different graph-based anomaly detection algorithms to the data: 1) the SUBDUE algorithm,

2) spectral clustering, and 3) ranked pairs of prescribers and dispensers. For each method, we implemented the same five-fold cross-validation approach as discussed above, and compared the resulting patient populations identified in the testing data. The methodology and our specific implementation of each algorithm is discussed briefly below.

The SUBDUE algorithm uses the minimum description length (MDL) principle to identify patterns that minimize the description length of a directed graph (Eberle and Holder 2007). SUBDUE is a well-known method for graph-based anomaly detection, having been successfully applied to many problem domains including counter-terrorism, network communications, earthquakes, aviation safety, and biomedicine (Holder 2021). We implemented SUBDUE on our data by executing its publicly available source code. Due to computational challenges with the algorithm, it was necessary to reduce the number of edges in our graph prior to implementation. We converted the directed graph to an undirected graph, and only added edges between nodes if they represented more than 1% of each node’s total prescriptions.

Spectral clustering is a popular clustering method that is found in the scikit-learn library and can be applied to graph-based anomaly detection problems. The method requires users to specify the number of clusters, which we set to 100 after testing a range of values. Similar to NPHGS, we calculated scores for each cluster using the Berk-Jones statistic and an α_{\max} value of 0.01. The top-scoring clusters were iteratively selected based on a threshold that matched the number of patients selected in the testing data for NPHGS for that quarter and fold, so that similar numbers of patients could be compared.

In our third comparison algorithm, we simply paired sets of prescribers and dispensers that were associated with at least nine prescriptions each (roughly one prescription per week) and determined the average MME per patient for the pair’s shared prescriptions. The prescriber-dispenser pairs were then ranked in descending order, and a threshold was chosen as for spectral clustering above.

Results

As shown in Table 2, the original NPHGS method—using the first minimum p-value—is outperformed by at least one other method in every single attribute. In a few categories, such as average and maximum patient MME, it ranks only above LOF. In contrast, NPHGS using Isolation Forests significantly outperforms the other variants, identifying clusters with the most anomalous values for 7 attributes. We therefore opted to use Isolation Forests as the first stage before generating node-level p-values, both for comparing the performance of NPHGS to other graph-based methods, and for implementing NPHGS on the full data for our case studies.

Table 3 compares NPHGS to the three competing methods. We observe that NPHGS identifies clusters with more anomalous values for all attributes used to compare the patient populations (using the held-out test sets). For some metrics, including the MME patient average and maximum, the percentage of prescriptions paid for in cash, and the percentage of patients with concurrent opioid and benzo prescriptions, the discovered clusters for NPHGS have values

p-value method	1st min	2nd min	3rd min	LOF	OC SVM	IF
# prescribers	1.65	1.60	1.73	1.54	1.81	1.84
# dispensers	1.44	1.42	1.47	1.40	1.46	1.51
MME avg	56	71	85	39	76	91
MME max	159	202	234	132	228	255
% opi. benzo	0.27	0.27	0.37	0.19	0.38	0.41
% Rx cash	0.47	0.54	0.35	0.18	0.35	0.39
% opi. cash	0.39	0.46	0.33	0.18	0.33	0.37
% opi. Rx	0.62	0.55	0.65	0.64	0.65	0.65
% 90+ MME	0.29	0.35	0.37	0.24	0.35	0.39

Table 2: Comparison of detected subgraphs (patient populations) among p-value selection methods for NPHGS.

detection method	base line	NPHGS	SUB-DUE	Spec. Clust.	Paired MME
# prescribers	1.27	1.87	1.41	1.54	1.53
# dispensers	1.14	1.51	1.32	1.30	1.32
MME avg	19	94	21	26	49
MME max	68	261	72	89	134
% opi. benzo	0.11	0.43	0.14	0.16	0.23
% Rx cash	0.09	0.41	0.22	0.10	0.07
% opi. cash	0.10	0.39	0.26	0.10	0.07
% opi. Rx	0.52	0.66	0.57	0.56	0.62
% 90+ MME	0.14	0.39	0.09	0.17	0.25

Table 3: Comparison of detected subgraphs (patient populations) between NPHGS and competing methods.

that exceed the competing methods by double or more. It is particularly notable that NPHGS outperforms SUBDUE, a well-known method for graph-based anomaly detection that has been successfully applied to multiple problem domains.

Case Studies

Following our evaluation study, we implemented NPHGS (with Isolation Forests) on the full dataset for each of the eight quarters of data. Of the 80 seed nodes (10 seed nodes per quarter), 41 expanded to include one or more neighbors. Of the clusters that expanded, there were 30 unique combinations of nodes; several clusters appeared more than once throughout the eight quarters of data. The clusters consisted of 82 distinct nodes, with 64 nodes belonging to prescribers and 18 nodes belonging to dispensers. 49 of the 82 nodes (60%) appeared in more than one cluster. The clusters had an average edge density of 0.82. We will discuss two of the detected clusters in detail as case studies below.

The cluster corresponding to Case Study A first appeared in the third quarter of 2014, and consists of four prescribers and two dispensers. The prescribers and dispensers in this cluster share a total of 498 prescriptions and 101 patients. In the visualization of this cluster, shown in Figure 1, the black and white circles indicate prescriber and dispenser nodes respectively. For this visualization, nodes are connected by edges if the providers have at least one patient in common, with numbers on each edge representing the number of shared patients.

As shown in Table 4, the four prescribers individually demonstrate extremely high-risk prescribing behavior, with

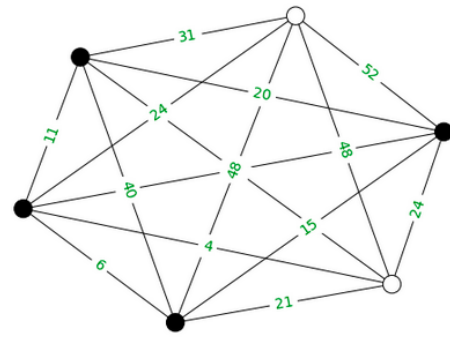


Figure 1: Visualization of prescribers (black nodes) and dispensers (white nodes) for Case Study A.

	Max MME per patient	Avg MME per patient	% Rx cash	% Rx opioid
Baseline	68	19	9%	51%
Node P1	323	86	65%	65%
Node P2	390	103	21%	77%
Node P3	210	76	96%	80%
Node P4	151	61	94%	73%

Table 4: Notable attributes of the four prescriber nodes for the detected cluster in Case Study A, for prescriptions filled at the two dispensaries shown in Figure 1.

MME dosages that are 2-4x the average, and up to 10x more prescriptions paid for in cash than the general population. The two dispensaries show similar patterns of individually high-risk behavior, as they both filled prescriptions with roughly double the typical MME dosage and had 99% of prescriptions paid for in cash. Additionally, the two dispensaries are not located near the four prescribers, three of whom share a zip code; they are not even located in the state of Kansas. One of the dispensaries is based in Boulder, Colorado; the other is in suburban New Jersey.

Because the providers in this cluster are prescribing and dispensing high dosages of opioids, we must also consider the possibility that they are providing inpatient or oncology care rather than engaging in high-risk prescribing behaviors. For this cluster, however, the possibility is extremely unlikely. Prescribers and dispensers that provide inpatient care generally share the same identifier (associated with the given hospital), and all six providers in this cluster have distinct identifiers; we can also assume that prescribers providing inpatient care would not order prescriptions out of state. Further, we can assume that none of the prescribers specialize in oncology or hospice treatment, as they all prescribe a wide range of controlled substances, including benzodiazepines (e.g., Valium, Xanax), sleep medication, and amphetamines.

While the behavior of each provider in the cluster can be considered individually anomalous, a closer examination uncovers strong connections between the six providers. 100% of all Kansas prescriptions filled by the cluster's two dispensaries come from the four prescribers in the cluster. Likewise, the four prescribers direct 19%, 50%, 95%, and

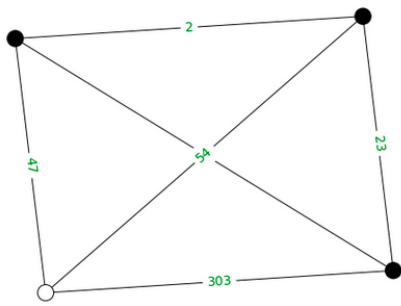


Figure 2: Visualization of prescribers (black nodes) and dispensers (white nodes) for Case Study B.

96% of prescriptions toward the cluster’s dispensaries, respectively. The prescribers also demonstrate potential coordination in terms of their choice of dispenser. During the first two months of the quarter (July and August), all four prescribers sent their prescriptions to only one of the dispensaries, Pharmacy A. However, on September 3, 2014, prescriptions from all four prescribers suddenly started to be filled at the other dispensary in the cluster, Pharmacy B.

Given the individually anomalous behavior of each provider, the strong connections between them, and the combined impact of their behavior (498 prescriptions dispensed in three months), this cluster would be of strong interest to state law enforcement and regulatory agencies. Yet without a network-based approach, the providers in this cluster would not necessarily be detected. Currently, the most common approach taken by state agencies to detect potentially illicit behavior is to pick a specific suspicious behavior and investigate the top handful of offenders. In this case, if state agencies were to examine the top prescribers for the ten “red flag” attributes examined in this work, none of the four prescribers would show up in the top 10. For eight of the ten attributes, none of the four prescribers would even show up in the top 100. However, if the prescriptions of the four prescribers in this cluster were combined and examined as a single entity, that entity would score in the 95th percentile or higher for seven of the ten “red flag” attributes. This demonstrates the power of the network-based approach to identify patterns of anomalous prescribing behavior that may not be obvious from any single prescriber.

This cluster only appears once in the dataset, but the pattern reappears in the 2nd, 3rd, and 4th quarters of 2015. Although the prescriber and dispenser nodes have different identifiers in these clusters, they share 70-74% of the same patients, and are located in the same zip codes as the original cluster. It is a strength of NPHGS that it is able to identify highly anomalous connected networks even when the individual actors may have attempted to create new identities.

Case Study B, shown in Figure 2, consists of three prescribers and one dispensary, and appeared in the first quarter of 2015. The prescribers and dispensers in this cluster share a total of 2,063 prescriptions and 363 patients. The cluster has a density of 1, indicating that all nodes have at least one patient in common. Both the prescribers and dispenser exhibit individually high-risk behavior. Specifically, as shown

	Max MME per patient	Avg MME per patient	% Rx cash	opioid + benzo
Baseline	68	19	9%	11%
Node P1	221	34	100%	63%
Node P2	170	30	100%	25%
Node P3	138	62	100%	55%

Table 5: Notable attributes of the three prescriber nodes for the detected cluster in Case Study B, for prescriptions filled at the dispensary shown in Figure 2.

in Table 5, 100% of prescriptions in the cluster were paid for in cash (compared to 9% overall), and the percentage of patients with concurrent opioid and benzo prescriptions was 2-6x higher than that of the general population. The MME patient maximum for both the prescribers and the dispenser was also more than double that of the general population. Additionally, the prescribers and dispenser were highly interconnected: 80% of all prescriptions filled by the dispenser were filled by prescribers in the cluster, and the prescribers sent 90% of their collective prescriptions to the dispensary.

Discussion

This project presented a novel approach to an under-explored question in the prescription drug literature: how to identify high-risk and potentially illicit patterns of behavior among opioid prescribers and dispensers. Rather than identifying individually anomalous providers, we modeled the prescription data as a directed graph consisting of both prescriber and dispenser nodes, and used a non-parametric scan statistic to detect connected networks of opioid prescribers and dispensers engaging in high-risk patterns of behavior. To our knowledge, this is the first work to approach the problem of monitoring prescribers and dispensers from a heterogeneous graph perspective. We showed that NPHGS significantly outperformed competing graph-based methods, including a well-known anomaly detection algorithm that has been successfully applied to many domains. As case studies for the potential utility of the method, we described two clusters generated by NPHGS, providing both qualitative context and quantitative data about each cluster.

Methodologically, the project extended NPHGS in two important ways: a method to integrate anomaly detection algorithms into the two-stage empirical p-value calibration process, which generated more effective node-level p-value scores, and a method to reduce the density of the graphs by instituting a minimum threshold for inclusion of graph edges. Reducing the density of the graph enabled NPHGS to find more meaningful patterns in the data during the iterative graph expansion. These two extensions create opportunities for more flexibility and customization when implementing NPHGS on different types and volumes of data.

Our future work will apply this approach to more states’ data, to ensure that the results from Kansas are generalizable to other states. We also plan to develop an open-source tool that could simplify implementation and accelerate adoption of NPHGS for public health and regulatory agencies.

Acknowledgments

We wish to thank Fan Xiong (Kansas Department of Health) and Reyne Kenton (Kansas Board of Pharmacy) for providing data and domain expertise for this project. K.R. was supported by the National Science Foundation Graduate Research Fellowship Program, Grant No. DGE-1839302.

References

- Akoglu, L.; Tong, H.; and Koutra, D. 2015. Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, 29(3): 626–688.
- Berk, R. H.; and Jones, D. H. 1979. Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 47: 47–59.
- Bertsimas, D.; Fazel-Zarandi, M.; and Ivanhoe, J. 2020. Early Detection of Opioid Over-Procurement: A Machine Learning Approach. <https://dbertsim.mit.edu/pdfs/papers/2020-fazel-zarandi-early-detection-of-opioid-overprocurement.pdf>. Accessed: 2022-10-20.
- Breunig, M. M.; Kriegel, H.-P.; Ng, R. T.; and Sander, J. 2000. LOF: identifying density-based local outliers. *ACM SIGMOD Record*, 29(2): 93–104.
- Centers for Disease Control and Prevention. 2021. Prescription Opioid Overdose Death Maps. <https://www.cdc.gov/drugoverdose/deaths/prescription/maps.html>. Accessed: 2023-04-20.
- Cepeda, M. S.; Fife, D.; Chow, W.; Mastrogiovanni, G.; and Henderson, S. C. 2013. Opioid shopping behavior: how often, how soon, which drugs, and what payment method. *Journal of Clinical Pharmacology*, 53: 112–117.
- Chen, F.; and Neill, D. B. 2014. Non-parametric scan statistics for event detection and forecasting in heterogeneous social media graphs. In *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*, 1166–1175. Association for Computing Machinery.
- Department of Justice. 2019. Appalachian Regional Prescription Opioid (ARPO) Strike Force Takedown Results in Charges Against 60 Individuals, Including 53 Medical Professionals. <https://www.justice.gov/opa/pr/appalachian-regional-prescription-opioid-arpo-strike-force-takedown-results-charges-against>. Accessed: 2023-04-20.
- Dowell, D.; Haegerich, T. M.; and Chou, R. 2016. CDC Guideline for Prescribing Opioids for Chronic Pain—United States. *MMWR Recommended Report*, 65(1): 1–49.
- Eberle, W.; and Holder, L. 2007. Anomaly detection in data represented as graphs. *Intelligent Data Analysis*, 11(6): 663–689.
- Fuiden, J. 2017. Don't ignore these opioid abuse red flags. [https://www.pharmacytimes.com/view/opioid-red-flags-for-consideration-.](https://www.pharmacytimes.com/view/opioid-red-flags-for-consideration-) Accessed: 2023-04-20.
- Holder, L. 2021. Subdue Graph Miner. <https://github.com/holderlb/Subdue>. Accessed: 2023-04-20.
- Hu, X.; Gallagher, M. R.; Loveday, W.; Connor, J. P.; and Wiles, J. 2015. Detecting anomalies in controlled drug prescription data using probabilistic models. In *Proc. Austral. Conf. on Artificial Life and Computational Intelligence*.
- Kansas State Board of Pharmacy. 2020. An Overview of K-TRACS: Kansas Prescription Drug Monitoring Program. <https://www.youtube.com/watch?v=UTm2OsV323M>. Accessed: 2023-04-20.
- Krivelevich, M. 2016. The phase transition in site percolation on pseudo-random graphs. *Electronic Journal of Combinatorics*, 23(1): P1.12.
- Lankenau, S. E.; Teti, M.; Silva, K.; Jackson Bloom, J.; Harocopos, A.; and Treese, M. 2012. Initiation into prescription opioid misuse amongst young injection drug users. *International Journal of Drug Policy*, 23(1): 37–44.
- Liu, F. T.; Ting, K. M.; and Zhou, Z.-H. 2008. Isolation forest. In *Proc. 8th IEEE Intl. Conf. on Data Mining*. IEEE Press.
- National Institutes of Health. 2022. Overdose Death Rates. <https://nida.nih.gov/drug-topics/trends-statistics/overdose-death-rates>. Accessed: 2023-04-20.
- New York State Attorney General. 2018. A.G. Underwood Announces Takedown of 12-Person Hudson Valley Narcotics Distribution Ring in Operation Candy Crush. <https://ag.ny.gov/press-release/2018/ag-underwood-announces-takedown-12-person-hudson-valley-narcotics-distribution>. Accessed: 2023-04-20.
- PDMP Training and Technical Assistance Center. 2022. State PDMP Profiles and Contacts. <https://www.pdmpassist.org/State>. Accessed: 2023-04-20.
- Schell, R. C.; Allen, B.; Goedel, W. C.; Hallowell, B. D.; Scagos, R.; Li, Y.; Krieger, M. S.; Neill, D. B.; Marshall, B.; Cerda, M.; and Ahern, J. 2022. Identifying predictors of opioid overdose death at a neighborhood level with machine learning. *American Journal of Epidemiology*, 191(3): 526–533.
- Schölkopf, B.; Williamson, R.; Smola, A.; Shawe-Taylor, J.; and Platt, J. 1999. Support vector method for novelty detection. In *Proc. 12th Intl. Conf. on Neural Information Processing Systems*, 582–588. MIT Press.
- Substance Abuse and Mental Health Services Administration. 2017. Prescription Drug Monitoring Programs: A Guide for Healthcare Providers. <https://store.samhsa.gov/sites/default/files/d7/priv/sma16-4997.pdf>. Accessed: 2023-04-20.
- Wang, C.; Neill, D. B.; and Chen, F. 2022. Calibrated non-parametric scan statistics for anomalous pattern detection in graphs. In *Proc. 36th AAAI Conference on Artificial Intelligence*, 4201–4209. AAAI.