"Nip it in the Bud!"

Managing the Opioid Crisis: Supply Chain Response to Anomalous Buyer Behavior

Annie L. Shi

Doctoral Student in Marketing, Washington University in St. Louis

P.B. (Seethu) Seetharaman †

W. Patrick McGinnis Professor of Marketing, Washington University in St. Louis

Anthony Sardella

Senior Research Advisor, Center for Analytics and Business Insights (CABI),
Washington University in St. Louis

Michael Wall

Professor of Practice, Washington University in St. Louis

Chenthuran Abeyakaran

Master Student in Data Analytics and Statistics, Washington University in St. Louis

December 1, 2021

† Corresponding Author: Seethu Seetharaman, email: seethu@wustl.edu; Ph: 314-935-3574. The authors thank the Bellwether Foundation for funding this research with their generous financial support. The authors also thank the Olin-Brookings Commission, an advisory group comprised of government and industry stakeholders, for their advice on the policy implications of our research.

"Nip it in the Bud!"

Managing the Opioid Crisis: Supply Chain Response to Anomalous Buyer Behavior

Abstract

Over the past three decades, the opioid epidemic has wreaked havoc upon thousands of communities across the US. In this study, we provide a supply-chain perspective to manage the ongoing opioid crisis. Using the ARCOS database -- which tracks opioid drug shipments across the entire supply chain in the US, spanning the period 2006 to 2012 -- we employ a novel anomaly detection algorithm to detect suspicious buyer activity. Our algorithm is non-intrusive on patient privacy, in that it does not rely on prescription-level data (from drug retailers or physicians). Using a random sample of 50,000 drug retailers, plus a set of 188 retail buyers who are *labeled*, using observed convictions from the Drug Enforcement Administration (DEA) website, as "suspicious", we train our anomaly detection algorithm to detect suspicious retail buyers based on their historical opioid buying patterns. Our anomaly detection algorithm, which is built on a training set of 25,000 drug retailers, yields an F-1 score of 61 % (with a precision of 100 %) in terms of correctly detecting suspicious retail buyers in a validation set of 25,188 drug retailers (which includes the 188 convicted retail buyers). While we employ a total of 40 input variables to train the anomaly detection algorithm, it ultimately relies upon only 7 input variables to achieve its impressive predictive accuracy. By applying our algorithm on real-time opioid shipments data as and when orders are placed by drug retailers around the country, manufacturers and distributors within the supply chain, as well as the DEA, can flag those that are tagged as suspicious for further investigation. By halting large shipments of opioids through early identification of suspicious orders placed by either (willfully or otherwise) negligent, or outright criminal, activities, these dangerous drugs can be prevented from reaching vulnerable communities, thus saving lives.

Keywords: Opioid Crisis, Opioid Diversion, Suspicious Shipments, ARCOS Data, Anomaly Detection.

Introduction

While new vaccines offer hope in the battle against the COVID-19 pandemic, the destructive effect of the opioid epidemic continues to be felt across thousands of vulnerable communities across the US, affecting the lives of millions of US families and tens of millions of US citizens. According to the Center for Disease Control (CDC), from 1999 – 2019, nearly 500,000 Americans have died from an overdose involving any opioid, including prescription and illicit opioids (see Figure 1 for a chronological timeline of opioid overdose deaths in the US). The number of drug overdose deaths in 2019 was 70,630 of which 70 % involved an opioid. The US contains 4.4 % of the world's population, but consumes 30.2 % of the world's opioids, and accounts for 27 % of the world's opioid overdose deaths. Aggressive marketing practices have been implicated as being complicit in the outbreak of drug abuse (Meier 2018). Recently, the Drug Enforcement Administration (DEA) launched an aggressive effort to monitor physician prescriptions and pharmacist fulfillment. In 2019, by order of the sixth circuit court of appeals, a database called ARCOS, established by the Drug Enforcement Administration (DEA) to track the manufacturing and distribution of all prescription drugs, was made available to the public spanning the period 2006-2012. In this research, we aim to address the opioid scourge in our communities by focusing on the question of whether suspicious orders of opioid drugs could be stopped at the source – even before they get in to the hands of those at risk. We employ a novel anomaly detection algorithm on the ARCOS database to detect suspicious activity to tackle the pernicious issue of illegal drug diversion by retailers.

[Insert Figure 1]

Our algorithm is non-intrusive on patient privacy, in that it does not rely on personally identifiable, prescription-level data from drug retailers or physicians.[1] Instead, we analyze longitudinal buying patterns of retail buyers, in terms of how much of a given opioid drug that each retail drug buyer orders from a given drug wholesaler at any given point in time. Using a random sample of 50,000 drug retailers, plus a set of

[1] Prescription drug monitoring programs (PDMPs), in contrast, are state-run electronic databases that track opioid prescriptions. They help health providers identify specific patients at risk of opioid misuse.

188 retail buyers who are *labeled*, using observed convictions from the DEA website, as "suspicious", we train our anomaly detection algorithm to detect suspicious buyers based on their historical opioid buying patterns. Our anomaly detection algorithm, which is built on a training set of 25,000 drug retailers, yields an F-1 score of 61 % (with a precision of 100 %) in terms of correctly detecting suspicious buyers in a validation set of 25,188 drug retailers (which includes the 188 convicted buyers). While we employ a total of 40 input variables to train the anomaly detection algorithm, it ultimately relies upon only 7 input variables to achieve its impressive predictive accuracy. Our algorithm outperforms other, more traditionally employed, AI algorithms on anomaly detection.

By applying our algorithm on real-time opioid shipments data as and when orders are placed by drug retailers around the country, manufacturers and distributors within the supply chain, as well as the DEA, can flag those that are tagged as suspicious for further investigation. By halting large shipments of opioids through early identification of suspicious orders placed by either (willfully or otherwise) negligent, or outright criminal, activities, these dangerous drugs can be prevented from easily reaching vulnerable communities, thus saving lives.

The rest of the paper is organized as follows. Section 2 presents a historical background on the opioid crisis in the US. In section 3, we offer a brief literature review. Section 4 explains the ARCOS data. In section 5, we explicate our novel anomaly detection algorithm, in relation to four other popular anomaly detection algorithms from the AI literature, as a predictive tool to detect suspicious buyers (drug retailers) based on their historical buying patterns. In section 6, we document the empirical performance of our predictive algorithm in terms of detecting suspicious buyers. Section 7 concludes with policy implications.

Background on the Opioid Crisis

During the nineteenth century, morphine was isolated from opium and used to treat battlefield injuries. Morphine also came to be used to treat joint pain, menstrual cramps, etc. Eventually, morphine addiction rates spiraled out of control and its use was prohibited. A new, ostensibly less addictive, opioid

was born - heroin. History repeated itself when heroin addiction then became a serious concern. The US government passed the Harrison Act in 1914, which restricted the use of heroin and other opioids. Over the next several decades, the US successfully weaned itself out of opioid use. By the 1980s, with the advent of the hospice movement, there was a feeling that the elderly, cancer patients, people seriously injured in accidents, etc. were being horrendously under-treated by not being prescribed opioids for pain. The medical community once again turned to opioids. In the early 1990s, a group of influential academics pushed and advocated the use of strong opioids to treat even ordinary kinds of pain. The watershed moment was Purdue Pharma's introduction in 1996 of OxyContin[2], which was heralded as a "wonder" drug for pain treatment.[3] OxyContin set the stage for the opioid overdose epidemic. Purdue Pharma argued that millions of people were suffering unnecessarily because doctors had exaggerated public fears about the addictive potential of prescription painkillers. In order to distance the drug from the bad connotations associated with the term "narcotic" the word "opioid" was coined as part of OxyContin's branding effort. Purdue Pharma launched an aggressive marketing campaign that targeted doctors to influence their prescribing behavior for OxyContin. Data analysis was used to identify which doctors had the most chronic pain patients, which doctors were more likely to prescribe pain medication even for mild pain etc. Purdue dispatched large sales forces specifically to these doctors.

Over the next decade, the Rust Belt and Appalachian states (PA, OH, WV) became most badly hit by opioid overdose deaths primarily because of OxyContin use. Huntington, WV, came to be called the overdose capital of America. Many of these regions were blue-collar, with many workers getting injured in their line of work (construction, oil refining, etc.) and, therefore, being prescribed opioids to treat their pain. The oversupply of prescription painkillers created increasing demand which led to unethical doctors prescribing opioids for cash payments (see the Netflix documentary, "The Pharmacist" which features a pediatrician in Louisiana, whose license was eventually revoked, who wrote 183,000 pain prescriptions for

[2] Purdue Pharma's main source of revenue, MSContin, a morphine pill for cancer patients, saw its patent expiring which led to the development of OxyContin.
[3] Unlike Motrin, which was an over-the-counter (OTC) painkiller, OxyContin was a prescription painkiller. It replaced Vicodin as the most abused prescription painkiller in the US.

adults over a year). This led to more overdose deaths. OxyContin was fundamentally different from older, classic opioid pain prescription medications, such as Percocet and Vicodin. While older drugs had a dosage range per pill of 2.5 – 10 mg of Oxycodone (an opioid 50% stronger than morphine), OxyContin had a dosage range of 10 – 80 mg of Oxycodone. The extremely potent Oxycodone, in addition to blocking pain signals in the nerves, bound to opioid receptors in the reward parts of the brain, thus elevating dopamine levels and giving users an immediate rush of energy and a drug-induced high. Taking Oxycodone daily was like taking heroin daily and, unsurprisingly, led to severe addiction. At its peak, OxyContin sales exceeded Viagra sales. Purdue's profits hit $3b by 2010 before they paid fines of $700m after being convicted of misrepresenting the opioid abuse liability. By 2012, opioid prescriptions covered 80 % of Americans.

Opioids are currently regulated under Schedule 2 of the Controlled Substances Act (CSA), which represents substances deemed to have a high potential for abuse which may lead to severe psychological or physical dependence. Also called narcotics, opioids include hydromorphone, methadone, meperidine, oxycodone, fentanyl, morphine, opium, codeine, hydrocodone etc. From 2006 to 2014, 100b prescription hydrocodone and oxycodone pills were distributed in the US. The opioid epidemic has cost the US $1t from 2001 to 2017 (Altarum 2021).

With the rising backlash against opioid manufacturers in the wake of the opioid epidemic, two opioid manufacturers – Mallinckrodt, Purdue Pharma – have filed for bankruptcy. The US government recently reached a $26b settlement with three opioid distributors – Cardinal Health, McKesson, AmeriSource Bergen – and an opioid manufacturer – Johnson & Johnson (J&J) – in August 2021. The three opioid distributors also agreed to pay $1.18b to NY, and Nassau and Suffolk counties. J&J agreed to pay $230m to settle opioid claims in NY. These lawsuits alleged that distributors failed to flag and halt a rising tide of suspicious orders of pain pills. *Addressing this issue, and enabling distributors to have a predictive system that can be used to flag and halt suspicious orders of opioid drugs, is the central focus of this study.*

Literature Review

A handful of researchers in the management sciences and economics have studied drug abuse. Zaric, Brandeau and Barnett (2000) assess the cost effectiveness of maintenance treatment for heroin addiction, with emphasis on its role in preventing HIV infection. The authors measure the health benefits of methadone maintenance in terms of life years gained. Corman and Mocan (2000) study whether drug abuse leads to an increase in violence and other crimes, as is commonly believed by many policy makers.[4] Using a unique dataset from the Crime Analysis Unit of the NYPD that tracks both monthly crimes and drug-related deaths, the authors find that drug usage has only a small effect on some property crimes.

Liu and Bharadwaj (2020) study the impact of digital platforms in contributing to the ongoing drug overdose epidemic. Relying on the phased rollout of Craigslist as an experimental setup, and applying a difference-in-differences approach on a national panel dataset for all counties in the US from 1997 to 2008, they find a 14.9 % increase in drug abuse treatment admissions, a 5.7 % increase in drug abuse violations, and a 6 % increase in drug overdose deaths after Craigslist's entry. The impacts of Craigslist's entry are found to be larger among women, whites, Asians, and the more educated. Further, the unintended consequences of Craigslist are found to be more likely to accrue in larger, wealthier areas with initially low levels of drug abuse.

Zheng and Alba (2021) argue that since drug addiction has been established as a biological morbidity by neuroscience research, the public now feels increasing sympathy for the drug abuse problem rather than viewing it as reflecting moral failures in self-control. This evolution in societal understanding of drug abuse, therefore, has important marketing implications in terms of how to communicate public policy solutions for the problem to the general public. Puntoni et al. (2021) review the role of AI technologies in improving consumers' lives in very concrete and relevant ways these days. For example, Express Scripts is using wearable devices to monitor whether and when their patients are at risk of drug

[4] There are three testable hypotheses underlying this belief: (1) drug use increases aggression and, therefore, violent crime; (2) drug users turn to crime to finance expenditures on drugs; (3) violence occurs in the drug market because participants cannot rely on contracts and courts to resolve disputes.

overdose so that remedial actions can be taken immediately. Zhang and King (2021) study the effects of physicians' contentious prescribing practices with regard to opioids – specifically, benzodiazepines (a key contributor to the drug epidemic) -- on the strengths of their social ties within physician networks. Bobroske et al. (2021) study the impact of operational interventions on long-term opioid use. Using a nationwide US database of medical and pharmaceutical claims, the authors find that follow-up appointment within 30 days of opioid initiation with a clinician other than the initial prescriber reduces the likelihood of long-term opioid use by 31 %.

In a recent article, Chandy et al. (2021) exhort the marketing discipline to recognize that marketers are an integral part, either as a cause or as a source of a solution, for grave societal problems. The authors discuss the growing importance for today's marketing academics and practitioners to redefine marketing's role in creating a better world. Our research responds to their call for action.

Data

The DEA maintains a database, Automated Reports and Consolidated Ordering System (ARCOS), to track the manufacturing and distribution of all prescription drugs, inclusive of opioids. The database tracks the path of every single pain pill sold in the US – by manufacturers to distributors, and by distributors to pharmacies and practitioner clinics – in every town and city in the United States. HD Media won a year-long battle in 2019 for access to the database from court filings resulting in the release of all activity from the years 2006 to 2012. The Washington Post published part of the data. In Figure 2, we present the distribution channel system for opioids. Manufacturers of opioid drugs include Purdue, Endo, Pfizer, Janssen, Mallinckrodt etc. Our focus is not on opioid shipments from manufacturers to distributors. Instead, *we focus on opioid shipments from distributors to retailers*. (Note: transactions from retailers to patients are not observed in the data).

[INSERT FIGURE 2]

There are a total of 9372 unique *opioid stock keeping units* (SKUs), also referred to as *opioid products*, in the ARCOS data. If one considered the *active process ingredients* (APIs) that are contained within these opioid products, there are 170 unique APIs in the data. In other words, the 9372 unique opioid products are simply different chemical combinations of these 170 unique APIs. If one went down one more level and considered the *drug compounds*, also called *drug products* (DPs), which are contained within the APIs, there are 14 unique DPs in the data. This breakdown from opioid products to APIs to DPs are diagrammatically explained in Figure 3.

In Figure 3, we also present the relative shares of the 14 DPs in terms of their observed sales volumes, represented in *morphine milligram equivalents (MME)*, which are calculated as follows: MME = (Drug Dosage Units) * (Drug Dosage Strength per Unit) * (MME Conversion Factor from Drug Dosage to Morphine Milligrams).[5] Hydrocodone enjoys the largest share (40 %), with Oxycodone coming second (23 %). Out of the 14 DPs, just 4 (Hydrocodone, Oxycodone, Fentanyl, Morphine) account for 84 % of all sales volume in the data.

[INSERT FIGURE 3]

Table 1 lists the top 50 best-selling opioid products during the period of study (2006-2012). One can see that 23 out of the top 50 opioid products contain hydrocodone, while 11 contain oxycodone, as the ingredient drug. This is not surprising since hydrocodone has been, by far, the most prescribed drug product for acute and chronic pain. The infamous OxyContin pill figures thrice in the top 50 list: 40 mg Oxycodone HCl form is ranked 33rd, Oxycodone HCl controlled form is ranked 35th, and 80 mg Oxycodone HCl form is ranked 43rd.

[INSERT TABLE 1]

The ARCOS database contains a total of 277,000 + unique buyers (drug retailers, physician clinics etc.) making over 400 million transactions over the six-year period. We tag a subset of 188 among these

[5] Since different opioid products are sold in different form (e.g., pill versus cap) and dosage strength (e.g., 20 mg versus 40 mg), MME, which represents the effective amount of morphine milligrams contained in the product, represents a standardized unit of sales volume across opioid products.

buyers, using observed convictions and revocations of registration that are published on the DEA website and cross-referencing the convicted buyers' DEA IDs with the corresponding buyer IDs in the ARCOS data, as *suspicious buyers*. From the remaining buyers in the ARCOS data, we randomly sample 50,000 buyers and split them in to two sets, a *training set* and a *validation set*. We add the 188 suspicious buyers to the validation set. This yields a training set with 25,000 (unlabeled) buyers, and a validation set with 25,188 buyers, i.e., 25,000 (unlabeled) buyers plus 188 (labeled) *suspicious buyers*. This is explained in diagrammatic form in Figure 4. Suspicious buyers can be considered to be a very small percentage of the general population of retail buyers of opioid products, so it would be fair to assume that unlabeled buyers largely represent legitimate buyers of opioid products.

[INSERT FIGURE 4]

We overlay the 188 suspicious buyers on the US map in Figure 5. Unsurprisingly a majority of suspicious buyers fall within the Rust Belt and the Appalachian Belt, two regions that were severely afflicted with opioid addiction and overdose deaths during the period of our data. A vast majority (138) of the suspicious buyers are found to be medical practitioners, which is consistent with popular reports in the press that "pill mills" operated by unethical physicians have played an instrumental role in the drug diversion problem. The second largest category of suspicious buyers are retail (mom-and-pop) pharmacies, which is consistent with the Netflix documentary, "The Pharmacist" that features small pharmacies that dispensed huge amounts of pain pills to addicts.

[INSERT FIGURE 5]

In Table 2, we present some descriptive statistics on three groups of buyers: (1) training set - unlabeled, (2) validation set – unlabeled, (3) validation set – suspicious. Some interesting findings are evident in the table. In terms of the average MME purchased per transaction, suspicious buyers buy almost 10 times (~ 2.4 million) as much as unlabeled buyers (~ 250,000). In terms of median MME purchased per transaction, suspicious buyers buy almost 20 times (~ 700,000) as much as unlabeled buyers (~ 37,000). In terms of the standard deviation of MME across transactions, suspicious buyers exhibit much more variation (~ 39 million) than unlabeled buyers (~ 15 million). Therefore, suspicious buyers not only buy more MME

per transaction, on average, but also display greater temporal variation in their purchase quantities of MME. If we look at the share of each opioid product within the total quantity of MME purchased by each buyer over the 6-year period (2006-2012), two interesting findings are observed in the table: one, Oxycodone accounts for about 20 % of a suspicious buyer's MME purchases, on average, but only about 4 % for an unlabeled buyer; two, Fentanyl accounts for only 8 % of a suspicious buyer's MME purchases, despite accounting for 30 % of an unlabeled buyer's MME purchases. Another interesting finding in Table 2 is that suspicious buyers buy fewer unique opioid SKUs (~ 38) than unlabeled buyers (~56).

[INSERT TABLE 2]

Based on Table 2, it is clear that suspicious buyers are different from unlabeled buyers in some discernible ways with regard to their observed purchase behavior of opioids. This suggests that an anomaly detection algorithm that is trained to distinguish suspicious buyers from unlabeled buyers may be well informed by the inclusion of *input variables* such as those listed in Table 2. In order to be comprehensive in our inclusion of relevant input variables, we construct a total of 40 input variables which are different variations of the basic set of input variables that are listed in Table 2. The 40 input variables are listed below:

1. Average MME per Transaction (X1),
2. Standard Deviation of MME per Transaction (X2),
3. Median MME per Transaction (X3),
4. Maximum MME per Transaction (X4),
5. Maximum Minus Median MME per Transaction (X5)
6. Coefficient of Variation of MME per Transaction (X6), i.e., X2 / X1,
7. Average MME per Day (X7),
8. Standard Deviation of MME per Day (X8),
9. Median MME per Day (X9),
10. Maximum MME per Day (X10),
11. Maximum Minus Median MME per Day (X11),

12. Average MME per Day per Person (X12),
13. Standard Deviation of MME per Day per Person (X13),
14. Median MME per Day per Person (X14),
15. Maximum MME per Day per Person (X15),
16. Maximum Minus Median MME per Day per Person (X16),
17. Average MME per Day per Old Person (X17),
18. Standard Deviation of MME per Day per Old Person (X18),
19. Median MME per Day per Old Person (X19),
20. Maximum MME per Day per Old Person (X20),
21. Maximum Minus Median MME per Day per Old Person (X21),
22. Hydrocodone Share (X22),
23. Oxycodone Share (X23),
24. Fentanyl Share (X24),
25. Morphine Share (X25),
26. Hydrocodone + Oxycodone Share (X26),
27. Hydrocodone + Oxycodone + Fentanyl Share (X27),
28. Hydrocodone + Oxycodone + Fentanyl + Morphine Share (X28),
29. # Unique SKUs (X29),
30. # SKUs in 80 % of all MME Purchases (X30),
31. # Unique Distributors (X31),
32. # Distributors in 80 % of all MME Purchases (X32),
33. Average # Opioid Products per Transaction (X33),
34. Median # Opioid Products per Transaction (X34),
35. Standard Deviation # Opioid Products per Transaction (X35),
36. Maximum # Opioid Products per Transaction (X36),
37. Average # Distributors per Transaction (X37),

38. Median # Distributors per Transaction (X38),

39. Standard Deviation # Distributors per Transaction (X39),

40. Maximum # Distributors per Transaction (X40),

In the above list of variables, X7 – X11 are identical to X1 – X5 except that X7 – X11 are additionally scaled by the number of days between successive transactions of the buyer, thus reflecting average purchase amount per day rather than per transaction. Further, X12 – X16 are identical to X7 – X11 except that X12 – X16 are additionally scaled by the population of the county in which the retail buyer is located, thus reflecting average purchase amount per day per resident. In contrast to X12 – X16, which use the county population, X17 – X21 use the county population that is 65 or older (i.e., relevant population for pain medication), thus reflecting the average purchase amount per day per old resident. Variables X22 – X28 track the shares of the buyer's MME purchases that correspond to the 4 main opioid drugs and combinations thereof. X29 represents the total number of unique opioid products bought by a buyer over the 6-year period, while X30 represents the subset of SKUs within X29 that collectively account for 80 % of the buyer's MME purchases. Similarly, X31 represents the total number of unique distributors from whom the buyer buys over the 6-year period, while X32 represents the subset of distributors within X31 who collectively account for 80 % of the buyer's MME purchases. The remaining variables, X33 – X40, are self-explanatory. In Table 3, we present several descriptive statistics on the full set of 40 input variables for all three groups of buyers: (1) training set - unlabeled, (2) validation set – unlabeled, (3) validation set – suspicious.

[INSERT TABLE 3]

Suspicious buyers buy far fewer SKUs (~ 38), on average, than unlabeled buyers (~ 56). In other words, suspicious buyers concentrate their opioid buying on a smaller number of SKUs. Suspicious buyers also buy from more distributors (~ 5), on average, when compared to unlabeled buyers (~ 3.4). Armed with these preliminary insights, we next build an anomaly detection algorithm to detect suspicious buyers ("anomalies") in a buyer population using the full set of 40 input variables.

Anomaly Detection Algorithm

Let *i* refer to a retail buyer. Define $Y_i = 1$ if *i* is a suspicious buyer, and $Y_i = 0$ if *i* is an unlabeled buyer. This is the *predictive outcome* of interest. Let $X_i = [X_{i1}\ X_{i2}\ \ldots\ X_{i,40}]$' be a 40-dimensional column vector denoting the *input variables* of buyer *i*. These input variables, as listed in the previous section, represent *purchase patterns* of buyer *i*. The goal of the anomaly detection algorithm is to predict Y_i based on X_i. The anomalous outcome, $Y_i = 1$, is specifically of interest. Since the outcome of interest represents only 0.37 % of the 50,188 randomly sampled buyers in our dataset (188 out of 50,188), a *supervised learning algorithm* will not work well for our predictive goal because of the scarcity of data records representing suspicious buyers ($Y_i = 1$). An *anomaly detection algorithm*, instead, would treat the scarce outcome as an "outlier" and learn how to predict outlier data records in the data using "regular" data records, which represent unlabeled buyers ($Y_i = 0$). Further, when there are many different types of suspicious buyers (representing *sub-classes* of suspicious buyers), as is quite likely in our case, a supervised learning algorithm will be rendered even more infeasible. This is why we use an anomaly detection algorithm in our predictive application. In doing this, *we introduce anomaly detection to the marketing literature.*

We have two mutually exclusive sets of data, the *training set* which contains 25,000 unlabeled buyers ($i = 1, \ldots, 25000$), and the *validation set* which contains 25,000 unlabeled buyers and 188 suspicious buyers ($i = 25001, \ldots, 50188$). On the training set, $X_1, \ldots, X_{25000}$ represent the input variable vectors. On the validation set, while $X_{25001}, \ldots, X_{50188}$ represent the input variable vectors, $Y_{25001}, \ldots, Y_{50188}$ represent the outcomes of interest. In other words, heterogeneous outcomes of interest are only observed within the validation set, taking the value $Y_i = 1$ for 188 buyers, and the value $Y_i = 0$ for 25,000 buyers. In the training set, however, since $Y_i = 0$ for all buyers, the outcome variable is not useful from an analysis standpoint. We explain our anomaly detection algorithm (which is a *density-based* algorithm) next.[6]

[6] We compare our algorithm, which is our original modeling contribution, to four popular anomaly detection algorithms from the machine learning literature: (1) Multivariate Gaussian Density (MGD), (2) Independent Gaussian Density with Principal Components (IGD w/ PCA), (3) Isolation Forest (IFOR), and (4) IFOR with Principal Components (IFOR w/ PCA). Our algorithm outperforms those four comparison algorithms in terms of predictive accuracy.

Step 1 (Training): For each input variable, X_j (j = 1, …, 40), we calculate the sample estimates of the mean (μ_j) and standard deviation (σ_j) across the 25,000 buyers in the training set, as shown below.

$$\mu_j = \frac{1}{25000}\sum_{i=1}^{25000} X_{ij},$$

$$\sigma_j = \frac{1}{25000-1}\sum_{i=1}^{25000}(X_{ij} - \mu_j)^2 ,$$

(1)

Step 2 (Validation): For each buyer, i (i = 25001, …, 50188) in the validation set, we calculate the univariate Gaussian density, $f(X_{ij})$, that is associated with each input variable, X_{ij} (j = 1, …, 40), under the assumption that the mean and standard deviation of the Gaussian density are (μ_j) and (σ_j), respectively, as calculated in step 1.[7] The calculation of the Gaussian density is shown below.

$$f(X_{ij}) = \frac{1}{\sigma_j\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X_{ij}-\mu_j}{\sigma_j}\right)^2} ,$$

(2)

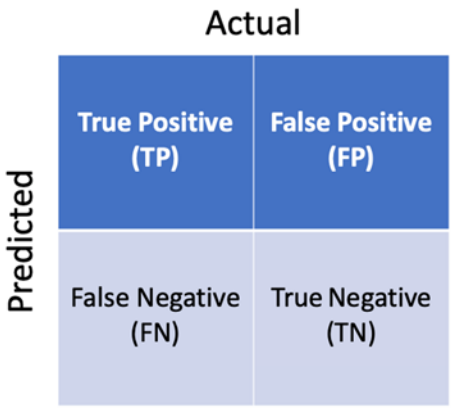
Step 3 (Validation): We assume arbitrary non-negative (small-valued) threshold parameters ε_1, …, ε_{40} for the 40 input variables, X_j (j = 1, …, 40). Then for each buyer, i (i = 25001, …, 50188) in the validation set, we predict whether they are a suspicious buyer (Y_i = 1) or not (Y_i = 0), according to the following disjunctive rule across the 40 input variables.

$$\begin{cases} Y_i = 1 & if\ f(X_{i1}) \le \varepsilon_1\ or\ f(X_{i2}) \le \varepsilon_2\ or \ldots f(X_{i40}) \le \varepsilon_{40} \\ Y_i = 0 & Otherwise \end{cases}$$

(3)

[7] This procedure is valid if the frequency histogram of each input variable looks symmetric and bell-shaped (i.e., Gaussian), otherwise one must transform the input variable, using a suitable transformation function (logarithm, square root etc.), to yield a Gaussian shape. In our case, we find that such transformations do not improve the predictive accuracy of the algorithm.

Step 4 (Validation): Having predicted the outcome of interest, Y_i, for all 25,188 buyers in the validation set, we summarize the predicted versus actual values of Y_i in the form of a 2*2 confusion matrix, as shown below.



(4)

Step 5 (Validation): We calculate the precision and recall measures of predictive accuracy that are associated with the confusion matrix of Step 4, as shown below.

$$Precision = \frac{TP}{TP+FP},$$

$$Recall = \frac{TP}{TP+FN},$$

(5)

Step 6 (Validation): We calculate the harmonic mean of the precision and recall measures of predictive accuracy, as calculated in step 5, to yield the F-1 score, as shown below.

$$F-1 = \frac{2\times(Precision\times Recall)}{(Precision+Recall)},$$

(6)

The F-1 score is a composite measure of predictive accuracy that takes in to account the dual abilities of our algorithm to both correctly predict observed anomalies (recall), as well as yield anomaly predictions that are accurately vindicated among validation buyers (precision). This is a commonly used accuracy metric for implementable machine learning algorithms.

Step 7 (Validation): We check to see if the F-1 score calculated in step 6 has reached its maximum possible value. If not, we go back to step 3 and change our assumed values of threshold parameters $\varepsilon_1, \ldots, \varepsilon_{40}$ for the 40 input variables, X_j ($j = 1, \ldots, 40$) and then cycle through steps 4-7 until the F-1 score reaches its maximum possible value. Once the best algorithmic performance has been achieved, we summarize the predictive performance of the anomaly detection algorithm. We report not only *precision*, *recall*, and *F-1 score* (as explained above in equations (5) and (6)), but also *specificity* and *negative predictive value*, as summarized below.

$$Specificity = \frac{TN}{TN+FP},$$

$$Negative\ Predictive\ Value = \frac{TN}{TN+FN}, \tag{7}$$

Our modeling goal is to estimate the threshold parameters, $\varepsilon_1, \ldots, \varepsilon_{40}$, to enable anomaly detection. This is done so as to achieve the highest level of predictive accuracy for our anomaly detection algorithm (which is built using data on unlabeled training buyers, in step 1 above), in terms of successfully maximizing the F-1 score metric among validation buyers (as explained in steps 2-7 above). However, this raises a key implementation challenge: Given the discontinuity of the disjunctive rule of step 3, gradient-based optimization cannot be used. We must, instead, sample an arbitrarily large number of (non-negative) values for each threshold parameter, and then figure out which combination of threshold parameters maximizes the F-1 score. This, in turn, raises the question of what range of values to sample each threshold parameter from. Even if valid ranges can be identified, implementing the algorithm becomes computationally prohibitive since the number of combinatorial possibilities becomes very large (even with as few as 10 grid points per threshold parameter, we have 10^{40} possibilities to consider). This challenge is uniquely faced by our algorithm compared to, say, the MGD, which involves the estimation of just one threshold parameter, a much easier task. However, we believe that our algorithm's flexibility for anomaly detection is an asset compared to traditional machine learning algorithms such as MGD, IGD w/ PCA, IFOR and IFOR w/ PCA.

We use the following *heuristic procedure*, based on the principle of *forward selection* of input variables, in order to implement our algorithm in a computationally convenient manner.

1. We use just one input variable at a time, X_j (j = 1, …, 40), and cycle through steps 1-7 above to maximize the predictive F-1 score (which is easy to do since it involves one-dimensional optimization). This yields 40 different values of maximized F-1 score. Choose the largest value among these and identify the corresponding variable as the most important input variable for the anomaly detection task. Its threshold parameter is set at its estimated value.
2. Now, we add a second input variable to the input variable that has been identified as first-best in the previous step. Again, we add just one input variable at a time, X_j (j = 1, …, 40), but not counting the first-best variable identified in the previous step, so allowing for 39 possibilities, and cycle through steps 1-7 above to maximize the predictive F-1 score (which is again one-dimensional optimization since the first-best variable's threshold parameter has already been estimated). This yields 39 different values of maximized F-1 score. Choose the largest value among these (which must be at least as large as the maximized F-1 score of step 1) and identify the corresponding variable as the second-most important input variable for the anomaly detection task. Its threshold parameter is set at its estimated value.
3. Now, we add a third input variable to the two input variables that have been identified as first-best and second-best in the previous two steps. Again, we add just one input variable at a time, X_j (j = 1, …, 40), but not counting the first-best and second-best variables identified in the previous two steps, so allowing for 38 possibilities, and cycle through steps 1-7 above to maximize the predictive F-1 score (which is again one-dimensional optimization since the threshold parameters for the first-best and second-best variables have already been estimated). This yields 38 different values of maximized F-1 score. Choose the largest value among these (which must be at least as large as the maximized F-1 score of step 2) and identify the corresponding variable as the third-most important input variable for the anomaly detection task. Its threshold parameter is set at its estimated value.
4. Etc.

In line with the above procedure, we sequentially keep adding one input variable at a time in order to improve the F-1 score until a point comes when the F-1 score does not improve from adding another input variable. We stop at that point. This procedure works well to identify the combination of input variables, from our full set of 40, which is sufficient to maximize the predictive accuracy of our algorithm. While we acknowledge that the forward selection procedure may not yield a globally optimal solution (that corresponds to the situation of simultaneously estimating the 40 threshold parameters), in our empirical application, using our proposed procedure leads to a predictive accuracy that is much higher than that of four popular anomaly detection algorithms from the AI literature, MGD, IGD w/ PCA, IFOR, and IFOR w/ PCA (see Ruff et al. 2021 for a current review of machine learning algorithms for anomaly detection). This completes our exposition of our proposed anomaly detection algorithm.

Empirical Results

First, we look at how our anomaly detection algorithm fares, relative to four comparison machine-learning algorithms (MGD, IGD w/ PCA, IFOR, IFOR w/ PCA), in predicting suspicious buyers in the validation set. The predictive accuracy comparisons are given in Table 4. In terms of five predictive metrics – precision, recall, F-1 score, AUC, Lift -- we see that our proposed algorithm outperforms the four comparison algorithms. We test a special case of our proposed algorithm where we first undertake PCA in order to reduce the 40 input variables to a small number (~15) of principal components, and then use the principal components as the bases for anomaly detection. This special case algorithm does not do as well as our algorithm which works directly with input variables. Interestingly, however, the empirical performance of IFOR w/ PCA is superior to that of IFOR. Our algorithm yields maximum possible precision (100 %), with an impressive F-1 score (61 %). For policy makers and drug distributors (such as McKesson), what a precision of 100 % means is that if our anomaly detection algorithm is used to flag suspicious buyers, the probability of the algorithmic prediction being correct would be 100 %. *Given that anomalies are*

extremely rare outcomes in the data, plus since the US justice system tends to err on the side of minimizing false positive predictive errors (falsely accusing innocent retail buyers of suspicious orders), having 100 % precision is a very important feature of our predictive algorithm. We are willing to sacrifice some recall (and increase false negative errors) in order to enable the practical adoption of our proposed algorithm. We plot the receiver operating characteristics (ROC) curve, i.e., recall versus 1-specificity, for each algorithm and calculate the area under the curve (AUC) as an additional metric of predictive accuracy. We also plot the precision versus recall curve for each algorithm and calculate the area under the curve (Lift) as a final metric of predictive accuracy. In terms of both AUC (72 %) and Lift (72 %), our algorithm outperforms the comparison algorithms. *Overall, therefore, our proposed anomaly detection algorithm does very well in terms of its ability to flag suspicious orders in the distribution channel before the drugs actually get shipped to the retail buyer.*

[Insert Table 4 here]

We present the confusion matrix associated with our anomaly detection algorithm in Table 5. Consistent with our above discussion about the tradeoff between precision and recall, we see that there are zero false positive errors made by our algorithm. However, the false negative error rate is moderately high (recall of 44 %) in order to enable this maximum precision (100 %). Our maximized F-1 score compares well to F-1 scores that have been recovered in successful anomaly detection applications (e.g., fraud detection, machine failure etc.) in the AI literature.

[Insert Table 5 here]

We report the input variables that assist our detection of suspicious buyers, in decreasing order of importance, in Table 6. First, we notice that only 7 out of our full set of 40 input variables are needed to enable the high predictive accuracy of our algorithm. The most important input variable turns out to be *median MME per day (X9).* MME per day is calculated as follows: each time a buyer places an order, we look at the quantity of MME ordered by the buyer during their previous transaction, and then divide that quantity by the number of days elapsed since the buyer's previous transaction until the current transaction.

This measure reflects the (possibly time-varying) rate at which the buyer sells opioids to their patients. It is not surprising that this variable is most predictive of suspicious buyer behavior since the fundamental product that is bought by pain patients and, therefore, the retail buyer who sells to pain patients, is morphine (rather than the specific drug product – oxycodone or hydrocodone -- that treats the patients' pain). The second most important variable is *Oxycodone share (X23)*, i.e., the share of a retail buyer's past purchases of MME that is accounted for by Oxycodone. Given that our study period spans 2006-2012, when OxyContin was a widely abused drug across the country, this finding is also not surprising. Next to tracking the amount of morphine equivalents in a retail buyer's order, tracking specifically the amount of Oxycodone in that order would play a vital role in flagging an order as possibly suspicious. The third most important variable is *Morphine share (X25)*, i.e., the share of a retail buyer's past purchases of MME that is accounted for by Morphine (the actual drug product itself, rather than the chemical morphine equivalents delivered therein). The fourth most important variable is *# Unique Distributors (X30)*, i.e., the total number of distributors from whom a retail buyer has sourced their past opioid purchases. Suspicious buyers tend to distribute their purchases across many distributors in order to decrease the chance of being flagged for placing suspiciously large orders for opioids from a single distributor. The fifth and sixth most important variables are *Median MME per Transaction (X3)* and *Average MME per Transaction (X1)*, respectively. That these variables are closely related to, yet additionally contribute beyond, *median MME per day (X9)*, in predicting suspicious orders, reiterates the fact that tracking a retail buyer's historical purchases of morphine equivalents in multiple ways is critical for early identification of suspicious orders. The seventh and final input variable of interest is *# Distributors in 80 % of MME Purchases (X32)*, which is a variant of X30 discussed above. However, we find that suspicious buyers tend to concentrate their MME buying among a very small (typically, one) distributor. This suggests an interesting combination of seemingly opposite tactics used by suspicious buyers to evade detection by the justice department. On the one hand, they place orders from a large number of distributors overall, which effectively camouflages their detection as suspicious buyers using the transactional data of large distributors (who are possibly more monitored by the DEA). On the other hand, these buyers tend to concentrate their real buying of opioids from just one

distributor, possibly a small distributor who is less likely to fall under the surveillance dragnet of the justice department, who is less scrupulous in terms of reporting suspicious buying activity to the DEA.

Overall, what Table 6 shows is that as long as one diligently tracks the mentioned 7 variables at the buyer-level in real-time, updating the value of each variable whenever a new order gets placed by a buyer, our algorithm can be efficiently used to flag suspicious orders in real time by drug distributors so that they can swiftly report them to the DEA. In order to test the practical efficacy of our algorithm, we need drug distributors and the DEA to partner and work with us in a pilot study application. We are hopeful that we can achieve this in the future.

[Insert Table 6 here]

In order to further elucidate the substantive implications of our proposed anomaly detection algorithm, we report in Table 7 the top 25 retail buyers in the validation set who are predicted most likely to be suspicious buyers by our algorithm. All of these buyers are, in fact, suspicious buyers (i.e., actual Y_i = 1), as noted earlier when we reported that our algorithm achieves 100 % precision in identifying suspicious buyers. Interestingly, we see that all these buyers are associated with the very high *median MME per day (X9)*, i.e., ~ 2344665 MME per day, and *Oxycodone share (X23)*, i.e., ~ 100 %, among all buyers in the validation set. They are also associated with the lowest possible value of *# Distributors in 80 % of MME Purchases (X32)*, i.e., 1. On the remaining 4 input variables, there is more heterogeneity among these 25 buyers. Interestingly, we find that these 25 buyers show suspicious activity on all 7 input variables, i.e., the values of all 7 variables for these buyers fall in to the tail regions of their frequency distributions such that all 7 densities are smaller than their estimated density thresholds as reported in Table 6. *In other words, the signals of anomaly detection are very strong for these egregiously suspicious buyers. Suppose the DEA invests human and financial resources to perform further due diligence and investigation on the top 25 suspicious buyers, as flagged by our algorithm, in any given month. Such investigation efforts may well pay off since the chance of a false positive error in these predictions is practically non-existent given the strengths of the signals in Table 7.This renders our algorithm very valuable for practical use.*

[Insert Table 7 here]

In order to demonstrate the practical use of our proposed anomaly detection algorithm, we report in Table 8 the top 25 retail buyers in the training set who are predicted most likely to be suspicious buyers by our algorithm. Since the buyers in the training sample are all unlabeled, this prediction exercise is an example of what the DEA can do with our predictive tool. Interestingly, we see that all these buyers are associated with the same high value of *median MME per day (X9)*, i.e., 80141 MME per day, as well as *Oxycodone share (X23)*, i.e., ~ 33 %. They are also associated with the lowest possible value of *# Distributors in 80 % of MME Purchases (X32)*, i.e., 1. On the remaining 4 input variables, there is quite a bit of heterogeneity among these 25 buyers.

[Insert Table 8 here]

In order to test the quality of these predictions, we performed Google searches on these 25 retail buyers. We summarize the results of our Google searches in Table 9. We find that 4 out of these 25 buyers (shaded in yellow in Table 9) committed clear cases of opioid fraud. We also uncover news articles associated with opioid fraud that can be somewhat tied, but not conclusively, to another 4 out of these 25 buyers. This lends further faith to the practical viability of our anomaly detection algorithm to identify suspicious retail buyers.

[Insert Table 9 here]

Conclusions

Over the past 20 years, about 500,000 Americans have died of opioid overdoses. This has been the most pressing health epidemic in the US. According to the US Surgeon General, 250 million prescriptions for opioids are written in the US each year. Our research was spurred by the recent availability of a massive database, called ARCOS, released by the DEA in 2019, which tracks the shipments of opioids through the supply chain. We propose a predictive capability to detect suspicious orders of opioid drugs at the source before drug distributors can release drug shipments to drug retailers. Our predictive approach achieves a precision of 100 % on the ARCOS data in terms of identifying suspicious buyers -- as revealed by observed

convictions on the DEA website -- by eliminating false positive errors. Our approach offers new hope in our nation's fight against the opioid epidemic.

We believe that our proposed predictive algorithm can also be leveraged across more than 100 other controlled substances, beyond opioid drugs, that are tracked by ARCOS. Our goal is to open the door to a breakthrough private-public pilot to test the creation of a real-time supply-chain detection and alerting system. Supported by both government and industry, it would address an ongoing blind spot in the opioid supply chain that continues to leave citizen communities at risk. An advanced detection and alert system will require rethinking of existing public policy in several areas, such as the following: (1) Data sharing and cross-agency communication; (2) Revised and modernized data reporting; (3) Funding sources and spending needs for system maintenance; (4) Response guidance when transactions are flagged.

Some important caveats are in order. First, as prescription opioids become harder to come by, users turn to heroin and other more dangerous street drugs. Therefore, an unintended consequence of better monitoring of suspicious legal shipments of opioids may be an increased incidence of unmonitored drug distribution on the streets. This is not unlike alcohol prohibition in the 1930s leading to bootlegging and increased street crime in the streets of Chicago. Second, the recent appearance of counterfeit versions of a particularly potent synthetic opioid, Fentanyl, is further driving up the body count of drug overdose deaths. This means that a multi-pronged policy approach, which tracks not only official shipments of opioids (as represented in ARCOS) but also unofficial shipments of counterfeit opioids and illegal drugs on the street, would be necessary to effectively address the national drug overdose epidemic. We leave these important issues for future research. Our research is an important first step. Aggressive marketing practices by Purdue Pharma, which are endemic to marketing practices across both the pharmaceutical industry and other industries where personal selling plays an important role, were largely responsible for the indiscriminate use of powerful opioids to treat not just chronic pain but also mild to moderate pain. This is what created and fueled the opioid overdose epidemic. Given this somber history, it is incumbent on the marketing community to acknowledge, as well as make amends for, its role in having created this societal problem. As values-based research scholars in marketing, we present our research as a humble effort in this direction.

References

Bobroske, K., Freeman, M., Huan, L., Cattrell, A, Scholtes, S. (2021). Curbing the Opioid Epidemic at its Root: The Effect of Provider Discordance after Opioid Initiation, Management Science, Forthcoming.

Chandy, R. K., Johar, G. V., Moorman, C., Roberts, J. H. (2021). Better Marketing for a Better World, Journal of Marketing, 85, 3, 1-9.

Corman, H., Mocan, H. N. (2000). A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City, The American Economic Review, 90, 3, 584-604.

Liu, J., Bharadwaj, A. (2020). Drug Abuse and the Internet: Evidence from Craigslist, Management Science, 66, 5, 2040-2049.

Meier, B. (2018). Pain Killer: An Empire of Deceit and the Origin of America's Opioid Epidemic, Columbia University Press, New York: Random House.

Puntoni, S., Reczek, R. W., Giesler, M., Botti, S. (2021). Consumers and Artificial Intelligence: An Experiential Perspective, Journal of Marketing, 85, 1, 131-151.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., Muller, K. R. (2021). A Unifying Review of Deep and Shallow Anomaly Detection, Proceedings of IEEE, 109, 5, 756-795.

Zaric, G. S., Brandeau, M. L. (2000). Methadone Maintenance and HIV Prevention: A Cost-Effectiveness Analysis, Management Science, 46, 8, 1013-1031.

Zhang, V. S., King, M. D. (2021). Tie Decay and Dissolution: Contentious Prescribing Practices in the Prescription Drug Epidemic, Organization Science, 32, 5, 1149-1173.

Zheng, Y., Alba, J. W. (2021). Consumer Self-Control and the Biological Sciences: Implications for Marketing Stakeholders, Journal of Marketing, 85, 4, 105-122.

TABLE 1: TOP 50 OPIOID STOCK KEEPING UNITS (SKUs)

SKU	Manufacturer	API	Drug Product
HYDROCODONE BIT. 10MG/ACETAMINOPHEN	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT/ACETAMINOPHEN 5MG/50	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
OXYCODONE HCL/ACETAMINOPHEN 5MG/325M	SpecGx LLC	OXYCODONE HYDROCHLORIDE	OXYCODONE
BUPRENORPHINE/NALOXONE 8MG/2MG (SUBO	Indivior Inc.	BUPRENORPHINE	BUPRENORPHINE
HYDROCODONE BIT 5MG/ACETAMINOPHEN 50	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT/ACETA 7.5MG/500MG US	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODO.BIT/APAP 7.5MG/750MG USP T	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT/ACETA 10MG/325MG USP	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
METHADONE HYDROCHLORIDE TABLETS, USP	SpecGx LLC	METHADONE HYDROCHLORIDE	METHADONE
HYDROCODONE BIT/ACETA 10MG/500MG USP	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT/ACETA 7.5MG/325MG US	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE.BIT./ACET.,10MG & 325MG/	Par Pharmaceutical	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT/ACETA 5MG/325MG USP	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
TUSSIONEX 10MG/5ML HYDROCODO.BIT	Unither Manufacturing LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
OXYCODONE HYDROCHLORIDE TABLETS 5MG	SpecGx LLC	OXYCODONE HYDROCHLORIDE	OXYCODONE
ENDOCET - 10MG OXYCODONE.HCL/325MG A	Par Pharmaceutical	OXYCODONE HYDROCHLORIDE	OXYCODONE
OXYCODONE HCL/ACETAMINOPHEN 10MG/325	SpecGx LLC	OXYCODONE HYDROCHLORIDE	OXYCODONE
HYDROCODONE BITARTRATE AND ACETA 5MG	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODO.BIT 10MG&AC USP TAB	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BITARTRATE AND ACETA 7.5	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT.7.5MG/ACETAMINOPHEN	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BIT./ACETAMINOPHEN TABS.	Amneal Pharmaceuticals LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCODONE BITARTRATE 7.5MG/ACETAMI	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
OXYCODONE.HCL/APAP 10MG/325MG TABS	Actavis Pharma, Inc.	OXYCODONE HYDROCHLORIDE	OXYCODONE
METHADONE HCL USP 10MG TABLET	West-Ward Pharmaceuticals Corp.	METHADONE HYDROCHLORIDE	METHADONE
FENTANYL 72HR 25MCG/HR TDS (2.55MG T	Mylan Pharmaceuticals, Inc.	FENTANYL BASE	FENTANYL
FENTANYL 72HR 50MCG/HR TDS (5.10MG T	Mylan Pharmaceuticals, Inc.	FENTANYL BASE	FENTANYL
SUBOXONE - BUPRENORPHINE 8MG/NALOXON	Indivior Inc.	BUPRENORPHINE	BUPRENORPHINE
HYDROCODONE BIT. 7.5MG/ACETAMINOPHEN	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROCOD.BIT./ACET.,7.5MG&500MG/15ML	Par Pharmaceutical	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
FENTANYL 72HR 100MCG/HR TDS (10.2MG	Mylan Pharmaceuticals, Inc.	FENTANYL BASE	FENTANYL
HYDROMORPHONE.HCL USP;4MG/TAB	SpecGx LLC	HYDROMORPHONE HYDROCHLORIDE	HYDROMORPHONE
OXYCONTIN - 40MG OXYCODONE.HCL CONTR	Purdue Pharma LP	OXYCODONE HYDROCHLORIDE	OXYCODONE
HYDROCOD.BIT.& APAP,10MG/660MG/TAB	SpecGx LLC	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
OXYCONTIN (OXYCODONE.HCL) CONTROLLED	Purdue Pharma LP	OXYCODONE HYDROCHLORIDE	OXYCODONE
CODEINE PHOSPHATE/ACETAMINOPHEN 30MG	Teva Pharmaceuticals USA, Inc.	CODEINE PHOSPHATE.1/2H20	CODEINE
HYDROCODONE BIT./IBUPROFEN;7.5MG & 2	Teva Pharmaceuticals USA, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROMET 5MG/5ML SYR.	Actavis Pharma, Inc.	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
HYDROMORPHONE.HCL USP;2MG/TAB	SpecGx LLC	HYDROMORPHONE HYDROCHLORIDE	HYDROMORPHONE
FENTANYL 72HR 75MCG/HR TDS (7.65MG T	Mylan Pharmaceuticals, Inc.	FENTANYL BASE	FENTANYL
OXYCODONE HYDROCHLORIDE 30MG TABLET	SpecGx LLC	OXYCODONE HYDROCHLORIDE	OXYCODONE
HYDROCODONE.BIT & ACETA 10MG & 500M	Par Pharmaceutical	HYDROCODONE BITARTRATE HEMIPENTAHYDRATE	HYDROCODONE
OXYCONTIN - 80MG OXYCODONE.HCL CONTR	Purdue Pharma LP	OXYCODONE HYDROCHLORIDE	OXYCODONE
OXYCODONE.HCL/APAP 7.5MG/325MG TABS	Actavis Pharma, Inc.	OXYCODONE HYDROCHLORIDE	OXYCODONE
FENTANYL 50MCG/ML INJECTABLE SOLUTIO	Hospira, Inc.	FENTANYL BASE	FENTANYL
ACETAMINOPHEN AND CODEINE PHOSPHATE	Teva Pharmaceuticals USA, Inc.	CODEINE PHOSPHATE.1/2H20	CODEINE
OXYCODONE HCL/ACETAMINOPHEN 7.5MG/32	SpecGx LLC	OXYCODONE HYDROCHLORIDE	OXYCODONE
MORPHINE SULFATE 30MG ER TABLET; 100	SpecGx LLC	MORPHINE SULFATE PENTAHYDRATE(I.E.,5H20)	MORPHINE
CODEINE PHOSPHATE/ACETAMINOPHEN 30MG	SpecGx LLC	CODEINE PHOSPHATE.1/2H20	CODEINE

TABLE 2: DESCRIPTIVE STATISTICS ON BUYERS

	Training Set (Unlabeled)	Validation Set (Unlabeled)	Validation Set (Suspicious)
Average MME per transaction	252,718	253,506	2,390,467
Median MME per transaction	36,764	37,180	700,326
Maximum MME per transaction	14,512.659	14,762,864	39,291,227
Std. Dev. MME across transactions	1,310,314	1,414,983	5,679,642
Hydrocodone Share	0.342	0.345	0.388
Oxycodone Share	0.042	0.042	0.203
Fentanyl Share	0.302	0.301	0.081
# Unique SKUs	56.55	55.31	37.60
# Distributors	1.05	1.05	1.09

TABLE 3 (a): DESCRIPTIVE STATISTICS ON TRAINING SET – UNLABELED BUYERS

Statistic	Mean	St. Dev.	Min	Median	Max
avg MME	252,718.500	601,903.900	421.823	16,356.250	2,491,840.000
std MME	1,310,314.000	4,049,320.000	0	11,978.5	18,386,964
median MME	36,764.340	73,248.700	355.963	5,750.000	293,786.200
coef variation	1.790	2.178	0	1.0	9
max MME	14,512,659.000	42,925,512.000	500.000	54,000.000	182,356,333.000
max-med MME	13,995,122.000	41,457,175.000	0	28,800	175,615,246
avg MME per day	65,014.180	144,130.900	0	598.7	580,734
std MME per day	211,527.500	471,544.300	0	779.6	1,912,014
median MME per day	8,548.308	19,314.770	0	212.5	80,141
max MME per day	2,401,320.000	4,971,878.000	0	3,057.3	15,708,440
max-med MME per day	2,377,532.000	4,924,906.000	0	2,456.1	15,682,190
avg MME per day per person	1.762	4.512	0	0.01	19
std MME per day per person	5.448	14.180	0	0.01	61
median MME per day per person	0.232	0.597	0	0.003	3
max MME per day per person	58.986	153.996	0	0.04	646
max-med MME per day per person	58.529	152.971	0	0.03	641
avg MME per day per old person	8.943	23.136	0	0.04	100
std MME per day per old person	31.413	84.522	0	0.1	380
median MME per day per old person	0.844	2.220	0	0.01	10
max MME per day per old person	351.859	924.189	0	0.2	3,960
max-med MME per day per old person	350.038	920.233	0	0.2	3,944
Hydrocodone Share	0.342	0.432	0	0.1	1
Oxycodone Share	0.042	0.088	0	0	0
Fentanyl Share	0.302	0.393	0	0	1
Morphine Share	0.017	0.040	0	0	0
Hydro Oxy Share	0.398	0.424	0	0.2	1
Hydro Oxy Fen Share	0.700	0.394	0	0.9	1
Hydro Oxy Fen Mor Share	0.739	0.377	0	1.0	1
Num Unique SKU	56.552	87.445	1	5	268
Num SKUs 80	3.270	3.834	1	1	14
Num Unique Dist	3.452	2.708	1	2	10
Num Dist 80	1.080	0.271	1	1	2
Avg Num Products	1.995	1.426	1	1.2	6
Median Num Products	1.700	1.175	1	1	5
Std Num Products	1.044	1.372	0	0.4	4
Max Num Products	7.334	9.202	1	2	30
Avg Num Distributors	1.047	0.091	1	1	1
Median Num Distributors	1.000	0.000	1	1	1
Std Num Distributors	0.112	0.171	0	0	1
Max Num Distributors	1.625	0.917	1	1	4

TABLE 3 (b): DESCRIPTIVE STATISTICS ON VALIDATION SET – UNLABELED BUYERS

Statistic	Mean	St. Dev.	Min	Median	Max
avg MME	253,506.800	605,208.900	421.823	16,771.480	2,491,840.000
std MME	1,414,983.000	4,109,804.000	0.000	90,356.130	18,386,964.000
median MME	37,180.430	73,569.920	356	5,875	293,786
coef variation	2.027	2.072	0.000	1.642	8.984
max MME	14,762,864.000	43,409,601.000	500	53,430	182,356,333
max-med MME	14,207,166.000	41,812,684.000	0	27,000	175,615,246
avg MME per day	72,317.310	140,814.200	3.629	8,638.332	580,733.600
std MME per day	248,819.400	455,694.400	5.336	81,098.370	1,912,014.000
median MME per day	9,586.705	19,199.290	3	1,666.7	80,141
max MME per day	2,847,558.000	4,834,512.000	5	118,877.5	15,708,440
max-med MME per day	2,795,802.000	4,785,166.000	0	112,377.5	15,682,190
avg MME per day per person	1.937	4.440	0.00003	0.086	18.820
std MME per day per person	6.208	13.801	0.00005	0.816	60.910
median MME per day per person	0.254	0.588	0.00002	0.018	2.532
max MME per day per person	63.468	149.449	0.00004	1.192	645.838
max-med MME per day per person	62.907	148.390	0.000	1.092	641.488
avg MME per day per old person	10.119	23.330	0.0001	0.585	100.074
std MME per day per old person	36.562	84.533	0.0002	6.147	379.678
median MME per day per old person	0.970	2.261	0.0001	0.082	9.761
max MME per day per old person	395.020	925.557	0.0001	9.117	3,960.028
max-med MME per day per old person	392.700	921.351	0.000	8.519	3,943.760
Hydrocodone Share	0.345	0.433	0.000	0.056	1.000
Oxycodone Share	0.042	0.088	0.000	0.000	0.325
Fentanyl Share	0.301	0.393	0.000	0.000	0.997
Morphine Share	0.017	0.040	0.000	0.000	0.163
Hydro Oxy Share	0.401	0.425	0.000	0.192	1.000
Hydro Oxy Fen Share	0.702	0.394	0.000	0.930	1.000
Hydro Oxy Fen Mor Share	0.741	0.376	0.000	0.960	1.000
Num Unique SKU	55.309	86.662	1	5	268
Num SKUs 80	3.239	3.819	1	1	14
Num Unique Dist	3.444	2.723	1	2	10
Num Dist 80	1.079	0.269	1	1	2
Avg Num Products	1.978	1.418	1.000	1.227	5.720
Median Num Products	1.687	1.169	1	1	5
Std Num Products	1.194	1.299	0.000	0.802	4.469
Max Num Products	7.216	9.125	1	2	30
Avg Num Distributors	1.047	0.090	1.000	1.000	1.317
Median Num Distributors	1.000	0.000	1	1	1
Std Num Distributors	0.129	0.164	0.000	0.092	0.524
Max Num Distributors	1.619	0.914	1	1	4

TABLE 3 (c): DESCRIPTIVE STATISTICS ON VALIDATION SET – SUSPICIOUS BUYERS

Statistic	Mean	St. Dev.	Min	Median	Max
avg MME	2,390,467.000	5,014,590.000	773.100	36,533.330	18,806,502.000
std MME	5,679,642.000	12,597,497.000	0.000	169,970.400	47,819,544.000
median MME	700,326.800	1,600,666.000	500	26,950	5,810,925
coef variation	1.492	1.098	0.000	1.464	4.485
max MME	39,291,227.000	93,569,579.000	802	128,325	374,864,080
max-med MME	37,782,160.000	91,490,910.000	0	78,885	367,395,209
avg MME per day	1,029,558.000	2,243,642.000	5.934	34,367.030	7,904,894.000
std MME per day	2,272,611.000	4,785,729.000	8.763	182,113.400	16,509,210.000
median MME per day	239,121.300	587,539.300	5.748	6,833.034	2,344,665.000
max MME per day	16,274,176.000	32,009,255.000	6.020	371,875.000	99,530,880.000
max-med MME per day	15,934,573.000	31,391,557.000	0.000	328,675.600	98,521,630.000
avg MME per day per person	8.457	17.303	0.00003	0.375	67.628
std MME per day per person	23.401	53.669	0.0001	3.749	230.123
median MME per day per person	1.646	3.586	0.00002	0.109	13.397
max MME per day per person	150.688	326.748	0.00004	4.499	1,282.881
max-med MME per day per person	148.518	323.094	0.000	4.461	1,268.501
avg MME per day per old person	20.252	35.931	0.0001	1.570	125.373
std MME per day per old person	54.395	96.327	0.0003	18.759	354.481
median MME per day per old person	3.619	7.289	0.00005	0.313	26.441
max MME per day per old person	439.798	832.322	0.0001	25.094	3,019.215
max-med MME per day per old person	435.878	829.004	0.000	19.752	3,012.595
Hydrocodone Share	0.388	0.460	0	0.1	1
Oxycodone Share	0.203	0.362	0	0	1
Fentanyl Share	0.081	0.195	0	0	1
Morphine Share	0.016	0.049	0	0	0
Hydro Oxy Share	0.591	0.442	0	0.9	1
Hydro Oxy Fen Share	0.682	0.409	0	0.9	1
Hydro Oxy Fen Mor Share	0.722	0.388	0	1.0	1
Num Unique SKU	37.609	66.791	1	6	231
Num SKUs 80	2.255	2.458	1	1	10
Num Unique Dist	4.809	4.154	1	3	15
Num Dist 80	1.250	0.553	1	1	3
Avg Num Products	1.930	1.179	1	1.4	5
Median Num Products	1.701	1.060	1	1	4
Std Num Products	1.130	1.157	0	0.9	4
Max Num Products	6.279	7.624	1	3	28
Avg Num Distributors	1.094	0.164	1	1	2
Median Num Distributors	1.000	0.000	1	1	1
Std Num Distributors	0.200	0.247	0	0.1	1
Max Num Distributors	1.915	1.251	1	1	5

TABLE 4: PREDICTIVE ACCURACY COMPARISON

	Precision	Recall	F-1 Score	AUC	Lift
Our Algorithm	1.00	0.44	0.61	0.72	0.72
Our Algorithm w/ PCA	0.77	0.27	0.40	0.63	0.52
Comparison Algorithms					
MGD	0.94	0.35	0.51	0.68	0.65
IGD w/ PCA	1.00	0.17	0.29	0.58	0.59
IFOR	0.33	0.18	0.23	0.59	0.26
IFOR w/ PCA	0.85	0.27	0.41	0.64	0.56

TABLE 5: CONFUSION MATRIX YIELDED BY OUR ANOMALY DETECTION ALGORITHM

	Actual Y = 1	Actual Y = 0
Predicted Y = 1	83 (TP)	0 (FP)
Predicted Y = 0	105 (FN)	25000 (TN)

TABLE 6: INPUT VARIABLES THAT ASSIST THE DETECTION OF SUSPICIOUS BUYERS

Variable	**Estimated density threshold (ε_j)**
1. Median MME per Day (X9)	2.07E-08
2. Oxycodone Share (X23)	0.009058
3. Morphine Share (X25)	0.000114
4. # Unique Distributors (X30)	0.003067
5. Median MME per Transaction (X3)	1.09E-08
6. Average MME per Transaction (X1)	4.98E-17
7. # Distributors in 80 % of MME Purchases (X32)	1.98E-11

TABLE 7: TOP 25 SUSPICIOUS BUYERS IDENTIFIED AMONG LABELED VALIDATION BUYERS

DEA Number	Buyer Name	X9 (MME/day)	X23 (Oxy)	X25 (Morp.)	X30 (Dist.)	X3 (Med-MME)	X1 (Avg-MME)	X32 (Dist80)
FB0254918	Beau Boshers, MD	1481143.00	1.000	0.0000	9	3033000.00	388.81	2
FW1453757	Randall Wolff	1728000.00	1.000	0.0000	9	5670000.00	284.69	1
AD7585865	Jacobo Dreszer, MD	1350000.00	1.000	0.0000	8	2981250.00	346.00	2
BC8112637	Cynthia Cadet, MD	1481143.00	1.000	0.0000	8	3198375.00	314.84	2
BA6733578	Michael Aruta, MD	2073600.00	1.000	0.0000	8	3456000.00	307.04	1
FD1201196	Roni, Dreszer	1883250.00	1.000	0.0000	7	2787750.00	388.81	1
FB0003943	Alfred Boyce, DO	2344665.00	1.000	0.0000	6	5810925.00	281.72	1
FD1749057	Jack Alan Danton, DO	2344665.00	1.000	0.0000	5	5810925.00	284.69	1
FP1312406	Zvi H Perper, MD	2344665.00	1.000	0.0000	7	5810925.00	91.20	1
BT5598214	Margy Temponeras, MD	1706400.00	0.995	0.0012	7	4209000.00	1282.88	1
FH0772257	Hills Pharmacy	2344665.00	0.970	0.0014	15	3770400.00	416.34	3
BS1314210	Barry M Schultz, MD	2344665.00	0.906	0.0028	9	5810925.00	388.81	1
AY1916103	Your Gruggist	2344665.00	0.598	0.0007	14	5810925.00	284.87	1
BC5289055	Holiday CVS	2344665.00	0.123	0.0027	8	3093850.00	1218.36	1
RH0208567	Houston Maintenance Clinic	824338.20	0.000	0.0000	2	5810925.00	31.39	1
BU6696073	United Prescription Services	2344665.00	0.000	0.0000	11	3600000.00	429.67	1
BN3795892	Newcare Home Health Serv.	2344665.00	0.000	0.0000	6	3140125.00	506.83	2
BW8625785	Wayne Pharmacy	925975.00	0.011	0.0008	9	924250.00	1021.74	1
FC1881211	Rene Casanova, MD	739607.10	0.968	0.0000	8	5328756.50	39.27	3
BK4015334	Algirdas J Krisciunas, MD	720000.00	0.770	0.0000	5	2592000.00	37.77	1
FB1490349	Steven B Brown, MD	660461.50	0.989	0.0000	4	2592000.00	281.72	1
BT9856002	Treasure Coast Specialty Ph.	620000.00	0.936	0.0011	15	843400.00	1282.88	2
FM0624139	TJ Mcnichol, MD	550000.00	0.780	0.0000	5	650000.00	48.40	1
BB0816441	Harriston L Jr Bass, MD	506250.00	0.000	0.0000	1	5810925.00	25.91	1
AP8271138	Fred J Powell, MD	478636.40	0.997	0.0000	9	1408500.00	1282.88	2

TABLE 8: TOP 25 SUSPICIOUS BUYERS IDENTIFIED AMONG UNLABELED TRAINING BUYERS

DEA Number	Buyer Name	X9 (MME/day)	X23 (Oxy)	X25 (Morp.)	X30 (Dist.)	X3 (Med-MME)	X1 (Avg-MME)	X32 (Dist80)
BV3682514	VA Medical Center	80141	0.3246	0.1632	10	293786	1.00E-09	1
AK3360904	Westside Neighborhood	80141	0.3246	0.1632	4	293786	1.64E+02	1
BF3905518	John Bradley VA Clinic	80141	0.3246	0.1632	3	293786	1.00E-09	1
AV4317447	Department of Veterans	80141	0.3246	0.1273	10	293786	8.07E+00	1
FP2478825	Prescott Valley Pharmacy	80141	0.3246	0.1236	6	202000	1.77E-03	1
AV4674556	Veterans Admin Med Center	80141	0.3246	0.1111	10	293786	1.14E+02	1
AV6023903	Department of Veterans Affairs	80141	0.3246	0.1017	8	293786	2.38E+00	1
BV6917958	Vaden Corp.	80141	0.3246	0.0892	7	259825	8.84E+01	1
BP5432365	Primary Care Center Pharmacy	80141	0.3246	0.0833	3	293786	7.23E-04	1
AV4291869	VA Medical Center	80141	0.3246	0.0828	5	293786	6.62E+00	1
BF7800572	Fred's Pharmacy	80141	0.3246	0.0826	6	238751	2.65E-03	1
BS7087237	Smith's Food & Drug	80141	0.3246	0.0822	4	135000	1.77E-02	1
AG1001988	Kaiser Foundation Health Plan, WA	80141	0.3246	0.0776	3	271875	1.00E-09	1
FU1700512	Upstate Pharmacy Cross Creek	80141	0.3246	0.0720	4	293786	6.46E+02	1
AU6389224	USAF – 377th Medical Group / SGSL	80141	0.3246	0.0720	8	215000	3.53E+01	1
FT1135385	Town and Country Drugs	80141	0.3246	0.0628	9	148512	5.82E+01	1
BF6708905	Family Pharmacy	80141	0.3246	0.0488	10	166862	8.23E-02	1
BP9638810	Professional Pharmacy	80141	0.3246	0.0479	10	179937	3.05E+01	1
BG7457294	Giant Pharmacy #363	80141	0.3246	0.0476	6	179500	2.80E-03	2
FA2626553	Asheville Highway Pharmacy, Inc.	80141	0.3246	0.0469	4	293786	1.71E-03	1
BV8122739	VA Medical Center	80141	0.3246	0.0462	4	266875	3.38E+01	1
RL0383276	Hui-Yin Li	80141	0.3246	0.0455	1	293786	1.00E-09	1
AV4593287	VA Medical Center	80141	0.3246	0.0438	8	293786	8.54E+01	1
FS2423678	St. Mina and Pope Kyrillos LLC	80141	0.3246	0.0430	1	293786	4.42E-02	1
BJ7649152	Joe's Pharmacy	80141	0.3246	0.0424	9	168525	2.91E-03	1

TABLE 9: GOOGLE SEARCH RESULTS ON THE TOP 25 IDENTIFIED SUSPICIOUS BUYERS

DEA Number	Buyer Name	Infraction
BV3682514	VA Medical Center	Healthcare inspection for alleged improper prescription in November 2013. http://www.va.gov/oig/pubs/VAOIG-13-00133-12.pdf
AK3360904	Westside Neighborhood	
BF3905518	John Bradley VA Clinic	
AV4317447	Department of Veterans	
FP2478825	Prescott Valley Pharmacy	News report says nearly 7 million opioid pills were dispensed at a Walgreens in Prescott from 2006 – 2012. https://www.azcentral.com/story/news/local/arizona-health/2019/07/25/millions-opioid-pills-flowed-arizona-pharmacies-prescription-drug-boom/1801196001
AV4674556	Veterans Admin Med Center	
AV6023903	Department of Veterans Affairs	
BV6917958	Vaden Corp.	
BP5432365	Primary Care Center Pharmacy	
AV4291869	VA Medical Center	Many controversies with quality of care; complaints about painkillers; easy-to-fill prescriptions; theft of opioids. https://www.daytondailynews.com/specials/left-behind-scandal-at-the-va/
BF7800572	Fred's Pharmacy	
BS7087237	Smith's Food & Drug	
AG1001988	Kaiser Foundation Health Plan, WA	
FU1700512	Upstate Pharmacy Cross Creek	Wife arrested after issuing blank prescriptions for opioids. https://www.foxcarolina.com/news/warrants-upstate-doctor-wife-arrested-after-issuing-blank-prescriptions-for-opioids/article_6e1a99ad-9c21-5d64-8bbe-84714f827a63.html
AU6389224	USAF – 377th Medical Group / SGSL	
FT1135385	Town and Country Drugs	
BF6708905	Family Pharmacy	
BP9638810	Professional Pharmacy	
BG7457294	Giant Pharmacy #363	
FA2626553	Asheville Highway Pharmacy, Inc.	
BV8122739	VA Medical Center	VA theft of opioids; https://www.nbcnews.com/storyline/va-hospital-scandal/opioid-theft-missing-prescriptions-prompts-investigation-va-hospitals-staff-n723291
RL0383276	Hui-Yin Li	80141 0.3246 0.0455 1 293786 1.00E-09 1
AV4593287	VA Medical Center	VA theft of opioids; https://www.nbcnews.com/storyline/va-hospital-scandal/opioid-theft-missing-prescriptions-prompts-investigation-va-hospitals-staff-n723291
FS2423678	St. Mina and Pope Kyrillos LLC	They have closed down.
BJ7649152	Joe's Pharmacy	Shut down after opioid investigations; https://news-bulletin.com/joes-pharmacy-shuts-down-after-opioid-investigations

Shaded in yellow are cases of retail buyers committing opioid fraud; shaded in grey, while also involving fraud, cannot be conclusively tied to the retail buyer in question.

FIGURE 1: OPIOID OVERDOSE DEATHS IN THE US FROM 1999 TO 2019

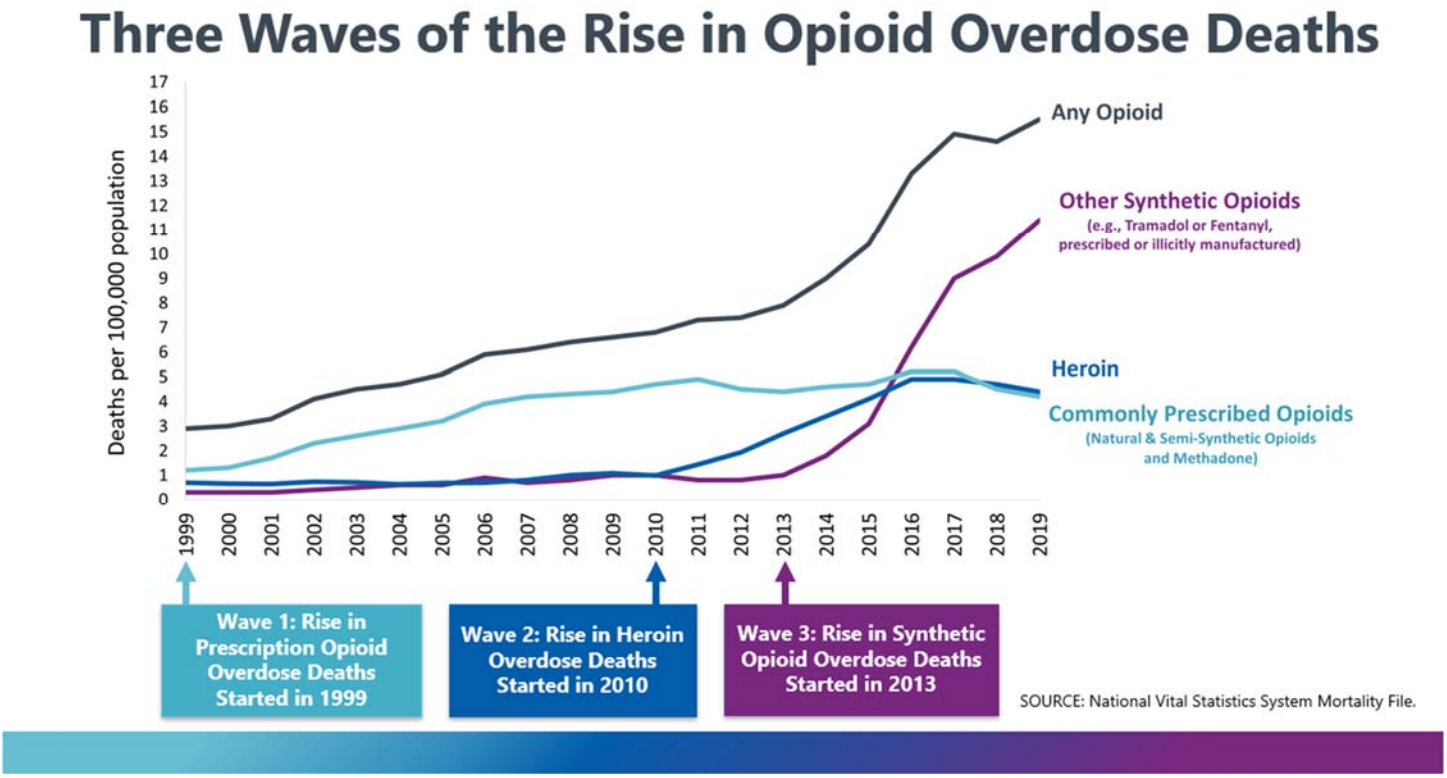


FIGURE 2: DISTRIBUTION CHANNEL STRUCTURE FOR OPIOIDS

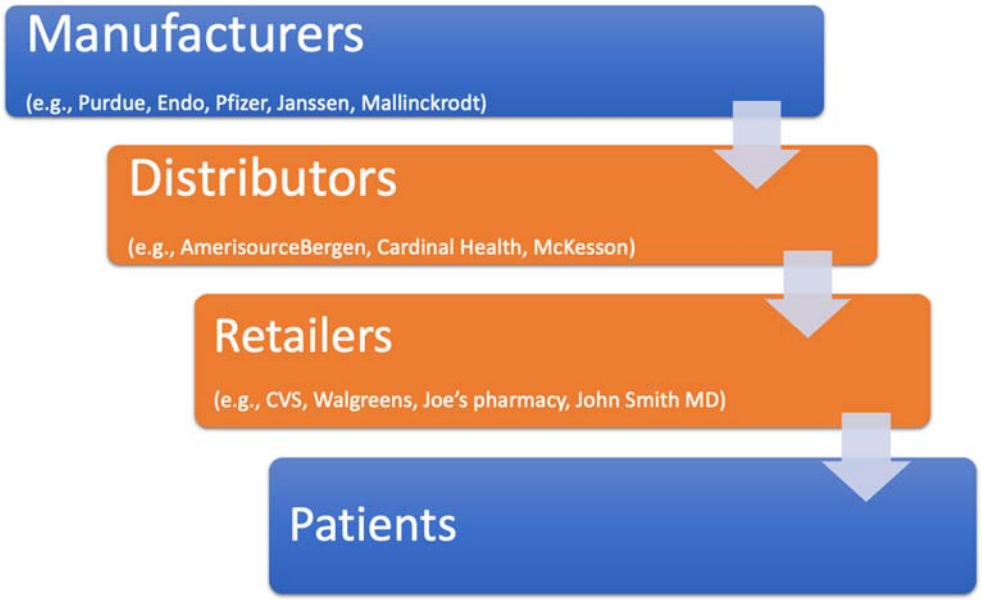


FIGURE 3: OPIOID DRUGS – A CATEGORY BREAKDOWN

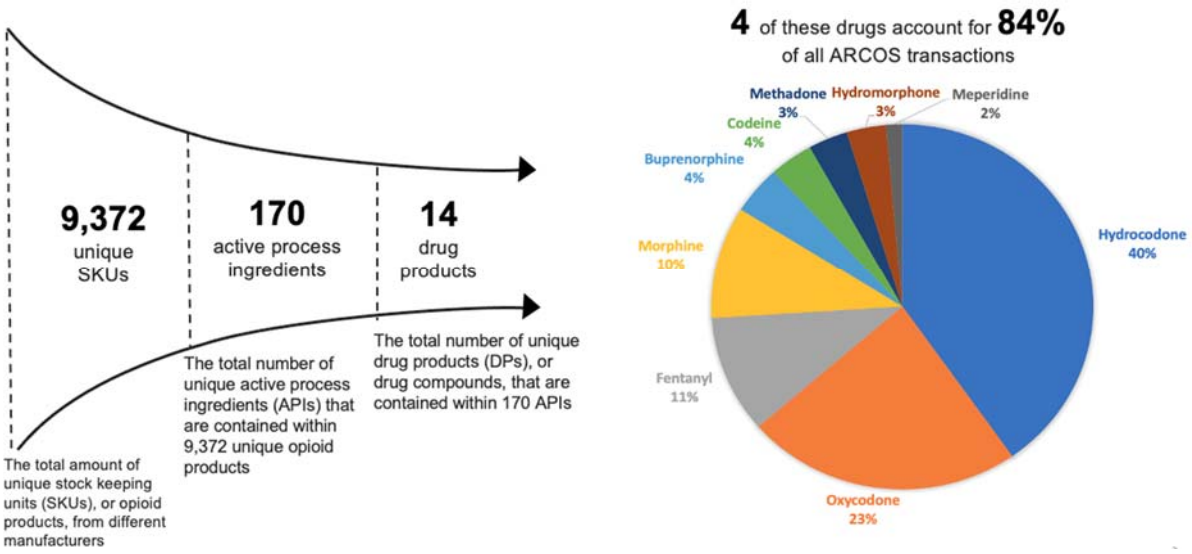


FIGURE 4: SAMPLING BUYERS FROM THE ARCOS DATASET

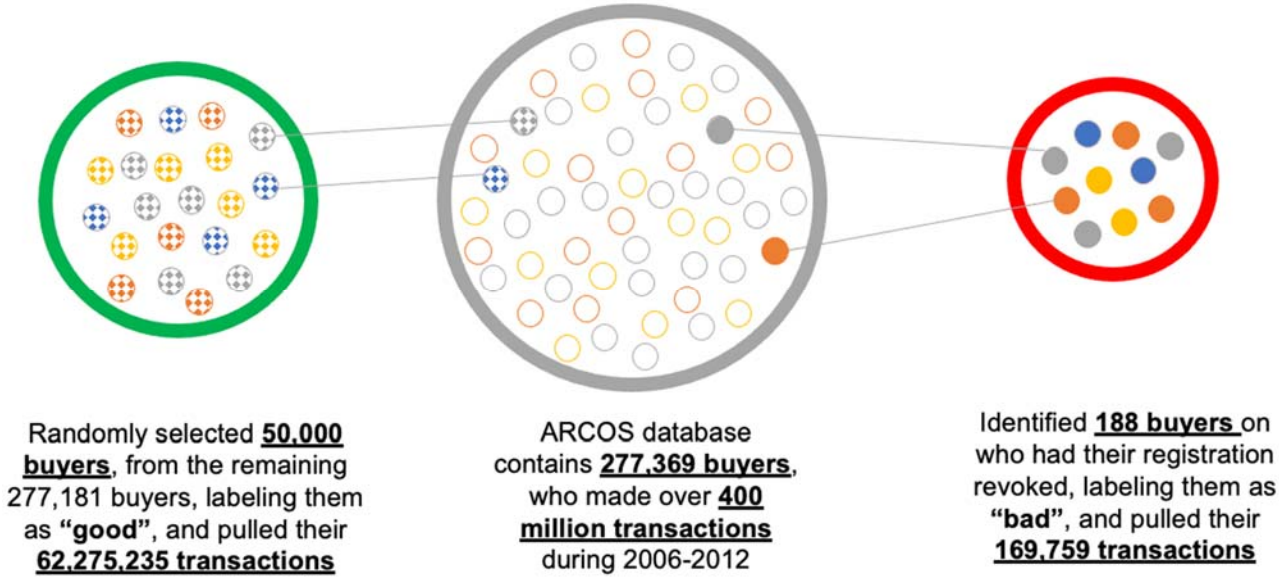


FIGURE 5: SUSPICIOUS BUYERS OVERLAID ON THE US MAP