Exploring the Causal Relationships between Initial Opioid Prescriptions and Outcomes

Jinghe Zhang, MS,[1,2] Vijay Iyengar, PhD,[2] Dennis Wei, PhD,[2] Bhanukiran Vinzamuri, PhD,[2] Hamsa S. Bastani, PhD,[2] Alexander R. Macalalad, MD,[3] Anne E. Fischer, BS,[3] Gigi Yuen-Reed, PhD,[3] Aleksandra Mojsilovic, PhD,[2] Kush R. Varshney, PhD[2]
[1]University of Virginia, Charlottesville, VA; [2]IBM Research, Yorktown Heights, NY; [3]IBM Watson Health, Cambridge, MA

Abstract

Addiction to opioid medications has become an epidemic affecting millions of Americans; drug overdoses accounted for more deaths than car crashes or gun violence in 2015. Oftentimes, the addiction begins when patients continue to take opioids for pain management beyond the required need. To combat the epidemic, it is vital to understand the exact circumstances under which medically-sanctioned treatments can devolve into addiction. In this research, we analyze the relationship between factors surrounding an initial opioid prescription and subsequent long-term use or addiction. In particular, we focus on causal relationships between characteristics of the initial opioid prescription and the outcome of long-term use or seeking addiction treatment among commercially insured individuals. The analytical results indicate that longer days of supply and synthetic opioids increase the risk of unfavorable outcomes. Ultimately, the findings of this research could provide knowledge to the public and guide physicians in opioid prescribing decisions.

Introduction

The United States is in the midst of an unprecedented opioid epidemic. Since 1999, both the amount of opioids sold and deaths from opioid overdose have nearly quadrupled[1]. In 2015, over 33,000 people died from opioid overdose and half of the deaths involved prescription opioids[2]. The Centers for Disease Control and Prevention (CDC) estimated the total economic burden of prescription opioid abuse to be $78.5 billion a year, including healthcare costs, lost productivity, and criminal justice involvement[3]. Oftentimes, the addiction begins when patients continue to take opioids for pain management beyond the required need. Physicians could be indirectly contributing to the epidemic by overprescribing opioids and one of the reasons is because they have generally lacked data to guide opioid prescribing decisions[4].

The ultimate goal of this research is to inform physicians of potential risk factors in opioid prescriptions, not only to avoid overprescribing, but also to help physicians make appropriate prescribing decisions. Thus, we investigate relationships among opioid prescriptions, long-term use, seeking addiction treatment, and other factors as represented in health insurance claims and enrollment data. As the first step, the focus of this work is the causal relationships between characteristics of the initial opioid prescription and the outcome of long-term use or seeking addiction treatment.

Related Work

The opioid epidemic in the United States has attracted tremendous attention by the news media who have reported many facts and consequences of the crisis. To combat the epidemic, CDC has published a few studies on opioid abuse and overdose in the nation[5] as well as guidelines for prescribing opioids for chronic pain[6]. A recent article discusses the correlation between the characteristics of the initial opioid episode and long-term use, focusing on the days of supply, cumulative dose, and categories of the initial opioid prescriptions such as long-acting and oxycodone short-acting[7]. Additionally, other researchers have explored the prediction of opioid addiction based on the patients' demographic factors, health conditions, history of drug abuse, and clinical and medical utilization measures[8–10]. However, little research has been done to study the causal relationships between opioid prescriptions and outcomes using large-scale observational health data, which could be more directly used to guide the prescribing decisions.

However, observational data often introduce treatment-selection bias by confounding factors when used in causal inference studies. Propensity score models have been widely used to address this issue in various fields, especially in epidemiological studies[11–16]. Many variants have been proposed to adjust for treatment assignment, such as weighting and stratification[17–19]. For example, propensity score modeling is used to understand the survival benefits of preoperative β-blocker use in patients undergoing coronary artery bypass graft surgery[12]. Hence, we also adopted propensity

score methods to study the causal relationships between the opioid prescriptions and the outcomes.

Table 1: Population characteristics of all patients and two cohorts with different outcomes

Characteristic	**All Patients** ($n = 205,468$)(%)	**Class I** ($n = 9,329$)(%)	**Class II** ($n = 196,139$)(%)
Age (yr), mean ±SD	38.3 ± 16.0	44.4 ± 12.9	38.0 ± 16.1
Male sex	98,110(48)	4,174(45)	93,936(48)
Region			
Northeast	31,108(15)	1,132(12)	29,976(15)
North central	44,685(22)	2,065(22)	42,620(22)
South	90,447(44)	4,449(48)	85,998(44)
West	39,228(19)	1,683(18)	37,545(19)
Industry			
Oil & gas extraction, mining	2,767(1)	130(1)	2,637(1)
Manufacturing	47,370(23)	2,102(23)	45,268(23)
Transportation, communication, utilities	24,006(12)	1,063(11)	22,943(12)
Retail trade	6,545(3)	341(4)	6,204(3)
Finance, insurance, real estate	19,220(9)	703(8)	18,517(9)
Services	31,841(15)	1,322(14)	30,519(16)
Agriculture, forestry, fishing	165(0.1)	9(0.1)	156(0.1)
Construction	551(0.3)	27(0.3)	524(0.3)
Wholesale	1,782(1)	83(1)	1,699(1)
Unknown	71,221(35)	3,549(38)	67,672(35)
Rural	30,981(15)	1,673(18)	29,308(15)
Healthcare utilization (# visits/yr), mean ±SD	7.7 ± 9.4	10.4 ± 11.7	7.6 ± 9.2
Procedures associated with initial opioid prescription			
Musculoskeletal surgery	30,722(15)	1,404(15)	29,318(15)
Anesthesia services	8,946(4)	302(3)	8,644(4)
Minor skin & breast surgery	3,003(1)	130(1)	2,873(1)
Psychotherapy	2,154(1)	76(1)	2,078(1)
Arthrocentesis, large joint	1,910(1)	68(1)	1,842(1)
Incision & drainage of cyst	1,693(1)	195(2)	1,498(1)
Diagnoses associated with initial opioid prescription			
Calculus of urinary tract	32,831(16)	1,964(21)	30,867(16)
Diabetes mellitus, type II	15,264(7)	850(9)	14,414(7)
Bursitis	5,196(3)	373(4)	4,823(2)
Spinal and back disorders: low back	7,134(3)	122(1)	7,012(4)
Spinal and back disorders: cervical	4,102(2)	162(2)	3,940(2)
Osteoarthritis, except spine	4,928(2)	139(1)	4,789(2)
Osteoarthritis, cervical spine	2,970(1)	399(4)	2,571(1)
Depression	3,911(2)	249(3)	3,662(2)
Injury: spine & spinal cord: low back	2,596(1)	116(1)	2,480(1)
Injury, open wound, or blunt trauma: abdomen or trunk	1,697(1)	58(1)	1,639(1)
Intervertebral disc disorders: lumbar & lumbosacral	2,147(1)	194(2)	1,953(1)
Arthropathies, bone & joint disorders	1,253(1)	127(1)	1,126(1)
Headache	926(0.5)	88(1)	838(0.4)
Daily MME*, mean ±SD	47.6 ± 67.4	46.1 ± 67.5	47.6 ± 67.4

* MME: morphine milligram equivalent.

Data Description

In this research, we use the health insurance claims and enrollment data of 5 million individuals from January 2011 to December 2015, a sample from the Truven Health Analytics MarketScan® Commercial Database containing 80 million enrollees. Patients with at least three years of continuous enrollment are selected in order to get complete claims information. Since the focus of this analysis is the initial opioid prescription, a "wash in" period is defined to ensure that selected patients had no observed opioid addiction history at least 6 months prior to the initial opioid

prescription, and patients with a diagnosis of cancer in this time range are also excluded. In this work, we define two outcome classes: Class I (continued opioid use after 1 year or sought addiction treatment) and Class II (discontinued opioid use). Here, discontinued opioid use is defined as opioid-free days ≥ 180 and the final selected patients must have more than 1 year of continuous enrollment after the initial opioid prescription to allow observation of the outcome. In the end, there are $9,329$ and $196,139$ distinct patients in Class I and II, respectively. Table 1 presents the population characteristics of all patients, as well as the two cohorts with different outcomes.

Propensity Score Modeling & Weighting

This study focuses on drug classes and days of supply of the initial opioid prescriptions and their relationships with the two treatment outcomes. According to the mode of synthesis, there are three classes of opioids: natural, semi-synthetic, and synthetic[20,21]. Thus, four initial treatment types are constructed: 1) natural or semi-synthetic opioids for less than or equal to 7 days; 2) synthetic opioids for less than or equal to 7 days; 3) natural or semi-synthetic opioids for greater than 7 days; and 4) synthetic opioids for greater than 7 days. Here, natural and semi-synthetic opioids are grouped together since the latter ones are derived from naturally occurring opiates and opium alkaloids[22]. Figure 1 presents the percentages of the four treatments in the two outcome classes. Treatment 1 is the most prevalent in both classes, while opioids with longer days of supply appears to be more common in Class I than in Class II.

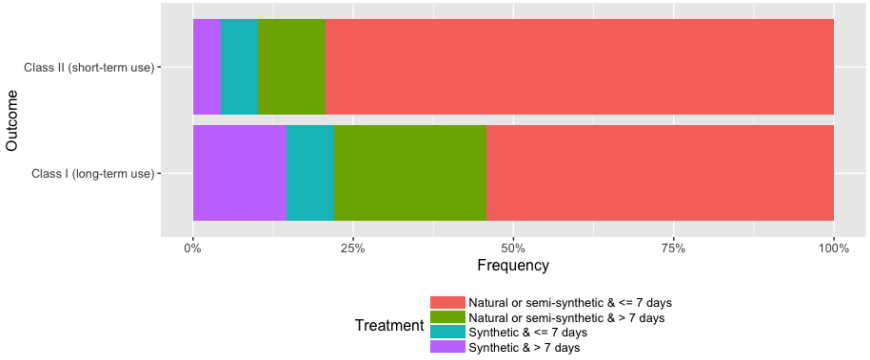
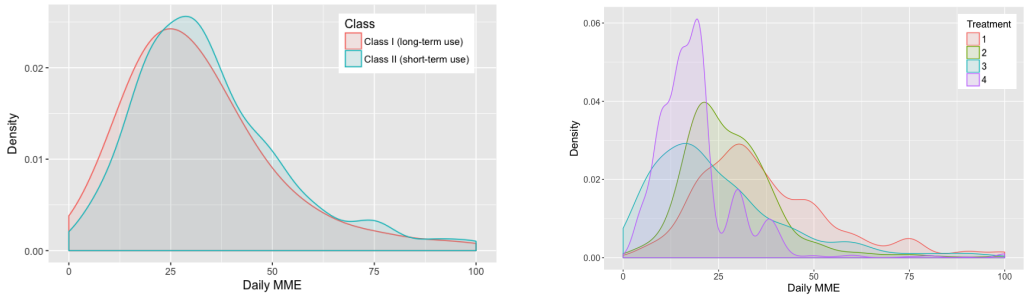


Figure 1: Treatment frequencies in two outcome classes

To allow the analysis of causal relationships using observational data, we balance confounding covariates between different treatment groups prior to average treatment effect estimation. To ensure that the effects of the treatments are not biased by the confounding factors, we first select the covariates related to both outcome and treatment. The confounders include non-medical variables, such as age, gender, industry, region, rural location, and average number of visits per year as a surrogate for healthcare utilization, as well as some medical events. We select the medical confounders by their frequencies in the population as well as lasso logistic regression[23], which models the likelihood of outcome Class I given the medical events associated with the initial opioid prescription. The selected medical confounders are: incision & drainage of cyst, musculoskeletal surgery, spinal and back disorders – low back, injury (open wound or blunt trauma), calculus of urinary tract, and so forth, listed as diagnoses and procedures in Table 1. Additionally, some general symptoms and factors influencing health status are selected by the lasso model, such as generalized pain and tobacco use.

Furthermore, we introduce the daily morphine milligram equivalent (MME), i.e., cumulative MME of the initial opioid prescription normalized by days of supply, as an additional confounder. Here, cumulative MME is computed based on the quantity of the prescribed opioid and its conversion factor to morphine. Figure 2 illustrates the densities of daily MME by outcome and treatment, respectively. There are significant differences in the daily MME distributions across the four treatment groups. Thus, daily MME is added as a confounder to balance between treatment groups.

With the selected confounders, we build a propensity score model based on the gradient boosting machine (GBM) using the twang package in R to estimate a patient's probability of receiving a certain treatment[24,25]. Then, we introduce the inverse-probability of treatment weighting (IPTW) to reduce confounding due to observed covariates[18]. Standardized mean differences of confounders between treatment groups, i.e., mean difference normalized by standard



(a) Daily MME by outcome class

(b) Daily MME by treatment group

Figure 2: Density Plots of Daily MME

deviation of the pooled sample, are used to measure the balancing results. The mean and maximum standardized mean differences across all confounders are 0.02 and 0.15, respectively. Furthermore, Figure 3 presents the density plots of the standardized mean differences before and after balancing. Accordingly, there is little difference in confounding variables between treatment groups after balancing[18,26].

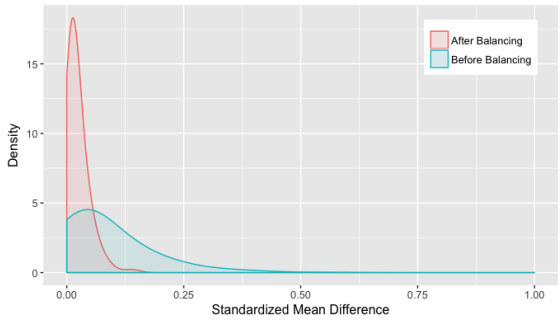


Figure 3: Density plots of standardized mean difference before and after balancing

Average Treatment Effect Estimation & Interpretation

With the weighted population, we build a logistic regression model to estimate the probability of outcome class I given a certain treatment. The estimates as well as the 95% and 99% confidence intervals (CIs) of the probabilities are presented in Table 2.

Table 2: Four treatments and the corresponding probabilities of Class I

Treatment	**Size**	**Probability**	95% **CI**	99% **CI**
Natural or semi-synthetic & $<= 7$ days	$160,563$	0.033	$(0.032, 0.034)$	$(0.032, 0.035)$
Synthetic & $<= 7$ days	$12,225$	0.052	$(0.047, 0.058)$	$(0.045, 0.059)$
Natural or semi-synthetic & > 7 days	$22,862$	0.095	$(0.090, 0.100)$	$(0.088, 0.102)$
Synthetic & > 7 days	$9,818$	0.142	$(0.122, 0.161)$	$(0.116, 0.167)$

According to Table 2, longer days of supply of the initial opioid prescription is a driving force of long-term use. Furthermore, there is a significant difference in average treatment effect between synthetic opioids and natural or semi-synthetic ones.

- In patients who are given synthetic opioids for no more than 7 days, using natural or semi-synthetic opioids instead could potentially reduce the risk of long-term use or addiction by 36.5%.
- For natural or semi-synthetic opioids, a shorter days of supply could potentially reduce the risk by 65.3% in patients with longer days of supply.

With the weighted population, we also build a logistic regression model with confounding variables included and the corresponding average treatment effects are consistent with the estimates shown in Table 2.

Conclusions & Future Work

In this research, we studied the causal relationships between characteristics of the initial opioid prescription and the outcomes of long-term use and seeking addiction treatment using health insurance claims data. The analytical results indicate that longer days of supply and/or synthetic opioids are more likely to lead to unfavorable outcomes in a commercially insured population. This work focused on the initial opioid prescription and the associated diagnoses and procedures, while other medical events that occurred afterwards might also contribute to the outcomes and will be considered as post-treatment confounders in our future work. Moreover, we will investigate later opioid prescriptions in addition to the initial ones and their relationships to the outcomes. One limitation of propensity score methods is that they can only adjust for observed confounding covariates. Thus, more sophisticated methods, such as latent variable models, will be explored to reduce the bias of such unobserved confounders. Ultimately, the findings of this research could provide knowledge to the public and could be used to help physicians in opioid prescribing decisions.

Acknowledgement

The authors would like to thank our colleagues Aliza Heching, Adam Perer and Jing Sha at IBM Research, and Yajuan Wang and Mustafa Karakus at IBM Watson Health for valuable discussion, feedback and support. This work was conducted under the auspices of the IBM Science for Social Good initiative.

References

1. Centers for Disease Control and Prevention. Drug overdose deaths in the United States continue to increase in 2015. [Online]. Available from: https://www.cdc.gov/drugoverdose/epidemic/index.html. [Accessed July 28, 2017].

2. Centers for Disease Control and Prevention. Drug overdose. [Online]. Available from: https://www.cdc.gov/drugoverdose/index.html. [Accessed July 28, 2017].

3. Harold A Pollack. So prescription opioid disorders are a $78.5 billion problem. *Medical care*, 54(10):899, 2016.

4. Daniel Cook. Post-op opioids overprescribed at Mayo Clinic. [Online]. Available from: http://www.outpatientsurgery.net/newsletter/eweekly/2017/07/183. [Accessed July 28, 2017].

5. Rose A Rudd. Increases in drug and opioid-involved overdose deathsUnited States, 2010–2015. *MMWR. Morbidity and mortality weekly report*, 65(50-51), 2016.

6. Deborah Dowell, Tamara M Haegerich, and Roger Chou. CDC guideline for prescribing opioids for chronic pain United States, 2016. *JAMA*, 315(15):1624–1645, 2016.

7. Anuj Shah, Corey J Hayes, and Bradley C Martin. Characteristics of initial prescription episodes and likelihood of long-term opioid use-United States, 2006-2015. *MMWR. Morbidity and mortality weekly report*, 66(10):265, 2017.

8. Dennis C Turk, Kimberly S Swanson, and Robert J Gatchel. Predicting opioid misuse by chronic pain patients: a systematic review and literature synthesis. *The Clinical journal of pain*, 24(6):497–508, 2008.

9. Peer Exchange. Understanding predictors of opioid abuse: predictive model development and validation. *Am J Pharm*, 6(5):208–216, 2014.

10. Bryan N Cochran, Annesa Flentje, Nicholas C Heck, Jill Van Den Bos, Dan Perlman, Jorge Torres, Robert Valuck, and Jean Carter. Factors predicting development of opioid use disorders among individuals who receive an initial opioid prescription: mathematical modeling using a database of commercially-insured individuals. *Drug and alcohol dependence*, 138:202–208, 2014.

11. Sherry Weitzen, Kate L Lapane, Alicia Y Toledano, Anne L Hume, and Vincent Mor. Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiology and drug safety*, 13(12):841–853, 2004.

12. T Bruce Ferguson Jr, Laura P Coombs, Eric D Peterson, and for the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. Preoperative β-blocker use and mortality and morbidity following CABG surgery in North America. *JAMA*, 287(17):2221–2227, 2002.

13. Sharon-Lise T Normand, Mary Beth Landrum, Edward Guadagnoli, John Z Ayanian, Thomas J Ryan, Paul D Cleary, and Barbara J McNeil. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. *Journal of clinical epidemiology*, 54(4):387–398, 2001.

14. Carisi A Polanczyk, Luis E Rohde, Lee Goldman, E Francis Cook, Eric J Thomas, Edward R Marcantonio, Carol M Mangione, and Thomas H Lee. Right heart catheterization and cardiac complications in patients undergoing noncardiac surgery: an observational study. *JAMA*, 286(3):309–314, 2001.

15. Sebastian Schneeweiss, Jeremy A Rassen, Robert J Glynn, Jerry Avorn, Helen Mogun, and M Alan Brookhart. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology (Cambridge, Mass.)*, 20(4):512, 2009.

16. Ralph B D'Agostino. Propensity scores in cardiovascular research. *Circulation*, 115(17):2340–2343, 2007.

17. Peter C Austin and Elizabeth A Stuart. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in medicine*, 34(28):3661–3679, 2015.

18. Melissa M Garrido, Amy S Kelley, Julia Paris, Katherine Roza, Diane E Meier, R Sean Morrison, and Melissa D Aldridge. Methods for constructing and assessing propensity scores. *Health services research*, 49(5):1701–1720, 2014.

19. Paul R Rosenbaum and Donald B Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38, 1985.

20. News Medical. Opioid types. [Online]. Available from: https://www.news-medical.net/health/Opioid-Types.aspx. [Accessed August 10, 2017].

21. Hasan Pathan and John Williams. Basic opioid pharmacology: an update. *British Journal of Pain*, 6(1):11–16, 2012.

22. Christopher E French, Anne M Hailes, Deborah A Rathbone, Marianne T Long, David L Willey, and Neil C Bruce. Biological production of semisynthetic opiates using genetically engineered bacteria. *Nature Biotechnology*, 13(7):674–676, 1995.

23. Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

24. Lane Burgette, Beth A Griffin, and Dan McCaffrey. Propensity scores for multiple treatments: A tutorial for the mnps function in the twang package. *R package. Rand Corporation*, 2017.

25. Jerome H Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 29(5):1189–1232, 2001.

26. Elizabeth A Stuart, Brian K Lee, and Finbarr P Leacy. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of clinical epidemiology*, 66(8):S84–S90, 2013.