Predicting Post-surgical Opioid Consumption using Perioperative Surgical Data

by

Justin Yu

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2020

Author . . . .
Department of Electrical Engineering and Computer Science

Certified by . . . .
Peter Szolovits
MIT
Thesis Supervisor

Accepted by . . . .
Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Predicting Post-surgical Opioid Consumption using Perioperative Surgical Data

by

Justin Yu

Submitted to the Department of Electrical Engineering and Computer Science
on May 20th, 2020, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Improper consumption of prescription opioids is a massive public health issue in the United States currently. Here, we propose one approach of tackling this issue through using machine learning techniques to predict opioid consumption post discharge for surgical patients. Through the data collected from surgical patients at BIDMC, relevant features will be identified and used to predict if patients high, outlier consumption. Using logistic regression and gradient boosted decision trees, model performance were evaluated at AUCs of 0.7270 and 0.7289 respectively.

Thesis Supervisor: Peter Szolovits
Title: MIT

‘

Acknowledgments

Thank you to Dr. Gabriel Brat and Dr. Pete Szolovits for being the most wonderful of supervisors. I could not have done this without your patient guidance and insight.

To my housemates Nick, Luis, Haris, Evan, Henry, and Tarun, thank you for keeping me sane and for a wonderful year. To me friends far and near, thank you for the memories and those to come. Game nights over Zoom were certainly quite the time.

Although four years of undergrad was a long time, I felt like there was still so much more to learn and do before leaving this wonderful institution. Taking the time to do my M.Eng has been worth every minute. Life is certainly full of surprises and not being able to thank the people who have made these past few years possible, in person, stings. However, I know that we will find the time to reconnect and meet again in the future. So until then, please take care and always take the time to be kind to yourself and those around you.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

1.1.1 US Opioid Situation

The current opioid crisis in the United States is dire but not without hope. Over the course of a year, there are over 10 million misusers of prescription opioids [Substance Abuse and Mental
The estimated total cost of opioid misuse in the US is $78.5 billion a year due to a myriad of factors including, but not limited to, loss of productivity, substance abuse treatment, and general healthcare needs [Florence *et al.* , 2016]. One of the driving factors of this problem is over-prescription of opioids. Over-prescribing can have deleterious effects on not only patients themselves but also those around them.

Around 53% of prescription opioid misusers obtain the drugs from friends and family [Substance Abuse and Mental Health Services Administration., 2018]. This may be due to consumption of leftover pills that are not disposed of properly. Whatever the case, a majority of misusers are not even direct recipients of the prescription drugs. For our study, we will be focusing on the 9.8% of prescription opioids that go to surgical patients [Levy *et al.* , 2015].

1.1.2 Addressing the Problem

It is only recently that real scrutiny into prescribing and usage practices has begun. It was oftentimes unclear where exactly the responsibility for the opioid crisis fell amongst the many stakeholders. Healthcare providers, policy makers and pharmaceutical representatives all had a role in the process, but who was ultimately responsible for affecting change? In short, there must be more accountability from everyone involved. Cases like Oklahoma v. Johnson & Johnson, which ended up with a $572 million dollar payout, are opening the doors for enforcing more responsible practices from manufacturers. At the prescriber level, more must be done as well. Providers must understand the dangerous long-term potential effects that over-prescribing can have and balance that with the risks of under-prescribing. On the policy end, efforts such as the Prescription Drug Monitoring Program (PDMP) along with other changes saw a greater than 50% decrease in oxycodone overdose deaths in Florida in 2012 [CDC, 2019]. This shows that efforts to responsibly curb opioid misuse can have strong, positive effects on communities.

1.1.3 Opioids in Patient Care

Current Guidelines

Current opioid prescription guidelines provide insufficient specific guidance for providers to follow. According to a study published in 2018, the American Pain Society does not provide any recommendations on determining the length of a prescription at all, while The Institute for Clinical Systems Improvement only suggests limiting initial prescriptions across all surgeries [Scully *et al.* , 2018]. Neither of these sources take into account a patient's medical history or in-hospital trajectory.

However, a number of studies over the past years have begun to investigate what best practices might look like. Specifically, we need guidance based on empirical evidence that takes into account the individual characteristics of each patient as well as the surgery that they underwent.

In a 2016 study, post-surgical opioid consumption was assessed in two cohorts: one with no intervention and a second where providers were given recommendations for reducing opioid prescription sizes. They found a 63% reduction in prescription sizes with no corresponding increase in refill rates [Howard *et al.* , 2018]. The key observation here is that there was no corresponding increase in refill rates. Refills indicate that the initial prescription size was insufficient for pain management. While there are many reasons why a patient may choose not to refill their medication, this observation suggests that one can safely reduce prescription sizes without compromising patient pain management.

Effects of Prescription Size

Opioids are often prescribed with instructions to take pills 'as needed' or PRN. However, what patients perceive to be necessary for adequate pain control may be affected by the size of the prescription they were given even after controlling for other variables. The evidence points toward a positive association between the quantity of opioids prescribed and the quantity consumed with roughly 0.53 more pills consumed per unit increase in pills prescribed [Howard *et al.* , 2019]. The study also indicated that as many as 92% of patients have leftover opioid pills, suggesting that the issue of increased consumption after higher prescribing affects a majority of surgical patients. In effect, at least 92% of patients could have had reduced prescription sizes without compromising their pain management.

In another study population, it was found that a longer supply period of prescribed opioids increases the likelihood of adverse outcomes such as addictive use [Zhang *et al.* , 2017]. Specifically, they identified patients that were either still consuming opioids a year later or had sought treatment for addiction. Thus, effects of higher consumption may last far beyond their surgical recovery period.

The evidence points towards a need for more accurate initial prescribing of opioids. This is not only beneficial for the patient but also reduces the chance that extra pills make their way to other consumers.

Under-prescription

Under-prescription, however, is not without potential risks. Broadly speaking, inadequate acute pain management may increase the likelihood of chronic pain issues later on [Sinatra, 2010]. While refills are a possibility, they require extra effort on the patient's part. This may not be ideal for certain patients recovering from surgeries.

Under-prescription also disproportionately affects certain ethnic minorities due to implicit biases [Santiago, 2019]. This further points to a need for a more standardized system for prescribing to also avoid cases of under-managed pain.

Chapter 2

Related Work

2.1 Related Work

2.1.1 Claims Data Predictions

One recent study looked at predicting the likelihood of opioid overdose using patient data and medical claims for over 550,000 patients. Specifically, they looked through claims data for overdose related ICD codes in the 3 month period following initial opioid prescription. Using gradient boosted machines (GBM) and deep neural nets (DNN), they performed risk prediction and stratification. For predicting opioid overdoses, GBM and DNN performed with AUC scores of 0.90 and 0.91 respectively.

One detail of note is that this data does not include consumption data. One of our aims is to determine the impact of collecting and analyzing consumption data in addition to prescription data. However, it does show that there are identifiable features that are correlated with a patient's trajectory post-discharge. From the GBM model, features such as total milligram morphine equivalents (MME) prescribed, history of substance use disorder, and age were among the strongest predictors.

2.1.2 Factors Affecting Opioid Use

To begin to develop better prescribing guidelines, it is necessary to determine which factors affect opioid consumption needs.

Demographics and Comorbidities

Different demographic factors were found to be correlated with higher or lower consumption in one study [Howard *et al.* , 2019]. Each year increase in age was associated with 0.77 fewer oral morphine equivalents (OMEs), while biological sex had no observed association. Tobacco users used 21.6 more OMEs, while obese patients used 7.38 more OMEs when compared with non-obese patients. So it appears that understanding a patient's demographics and comorbidities may give us insight into their opioid consumption levels.

Behavioral

In a 2018 study, behavioral factors were assessed in relation to chronic opioid use after orthopedic surgery [Rhon *et al.* , 2018]. They found that factors such as high health-seeking behavior, insomnia and a history of mental health disorder were associated with chronic opioid usage. Given the problem that high-health seeking behavior plays, providers may develop implicit biases that affect how they prescribe opioids for patients. Thus, it is important to understand how all of these factors come into play together.

Prior Pharmaceuticals Use

A 2019 study investigated the relationship between patient characteristics and adverse opioid related outcomes within a 5 year period of the initial prescription using a patient cohort of over 80,000 members [Hastings *et al.* , 2020]. Specifically, these were outcomes such as opioid poisoning, opioid abuse, opioid dependence, or treatment for any of these conditions. Using an approach of finding odds ratios using bootstrapped logistic regression, prior prescriptions for benzodiazepines, centrally acting skeletal muscle relaxants, and opiate agonists increased the odds of an adverse outcome by over 1.25 times.

Chapter 3

Methods

3.1 Data Collection and Preprocessing

3.1.1 Data Source

Data were obtained from three main sources: our own study, the BIDMC pharmacy, and an ICD codes database. The data dictionary is available in the appendix in Table A.1. All data collection and use were approved by BIDMC IRB.

Study Source

The first data source was from data previously collected by our group from 3322 surgical patients at Beth Israel Deaconess Medical Center (BIDMC), from October 2017 through June 2018. Out of these, 1984 later opted into surveys about their opioid consumption post-discharge. This source included information such as patient demographics, medication history, medical history, surgical characteristics, and post-discharge prescribed and consumed opioid quantities.

Pharmacy Data

The second source was the BIDMC pharmacy, which included opioid orders data used as a proxy for in-hospital consumption. This source included orders for opioids listed by medical record numbers and admission dates. Each entry contains the National

NDC	Units billed	Strength per unit	Unit of measure	Generic Name	MME conversion factor
904644461	8	5	mg	Oxycodone HCl	1.5

Table 3.1: Here we see eight 5mg tablets of oxycodone ordered. This is 40mg in total. Per the CDC guidelines for opioids research, 1mg of oxycodone can be converted to 1.5 MMEs. Thus, this entry translates to 60 MMEs total.

Drug Code (NDC), generic name, strength per unit, unit of measure, and MME conversion factor. An example of how the total MMEs was then calculated is shown below for an abbreviated sample entry.

ICD Data

International Classificiation of Diseases (ICD) codes were obtained through the BIDMC Clinical Data Repository's Clinical Query 2 (CQ2) tool. These contained both older entries in ICD-9 format and newer ones in ICD-10 format. For each entry, a medical record number, date assigned and ICD code were given.

3.1.2 Surgical Categorization

Surgeries were grouped into categories if they were determined to be the same or similar procedures. For example, central pancreatectomies and distal pancreatectomies were grouped as pancreatectomies. Our intent was to group together similar surgeries to be able to best compare patient opioid consumption. Certain surgeries are associated with higher levels of post-operative pain, while others are more minor with faster recoveries. As it would only make sense to compare patients that received similar procedures, manual categorization of all of the patients was completed. 79 unique surgical categories were established. The full list of surgical categories with counts is shown in the appendix figure A.2.

We then took the categories and found the median amount of consumed opioids in milligram morphine equivalents (MMEs). We labeled categories as low if their median was below 100, mid if their median was between 100 and 200, and high if their median MME consumed was over 200. This categorization was then included as a feature.

3.1.3 Sample Inclusion/Exclusion

After excluding 1338 patients that did not complete post-discharge opioid consumption surveys, 1984 patients remained. 113 of them were dropped if they belonged to surgical categories that had fewer than 10 patients. This was done to allow for proper statistical comparison. 4 were dropped due to missing pre-operative opioid consumption data. Finally, 1867 were entered into our analysis as shown in Table 3-1.

To determine if there were significant differences between our included and excluded cohorts, we looked at the base demographic information for both groups, shown in Table 3.2. Since the age distributions are not necessarily normal, we use a non parametric test, the Mann-Whitney U Test, to compare them. Comparing the excluded and included groups, we see that there is a difference between the age distributions, with excluded patients being slightly younger. Looking at gender and race, there were no statistically significant differences between the two groups using Fisher Exact Test. Knowing this, we have a better understanding of how applicable our results may be to the broader surgical population of BIDMC.

3.1.4 Data Imputation

21.1% of our patients were missing BMI data. BMI data for these patients were imputed using K-nearest neighbors with k = 5. While our original data included exact, numerical values for BMI, we converted these to broader BMI categories as per generally accepted standards shown in Table 3.3

3.1.5 Outcome Classes

Patients were then grouped into two outcome classes: those with post-discharge opioid consumption less than or equal to the 75th percentile for their surgery category, and

Characteristic	**Included in Study**	**Excluded from Study**
N	1867	1488
Age*	60 [47 - 69]	57 [45 - 66]
Gender		
Female	890 (47.67)	696 (46.77)
Male	940 (50.35)	764 (51.34)
Other	37 (1.98)	28 (1.88)
Race		
White	1457 (78.04)	1122 (75.4)
Black	169 (9.05)	146 (9.81)
Asian	47 (2.52)	26 (1.75)
American Indian / Alaska Native	5 (0.27)	11 (0.74)
Native Hawaiian / Other Pacific Islander	20 (1.07)	26 (1.75)
Other	37 (1.98)	36 (2.42)
Unknown	132 (7.07)	121 (8.13)

Table 3.2: Results are shown as *median [25th percentile – 75th percentile]* for continuous variables and N (%) for categorical variables (e.g., 890 female patients represent 47.67% of the 1867 patients included in the study). Basic demographics are shown above as a baseline for determining if there are marked differences between the included and excluded patient sets. Mann-Whitney U Test suggests a statistically significant difference between the age distributions of those included and excluded from the study. No significant differences were found within gender and race using Fisher Exact Test.

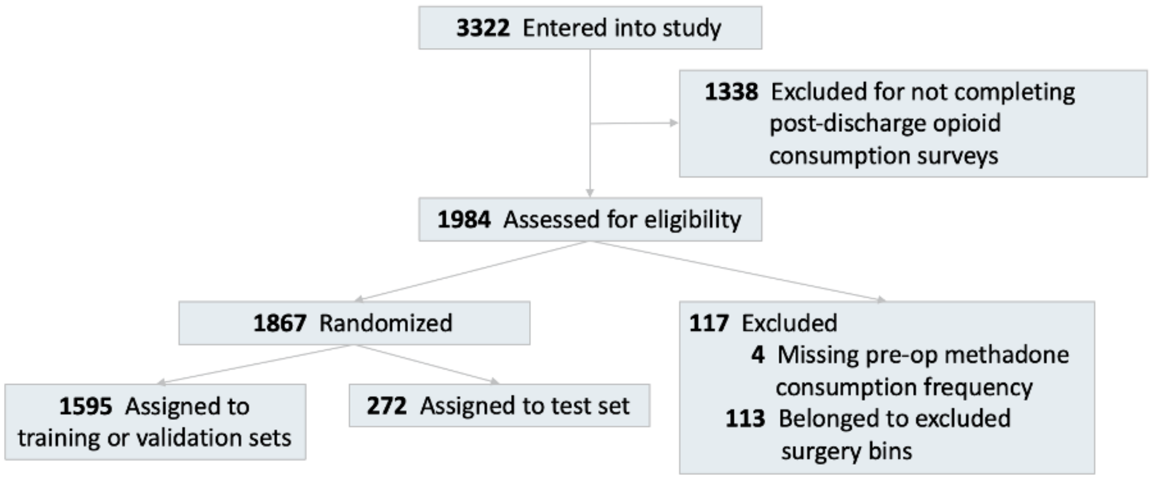

Figure 3-1: Dataflow of patient inclusion / exclusion

those with consumption greater than that. The motivation for identifying patients that will have higher consumption is to eventually be able to prescribe those patients larger prescription sizes while reducing prescription sizes for those patients who will not benefit from that. If there is no way to distinguish between these patients, they will likely receive larger prescriptions so as to reduce the chance of under-prescribing, which could lead to unmanageable pain and readmission to the hospital.

BMI Category	**BMI Range**
Underweight	$<$18.5
Normal	18.5 - 24.9
Overweight	25.0 - 29.9
Obese Class I	30.0 - 34.9
Obese Class II	35.0 - 39.9
Obese Class III	$>$40.0

Table 3.3: BMI Categories and corresponding BMI ranges.

3.2 Data Exploration

3.2.1 Characteristics of the Cohort

Our first analysis begins with looking at the distribution of patients across different groupings as seen in Table 3.4. Our aim was to determine if there were any demographics or elements of one's medical history that would be correlated with higher opioid consumption. As seen in the table, the Mann-Whitney U Test found a statistically significant difference between the age distributions of those with lower consumption and those with higher consumption. This is consistent with previous research that shows that older age is correlated with lower opioid consumption [Howard *et al.* , 2019]. We also see that those with pre-operative opioid exposure, current tobacco use, marijuana use, and a history of recreational drug / substance abuse were more likely to belong to the high percentile consumption group. These are all consistent with previous work as well.

3.2.2 Feature trends

To better understand broad patterns and trends in our data, an initial dive into the distribution of consumed MMEs post-surgery between different groupings was completed. boxplots were generated to compare how taken MMEs vary between different groupings. We also removed the highest 2.5% of outliers. In Figure 3-2, we see that the median and 75th percentile values for taken MMEs are lower in people without a mental health history.

In Figures 3-3, 3-4, 3-2, we see the variation of taken MMEs across different categories of patients based on their pre-surgical exposure to opioids, gender and mental health history. These were among the many data exploration graphs generated to determine if there were associations between consumed MMEs and demographic or medical data points.

Characteristic	Total	< 75th percentile for surgery	> 75th percentile for surgery
N	1889	1457	426
Age*	60 [47 – 69]	62 [49 - 70]	54 [43 - 64]
Gender			
Female	890	689 (77.42)	201 (22.58)
Male	940	727 (77.34)	213 (22.66)
Other	37	26 (70.27)	11 (29.73)
Race			
White	1457	1143 (78.45)	314 (21.55)
Black	169	122 (72.19)	47 (27.81)
Asian	47	41 (87.23)	6 (12.77)
American Indian / Alaska Native	5	3 (60.00)	2 (40.00)
Native Hawaiian / Other Pacific Islander	20	15 (75.00)	5 (25.00)
Other*	37	20 (54.05)	17 (45.95)
Unknown	132	98 (74.24)	34 (25.76)
Pre-op Opioid Status			
Naïve	1705	1345 (78.89)	360 (21.11)
Exposed*	155	93 (60.00)	62 (40.00)
Misuse History	7	4 (57.14)	3 (42.86)
Tobacco Use			
None	941	753 (80.02)	188 (19.98)
History	728	570 (78.30)	158 (21.7)
Current*	198	119 (60.10)	79 (39.9)
Marijuana Use			
None	1672	1322 (79.07)	350 (20.93)
Positive*	195	120 (61.54)	75 (38.46)
History of Recreational Drug/Substance Abuse			
None	1797	1398 (77.80)	399 (22.20)
Positive*	70	44 (62.86)	26 (37.14)
Depression			
None	1492	1167 (78.22)	325 (21.78)
Positive	375	275 (73.33)	100 (26.67)

Table 3.4: Results are shown as *median [25th percentile – 75th percentile]* for continuous variables and N (%) for categorical variables. For categorical variables, the first variable listed is the reference class (i.e. Naïve for Pre-Op Opioid Status). Statistically significant differences ($p < 0.05$) to the reference class between the non-outlier and outlier consumption groups are marked with an asterisk. Statistical significance was determined using Fisher Exact Test with Bonferroni Correction for categorical variables and Mann-Whitney U Test for continuous variables.

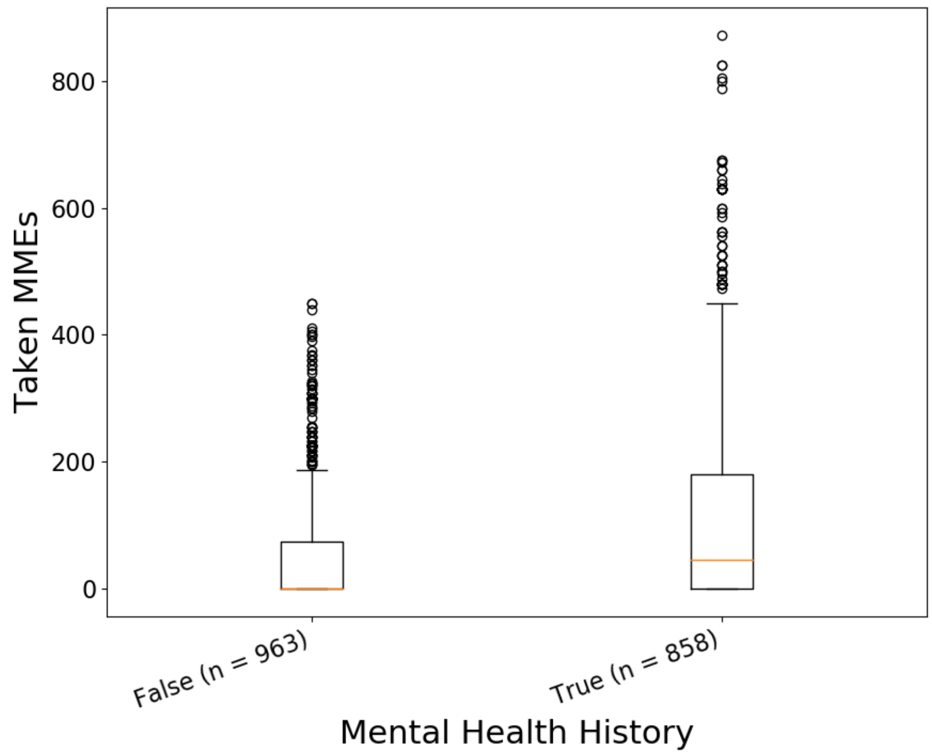

Figure 3-2: Taken MMEs vs mental health history status Top 2.5% of outliers were removed from graph. Outliers are marked with circles above each box plot.

3.3 Statistical Models

With a better understanding of the data trends and a set of established goals, we moved to statistical and machine learning models.

3.3.1 Data Pre-processing

After any data that were collected after discharge were eliminated, data were pre-processed and standardized before being used for analysis.

Continuous Variables

Continuous variables such as age and MMEs consumed in-hospital were standardized to a mean of 0 and standard deviation of 1.

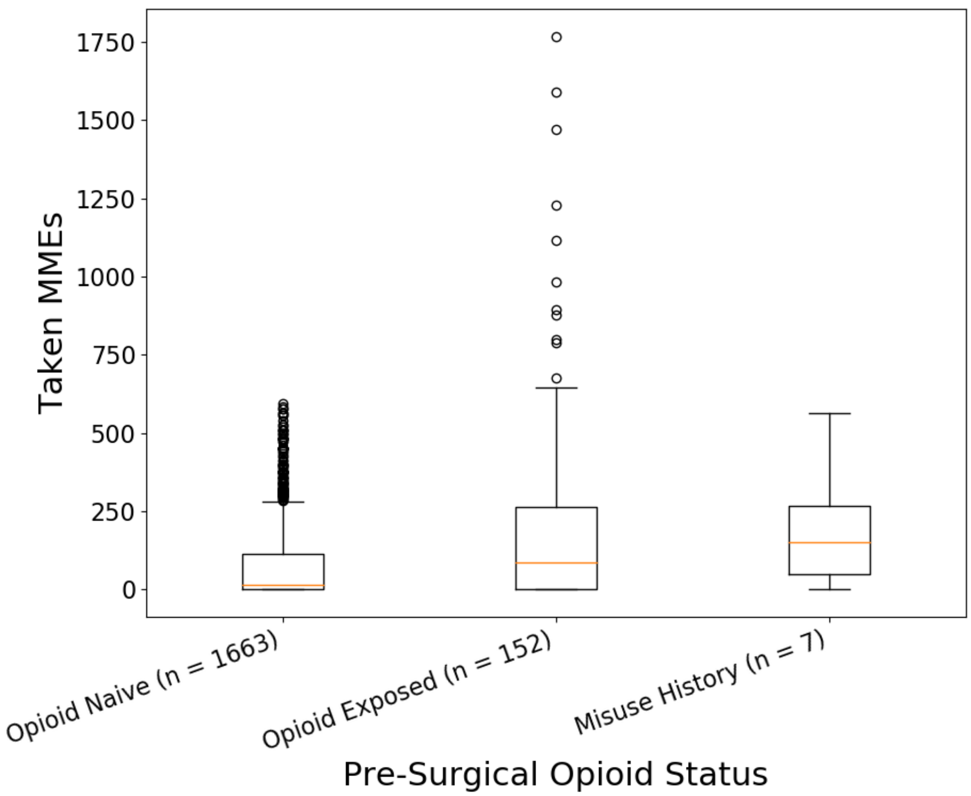

Figure 3-3: Taken MMEs vs pre-surgical opioid status. Top 2.5% of outliers were removed from graph. Outliers are marked with circles above each box plot.

Categorical Variables

For categorical variables such as gender, one-hot encoding was used. Three columns for male, female and other were generated.

Ordinal Variables

For categorical variables that might benefit from being represented as ordinals, different ordinal representations were tested using our training and validation sets. As an example, the feature representing alcohol abuse was represented as follows: 0 for no such history, 1 for a history of alcohol abuse, and 2 for current alcohol abuse.

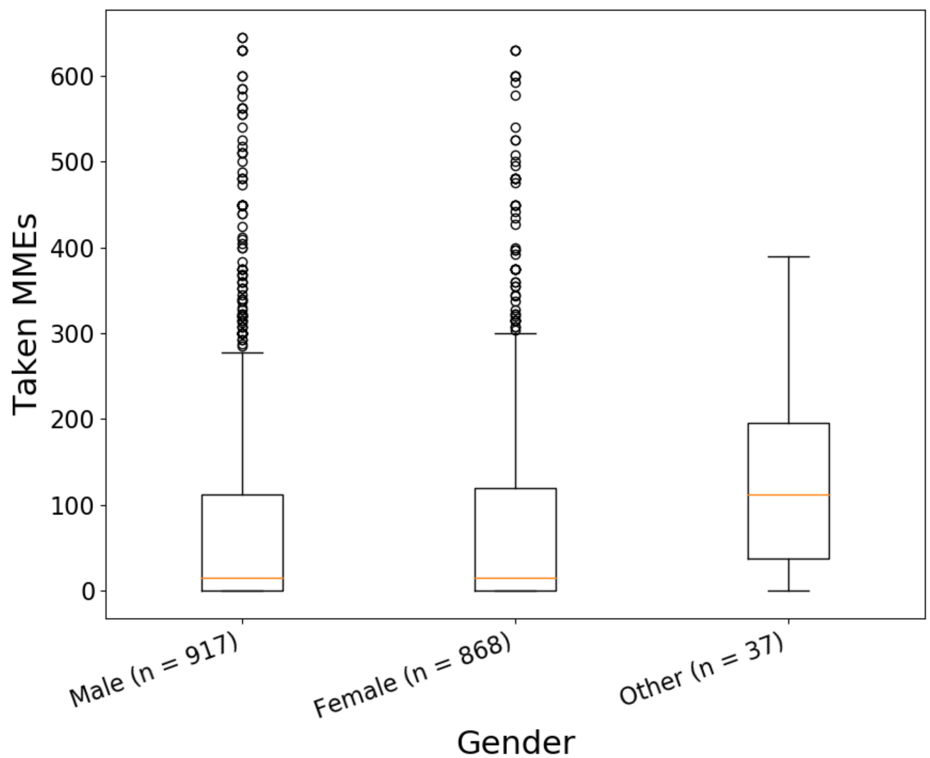

Figure 3-4: Taken MMEs vs gender. Top 2.5% of outliers were removed from graph. Outliers are marked with circles above each box plot.

3.3.2 Multivariate Logistic Regression

For predicting our binary outcome, logistic regression was first used. L1, L2 and L1/L2 mixed penalization were all tested. Different output thresholds were also tested to find an optimal balance between sensitivity and specificity.

Bootstrapping

To calculate odds ratios, bootstrapped logistic regression with 5000 replicates was completed using elastic net regularization. For each replicate, entries were randomly picked from our sample data with replacement to create a training set of the same original size. Coefficients from the logistic regression were converted into odds ratios. Odds ratios were then averaged over 5000 replicates to generate 95% confidence intervals as seen in table 4.1.

3.3.3 Gradient Boosted Decision Trees

To test if there were nonlinearities within our data that would benefit from a more complicated model, we also tested gradient boosted decision trees.

Feature Selection Different encodings will be used to find optimal forms for working with our data. For example, categories such as presurgical opioid status (opioid naive, opioid use, misuse history) will be represented in one-hot as well as ordinal forms with performance compared. Due to the large number of predictors present, our aim is to keep our feature count as small as possible to improve performance. Lasso regressions and ablation studies will be performed to identify these features.

3.4 Evaluation of Predictions

3.4.1 Binary Categorization Predictions

Logistic regression predictions were evaluated using area under the receiver operating characteristic curve (AUC), precision, sensitivity, and specificity metrics. Finding the right threshold for binary categorization is necessary in balancing the sensitivity and specificity performance. One might consider high sensitivity to avoid under-prescribing high consumers in comparison to a higher specificity model which would reduce the number of low consumers that receive an overly high prescription size.

Chapter 4

Results

4.1 Strongest Predictors

4.1.1 Odds Ratios

In Table 4.1, we see that for each standard deviation increase in age, patients were 0.728 times as less likely to be in the high opioid consumption group. Patients with higher BMIs, current tobacco use, or previous opioid exposure were all more likely to be in the higher consumption group. Lastly, we see that for each standard deviation increase in discharge day opioids, a patient is around 1.8 times as likely to be in the high consumption group after discharge.

Feature	**Odds Ratio (95% CI)**	**P value**
Age at surgery	0.728 (0.630 – 0.842)	< 0.0001*
BMI	1.198 (1.077 – 1.334)	0.00107*
Current Tobacco Use	1.888 (1.307 – 2.682)	0.000519*
Pre-Op Opioid Status	1.989 (1.091 – 3.531)	0.0167
Discharge Day IV Units	1.799 (1.222 – 2.499)	0.00105*
Discharge Day MME	1.787 (1.277 – 2.502)	0.000872*

Table 4.1: Odds ratios were calculated with 95% CI using bootstrapped elastic net logistic regression with 5,000 replicates. Across 95 features, 6 had 95% CI that did not encompass an odds ratio of 1. The five bolded p-values are statistically significant as determined by the Benjamini-Hochberg procedure using a false discovery rate of 0.05.

	Logistic Regression	**Logistic Regression Reduced Model**
Train AUC	0.7471	0.7159
Test AUC	0.7270	0.7148
Test Specificity A	93.43%	88.26%
Test Sensitivity A	32.2%	33.9%
Test Specificity B	34.27%	34.2%
Test Sensitivity B	91.53%	86.44%

Table 4.2: Logistic Regression Performance

	XGBoosted Tree	**XGBoosted Reduced Model**
Train AUC	0.7688	0.7533
Test AUC	0.7289	0.7165
Test Specificity A	91.54%	92.02%
Test Sensitivity A	33.33%	25.42%
Test Specificity B	36.62%	36.62%
Test Sensitivity B	93.22%	88.14%

Table 4.3: XGBoost Decision Tree Performance

4.2 Predictive Performance

In Table 4.2 and 4.3, we have our model performance for logistic regression and XG Boosted Tree. Two sets of sensitivity/specificity are given. One with a high threshold that gives higher specificity, and one with a lower threshold that gives lower specificity. Model performance is comparable between the train and validation due to increased levels of regularization. Higher levels of L1 penalization were used for all models to reduce overfitting. To generate the specificity and sensitivity, different thresholds were tested. The two results using high and low thresholds are listed as A and B respectively.

Confusion matrices are shown below in Table 4.4 for logistic regression models using threshold A. Given performance, it does not appear that the more complicated XGBoost model provides significant performance gains when compared with logistic regression.

		Prediction outcome	
		>75th	≤ 75th
Actual	>75th	TP 19	FN 40
	≤ 75th	FP 14	TN 199

		Prediction outcome	
		>75th	≤ 75th
Actual	>75th	TP 20	FN 39
	≤ 75th	FP 25	TN 188

Table 4.4: Logistic regression confusion matrix

4.2.1 Reduced Models

Reduced models limited to fewer than 10 features were tested to determine if predictive performance would be maintained. It seems that model performance for logistic regression is not impacted as greatly from dropping certain features. This is in part due to L1 penalization which already reduces the impact of more tangential features for the full model.

Chapter 5

Discussion

5.1 Refining for clinical use

In order for such a model to be used for clinical use, a few conditions must be met. The first, is a sufficient specificity and sensitivity that clinicians would be comfortable making decisions aided by the model. At this point, more work must be done to improve model performance in collaboration with clinicians to decide how to meet such standards. However, using our threshold that allows for higher sensitivity, we can capture around 9/10 high consuming patients while maintaining a specificity of around 1/3. While this does mean that patients without high consumption may be prescribed more than is necessary, this is already the case and thus may not be different from what is already being done.

Secondly, the model must not be a black box. Clinical decision making tools must be able to, at minimum at a high level, show to the clinician how any risk stratification is being done. Ideally, accurate risk stratification could be done using only a few features to improve explainability. For this reason, we tested a reduced model that used a reduced selection of features.

Lastly, test sets from other hospitals must be gathered to determine if there are dataset shifts that bias the models performance towards our BIDMC surgical cohort. If a model is to be used broadly across institutions, rigorous testing must be completed, which would require a coordinated effort to collect such data.

5.1.1 Role of reduced models

The role of reduced models was to test if predictive performance could be maintained while reducing the number of features per patient. If possible, this would mean that the final model could be simpler to use and test. Asking providers to gather a small selection of patient data points is much more realistic than expecting them to gather a large list of information for each patient. Given the performance of our reduced model, it seems promising that this could be refined to perform well with few features, increasing the likelihood of eventually being used in a clinical setting.

5.2 Strongest Predictors

Looking at our strongest predictors, age, BMI, current tobacco usage, pre-operative opioid status, and IV and non IV opioid consumption on the day of discharge, we find that our results are in line with previous studies' findings. However, one observation of note more unique to our study is the relevance of in-hospital opioid consumption. The motivation for obtaining this data was a hypothesis that patients who consumed fewer opioids in-hospital after their surgery would be less likely to consume high amounts of opioids after discharge. At the same time, we sought to test the hypothesis that patients consuming high levels of opioids up and until discharge would also take higher levels of opioids after discharge. Given our results, our hypothesis is currently supported.

Chapter 6

Limitations and Future work

6.1 In-hospital Opioid Consumption

6.1.1 Current Method

Currently, order data from the BIDMC pharmacy for opioids was used as a proxy for opioid consumption. This presents two problems. The first is that ordered opioids are not always administered to the patient. While it may be a close approximation, opioids are sometimes wasted or returned if they are no longer necessary. The second is specific to IV opioids and patient-controlled analgesia (PCA). In cases where patients control how much analgesia they will be administered based on their pain levels over time, it is not possible from the order data to know how much was actually consumed. While one patient may have used up the majority of the medication, another may experience little pain and use none of it. In future studies, understanding exactly how much opioid medication was consumed in-hospital may provide better predictive results given that our current results already show an association between in-hospital opioid consumption and opioid consumption post-discharge.

6.1.2 Electronic Medication Administration Record (eMAR)

One possibility of better gauging consumed opioids in-hospital is through the eMAR system. This system records the amount of opioids administered for a patient. While it may not be as revealing for PCA cases, it may give us a better understanding of IV opioid usage for provider administered cases. For non IV opioids such as oxycodone tablets, it may also paint a clearer picture of exactly how much was consumed since not everything that is ordered is taken by the patient.

6.2 Post-discharge Opioid Consumption

6.2.1 Accuracy

A second limitation is our measurement of opioids consumed after discharge as there is no guarantee that all of the pills consumed were taken by the patient. To proactively address this, patients were followed up within two weeks of discharge to most accurately assess patient consumption. To this end, we believe that our measurement of opioids consumed was a reasonable proxy for the ground truth value.

6.2.2 Patient Survey Response Rates

As shown in Figure 3-1, 1338 patients did not respond to our opioid consumption surveys. There is also a skew where younger patients were less likely to respond. Thus, our remaining population may not be perfectly representative of the broader population. For future studies, other methods of collecting this data such as through phone applications may be tested to determine if response rates might increase.

6.3 Missing Data points

6.3.1 Pain scores

A helpful piece to understanding a patient's pain management is their subjective pain score. While there are limitations in this too, given that people have varying pain tolerances, it may be of value to investigate. Pain scores were not routinely assessed in our study and thus we were unable to determine if this would have been a helpful feature. For future studies, having a system of collecting pain scores both before and after discharge may shed some further light on this topic.

6.3.2 Surgical Approach

Surgical approach (i.e. open, minimally invasive, etc.) is also relevant to understanding what trajectory a patient's pain might take. This was not systematically collected, which led to an inability to use this as a feature. Laparascopic surgeries, for example, are associated with higher levels of pain immediately after surgery but lower levels after 24 hours. Informing patients of what their pain levels might look like over time may help them calibrate medication usage beyond simply adjusting their prescription size.

These prediction models, if successful, could help providers more accurately tailor pain management plans for individual patients. By identifying cases where patients can have appropriate pain management with smaller prescriptions, we can potentially reduce consumption for that patient. This also reduces the amount of leftover pills, which could make their way to other consumers and result in addictive behaviors. Smarter opioid prescribing is an important step forward in tackling this public health issue. While we have a collection of data on each patient now, we will look to expand what data is available for each patient to maintain a wide scope of factors to consider. As time allows, more and more data per patient will be obtained from the various sources that store them.

Chapter 7

Conclusion

Our work here makes progress towards predicting outlier opioid consumption using information available prior to discharge. With further refinement, we may be able to tailor opioid prescription sizes based on an individual patient's characteristics and in-hospital opioid consumption. Being able to identify high opioid consumers would allow for providers to reduce prescription sizes for all other patients while limiting risk of under-prescription. Given that the demographic and health factors that we found to have predictive value in opioid consumption are in line with other studies, we believe that our model here shows promise in potentially being tested with surgical patients from other institutions.

Our main goal of determining if this prediction task is possible was met. Given our limitations and results, we also have a clearer idea of what work is necessary before such a model could be brought into a clinical setting. With work towards refining the accuracy of some of our proxies, we can get data that most closely represents the patients' opioid usage. The potential benefits of such a model are immense. The potential to reduce opioid prescriptions safely would help reduce the massive public health and financial strain that opioid misuse currently has in our country.

Appendix A

Tables

Feature	**Description**	**Form**
Study ID	Study identification number unique to each patient	Numerical
MRN	Medical Record Number unique to each patient for use with hospital databases	Numerical
Age	Age at date of surgery	Numerical
Gender	Male, Female, Other	One-hot
Ethnicity	Hispanic/Latino, Not Hispanic or Latino, Unknown	One-hot
Race	American Indian or Alaska Native, Asian, Black, Native Hawaiian or Other Pacific Islander, Other, Unknown, White	One-hot
BMI Class	Underweight, Normal, Overweight, Obese Class I, Obese Class II, Obese Class III	Ordinal (1-5)
Pre-Op OTC Pain Medications	Tylenol (Acetaminophen), Aspirin, Motrin (Ibuprofen), Excedrin, None, Other	One-hot

Pre-Op Opioid Abuse or Overdose?	Pre-surgical opioid abuse/overdose from patient survey	Binary
Pre-Op Antidepressant, Antipsychotic, Benzodiazepine, Antispasmodic, or ADD/ADHD medication use	Pre-operative use of any of above	One-hot
Pre-Op Miscellaneous Drug Use	Clonidine, Buspirone, Melatonin, Ambien (Zolpidem), Suboxone (buprenorphine/naloxone), Fioricet (butalbital/acetaminophen/caffeine), Voltaren (diclofenac), Neurontin (gabapentin), Lyrica (pregabalin), Baclofen, Robaxin (Methocarbamol), lidocaine (topical or gel patch), Prescription NSAIDS, Triptan, Chantix (Varenicline), Other, None	One-hot
Total Pre-Op Immediate Acting MMEs	Pre-operative Opioid consumption in MMEs/day for Hydrocodone, Oxycodone, Tramadol, Hydromorphone, Codeine	Numerical
Total Pre-Op Long Acting MMEs	Pre-operative Opioid consumption in MMEs/day for MS Contin, Oxycontin, Methadone	Numerical
Pre-Op Opioid Script Length	<1 week, 1-2 weeks, 2-3 weeks, 3-4 weeks, 1-3 months, 3-6 months, >6 months	Ordinal (1-7)
Alcohol Consumption	None, History of, Current	Ordinal (1-3)

Heavy Alcohol Consumption (>14 drinks / week for men, >7 drinks / week for women)	None, History of, Current	Ordinal (1-3)
Alcohol Abuse	None, History of, Current	Ordinal (1-3)
Pre-Operative Diagnoses	Current Tobacco Use, History of Tobacco Use, History of Recreational Drugs or Substance Abuse, Mood Disorder, Anxiety, Depression, PTSD, Chronic Pain >6 months, ADD/ADHD, Migraines, Insomnia, Fibromyalgia, None, Other	One-hot
Recreational Drug Use	Marijuana, Cocaine, Heroin, LSD, Methamphetamines/Amphetamines/Any Intravenously injected drugs, other	One-hot
Prior to surgery, were you were taking any medication for	Pain, Anxiety, Depression, Sleep, ADD/ADHD, Drug Depenendence, None of the above	One-hot
Surgery Category Median MME Category	low, mid, high	Ordinal (1-3)
Pre-surgical Opioid Status	Opioid Naïve, Opioid Exposed, Misuse History	Ordinal (1-3)
History of Mental Illness	Yes, No	Binary
Discharge Day IV Units	Units of IV opioids ordered on discharge day	Numerical

Discharge Day MMEs	MMEs of non IV opioids ordered on discharge day	Numerical
Discharge Day (-1) IV Units	Units of IV opioids ordered one day prior to discharge	Numerical
Discharge Day (-1) MMEs	MMEs of non IV opioids ordered one day prior to discharge	Numerical
Discharge Day (-2) IV Units	Units of IV opioids ordered two days prior to discharge	Numerical
Discharge Day (-2) MMEs	MMEs of non IV opioids ordered two days prior to discharge	Numerical
Pre-discharge Opioid MME Slope	Slope of MMEs of non IV opioid use over up to 3 days prior to discharge	Numerical
Hospital Stay Length Post-surgery	Number of days from the date of surgery until discharge	Numerical
Outcome indicator	Indicator variable for outcome classes: <= 75th percentile or >75th percentile	Binary

Table A.1: Data Dictionary for data used in analysis

Surgery Category	**Count**
Sternotomy	160
Inguinal Hernia	112
Thyroidectomy	89
Cholecystectomy	87
Video Assisted Thoracoscopy (VATs)	85
Partial Mastectomy	76
Gastric Sleeve	61
Knee Arthroscopy	59

Thoracic/Lumber/Sacral / Discectomy/Laminectomy/Fusion	59
Prostatectomy	54
Appendectomy	50
Mammoplasty	47
Total Hip Replacement (THP)	45
Colectomy	44
Deep inferior epigastric perforators (DIEP) Flap repair	42
Mastectomy- Gender Affirmation	39
Nephrectomy	38
Shoulder Arthroscopy	38
Total Knee Replacement (TKR)	35
Anterior Cervical Discectomy and Fusion	34
Ventral Hernia	34
Sigmoidectomy	34
Transurethral Resection of Bladder Tumors (TURBT)	31
Carotid Endarterectomy (CEA)	30
Stab Phlebectomy	30
Transurethral resection of prostate (TURP)	28
TAVR	28
Hemorrhoidectomy	27
Thoracic/Lumbar/Sacral Microdiscectomy	26
Lower Extremity Angiogram	23
Abdominal Exploration	23
Upper Extremity Dialysis Access	21
Fundoplication	21
Parathyroidectomy	21
Mastectomy w/Implant	18

Lower Extremity Bypass	17
Umbilical Hernia	17
Mastectomy	17
Carpal Tunnel Release	17
Component Separation for Hernia	16
Liver Resection	15
Low Anterior Resection (LAR)	15
Exam under anesthesia w/ fistulotomy	14
Whipple	14
Ileostomy Takedown	12
Pancreatectomy	11
Fenestrated Endovascular Aortic Repair (FEVAR)	11
Donor Nephrectomy	11
Endovascular Aortic Repair (EVAR)	11
Total Shoulder Replacement	10
Tonsillectomy	10
Revision of Upper Extremity Dialysis Access	9
Pilonidal Cyst Excision	8
Kidney Transplant Recipient	8
Exam under anesthesia, biopsy or polyp excision	8
Ileocecectomy	7
Duodenal Surgery	6
Esophagectomy	6
Thoracotomy	5
Abdominoperineal Resection	5
Mediastinoscopy	5
Lateral internal sphincterotomy	5
Exam under anesthesia with seton placement	5

Drainage of Perirectal Abscess	4
Non-ACDF Cervical Spine	4
Breast Augmentation	4
Transmetatarsal Amputation (TMA)	3
Tracheobronchoplasty	3
Total Abdominal Colectomy (TAC)	3
Liver Transplant Recipient	3
Tenckhoff Catheter Insertion	2
Ileal Conduit	2
Open Abdominal Aortic Aneurysm Repair	2
Latissimus Dorsi Flap	1
Laparoscopic Resection of Diaphragm Melanoma	1
Radical Neck Dissection	1
Groin Mass Excision	1
Foot Mass Excision	1
Laparoscopic Hepatic Cyst Unroofing	1

Table A.2: Surgery Categories

Bibliography

[CDC, 2019] 2019 (Jul). *Opioid Overdose: State Information.*

[Che *et al.* , 2018] Che, Zhengping, Sauver, Jennifer St., Liu, Hongfang, & Liu, Yan. 2018 (4). Deep Learning Solutions for Classifying Patients on Opioid Use. vol. 2017: 525-534. AMIA Annual Symposium.

[Ekstein *et al.* , 2006] Ekstein, Perla, Szold, Amir, Sagie, Boaz, Werbin, Nachum, Klausner, Joseph M., & Weinbroum, Avi A. 2006. Laparoscopic Surgery May Be Associated With Severe Pain and High Analgesia Requirements in the Immediate Postoperative Period. *Annals of Surgery*, **243**(1), 41–46.

[Florence *et al.* , 2016] Florence, Curtis S., Zhou, Chao, Luo, Feijun, & Xu, Likang. 2016. The Economic Burden of Prescription Opioid Overdose, Abuse, and Dependence in the United States, 2013. *Medical Care*, **54**(10), 901–906.

[Hastings *et al.* , 2020] Hastings, Justine S., Howison, Mark, & Inman, Sarah E. 2020. Predicting high-risk opioid prescriptions before they are given. *Proceedings of the National Academy of Sciences*, **117**(4), 1917–1923.

[Howard *et al.* , 2018] Howard, Ryan, Waljee, Jennifer, Brummett, Chad, Englesbe, Michael, & Lee, Jay. 2018. Reduction in Opioid Prescribing Through Evidence-Based Prescribing Guidelines. *JAMA Surgery*, **153**(3), 285.

[Howard *et al.* , 2019] Howard, Ryan, Fry, Brian, Gunaseelan, Vidhya, Lee, Jay, Waljee, Jennifer, Brummett, Chad, Campbell, Darrell, Jr, Seese, Elizabeth, Englesbe, Michael, & Vu, Joceline. 2019. Association of Opioid Prescribing With Opioid Consumption After Surgery in Michigan. *JAMA Surgery*, **154**(1), e184234–e184234.

[Levy *et al.* , 2015] Levy, Benjamin, Paulozzi, Leonard, Mack, Karin A., & Jones, Christopher M. 2015. Trends in Opioid Analgesic–Prescribing Rates by Specialty, U.S., 2007–2012. *American Journal of Preventive Medicine*, **49**(3), 409–413.

[Lo-Ciganic *et al.* , 2019] Lo-Ciganic, Wei-Hsuan, Huang, James L., Zhang, Hao H., Weiss, Jeremy C., Wu, Yonghui, Kwoh, C. Kent, Donohue, Julie M., Cochran, Gerald, Gordon, Adam J., Malone, Daniel C., Kuza, Courtney C., & Gellad, Walid F. 2019. Evaluation of Machine-Learning Algorithms for Predicting Opioid Overdose Risk Among Medicare Beneficiaries With Opioid Prescriptions. *JAMA Network Open*, **2**(3), e190968–e190968.

[Rhon *et al.* , 2018] Rhon, Daniel I, Snodgrass, Suzanne J, Cleland, Joshua A, Sissel, Charles D, & Cook, Chad E. 2018. Predictors of chronic prescription opioid use after orthopedic surgery: derivation of a clinical prediction rule. *Perioperative Medicine*, **7**(1).

[Santiago, 2019] Santiago, Christine. 2019. Opioids. *Academic Medicine*, **94**(1), 10–9.

[Scully *et al.* , 2018] Scully, Rebecca E., Schoenfeld, Andrew J., Jiang, Wei, Lipsitz, Stuart, Chaudhary, Muhammad Ali, Learn, Peter A., Koehlmoos, Tracey, Haider, Adil H., & Nguyen, Louis L. 2018. Defining Optimal Length of Opioid Pain Medication Prescription After Common Surgical Procedures. *JAMA Surgery*, **153**(1), 37–43.

[Sinatra, 2010] Sinatra, Raymond. 2010. Causes and Consequences of Inadequate Management of Acute Pain. *Pain Medicine*, **11**(12), 1859–1871.

[Substance Abuse and Mental Health Services Administration., 2018] Substance Abuse and Mental Health Services Administration. 2018 (09). *Key Substance Use and Mental Health Indicators in the United States: Results from the 2017 National Survey on Drug Use and Health.* HHS Publication No. SMA 18-5068, NSDUH Series H-53. Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration.

[Tan *et al.* , 2018] Tan, Wen Hui, Yu, Jennifer, Feaman, Sara, McAllister, Jared M., Kahan, Lindsey G., Quasebarth, Mary A., Blatnik, Jeffrey A., Eagon, J. Christopher, Awad, Michael M., & Brunt, L. Michael. 2018. Opioid Medication Use in the Surgical Patient: An Assessment of Prescribing Patterns and Use. *Journal of the American College of Surgeons*, **227**(2), 203 – 211.

[Zhang *et al.* , 2017] Zhang, Jinghe, Iyengar, Vijay, Wei, Dennis, Vinzamuri, Bhanukiran, Bastani, Hamsa, Macalalad, Alexander R., Fischer, Anne E., Yuen-Reed, Gigi Y., Mojsilovic, Aleksandra, & Varshney, Kush R. 2017. Exploring the Causal Relationships between Initial Opioid Prescriptions and Outcomes.