Human-Based Versus Algorithmic-Based Treatments: Evidence from the Opioid Epidemic

Alireza Boloori
Milgard School of Business, University of Washington, Tacoma, aboloori@uw.edu

Soroush Saghafian
Harvard Kennedy School, Harvard University, Soroush_Saghafian@hks.harvard.edu

Emisa Nategh
McDonough School of Business, Georgetown University, en471@georgetown.edu

Abstract. The U.S. opioid epidemic motivated several federal prescribing guidelines, which nonetheless relied heavily on one-size-fits-all, human-based clinical judgment to navigate the complex tradeoff between the harms of opioid use and those of inadequately managed pain. We study whether algorithmic-based approaches can be developed to learn personalized, multi-modal, and cost-effective pain treatment policies capable of improving on both observed clinical practice and CDC guidelines along this tradeoff. Using longitudinal claims data on over 2.7 million patients (2017–2022), we formulate multi-modal pain treatment as a sequential decision problem and apply two offline reinforcement learning algorithms to recommend opioid and non-opioid dose, days-supply, and use of non-pharmacologic treatments. We evaluate the resulting policies against three benchmarks: observed clinical practice, and counterfactual implementations of the CDC's 2016 and 2022 prescribing guidelines. Both algorithmic policies outperform all three benchmarks on several dimensions, including cost-effectiveness. However, for a subset of patients, human-based clinical judgment is the superior option. Robustness checks on a separate Medicaid population and under class rebalancing preserve the main conclusions, with the cost-effectiveness gap widening under population shift. Comparing CDC guidelines, we find the more restrictive 2016 guideline outperforms its less restrictive 2022 revision, indicating that CDC's efforts in revising the guideline have not been fruitful. Overall, our findings indicate algorithmic-based approaches are more powerful in providing pain treatment recommendations. The analysis, however, identifies a non-trivial subset of patients for whom no algorithmic policy improves upon human-based clinical judgment, providing an explicit boundary for real-world deployment.

Key words: opioid epidemic; multi-modal pain management; personalized medicine; human vs. machine; reinforcement learning
History: August 21, 2026

1. Introduction

According to the National Institute on Drug Abuse (NIDA), 915,515 drug-related deaths occurred in the U.S. between 2000 and 2020, among which opioid analgesics were the main contributing factor in 556,472 deaths (60.78% of total deaths). These opioids may result in patients switching to heroin, which, in turn, caused additional 140,792 deaths during the same period (NIDA 2022). The economic cost of the U.S. opioid epidemic was estimated to be over a trillion in 2017 (Luo et al. 2021), with about two million people being either dependent on prescription opioids or abusing them in 2016 and many more in the following years. (Phillips et al. 2017).

To address this crisis, the U.S. Centers for Disease Control and Prevention (CDC) proposed a set of guidelines in 2016 with the aim of reducing opioid prescriptions by clinicians (CDC 2016, Dowell et al. 2016). The CDC guidelines, however, were widely criticized for several reasons, including the fact they did not make use of existing clinical evidence to provide a clear level of specificity. For example, the American Medical Association (AMA) criticized the CDC guidelines by stating that *"we continue to have serious concerns that some of these guidelines either contain a degree of specificity not supported by the existing evidence or conflict with official Food and Drug Administration (FDA) approved product labeling for opioid analgesic products."* (AMA 2016)

In essence, the CDC guidelines relied on human judgment by recommending that *"Clinicians should consider opioid therapy only if expected benefits for both pain and function are anticipated to outweigh risks to the patient"* (Dowell et al. 2016). However, understanding when, and for which patients, the benefits outweigh the risks would not be an easy task for human clinicians. Furthermore, the lack of specificity in the CDC recommendations made them a "one-size-fits-all" approach. In particular, instead of providing recommendations tailored to patient characteristics, the CDC recommended fixed thresholds for all patients: no more than 90 morphine milligram equivalent (MME) per day for the strength of opioids or no more than 7 days of supply for acute pain (CDC 2016).[1] Lack of a clear treatment guideline tailored to each patient's characteristics, juxtaposed with *human-based perceptions*, potentially led clinicians to show wide variations in opioid prescriptions, which, in turn, is considered as another contributing factor to the opioid epidemic (Barnett et al. 2017). A main goal of our study is to use large-scale clinical evidence to (1) develop an analytics-driven *algorithmic-based* approach that can help clinicians (as the human component of the decision-making process) quantify when, and for which patients, the benefits of opioid prescriptions outweigh their risks, (2) use our algorithmic-based approach to enable policymakers to create recommendations that are personalized (i.e., adapted to each patient's characteristics) and not based on one-size-fits-all rules, and (3) quantify the impact of using our algorithmic-based approach vis-a-vis the human-based practices.

To this end, we make use of large-scale claims data, Merative™ MarketScan® Databases, which contains the longitudinal medical and prescription history of millions of U.S. patients over the period 2017–2022. From this data, we retrieve detailed information on the medical history of each individual patient, including diagnoses made by providers, prescriptions of pharmacologic treatments (e.g., opioid and non-opioid medications), and use of non-pharmacologic treatments (e.g., physical therapy and chiropractic). We formulate the problem of multi-modal pain treatment as a sequential

[1] In 2022, the CDC issued a revised set of guidelines, the most notable change from the 2016 version being the removal of the numeric thresholds for opioid dose and duration of supply (CDC 2022). The revision restored discretion to clinicians but, as critics have noted, replaced one form of underspecification with another: clinicians retain wide latitude with limited guidance on how to exercise it for any particular patient (NPR 2022).

decision problem and apply *offline reinforcement learning* (offline RL)[2] to obtain multi-modal pain treatment policies comprised of opioid and non-opioid strengths, days of supply, and use of non-pharmacologic treatments. Specifically, we implement two algorithmic-based approaches representing distinct points on the conservatism spectrum of offline RL: Conservative Q-Learning (CQL) and Implicit Q-Learning (IQL). We evaluate the resulting policies for their cost-effectiveness and their downstream effects on two competing risks: risk of opioid use disorder and risk of undertreated pain, representing the side effects and potential benefits of pain treatments, respectively. We further compare these algorithmic-based policies against three human-based benchmarks: observed clinical practice recovered directly from the data, as well as the CDC 2016 and 2022 guidelines.

Overall, our contributions are two-fold:

(1) We develop an algorithmic-based approach capable of dynamically recommending multi-modal treatments personalized to individual patients. Our approach is based on offline RL methods that learn directly from large-scale observational data while explicitly guarding against the distributional-shift pitfalls that have historically limited the deployment of data-driven policies in clinical settings (Kumar et al. 2020, Levine et al. 2020).

(2) We conduct extensive computational experiments and use our findings to provide important insights and implications for both policymakers and individual physicians:

- *Both algorithmic policies outperform observed practice and both CDC guideline regimes on aggregate cost-effectiveness.* Evaluated through cost-effectiveness analysis against three human-based benchmarks, both CQL and IQL achieve substantially higher benefits, with CQL the stronger of the two in aggregate terms. This indicates that offline RL can extract clinically meaningful gains from observational claims data and can yield policies superior to the widely adopted regulations designed by highly experienced human experts. A patient-level decomposition of net monetary benefit further reveals that 71.7% of patients are better served by at least one of the two algorithms than by observed practice, but 28.3% are not. The latter establishes a non-trivial subset of patients for whom algorithmic policies could not outperform clinical judgment. Our results, hence, suggest implementing a triaged-based human-AI approach in which some patients are routed to algorithmic-based treatment and others to a human-based one.
- *Lower opioid-use-disorder risk is achieved through dose-duration restructuring, not blanket opioid restriction.* Relative to the observed practice, the CQL policy recommends moderately higher daily opioid dose alongside substantially shorter days-supply, and yet attains a lower opioid-use-disorder risk. This is a pattern consistent with clinical literature identifying cumulative duration of opioid exposure, rather than daily dose alone, as a dominant predictor of opioid-related harm. Notably,

[2] This is a class of methods designed to learn high-quality policies from fixed observational datasets without further interaction with patients (Levine et al. 2020).

jointly considering the risks of opioid use disorder and undertreated pain does not translate into looser opioid recommendations relative to practice.

- *The more restrictive CDC 2016 guidelines outperform the looser CDC 2022 guidelines on aggregate cost-effectiveness.* Contrary to what the 2022 revision's removal of explicit dose and duration thresholds might suggest, the more restrictive 2016 guidelines achieve a higher level cost-effectiveness than the 2022 revised policy. The substantive lift from offline RL relative to both guideline regimes suggests that the relevant choice may not be between more or less restrictive fixed threshold (on the strength or duration of the opioids) imposed in human-based policies; instead, it is between imposing any kind of threshold (to remove the need for an algorithmic-based implementation) and benefiting from fully personalized treatments enabled via carefully designed algorithms.
- *The algorithmic-based approach is robust to threats that commonly undermine deployed healthcare models.* We probe robustness to such as population shift (re-estimating on a Medicaid population, distinct from our main commercially-insured and Medicare population) and to class imbalance (re-estimating on a class-rebalanced version of our main sample). In both cases, our main conclusions are preserved. Furthermore, we find that, under population shift, the cost-effectiveness gap relative to practice widens rather than narrows, suggesting that the value of an algorithmic-based policy may be especially pronounced in more vulnerable populations.

The rest of this paper is organized as follows. In §2, we provide a review of relevant literature. In §3, we discuss the data and study design. In §4, we present the analytical setting, including the offline RL formulation, the CQL and IQL algorithms, and the construction of the three benchmarks for policy comparison. In §5, we present our numerical experiments. In §6, we provide a summary of relevant insights and policy implications. Finally, in §7, we conclude the paper and briefly discuss avenues for future research.

2. Literature Review

2.1. Studies on Side Effects of Opioid Medications

A stream of literature related to our work focuses on establishing statistical associations between opioid-related side effects and various risk factors such as age, gender, duration of supply, high dose of opioid, and history of alcohol abuse, smoking, and mental health disorders (Ciesielski et al. 2016, Kendler et al. 2023). For reviews of studies on the impact of opioid painkillers on addiction, drug dependence, overdose, or death, we refer interested readers to Fishbain et al. (2007) and Nuckols et al. (2014). Within this stream, some studies apply machine learning to address the opioid-related side effects. Haller et al. (2016) employed Natural Language Processing techniques to predict risks of drug abuse and addiction before a prescription is written. Che et al. (2017) made use of a deep feed-forward neural network to predict the possibility of long-term opioid use. Crosier et al. (2017)

used random forests to predict opioid overdose. Vunikili et al. (2018) utilized an Extreme Gradient Boosting algorithm along with logistic regression to predict the risk of opioid abuse, overdose, and death. Bjarnadottir et al. (2019) applied LASSO, adaptive boosting, and random forests to establish risk factors associated with the chronic use of opioid painkillers. More recently, Lo-Ciganic et al. (2022) developed and validated a machine-learning algorithm to predict opioid overdose in Medicaid beneficiaries, and Xue et al. (2025) extended this line of work by training gradient-boosted models to predict acute-care utilization among individuals already engaged in opioid treatment programs. Compared to this stream, we not only analyze the risk of opioid dependence, abuse, overdose, and death, but also measure the risk of undertreated pain to account for potential benefits of pain treatments. In addition, our work is prescriptive as opposed to these predictive studies. Thus, our approach allows decision-making, and hence, developing suitable treatment policies. Furthermore, the foregoing studies are based on *cross-sectional* analyses, while our proposed policies are based on *longitudinal* approaches.

2.2. Studies on Measuring Pain and Efficacy of Pain Management

To measure the intensity of pain and evaluate the efficacy of managing it, there are evidence-based pain assessment scales, such as verbal rating scales, numerical rating scales, visual analogue scales, and behavioral pain scales (Huskisson 1974, Katz and Melzack 1999). These scales are typically evaluated using patient-based surveys, which are often impacted by patients' satisfaction (Wells et al. 2008). However, accounting for patients' satisfaction is known to inadvertently propell providers to prescribe opioid medications (Zgierska et al. 2014). Thus, CMS proceeded towards removing pain management questions from the Hospital Consumer Assessment of Healthcare Providers and Systems survey (Boloori et al. 2020b). To characterize the efficacy of pain management in our study, we instead focus on repeated visits due to an unresolved pain related condition, which is a proxy on whether the condition is being treated effectively (McPhillips-Tangum et al. 1998, Xiao and Barber 2008). In particular, we make use of information from our claims data about repeated visits of patients before and after an index opioid prescription (for more details, see §3.1).

2.3. Relevant Operations Research/Management Science Studies

Pitt et al. (2018) proposed a dynamic compartmental model based on a combination of pain, opioid use, and addiction status, and found that various resource allocation schemes (e.g., supplying naloxone as an opioid antagonist, supplying needles for addicted patients, and promoting medication-assisted treatments) could have a positive long-term impact on patients' quality of life. Gan et al. (2019) analyzed the impact of wearable devices on detecting opioid use disorder via urine tests, and proposed a partially observable Markov decision process model to optimize decisions on wearing these devices given various budget and patient adherence considerations. Bobroske et al. (2022)

conducted an empirical study and showed that getting a second opinion such as visiting another primary care provider for opioid prescription is associated with a lower rate of long-term opioid use. More recent contributions extend the OR/MS related studies in adjacent directions: Luo and Stellato (2024) develop a data-driven facility location and resource allocation framework for opioid treatment access with explicit equity constraints, and Attari et al. (2024) examine how supply-chain opacity has shaped the trajectory of the crisis. A broader set of issues in optimizing decisions regarding care delivery via wearables and mobile health (mHealth) devices is discussed in studies such as Saghafian and Murphy (2021), and the references therein.

From a methodology standpoint, the Operations Research/Management Science (OR/MS) literature has applied supervised learning in optimization frameworks. For example, Bertsimas et al. (2016) followed this approach to improve cancer chemotherapy regimens. However, this body of literature has primarily focused on single-period decision-making problems, where a machine learning method is applied to cross-sectional data. Reinforcement Learning (RL) is one of viable methods suited for modelling multi-period decision-making problems. For example, Saghafian (2024) shows how RL can be applied to longitudinal observational data sets in order to obtain optimal dynamic treatment regimes that yield causal improvements, even if the longitudinal observational data is subject to hidden confounders. Recent studies also show that well-designed RL approaches can be effective in utilizing longitudinal offline data from wearables to provide multi-dimensional personalized recommendations in variety of diseases and disorders (see, e.g., Lin et al. (2025), and the references therein). Other approaches in handling mutli-period settings include Markov decision processes and multi-armed bandit models, which have been applied to a variety of problems such as in optimizing warfarin dosing (Bastani and Bayati 2020) and joint immunosuppressive and diabetes medications (Boloori et al. 2020a). Despite their widespread applications, these methods could fall short in addressing the curse of dimensionality, nonstationary rewards, and history-dependency that may arise in problems like the one we study in this paper. In addition, the curse of ambiguity introduced in Saghafian (2018, 2024) often limits the applicability of these methods when applied to observational longitudinal data sets.

We contribute to this stream by developing a multi-period optimization model based on Offline RL. While allowing us to analyze the longitudinal data, the Offline RL framework also enables us to address some of the foregoing challenges such as dependency on history (Levine et al. 2020). Offline RL has recently begun to see application in clinical decision-making settings where online interaction with patients is infeasible; e.g., in safe sepsis treatment and mechanical ventilation (Kondrup et al. 2023). From an application standpoint, to the best of our knowledge, our proposed framework is the first to address the prescription of multi-modal pain treatments, while simultaneously addressing the side effects and potential benefits of these treatments.

Finally, a stream of literature suggests that "algorithm aversion" might limit the impact that using an algorithmic approach such as ours might have in practice, because human experts might be reluctant to make use of recommendations obtained from algorithms (Dietvorst et al. 2015). Some recent studies, however, show that humans do possess "algorithm appreciation" (Logg et al. 2019), and hence, in critical and complex decision-making settings such as the one we study in this work, they are likely to take into account the advice they receive from a well-designed algorithm. We believe improving the interpretability and removing the black-box nature of some of the algorithms can go a long way in this regard, and view this as an important direction for future work.

3. Data and Study Design

The data that we utilize in our study comes from the Merative MarketScan Databases for years 2017-2022, a large-scale data set of more than 3.7 million patients and more than 25 million number of patient visits. There are two main sources of information that we have retrieved from the MarketScan: (a) information about patients' encounters and diagnoses,[3] and (b) information about the history of prescriptions. Furthermore, after applying the inclusion criteria, described in Table 1, to this data, our study sample size includes 3,729,647 patients. Of note, among these patients, 2,738,118 were enrolled in Commercial or Medicare plans, for whom geographical covariates including residential region and population-area size are available. The remaining 991,529 patients were enrolled in Medicaid. Since the Medicaid file is multi-state and does not carry the geographical

[3] To determine the diagnoses, we follow the International Classification of Diseases, Tenth Revision, Clinical Modification (referred to as ICD-10 hereafter).

Table 1 Summary of inclusion criteria

Citerion	Patients count
(1) At least one episode of an opioid analgesic prescription (index prescription)	29,835,749
(2) Sustained enrollment: 6 and 24 months for pre and post index prescription, respectively [a]	13,312,228
(3) No history of cancer or end-of-life (palliative) status [b]	9,735,870
(4) Age $\geq$ 18 in the beginning of patients' recorded data	6,595,247
(5) No opioid prescription within the first 180 days of a patient's records [c]	4,216,521
(6) No history of opioid side effects within the first 180 days of a patient's records [d]	4,022,547
(7) At least one recorded encounter prior to (or at the time of) the index prescription [e]	3,729,647

[a] Instead of "full insurance enrollment" over the six-year study period, this choice, while retaining the continuity needed to characterize each patient's exposure and outcome trajectory, would avoid excluding patients with intermittent coverage, whose exclusion could bias the cohort toward populations with more stable insurance access (thus healthier populations).
[b] The CDC guidelines on opioid prescription do not apply to patients who suffer from cancer or are in their palliative stage (Dowell et al. 2016). We identify cancer patients if a patient has any ICD-10 diagnosis code from chapter 2 (neoplasms) or any other ICD-10 diagnosis code to identify cases with cancer (search terms like neoplasm, malignant, malignancy, benign, carcinoma, and palliative).
[c] Patients who satisfy this condition are called *opioid naïve*; i.e., for non-naïve patients, their prescription might be related to a medical condition occurring prior to the beginning of our data (see, e.g., Johnson et al. (2016) and Bobroske et al. (2022)).
[d] When an overdose occurs for a patient, we must know which medications, that might have had significant impacts on this overdose, s/he had used prior to the incidence. To be consistent with criterion (5), we make use of a 180-day window to exclude patients without sufficient information within our data set.
[e] We exclude patients that do not satisfy this criterion, because our aim is to identify a pain-inducing medical condition due to which an opioid was prescribed for the first time.

Table 2 Summary statistics

Variable	Description	Mean (SD) / Num (%) [a]
Dependent	Incidence of opioid-use disorder (Event 1) [b]	15,599 (0.6%)
	Incidence of undertreated pain (Event 2) [b]	313,870 (11.5%)
Independent	Daily opioid strength (MME)	29.07 (32.21)
	Daily non-opioid strength (mg)	1217.25 (943.54)
	Days of supply[c]	13.77 (22.15)
	Use of non-pharmacologic treatments	0.02 (0.16)
Control	Age	46.13 (15.64)
	State annual opioid dispense rate (per 100 residents)	50.21 (14.28)
	Number of unique days a patient makes visits in a window	0.07 (0.40)
	Number of unique overlapping prescriptions in a window	3.79 (3.05)
	History of substance use (excluding opioid) [b]	279,162 (10.2%)
	History of surgical procedures [b]	765,546 (28.0%)
	Gender: Male	1,154,563 (42.2%)
	Female	1,583,555 (57.8%)
	Pain pathology: Surgery	690,772 (25.2%)
	Acute	788,401 (28.8%)
	Chronic	1,258,945 (46.0%)
	Comorbidity (Elixhauser): <0	1,736,654 (63.4%)
	1–4	495,572 (18.1%)
	5+	505,892 (18.5%)
	Insurance: Commercial	2,521,429 (92.1%)
	Medicare	216,689 (7.9%)
	Region: Northeast	311,092 (11.4%)
	North Central	583,986 (21.3%)
	South	1,436,793 (52.5%)
	West	406,247 (14.8%)
	Population size: Metro ≥1 million	1,161,369 (42.4%)
	Metro 0.25-1 million	446,788 (16.3%)
	Metro < 250K	155,119 (5.7%)
	Non-metro <50K	342,008 (12.5%)
	Unknown [d]	632,834 (23.1%)

[a] Rates (in %) are reported out of 2,738,118 patients. [b] For this variable, the value reported is the number of patients who ever experienced an event or had a history of a medical condition. [c] Length of a time window = 90 days. [d] These are patients with known region/state but unknown Metropolitan Statistical Area (MSA) in the MarketScan data.

identifiers (e.g., region and population size), we reserve this subpopulation for a robustness check analysis (§5.4.3).[4] We conduct the primary analysis on the 2,738,118-patient Commercial and Medicare cohort. Table 2 shows the summary statistics.

3.1. Dependent Variables

Our dependent variables capture the occurrence of two main and medically relevant events: opioid use disorder (OUD) and undertreated pain (UTP), referred to as Events 1 and 2, respectively.

[4] Among the 991,529 patients in the Medicaid sample, 6.4% are dually eligible for Medicare and Medicaid; these patients are retained in the Medicaid sample only, so no patient contributes utilization to both samples.

DEFINITION 1 (EVENT 1). We say Event 1 has occurred if there is at least one diagnosis made for dependence, abuse, poisoning, or any adverse effect caused by an opioid. We obtain the ICD-10 codes used to identify the occurrence of Event 1 from NIH.gov (2016).

Next, we turn to Event 2. The lack of pain relief is commonly referred to as un-/undertreated pain.[5] When a patient is prescribed a pain medication to address a pre-existing *pain-inducing condition* (PIC), but has post-supply visits due to the same condition, it is likely that the visits are due to undertreated pain (McPhillips-Tangum et al. 1998, Xiao and Barber 2008, Nursing 2017, STAT 2022). We operationalize this via the following definition.

DEFINITION 2 (EVENT 2). Let DX_n be a primary diagnosis among visits in time window $n \geq 1$. Then, Event 2 has occurred in time window n if DX_n is among the pain-inducing conditions:

$$\text{Event 2 (in window } n) = \begin{cases} 1, & \text{if } DX_n \in \text{PICs}, \\ 0, & \text{if } DX_n \notin \text{PICs}. \end{cases} \tag{1}$$

A value of 1 indicates that a pain-inducing condition persists in window n, consistent with undertreated pain; a value of 0 is consistent with the pain treatment having addressed the underlying condition.

Pain-inducing condition (PIC). The identification of PICs from observational claims data is not straightforward, because encounters and prescription claims are typically filed by different parties (e.g., physicians vs. pharmacists), and the pain diagnoses underlying a prescription are not directly linked to the prescription itself in the data. We address this gap by inferring PICs from the diagnoses recorded during the *pre-supply period*, which we define as a short period preceding the index opioid prescription. We adopt a conservative threshold of 30 days for this time window (Bobroske et al. 2022). Within the pre-supply period, three categories of conditions can give rise to an opioid prescription: surgeries, acute conditions, and chronic conditions. In the presence of surgical procedures, the painkiller is likely prescribed for post-surgical pains (Overton et al. 2018). In the absence of surgical procedures, the prescription is likely to address acute or chronic conditions. Consistent with the CDC's recommendations (Dowell et al. 2016), we assume that an opioid prescription is attributed to an acute condition first and to a chronic condition only when no acute condition is present. Algorithm 1 summarizes the procedure for the identification of PICs.

Surgical procedures, post-surgical pain, and acute/chronic conditions. To identify surgical procedures, we utilize information on inpatient admission type, procedure type, provider type, and place of service (see Appendix A.1). We also use the diagnosis codes reported from ICD10 (2026b) for the identification of post-surgical pain diagnoses. Furthermore, to identify acute/chronic conditions,

[5] Pain relief is the first outcome measure in assessing the efficacy of pain treatments (Teater 2015, Smith et al. 2018).

We map each ICD-10 diagnosis code among a patient's pre-supply visits to an acute or chronic condition using the binary Chronic Condition Indicator (CCI) (AHRQ.gov 2025). In Appendix A.2, we have listed sample ICD-10 codes for chronic conditions (Mayhew et al. 2019).

Remark 1 (Impact of Repeated Visits). We characterize Event 2 based on the premise that repeated visits for a similar medical condition could indicate undertreated pain. However, repeated visits could also be due to other reasons such as (i) a doctor ordering some tests on the first visit, and then evaluating or discussing the results in a follow-up visit, or (ii) a patient visiting providers too often, which could, in turn, divulge the drug-seeking behavior of that patient. To address (i), we excluded cases with ICD-10 diagnosis codes (Z08 and Z09 series) used for follow-up examination. To remedy (ii), we account for two control variables that are indicative of doctor or pharmacy shopping behavior of patients (Delcher et al. 2021): number of overlapping prescriptions and number of different days a patient makes visits in a time window.[6]

3.2. Independent Variables

As part of our independent variables, we measure the strength and duration of supply for opioid and non-opioid medications via three steps:

Step 1: original strength. For drugs that were prescribed to patients, we decompose the generic name and differentiate the strength of opioid from that of non-opioid.[7] Out of 50 opioid and 51 non-opioid unique drugs that were identified from our data, we determine 21 opioid and 34 non-opioid unique

[6] A patient's propensity for such behavior may be more pronounced when repeated visits are made to different providers rather than the same provider. However, provider identifiers in MarketScan have substantial missingness, which precludes a reliable same-versus-different-provider distinction. In the absence of this information, we rely on the number of distinct visit-days within a time window as a proxy: a patient accumulating many visit-days within a window is exhibiting a pattern consistent with doctor- or pharmacy-shopping regardless of whether those visits are with one provider or spread across several.

[7] Example: 'Acetaminophen/Propoxyphene' with strength 325mg-50mg, where Propoxyphene (Acetaminophen) is a opioid (non-opioid) component with the strength of 50 (325) milligram per tablet.

Algorithm 1 Identification of pain-inducing conditions (PICs)

Input: For each patient, encounters within the pre-supply period with information such as ICD-10 diagnoses, admission type, provider type, and place of service.
Output: PICs for each patient

for all Patient within the pre-supply period **do**
  if there is any surgical procedure **then**
    PICs = {post-surgical pains} ▷ Category 1[a]
  else if there are acute conditions **then**
    PICs = {those acute conditions} ▷ Category 2[b]
  else if there are chronic conditions **then**
    PICs = {those chronic conditions} ▷ Category 3[c]
  end if
end for

[a] Category 1 (post-surgical pain): 690,751 patients (25.23%). All rates are reported from 2,738,118 cohort size.
[b] Category 2 (acute conditions): 788,404 patients (28.79%).
[c] Category 3 (chronic conditions): 1,258,962 patients (45.98%).

components (for a complete list of these components, see Appendix A.3). We then obtain the duration of supply and original strength:

$$\text{Duration of supply} = \big(\text{REFILL} + 1\big) * \text{DAYSUPP}, \tag{2a}$$

$$\text{Original strength (per day of supply)} = \text{STRENGTH} * \big(\text{QTY}/\text{Duration of supply}\big), \tag{2b}$$

where REFILL (number of refills), DAYSUPP (days supply for each refill), STRENGTH (strength per drug unit), and QTY (number of units dispensed per prescription) are available from our data.

Step 2: transformation of strength. We then transfer the original strength of opioids obtained from Equation (2b) to the Morphine Milligram Equivalent (MME) (Palliative Drugs 2009, CDC 2020). For non-opioid medications, we consider their strengths in milligram (no transformation involved).

Step 3: adjustment based on time windows. To be consistent with the recommended regular intervals of pain management reassessment (CDC 2016), we consider an average in-between visits time window of $W = 90$ days. We then compute, on aggregate, the strengths and durations of relevant prescriptions in each time window (Figure 1). Of note, if a supply starts during a time window but ends after it, we only consider the time portion within that window, while the remaining portion will count towards the duration in next time window. Furthermore, for prescriptions with overlapping durations, we consider the prescription with the longest duration (Bobroske et al. 2022).

Finally, we note that a multi-modal pain treatment also involves non-pharmacologic treatments (e.g., physical therapy, chiropractic, and acupuncture). We identify this from our data via information related to procedures conducted, procedure groups, provider types, and service categories (for more details, see Appendix A.4).

3.3. Control Variables

We account for control variables, on any observable factors from our data, that can make the relationship between prescription of pain treatments and incidence of Events 1–2 biased. As we describe next, these include a set of demographic variables, clinical characteristics, geographical factors, and indicators of doctor or pharmacy shopping behaviors.

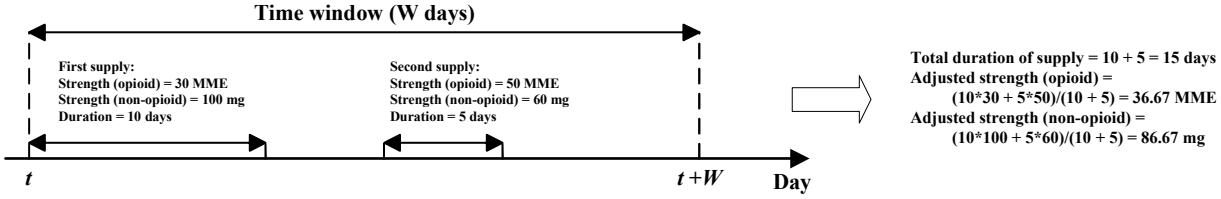


Figure 1 **Adjustment of strengths and duration of supply based on the number of prescriptions in a time window**

Demographics. Age (numerical) and gender (male vs. female).

Clinical characteristics. We account for four sets of clinical characteristics:

(i) History of substance use disorders (excluding opioid) (ICD10 2026a, DHCS.CA.gov 2026).

(ii) History of comorbidities (Elixhauser Comorbidity Index). We use ICD-10 diagnosis codes to measure the comorbidity index (Elixhauser et al. 1998, umanitoba.ca 2023), and then classify the index into $\leq$0, 1-4, and 5+ comorbidities (Menendez et al. 2014).

(iii) Pain pathology: surgery, acute, or chronic conditions (see §3.1 for more discussion).

(iv) History of surgical procedures and inpatient admission. Treatment of post-surgical pain with opioid medications has been shown to influence onsets of post-operative opioid addiction or long-term opioid usage after hospital discharge (Pletcher et al. 2008, Wu et al. 2019). In contrast to the pain-pathology variable in (iii), which is fixed at the pre-supply period, this variable updates dynamically across the treatment horizon to capture surgical events and hospitalizations occurring during therapy. We control for indicators of whether a patient has had a surgical procedure and/or an inpatient admission.

Geographical factors. We control for various geographical factors including the region (Northeast, North Central, South, and West) and whether a patient's residential area is located in a large, medium, or small metro area ($\geq$1 million population, 0.25-1 million population, and $<$250,000 population, respectively) or a non-metro area (fewer than 50,000 population). These metro/non-metro areas are identified by "Metropolitan Statistical Area" in our data. Information about their population size is retrieved from census.gov (2021). We then classify the population size based on categories defined by the Rural-Urban Continuum Codes for urban areas from ers.usda.gov (2020). To further reflect on the context of opioid prescription, we also control for the state annual opioid dispensed per 100 residents. The information was retrieved from CDC (2021) based on a patient's state (variable in our data: Geographic Location Employee) and the year when opioid was initiated.

Variables related to doctor or pharmacy shopping behaviors. As noted in Remark 1, repeated visits by a patient may imply a drug-seeking behavior. To this end, we control for variables that reflect doctor or pharmacy shopping behavior of patients: number of overlapping prescriptions and number of different days a patient makes visits in a time window.

4. The Analytical Setting

We formulate the personalized, multi-modal pain treatment recommendation as a sequential decision problem under uncertainty. In §4.1, we specify the cost-effectiveness objective that any candidate treatment policy is evaluated against. In §4.2, we describe the offline reinforcement learning

algorithms we use to learn such a policy from our observational data. In §5, we derive algorithmic-based policies and compare them with the human-based practice observed from our data as well as the CDC 2016 and 2022 guidelines.

4.1. Cost-Effectiveness Objective

We aim to determine multi-modal pain treatment plans that are most cost-effective. To this end, for any treatment plan, we measure the *net monetary benefit* (Drummond et al. 2015):

$$\text{Net monetary benefit} = \text{WTP} * \text{QALY} - \text{COST}, \tag{3}$$

where QALY is the *quality-adjusted life years* that a patient can accrue, WTP is the willingness to pay for an additional QALY, and COST is the cost of care as the total payments to providers and healthcare settings by payers and patients (e.g., reimbursement and co-payment).

Let $\mathcal{A} = \mathcal{A}^{os} \bigtimes \mathcal{A}^{nos} \bigtimes \mathcal{A}^{ds} \bigtimes \mathcal{A}^{nph}$ be the set of feasible actions representing possible multi-modal treatment options ($\bigtimes$: the Cartesian product). $\mathcal{A}^{os} = [0, 100]$ MME and $\mathcal{A}^{nos} = [0, 3000]$ mg are feasible intervals for the opioid and non-opioid strengths, respectively.[8] $\mathcal{A}^{ds} = \{0, 1, \ldots, 90\}$ days is the feasible set for the duration of supply of medications (90: length of a time window),[9] and $\mathcal{A}^{nph} = \{0 : \text{No non-pharmacologic treatment}, 1 : \text{Use non-pharmacologic treatment}\}$ is the set of possible actions with respect to non-pharmacologic treatments.[10] Also, let $\boldsymbol{a} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_T) \in \mathcal{A}^T$ be a sequence of multi-modal treatments (e.g., doses of opioid and non-opioid medications, their duration of supply, and use of non-pharmacologic treatments) taken throughout the time horizon T, and $\boldsymbol{H}_t = (\boldsymbol{V}_1, \boldsymbol{a}_1, \ldots, \boldsymbol{V}_{t-1}, \boldsymbol{a}_{t-1}, \boldsymbol{V}_t)$ be the history of the patient's covariates up to time window t (denoted by $\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_t$) as well as all the prior treatments.[11] For Event $k \in \{1, 2\}$ and a fixed history $\boldsymbol{H}_t$, we measure the cost and QALY accrued over the time horizon by:

[8] For opioid strength, the maximum threshold originally recommended by the CDC 2016 guidelines was 90 MME (Dowell et al. 2016); 95$^{\text{th}}$-percentile in our data across windows with dose > 0 is 60 MME. For non-opioid strength, 95$^{\text{th}}$-percentile in our data across windows with dose > 0 is 2785.71 mg. Of note, for the opioid and non-opioid strengths and the duration of supply, we consider a uniform grid-based discretization of the action space.

[9] Our data indicate that repeated visits within a single time window are rare. Specifically, 99.40% of patient-window observations in our data involve at most one distinct visit-day; only 0.60% involve two or more distinct visit-days within the same 90-day window. This suggests that our choice of a 90-day decision epoch, which is also consistent with the interval used in the CDC guidelines (CDC 2016), rarely obscures within-window treatment revisions of the kind a shorter epoch would be needed to capture.

[10] Among patients who ever used non-pharmacologic treatments, 83.08% used either physical therapy or chiropractic. Since the majority of non-pharmacologic treatment use relates to either physical therapy or chiropractic, we therefore consider a binary variable (i.e., not differentiate based on the specific type of non-pharmacologic treatment).

[11] We assume the patient's covariate trajectory $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_t$ is exogenous to the treatment sequence under evaluation; that is, held fixed at its realized values in the data regardless of which candidate action sequence $\boldsymbol{a}$ is substituted into $\boldsymbol{H}_t$. Equations (4a)–(7) therefore measure the net benefit that would have accrued had the patient instead received the deterministic treatment sequence $\boldsymbol{a}$ along that same realized covariate path, rather than under a fully closed-loop model in which treatment choices are allowed to alter the future covariate trajectory.

$$\text{COST}_k(\boldsymbol{a}) = \sum_{t=1}^{T} \beta^{t-1}\ P_k(\boldsymbol{a}_t, \boldsymbol{H}_t)\, C_k, \tag{4a}$$

$$\text{QALY}_k(\boldsymbol{a}) = \sum_{t=1}^{T} \beta^{t-1}\ r_k(\boldsymbol{a}_t, \boldsymbol{H}_t)\,, \tag{4b}$$

where $\beta \in [0,1)$ is a discount factor that allows us to prioritize the outcomes in the current time window over those in the future, $P_k(\boldsymbol{a}_t, \boldsymbol{H}_t)$ is the probability of experiencing Event k under treatment $\boldsymbol{a}_t$ and history $\boldsymbol{H}_t$, C_k is the total cost accrued due to an occurrence of Event k, and $r_k\left(\boldsymbol{a}_t, \boldsymbol{H}_t\right)$ is the immediate QALY impact under $\boldsymbol{a}_t$ and $\boldsymbol{H}_t$. We denote by q_{k0} (q_{k1}) the monthly quality-of-life score for a patient who does not (does) experience Event k, where $0 \le q_{k1} \le q_{k0}$. Using this notation, the immediate QALY impact is measured as:

$$\begin{aligned} r_k\left(\boldsymbol{a}_t, \boldsymbol{H}_t\right) &= \big(1 - P_k(\boldsymbol{a}_t, \boldsymbol{H}_t)\big) \times q_{k0} + P_k(\boldsymbol{a}_t, \boldsymbol{H}_t) \times q_{k1} \\ &= q_{k0} - (q_{k0} - q_{k1}) \times P_k(\boldsymbol{a}_t, \boldsymbol{H}_t) \qquad \text{for } t \le T, \end{aligned} \tag{5}$$

where the first (second) part of summation on the first line indicates the expected QALY for a patient who is not experiencing (experiencing) Event k.

Using Equations (3)-(5), we measure the net benefit for Event $k \in \{1,2\}$ under a fixed history $\boldsymbol{H}_t$ as:

$$\text{NB}_k(\boldsymbol{a}) = \text{WTP} \times \text{QALY}_k(\boldsymbol{a}) - \text{COST}_k(\boldsymbol{a}) \tag{6a}$$

$$= \sum_{t=1}^{T} \beta^{t-1} \Big[q_{k0} \times \text{WTP} - P_k(\boldsymbol{a}_t) \Big((q_{k0} - q_{k1}) \times \text{WTP} + C_k \Big) \Big], \tag{6b}$$

where, for brevity, we suppress the dependency of $P_k(\cdot)$ and subsequent notations on $\boldsymbol{H}_t$.

Using (6a)-(6b), the overall objective, for each patient, is to find the multi-modal treatment plan that maximizes the weighted average net benefit:

$$\max_{\boldsymbol{a} \in \mathcal{A}^T} \Big\{ \text{NB}(\boldsymbol{a}) = \sum_{k=1}^{2} w_k \text{NB}_k(\boldsymbol{a}) \Big\}, \tag{7}$$

where $w_k \in [0,1]$ is the weight assigned to Event k with the property that $\sum_k w_k = 1$.

Of note, the sequential, multi-dimensional nature of (7), where both treatments and covariates evolving dynamically over a multi-period horizon with continuous and discrete action components, would make direct solution methods computationally infeasible. We, therefore, adopt an offline reinforcement learning approach next, to learn a policy that approximates the solution to (7) directly from observational data.

4.2. Offline Reinforcement Learning

We approach the problem in (7) as a Markov Decision Process (MDP) with action $\boldsymbol{a}_t \in \mathcal{A}$ and per-window reward

$$R(\boldsymbol{s}_t, \boldsymbol{a}_t) = \sum_{k=1}^{2} w_k [\text{WTP} \cdot r_k(\boldsymbol{a}_t, \boldsymbol{H}_t) - P_k(\boldsymbol{a}_t, \boldsymbol{H}_t)\, C_k]\,. \tag{8}$$

Rather than the raw, unbounded history $\boldsymbol{H}_t$ introduced in §4.1, we represent the state $\boldsymbol{s}_t$ as a fixed-dimension feature vector. Two of its components are the patient's predicted probabilities of Event 1 (OUD) and Event 2 (UTP) at window t, produced by a sequence of offline RL models conditioned on the patient's realized trajectory through t. The remaining components are the current window's treatment intensity and largely time-invariant demographic and clinical covariates. This construction keeps $\boldsymbol{s}_t$ Markovian and fixed in dimension, while still reflecting the patient's accumulated history through the two model-based probability components, rather than through an unbounded concatenation of $\boldsymbol{H}_t$. These two probability components are produced by a transformer-based sequence model with self-attention over the patient's fixed time horizon, trained to predict each window's event probabilities from the realized history up to that window. This model plays two roles in our framework: it supplies the historical-risk components of $\boldsymbol{s}_t$ described above, and (separately, at evaluation time) it serves as the counterfactual outcome model used to score a candidate policy's downstream effects (see Figure 2, "common evaluation layer").

The cumulative discounted reward over the horizon, under this construction, recovers the net benefit $\text{NB}(\boldsymbol{a})$ in (7). A policy $\pi(\boldsymbol{a} \mid \boldsymbol{s})$ maps states to a distribution over actions, and the goal is to learn a policy π^* that maximizes the expected discounted return $\mathbb{E}_\pi\left[\sum_{t=1}^{T} \beta^{t-1} R(\boldsymbol{s}_t, \boldsymbol{a}_t)\right]$.[12] We note a formulation-level distinction from §4.1: the objective in (7) is a deterministic maximization over a fixed action sequence $\boldsymbol{a}$, evaluated along a fixed, realized covariate trajectory, whereas the MDP formulation above involves an expectation over a (possibly stochastic) policy $\pi(\boldsymbol{a} \mid \boldsymbol{s})$ and, at evaluation time, a stochastic environment model governing the transition to $\boldsymbol{s}_{t+1}$. Our offline RL algorithms, applied to this MDP, therefore yield an *approximate* solution to (7) rather than an exact one; we adopt this approximation because it allows us to learn a policy directly from observational data at the scale of our setting, as we discuss next.

[12] Alternative approaches to (7), such as approximate dynamic programming (Powell 2011), sequential predict-then-optimize via supervised event-probability estimation, and causal-inference-based policy evaluation (Robins et al. 2000), each could face limitations in our setting: the high-dimensional mixed continuous-discrete action space and longitudinal covariates in $\boldsymbol{H}_t$ strain value-function approximation; decoupling estimation from optimization can compound errors across the horizon and does not directly target the policy of interest (Elmachtoub and Grigas 2022); and importance-weighting-based causal-inference estimators are known to suffer extreme variance inflation in high-dimensional sequential action spaces (Levine et al. 2020).

4.2.1. Conservative Q-Learning (CQL). CQL (Kumar et al. 2020) addresses distributional shift by augmenting the standard Bellman backup with a penalty that pushes down the Q-values of actions not supported by the data. Specifically, CQL learns a Q-function parameterized by Q_θ that solves:

$$\min_\theta \ \alpha \left(\mathbb{E}_{\boldsymbol{s}\sim\mathcal{D},\boldsymbol{a}\sim\mu}[Q_\theta(\boldsymbol{s},\boldsymbol{a})] - \mathbb{E}_{(\boldsymbol{s},\boldsymbol{a})\sim\mathcal{D}}[Q_\theta(\boldsymbol{s},\boldsymbol{a})] \right) + \mathcal{L}_{\text{Bellman}}(\theta), \tag{9}$$

where $\mathcal{L}_{\text{Bellman}}$ is the standard temporal-difference loss, μ is a distribution over actions used to evaluate out-of-distribution candidates, and $\alpha > 0$ governs the strength of the conservatism penalty. The first term provably yields a Q-function that lower-bounds the true value of any policy whose action distribution diverges sufficiently from π_b, so the policy extracted from Q_θ is biased toward actions well-supported by the data.

4.2.2. Implicit Q-Learning (IQL). Rather than penalizing out-of-distribution actions, IQL (Kostrikov et al. 2021) avoids ever querying the Q-function on actions outside the dataset. IQL fits a state-value function V_ψ to the *expectiles* of $Q_\theta(\boldsymbol{s},\boldsymbol{a})$ over actions drawn from $\mathcal{D}$,

$$\min_\psi \ \mathbb{E}_{(\boldsymbol{s},\boldsymbol{a})\sim\mathcal{D}}[L_2^\tau(Q_\theta(\boldsymbol{s},\boldsymbol{a}) - V_\psi(\boldsymbol{s}))], \tag{10}$$

where $L_2^\tau(\cdot)$ is the expectile loss with parameter $\tau \in (0.5, 1)$ that emphasizes high-value in-distribution actions, and then updates Q_θ via a SARSA-style target $R(\boldsymbol{s},\boldsymbol{a}) + \beta V_\psi(\boldsymbol{s}')$ that involves no maximization over actions. The final policy is extracted via advantage-weighted regression on the in-distribution actions in $\mathcal{D}$. Because the Q-function is never evaluated on actions outside the data, the distributional-shift problem is sidestepped rather than penalized away.

The two algorithms embody distinct perspectives for handling the offline setting: explicit conservatism over the action space (CQL) versus implicit avoidance of out-of-distribution evaluation (IQL). Therefore, their relative behavior on a given problem is, in general, an empirical question (Levine et al. 2020). Reporting both allows us to characterize the robustness of our findings to this algorithmic choice (see §5).

5. Numerical Experiments

5.1. Comparing the Cost-Effectiveness of Algorithmic-Based Treatment Policy with the Human-Based Practice and Guidelines

We evaluate the economic performance of algorithmic treatment policies against three benchmarks: observed clinical practice, a guideline-constrained practice reflecting the 2016 CDC recommendations, and a guideline-informed practice reflecting the revised 2022 CDC recommendations (for the comparison framework, see Figure 2).

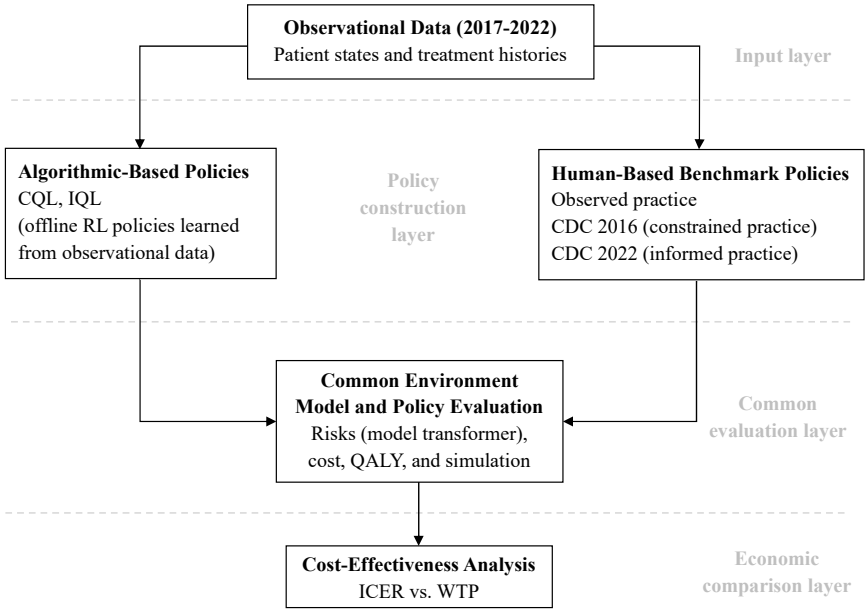


Figure 2 Overview of the policy construction and evaluation framework

Two focal algorithmic policies (CQL and IQL) and three human-based benchmark policies are derived from the same observational data but differ in how treatment decisions are constructed. All policies are evaluated under a common environment model and assessed using the same cost-effectiveness framework.

Benchmark 1: Observed Clinical Practice. This corresponds to treatment decisions as observed in our data. Because our sample spans the period 2017–2022, these decisions might reflect real-world prescribing behavior following the release of the CDC 2016 guideline, albeit including heterogeneous and incomplete adoption of its recommendations across providers and settings.

Benchmark 2: CDC 2016 (Counterfactual Constrained Practice). This imposes structured constraints motivated by the CDC 2016 guideline. In particular, opioid prescribing is subject to limits on dosage strength and days supply.[13] Relative to the observed practice, this benchmark represents a counterfactual scenario in which CDC 2016 guideline recommendations are implemented uniformly across patients rather than heterogeneously.

Benchmark 3: CDC 2022 (Counterfactual Unconstrained Practice). This reflects another benchmark based on the CDC 2022 guideline, which emphasized removing rigid prescribing thresholds imposed in the 2016 guideline and allowing patient-specific flexibility. In Appendix B, we provide details and extensively describe the operationalization and validation steps for these benchmarks.

[13] We operationalized this as follows: we first observed the pain treatments from our data. We then adjusted the strength of opioid and duration of supply by thresholds recommended by the CDC 2016 guidelines: maximum of 90 MME for opioid strength and a maximum total of 7 (90) days of supply for acute (chronic) pain (Dowell et al. 2016, CDC 2016, AZDHS 2017). Of note, there was no specific guideline for the non-opioid strength or use of non-pharmacologic treatments. Thus, for those treatment modes, we used the values observed directly from the data.

Cost-Effectiveness Evaluation. For each proposed policy, we evaluate cost-effectiveness by computing incremental costs and health outcomes relative to each benchmark. For any WTP > 0, our treatment policy is said to be more cost-effective than a benchmark if (Drummond et al. 2015):

$$\begin{aligned}&\text{Incremental Cost-Effectiveness Ratio (ICER)} = \\ &\frac{\text{Incremental Cost}}{\text{Incremental QALY}} = \frac{\text{Cost(our policy)} - \text{Cost(benchmark)}}{\text{QALY(our policy)} - \text{QALY(benchmark)}} \leq \text{WTP},\end{aligned} \tag{11}$$

where we measure Cost and QALY under each policy using Equations (4a)-(4b).

The percentage of patients satisfying the inequality in (11) determines the cost-effectiveness probability of our proposed policy against benchmarks (Fenwick et al. 2006). Based on our results in Figure 3, we make the following observation.

Observation 1. *(i) CQL achieves a significant cost-effectiveness improvement against all three human-based benchmarks across the full range of willingness-to-pay thresholds, while IQL improvement is modest. (ii) CQL's cost-effectiveness advantage is more pronounced against observed practice and the CDC 2022 benchmark than against the CDC 2016 benchmark. (iii) Relative to observed practice, CQL increases the average QALY by 6.6 days and reduces the average cost by $1,927 per patient over a three-year therapeutic course.*

Observation 1(i) establishes CQL as the cost-effective algorithmic-based policy: its CE probability against all three human-based benchmarks exceeds the conventional 50% acceptability threshold across the entire WTP range, reaching approximately 61.39% and 60.14% against the CDC 2016 benchmark at WTP = \$10K and \$100K per QALY, and approximately 64.96% and 63.57% against observed practice and the CDC 2022 benchmark at the same thresholds. IQL, by contrast, yields CE probabilities of 28–30% across all benchmarks.

Remark 2 (Conservative Exploration: CQL vs. IQL). The divergence between CQL and IQL arises from how each method handles limited support in observational data. CQL penalizes out-of-distribution actions, producing conservative value estimates that discount poorly supported treatments. In a heterogeneous treatment setting, this allows CQL to identify cost-effective policy adjustments that remain empirically grounded. IQL, by contrast, restricts value estimation to observed actions via expectile regression and does not evaluate unobserved treatments. While effective when the behavioral policy is near-optimal, this constraint is limiting here: meaningful improvements require considering actions beyond observed prescribing patterns. As a result, IQL remains tied to the observed action distribution and fails to generate economic gains, yielding a less cost-effective policy compared with benchmarks.

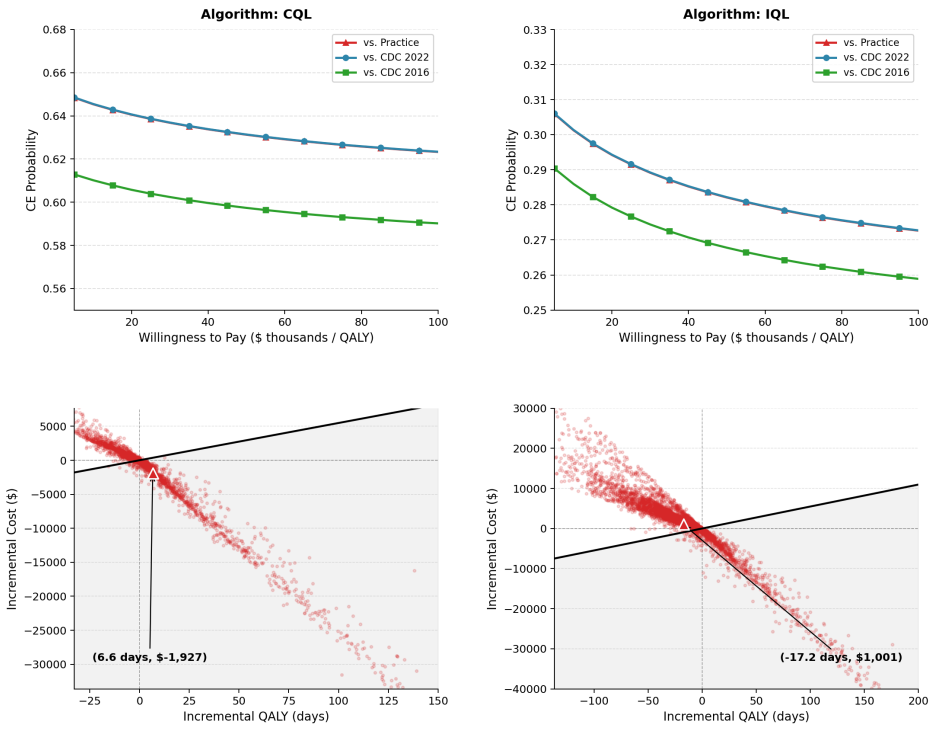


Figure 3 Cost-effectiveness of algorithmic-based versus the human-based policies

(Top) it shows the CE probability of CQL and IQL against the practice, measured as the fraction of patients satisfying the inequality $\Delta\text{Cost}(i) \leq \text{WTP} \times \Delta\text{QALY}(i)$, for WTP ranging from \$10,000 to \$100,000/QALY. **(Bottom)** it displays the corresponding scatter for the CQL and IQL policies. The black line in each panel represents the cost-effectiveness (CE) boundary at a WTP threshold of \$20,000/QALY. Points falling below this boundary are cost-effective. **(Pareto profile)** The CE probability curves exhibit limited sensitivity to WTP. The majority of simulated instances cluster along a negatively sloped trajectory (lower incremental cost accompanied by higher incremental QALY), which is consistent with a Pareto-dominant profile relative to the benchmarks. When incremental costs and QALYs covary in this manner, the share of instances falling below the WTP threshold line is relatively stable as the threshold shifts.The parameter values used for our numerical experiments are listed in Appendix B.1. **(Practice vs. CDC 2022)** The CE probability curves for CQL and IQL against observed practice and against CDC 2022 differ by less than 0.2% points throughout the WTP range shown (visually near-indistinguishable). This is consistent with CDC 2022's shift away from population-level thresholds toward individualized clinical judgment, which brings our counterfactual construction of CDC 2022 guideline close to observed practice for this cohort.

Observation 1(ii) indicates that CQL's cost-effectiveness improvement is lowest against the CDC 2016 benchmark. This is consistent with the notion that the 2016 guideline's rigid opioid prescription thresholds (e.g., no more than 90 morphine milligram equivalent (MME) per day for the strength of opioids or no more than 7 days of supply for acute pain) induced systematic undertreatment for a subset of patients, suppressing costs at the expense of health outcomes and thereby presenting a more favorable incremental cost profile. Notably, the CE probability curves against observed practice (Benchmark 1) and the CDC 2022 (Benchmark 3) are very comparable. This near-equivalence reflects the fact that for most patients identified as constrained by the 2016 cap, the simulated unconstrained dose (under the 2022 guidelines) would lie close to the threshold. This implies that

observed prescribing over 2017–2022 could already approximate the individualized flexibility later endorsed by CDC in the 2022 guideline.

As the second and more recent wave of CDC guidelines still emphasize the harms of opioids rather than the harms of poorly managed pain, some believe that these guidelines have the potential for human misunderstanding or misapplication. On the other hand, some other experts are wary against removing the rigid prescription thresholds imposed in the earlier guidelines, concerning that it could make it more difficult for physicians who had relied on these thresholds in their practice over the prior several years (NPR 2022). Our results summarized in Observation 1 indicate that lifting the rigid 2016 thresholds, as prescribed by the CDC 2022 guidelines, does not yield a material improvement over observed clinical practice in terms of cost-effectiveness. This may suggest that clinicians in our sample were already approximating the individualized flexibility endorsed by the 2022 revision. However, we find that neither the 2016 nor the 2022 guidelines are as cost-effective as what an algorithmic-based policy can achieve. However, we also observe some heterogeneity: Using an algorithmic-based approach is not necessarily the preferred choice across all patients. In the next section, we provide detailed insights into patients characteristics for which algorithmic-based policies perform particularly better compared to the observed human-based practice.

5.2. Patient Characterization by Algorithmic Performance

While the results in §5.1 favor CQL over IQL at the population level, aggregate statistics may not show meaningful heterogeneity in algorithmic performance across patient subgroups. We, therefore, examine whether algorithmic dominance is uniform across the patient population, or whether there exist clinically identifiable subgroups for whom the clinical judgment used in the observed practice cannot be overcome by the CQL or IQL algorithms. Furthermore, we investigate which of these two algorithms should be used (if either). This latter investigation has a justification for clinical adoption: a patient-adaptive algorithm assignment strategy that routes each patient to the algorithm best suited to their clinical profile.

For each patient i, we compute the incremental net monetary benefit (NMB) under CQL and IQL separately, relative to the observed practice. Each patient is then classified into one of four mutually exclusive categories based on the signs of $\text{NMB}_i^{\text{CQL}}$ and $\text{NMB}_i^{\text{IQL}}$:

- *CQL-favorable only*: $\text{NMB}_i^{\text{CQL}} > 0$ and $\text{NMB}_i^{\text{IQL}} \leq 0$
- *IQL-favorable only*: $\text{NMB}_i^{\text{IQL}} > 0$ and $\text{NMB}_i^{\text{CQL}} \leq 0$
- *Both favorable*: $\text{NMB}_i^{\text{CQL}} > 0$ and $\text{NMB}_i^{\text{IQL}} > 0$
- *Neither favorable*: $\text{NMB}_i^{\text{CQL}} \leq 0$ and $\text{NMB}_i^{\text{IQL}} \leq 0$ (i.e., practice outperforms both)

Figure 4 displays the distribution across the patient population. Approximately 41% of patients are classified as CQL-favorable, 7% as IQL-favorable, 23% as both favorable, and 29% as neither

favorable. These results lend to the question of whether the algorithmic performance can be linked to clinical characteristics. We next explore this question via two tiers, each serving a distinct purpose for clinical adoption: Interpretability (Tier 1) and Algorithm Applicability (Tier 2).

5.2.1. Tier 1: Algorithm Interpretability. We contrast CQL-favorable and IQL-favorable patients with the goal of examining whether the patient characteristics associated with algorithmic benefit are clinically coherent and observable at the time of treatment initiation. This can also help us address the *explainability* of our algorithmic policies. From Table 3, the two groups differ across several dimensions. With respect to pain pathology, CQL-favorable patients are predominantly chronic (67.5%), whereas IQL-favorable patients are predominantly acute (69.4%). CQL-favorable patients present with higher average opioid strength (28.26 vs. 23.13), higher average non-opioid strength (1221.83 mg vs. 1093.98 mg), and substantially shorter days supply (8.46 vs. 16.01). This can imply a more stable, non-opioid-anchored treatment regimen for CQL-favorable patients. IQL-favorable patients are older (avg age 47.96 vs. 43.44), disproportionately female (63.6% vs. 54.6%), more likely to have Medicare coverage (14.4% vs. 3.0%), and carry a higher comorbidity burden (ELX ≥ 5: 26.9% vs. 12.3%). State-level opioid dispensing rates are comparable across the two groups (49.58 vs. 50.21), which may rule out potential geographic confounding as a driver of the algorithms differences.

These patterns are consistent with the algorithmic mechanisms (see Remark 2). CQL's conservative out-of-distribution penalization can be beneficial for *homogeneous, low-complexity chronic patients* whose long treatment trajectories provide stable ground for incremental optimization. IQL's implicit conservatism (which implies staying close to the observed behavior from data) can be suitable for *complex, acute, high-comorbidity patients* for whom the observational data already provides rich

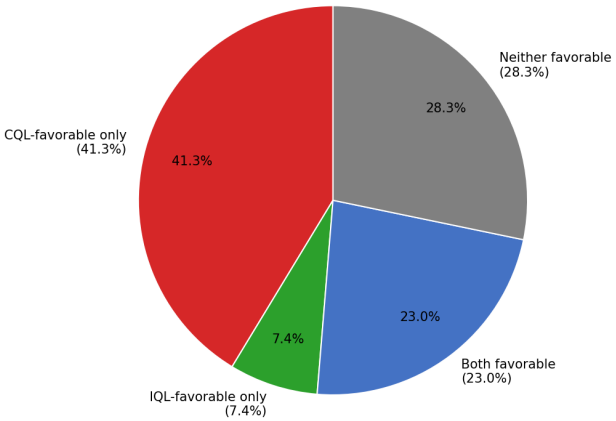


Figure 4 Patient classification by incremental NMB under CQL and IQL relative to observed practice

$N = 2,738,118$ patients. Results presented here are for WTP = \$20K/QALY. This classification is robust to the choice of WTP (see Appendix B.4 for more details).

Table 3 Covariate profiling by policy performance groups

Covariate	CQL-favorable	IQL-favorable	Both favorable	Neither favorable
	1,131,466 (41.3%)	202,035 (7.4%)	631,073 (23.0%)	773,544 (28.3%)
Daily opioid strength (MME)	28.26 (24.00)	23.13 (20.00)	35.60 (22.50)	26.46 (25.00)
Daily non-opioid strength (mg)	1221.83 (1300.00)	1093.98 (1200.00)	933.47 (812.50)	1474.24 (1625.00)
Days supply	8.46 (5.00)	16.01 (5.00)	28.87 (15.00)	8.65 (3.00)
Use of nonpharm. treatments	0.01 (0.00)	0.01 (0.00)	0.08 (0.00)	0.00 (0.00)
Age	43.44 (45.00)	47.96 (47.00)	54.44 (56.00)	42.81 (42.00)
State annual opioid dispense rate (per 100 residents)	49.58 (47.00)	50.21 (47.20)	51.42 (48.40)	50.00 (47.20)
# of unique days a patient makes visits in a window	0.04 (0.00)	0.11 (0.00)	0.15 (0.00)	0.04 (0.00)
# of unique overlapping prescriptions in a window	3.07 (2.00)	4.36 (4.00)	5.60 (5.00)	3.20 (2.00)
History of substance use (excluding opioid)[a]	10.0%	12.0%	11.0%	10.0%
History of surgical procedures[a]	16.0%	12.0%	32.0%	10.0%
Gender: Male	45.4%	36.4%	41.1%	39.8%
Female	54.6%	63.6%	58.9%	60.2%
Pain pathology: Surgery	24.3%	16.2%	41.7%	15.5%
Acute	8.1%	69.4%	22.9%	53.3%
Chronic	67.5%	14.4%	35.4%	31.3%
Comorbidity (Elixhauser): <0	72.7%	51.9%	48.0%	65.5%
1–4	15.1%	21.3%	21.8%	18.7%
5+	12.3%	26.9%	30.2%	15.8%
Insurance: Commercial	97.0%	85.6%	80.4%	96.1%
Medicare	3.0%	14.4%	19.6%	3.9%
Region: Northeast	11.1%	11.4%	12.7%	10.6%
North Central	19.9%	22.7%	24.2%	20.8%
South	53.1%	52.9%	48.4%	54.5%
West	15.9%	13.0%	14.7%	14.1%
Population size: Metro ≥1 million	58.1%	52.7%	51.9%	54.1%
Metro 0.25-1 million	21.0%	21.4%	21.7%	21.1%
Metro < 250K	6.9%	8.2%	7.9%	7.5%
Non-metro <50K	14.0%	17.7%	18.6%	17.3%

Continuous variables: mean (median). Categorical variables: %. Kruskal-Wallis and Chi-square tests for continuous and categorical variables, respectively. All estimates were statistically significant with $p < 0.001$. [a] For this variable, the value reported is the number of patients who ever experienced an event or had a history of a medical condition.

coverage of the relevant treatment space, making departures from observed practice unnecessary. Of note, caution should be exercised in following these recommendations, mainly because these insights are grounded in a descriptive nature, and we do not claim any causal relationship between patient characteristics and potential algorithmic benefits.

5.2.2. Tier 2: Algorithm Applicability. Here, we contrast two groups: algorithm-responsive patients, for whom at least one algorithm improves upon observed practice (71.7% of the population), and practice-responsive patients (28.3%), for whom observed practice would offer a more cost-effective treatment. This can help define the deployment boundary within which algorithmic treatment recommendations add value. It can also inform development and implementation of a *triage-based human-algorithm* approach in which only a subset of patients are routed to algorithmic-based treatments.

From Table 3, practice-responsive patients are distinct from the algorithm-responsive group in various respects. On aggregate, they are the youngest subgroup (avg age 42.81), present predominantly with acute pain (53.3%), and exhibit the highest average non-opioid treatment strength (1474.24 mg), and the lowest opioid strength (26.46 MME). This can be indicative of acute pain managed primarily

through non-opioid modalities. They also carry the lowest comorbidity burden (ELX < 0: 65.5%) and are predominantly commercially insured (96.1%), suggesting a relatively uncomplicated patient population for whom current prescribing practice might be already well-calibrated.

By contrast, algorithm-responsive patients tend to be older, carry higher opioid MME, have longer and more complex treatment histories (as evidenced by higher unique prescription days and days supply), and present more frequently with chronic pain or surgical indications. Notably, these conditions are also the ones under which our observational data contains sufficient structure for an offline RL algorithm to identify systematic improvements over practice.

The foregoing two tiers yield a coherent and clinically actionable picture: algorithmic policies add value for patients with complex, long-horizon treatment needs, and the specific algorithm to deploy depends on whether the patient's complexity is driven by chronicity and treatment stability (favoring CQL) or by acuity, comorbidity, and treatment intensity (favoring IQL). For straightforward acute patients managed with more favorable non-opioid treatments, observed clinical practice could remain the appropriate standard of care.

5.3. Algorithmic-Based Multi-Modal Pain Treatments

We compare the longitudinal treatment trajectories induced by the CQL and IQL policies across the four patient-policy subgroups identified in §5.2. From Figure 5, we observe the following:

Observation 2. *(i) Across all four subgroups, the CQL policy consistently recommends higher daily opioid strength but substantially shorter days-supply relative to both observed practice and the IQL policy. (ii) The two algorithms diverge most sharply on non-pharmacologic treatment use: relative to practice, CQL recommends near-zero non-pharmacologic treatment use across all four subgroups, whereas IQL recommends elevated and sustained non-pharmacologic treatment use in three of the four subgroups (the exception being the IQL-favorable subgroup, where the elevation over practice is more modest).*

The differences in treatment trajectories may not be merely algorithmic artifacts. Indeed, they may be aligned with patient complexity and baseline care structure. For the CQL-favorable subgroup, which consists of lower-complexity chronic pain patients with lower comorbidity burden, the relatively conservative treatment trajectories induced by CQL yield better cost-effectiveness relative to both observed practice and IQL. In this subgroup, the more intensive multimodal treatment patterns induced by IQL (e.g., sustained non-pharmacologic utilization and persistent pharmacologic support) may generate additional treatment burden and cost without proportional downstream benefit.

By contrast, the IQL-favorable subgroup exhibits substantially different longitudinal dynamics. Relative to observed practice, which is characterized by prolonged duration of supply, the IQL policy induces a more balanced multimodal restructuring strategy involving lower opioid exposure,

Figure 5 Multi-modal pain treatments obtained from algorithmic policies vs. the practice

Results are obtained for WTP = \$20K. x-axes represent time window (each window = 3 months). Shades represent 95% confidence intervals. Supply is zero when both opioid/non-opioid strengths are zero (and vice versa). non-pharmacologic treatments: 1 (0) indicates use (no use).

moderate duration compression, and increasing utilization of non-pharmacologic treatments over time. These trajectories may suggest that opioid prescription reduction under IQL does not occur in isolation, but rather alongside a sustained increase in supportive non-pharmacologic care, indicating a complementary substitution pattern across treatment modalities. Compared to the more compressed trajectories induced by CQL, the IQL policy appears to preserve a greater degree of longitudinal treatment support, which may be particularly valuable for patients with more acute, clinically dynamic, or comorbidity-intensive care trajectories.

For the both-favorable (i.e., algorithm-responsive) subgroup, both algorithms substantially improve upon observed practice, which itself exhibits the highest opioid burden, longest duration of supply, and greatest treatment intensity among all groups. Our results here suggest that structured longitudinal optimization alone, regardless of the specific offline RL paradigm employed, can yield improvements relative to existing prescribing patterns. This subgroup can, therefore, represent an "algorithmic opportunity zone" in which current clinical practice may be misaligned with economically favorable longitudinal care trajectories.

Finally, for the practice-responsive subgroup, observed practice exhibits relatively low opioid exposure, higher duration of supply, and low utilization of non-opioid and non-pharmacologic treatments. For this subgroup, neither RL policy yields meaningful improvement over practice, suggesting that algorithmic intervention may offer limited incremental value for lower-risk patients already receiving relatively efficient care.

Put together, these findings suggest that the value of an algorithmic-based treatment approach may depend not only on the quality of the underlying algorithm, but also on the alignment between the algorithm's induced treatment philosophy and the patient's clinical complexity regime. More broadly, our results indicate that algorithmic treatment policies are unlikely to be universally beneficial across all patient populations, and instead may be most effective when deployed selectively for patients whose longitudinal treatment trajectories contain sufficient complexity and opportunity for improvement relative to the observed clinical practice.

5.4. Robustness Checks

We check the robustness of our main findings against alternative modeling assumptions, temporal context, study population, and data preprocessing choices. Complete results for these robustness checks are presented in Online Appendix C.

5.4.1. Parameter Uncertainty: Quality-of-Life Scores, Costs, and Weights. We examine parameter uncertainty by perturbing the quality-of-life scores, event-related costs, and relative weights assigned to these events (Appendix C.1). First, across all perturbations of the quality-of-life scores, the cost-effectiveness probabilities remain nearly unchanged for both CQL and IQL.

This suggests that our main findings are not driven by the specific QALY values assumed in the baseline analysis. Second, increasing the cost assigned to an OUD incidence would improve the relative cost-effectiveness of policies, whereas increasing the cost assigned to a UTP incidence could weaken this advantage. Nevertheless, CQL remains consistently cost-effective across all cost perturbations, indicating that our main findings are robust against event-cost assumptions. Third, the weight-sensitivity analysis confirms that CQL remains cost-effective across most weighting scenarios; however, its advantage weakens when the objective is driven entirely by UTP, indicating that the value of CQL is rooted in balancing opioid-related harm against undertreated pain rather than optimizing either outcome in isolation. This result can further shed light on the CDC guidelines. To the extent that the guidelines primarily emphasize reducing opioid-related harms, our findings suggest that algorithmic policies can perform well under a similar policy objective. However, unlike threshold-based guideline approaches, our framework evaluates opioid-related harms jointly with undertreated pain, allowing the treatment policy to preserve a balance between reducing OUD risk and avoiding potential pain undertreatment.

5.4.2. Temporal Validity: Pre-COVID versus Post-COVID. Given the potential changes in healthcare utilization, pain management, and prescribing behavior during the COVID-19 pandemic (Lee et al. 2021), we assess temporal validity by repeating the analysis separately for the pre-COVID and post-COVID periods. Our results (Appendix C.2) show that cost-effectiveness declines in the post-COVID period under both algorithms. This pattern suggests that the post-COVID care environment might have reduced the incremental economic gains from algorithmic treatment policies, potentially due to pandemic-related disruptions in healthcare utilization, prescribing behavior, and access to non-pharmacologic pain-management modalities. Importantly, however, the qualitative ranking of the algorithms remains unchanged, as the CQL policy remains consistently more cost-effective than the human-based practices in both the pre-/post-COVID periods.

5.4.3. Population Generalizability: Medicaid Population. We examine population generalizability by replicating our analysis in the Medicaid population, which differs from the primary commercially insured population in patient mix, access patterns, and baseline risk. Our patient-level classification (Appendix C.3) reveals that: (i) the share of patients for whom CQL is only favorable compared to the practice falls from 41.3% to 24.8%; (ii) the share of patients for whom IQL is only favorable remains comparable (7.4% vs. 7.9%); (iii) the share of patients for whom either algorithm is favorable more than doubles, from 23.0% to 53.6%. The neither-favorable share (i.e., patients for whom the current practice performs better) contracts from 28.3% to 13.8%.

This reallocation toward algorithmic favorability also comes with a different treatment-intensity profile in the Medicaid cohort compared to the commercial and Medicare baseline (for more details,

see Appendix C.3). In the baseline cohort, IQL's opioid-strength and non-pharmacologic-treatment (NPH) recommendations diverge substantially from observed practice throughout the treatment horizon, while CQL's NPH recommendation remains near zero. In the Medicaid cohort, this pattern is reversed, with CQL recommending elevated NPH use early in the horizon while IQL tracks observed practice closely across opioid strength, days supply, and NPH use. This reversal might be related to how each algorithm's policy is anchored to population-specific empirical practice patterns. This reversal may be related to substantial differences in the underlying patient populations: the Medicaid cohort carries a markedly higher comorbidity burden and a higher prevalence of substance-use history than the baseline cohort (Appendix C.3), and a policy trained on this higher-risk population may anchor its recommendations to a different empirical practice pattern as a result. Of note, this as a structurally consistent hypothesis rather than an established mechanism. Specifically, we do not draw any causal link between this treatment-level divergence and the shift reported above, and leave a formal causal investigation of population-conditional policy behavior to future research.

5.4.4. Data Preprocessing Validity: Event 1 Incidence Rate. Event 1, captured by OUD, could be underdiagnosed or underreported in electronic health records due to a variety of reasons such as lack of specialized training in addiction medicine, undercoding in claims data, and stigma around these conditions (Palumbo et al. 2020, Cheetham et al. 2022). In our data, we observe a significant imbalance in the incidence rate of Event 1, which was observed to be 0.6%, compared to a more balanced incidence rate of 11.5% for Event 2. We improve the class imbalance via Synthetic Minority Over-sampling Technique (SMOTE) (Chawla et al. 2002), which is a combination of K-nearest neighbors classification and bootstrapping methods. We then replicate our analysis on the synthetic data. Our patient-level classification (Appendix C.4) reveals that addressing the class imbalance yields three noticeable shifts from our main analysis: (i) the share of patients for whom CQL is only favorable compared to the practice collapses from 41.3% to 14.5%; (ii) the share of patients for whom IQL is only favorable increases from 7.4% to 15.0%; (iii) the share of patients for whom either algorithms are favorable increases from 23.0% to 47.2%; and (iv) the neither-favorable share (i.e., portion of patients for whom the current practice performs better) remains comparable (28.3% vs. 23.2%).

The asymmetry across the two algorithms might be related to how each algorithm's inductive bias engages with rare-event signal. CQL's conservatism operates through an explicit penalty on actions unsupported in the data; the rare Event-1 cases already supply this signal, anchoring the policy to actions that empirically precede Event-1 outcomes. Synthetic oversampling amplifies Event-1 frequency but adds no new action-outcome information, so CQL's mechanism, already aligned with the event's structural rarity, gains little from rebalancing. IQL, by contrast, extracts its policy via

advantage-weighted behavior cloning over a value function fit by expectile regression (an estimator that is variance-sensitive in the upper tail), where Event-1 sparsity introduces noise into the value estimates driving its policy. Rebalancing stabilizes this distribution and improves the resulting signal-to-noise ratio. In short, the SMOTE results reflect a property of the algorithms, not the data: CQL remains robust because its conservatism needs no artificial rebalancing of the rare-event regime, while IQL is more responsive to the empirical incidence rate (for further discussion, see Levine et al. (2020), Kumar et al. (2020), Kostrikov et al. (2021)). A formal investigation of these mechanisms is beyond this paper's scope and left to future work.

5.4.5. Sensitivity to the Bunching Bandwidth for CDC 2022 Counterfactual. The construction of Benchmark 3 (CDC 2022 Counterfactual; §5.1) relies on a bunching bandwidth δ that determines which patient-visits are classified as constrained under the CDC 2016 caps (these are the candidates for upward adjustment under the CDC 2022 counterfactual). In the baseline analysis, we use $\delta = 0.15$. To assess whether our cost-effectiveness comparisons against Benchmark 3 are sensitive to this choice, we re-run the full CE analysis under narrower and wider bandwidths ($\delta = 0.05$ and 0.25, respectively). Across both alternatives, the qualitative ordering of CQL and IQL is preserved, and the magnitudes of their cost-effectiveness probabilities against Benchmark 3 only marginally change (Appendix C.5). Our conclusions regarding CDC 2022 counterfactual are therefore robust to the choice of bunching bandwidth.

6. Policy Implications

We now make use of our main findings and summarize a few important insights that could enable authorities impose more effective policies.

Avoiding the naive OUD–UTP tradeoff. Given that our framework jointly optimizes against both opioid use disorder (OUD) and undertreated pain (UTP), one can expect our recommended policies to favor more permissive opioid prescribing relative to practice. Our results, however, indicate the opposite mechanism. Compared to observed practice, our proposed algorithmic-based policy recommends a markedly shorter days-supply alongside a moderately higher daily opioid dose (§5.3), and yet achieves a better cost-effectiveness than the practice (§5.1). This pattern is consistent with a substantial body of clinical literature finding that cumulative duration of opioid exposure, rather than daily dose alone, is among the strongest predictors of progression to long-term use and opioid use disorder (Hadlandsmyth et al. 2020, Johnson et al. 2022). Indeed, daily dose has separately been identified as a contributing risk factor, but several studies find duration to carry comparatively greater and more independent predictive weight (Kurteva et al. 2021). Our policy's restructuring of the dose-duration profile, while permitting some increase along a comparatively

secondary dimension, could offer one plausible account of how a lower OUD risk is achieved without resorting to more conservative opioid dose prescriptions across the board.

Holistic pain treatments. Beyond opioid dose and duration, our proposed policies diverge from observed practice across non-opioid analgesic intensity and non-pharmacologic treatment use. Relative to practice, the CQL policy recommends little to no use of non-pharmacologic treatments throughout the therapeutic course, while the IQL policy recommends them at a substantially higher rate; the two algorithms' non-opioid analgesic recommendations likewise diverge from practice in different directions and at different points along the treatment horizon (§5.3). Rather than uniformly (dis)favoring any single treatment modality, each algorithm restructures the overall treatment profile along multiple dimensions at once. This shows that the value of an algorithmic-based policy lies not in adjusting any one lever in isolation, but in the joint and holistic reconfiguration of a complete pain treatment plan.

CDC guideline revision. The CDC has issued two sets of guidelines for opioid prescribing in pain management, in 2016 and 2022. A key distinction between them is that the 2022 guidelines relaxed the explicit dose and days-supply thresholds recommended in 2016. Our results show that, contrary to what this relaxation might suggest or have intended to, the more restrictive 2016 guidelines are more cost-effective than the 2022 guidelines (§5.1). At the same time, our proposed algorithmic-based policies outperform both sets of guidelines, indicating that the relevant choice is not merely between more or less restrictive fixed thresholds, but between fixed, population-wide thresholds of any kind and a personalized treatment policy. We hope these results can inform authorities in calibrating future guideline revisions.

Policy implementation considerations. When considering the implementation of our proposed algorithmic-based approach, the following factors should be taken into account. First, our multimodal pain treatments used both pharmacologic and nonpharmacologic options. However, in the clinical practice, these treatments are typically administered by different providers, and some of these treatments may not be covered by insurers. Further considerations should be taken to safeguard care coordination between different providers in charge as well as potential issues that might arise due to the lack of coverage by payers. Second, implementing our approach in practice requires addressing issues such as the "degree of automation" (how much should clinicians be able to override our algorithmic-based recommendations?) or "human-in-the-loop" (how much should the human judgment be combined with our algorithmic-based approach?). Finally, such implementation would also require additional efforts which might be costly, including collecting and hosting large-scale training data sets, educating providers, and ensuring adherence to the models' recommendations. Despite these, we believe the authorities could carefully consider adopting

recommendations stemming from our framework as they may not only improve health outcomes (e.g., as measured by the QALYs gained), but also yield significant cost savings. Notably, our results show that, over a three-year therapeutic course, while improving QALYs by 6.6 days per patient, the average cost saved per patient under our treatment policies compared to the observed practice of physicians is $1,927 (see §5.1). Given that we obtain these results from a pool of over 2.7 million U.S. patients, this would translate to more than $5.2 billion in cost and more than 48,800 years worth of lives anticipated to be saved for such population. Therefore, we believe the overall benefits obtained from following an approach like ours is well worth the costs.

7. Conclusion and Limitations

In this study, we develop an algorithmic-based approach for personalized, cost-effective, multi-modal pain treatment strategies by making use of data from over 2.7 million patients. Across extensive comparisons against both the human-based practice of physicians and the CDC's 2016 and 2022 guidelines, our proposed approach consistently outperforms these alternatives. Notably, our comparisons enable us to derive important policy implications that might be invaluable for policymakers, physicians, and various authorities. Future research can build on our study along several fronts: running randomized control experiments to further investigate the validity of our policy recommendations; developing alternative computational methods for addressing the curse of dimensionality inherent to multi-period, multi-modal, personalized treatment planning; exploring better ways of combining human intuition with algorithmic recommendations, an area in which human-algorithm "centaurs" have shown promise in healthcare and beyond (Orfanoudaki et al. 2022, Saghafian and Idan 2024); and exploring large language model (LLM)-based approaches to multi-modal treatment recommendation. The latter presents a promising but non-trivial direction: unlike our offline RL framework, foundation models are not typically equipped to emit calibrated continuous dosing and days-supply recommendations, nor do they come with an offline counterfactual evaluation machinery comparable to the Q-function-based estimates underlying CQL and IQL.

Our study is subject to limitations. (1) Our characterization of Event 2 (UTP) is a constructed proxy rather than a directly observed clinical outcome: unlike Event 1 (OUD), which we identify from diagnosis codes, undertreated pain is not recorded as such in claims data, and we do not have access to ground-truth clinical assessments (e.g., patient-reported pain scores) against which to validate our characterization. While we extensiely document our construction of this outcome in detail (§3.1), we cannot rule out that it might systematically over- or under-counts true clinical undertreatment. Thus, results pertaining to Event 2 should be interpreted with caution. (2) Our framework relies on standard identifying assumptions, sufficient coverage of the state-action space and the absence of unmeasured confounders jointly affecting treatment and outcomes, that cannot

be verified from claims data alone. Estimated policy performance is therefore a counterfactual construct, not an observed outcome, and where these assumptions are violated, estimated benefits relative to observed practice may be overstated. (3) We have intentionally avoided making causal claims throughout our study, and have left such claims to future work, since any related investigation might require running a randomized controlled experiment. (4) Our counterfactual benchmarks of CDC 2016 and 2022 guidelines are themselves modeled constructs rather than independently observed comparators. Comparisons against these benchmarks are therefore only as reliable as the modeling assumptions underlying their construction, and should not be interpreted as comparisons against a gold standard of an independently verified clinical trial. (5) Our state representation summarizes history into a fixed-dimension vector (§4.2), not the full, unbounded sequence; some information loss relative to both the full history and our outcome model's richer sequence-level information is inherent. Combined with the stochastic policy and environment model used at evaluation, CQL and IQL therefore solve an approximate, not exact, version of the deterministic objective in (7).

References

AMA (2016) AMA comments on CDC's Proposed Opioids Prescription Policy. Retrieved Apr 10, 2021, Link.

Attari I, Helm JE, Mejia J (2024) Hiding behind complexity: Supply chain, oversight, race, and the opioid crisis. *Production and Operations Management.* 34(4):725–743.

AZDHS: Arizona Department of Health Services (2017) Appendix B: State-by-state summary of opioid prescribing regulations and guidelines. Retrieved Dec 20, 2020, Link.

Barnett ML, Olenski AR, Jena AB (2017) Opioid-prescribing patterns of emergency physicians and risk of long-term use. *New England Journal of Medicine.* 376(7):663–673.

Bastani H, Bayati M (2020) Online decision-making with high-dimensional covariates. *Operations Research.* 68(1):276–294.

Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. *Management Science.* 62(5):1511–1531.

Bjarnadottir M, Anderson D, Agarwal R, Prasad K, Nelson A (2019) Continuity of care versus a second opinion: Evidence from the opioid crisis. *INFORMS Annual Conference.* Seattle, WA, USA. October 2019.

Bobroske K, Freeman M, Huan L, Cattrell A, Scholtes S (2022) Curbing the opioid epidemic at its root: The effect of provider discordance after opioid initiation. *Management Science.* 68(3):2003–2015.

Boloori A, Saghafian S, Chakkera HA, Cook CB (2020a) Data-driven management of post-transplant medications: An ambiguous partially observable Markov decision process approach. *Manufacturing & Service Operations Management.* 22(5):1066–1087.

Boloori A, Arnetz BB, Viens F, Maiti T, Arnetz JE (2020b) Misalignment of stakeholder incentives in the opioid crisis. *International Journal of Environmental Research and Public Health.* 17(20):7535.

Capitaine L, Genuer R, Thiébaut R (2018) Random forests for high-dimensional longitudinal data. Retrieved Nov 10, 2020, Link.

CDC: Centers for Disease Ccontrol and Prevention (2016) Pocket card for prescribing opioids for chronic pain. Retrieved Feb 16, 2020, Link.

CDC: Centers for Disease Ccontrol and Prevention (2020) Analyzing Opioid Prescription Data and Oral Morphine Milligram Equivalents (MME). Retrieved May 16, 2021, Link.

CDC: Centers for Disease Ccontrol and Prevention (2021) U.S. Opioid Dispensing Rate Maps. Retrieved Apr 02, 2023, Link.

CDC: Centers for Disease Control and Prevention (2022) Federal Register Notice: CDC's updated Clinical Practice Guideline for Prescribing Opioids is now open for public comment. Retrieved Apr 23, 2022, Link.

census.gov: The United States Census Bureau (2021) Metropolitan and Micropolitan Statistical Area Datasets. Retrieved Apr 04, 2023, Link.

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research.* 16:321–357.

Che Z, Sauver JS, Liu H, Liu Y (2017) Deep learning solutions for classifying patients on opioid use. *AMIA Annual Symposium Proceedings.* 2017:525–534.

Cheetham A, Picco L, Barnett A, Lubman DI, Nielsen S (2022) The impact of stigma on people with opioid use disorder, opioid treatment, and policy. *Substance Abuse and Rehabilitation.* 13:1–12.

Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J (2016) Doctor AI: Predicting clinical events via recurrent neural networks. *Machine Learning for Healthcare Conference.* 2016:301–318.

Ciesielski T, Iyengar R, Bothra A, Tomala D, Cislo G, Gage BF (2016) A tool to assess risk of de novo opioid abuse or dependence. *The American Journal of Medicine.* 129(7):699–705.

Crosier BS, Borodovsky J, Mateu-Gelabert P, Guarino H (2017) Finding a needle in the haystack: Using machine-learning to predict overdose in opioid users. *Drug and Alcohol Dependence.* 171:e49.

Delcher C, Harris DR, Park C, Strickler GK, Talbert J, Freeman PR (2021) "Doctor and Pharmacy Shopping": A fading signal for prescription opioid use monitoring?. *Drug and Alcohol Dependence.* 221:108618.

DHCS.CA.gov (2026) Substance Use Disorder ICD 10 Analysis Included Codes. Retrieved Apr 09, 2026, Link.

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General.* 144(1):114.

Dowell D, Haegerich TM, Chou R (2016) CDC guideline for prescribing opioids for chronic pain—United States, 2016. *Jama.* 315(15):1624–1645.

Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW (2015) Methods for the economic evaluation of health care programmes. *Oxford University Press.*

Elmachtoub AN, Grigas P (2022) Smart "predict, then optimize". *Management Science.* 68(1):9–26.

Elixhauser A, Steiner C, Harris DR, Coffey RM (1998) Comorbidity measures for use with administrative data. *Medical Care.* 36(1):8–27.

ers.usda.gov: Economic Research Service – U.S. Department of Agriculture (2020) Rural-Urban Continuum Codes. Retrieved Apr 04, 2023, Link.

Fenwick E, Marshall DA, Levy AR, Nichol G (2006) Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. *BMC Health Services Research.* 6(1):52.

Fishbain DA, Cole B, Lewis J, Rosomoff HL, Rosomoff RS (2007) What percentage of chronic nonmalignant pain patients exposed to chronic opioid analgesic therapy develop abuse/addiction and/or aberrant drug-related behaviors? A structured evidence-based review. *Pain Medicine.* 9(4):444–459.

Gan K, Scheller-Wolf AA, Tayur SR (2019) Personalized Treatment for Opioid Use Disorder. Retrieved Jul 25, 2021, Link.

Hadlandsmyth K, Mosher HJ, Vander Weg MW, O'Shea AM, McCoy KD, and others (2020) Utility of accumulated opioid supply days and individual patient factors in predicting probability of transitioning to long-term opioid use: An observational study in the Veterans Health Administration. *Pharmacology Research & Perspectives.* 8(2):e00571.

Haller IV, Renier CM, Juusola M, Hitz P, Steffen W, and others (2016) Enhancing risk assessment in patients receiving chronic opioid analgesic therapy using natural language processing. *Pain Medicine.* 18(10):1952–1960.

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. *Springer.*

Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project. (2025) Chronic Condition Indicator Refined (CCIR) for ICD-10-CM. Retrieved Apr 15, 2025, Link.

Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Computation.* 9(8):1735–1780.

Huskisson EC (1974) Measurement of pain. *The Lancet.* 304(7889):1127–1131.

ICD10 (2026a) ICD-10-CM Diagnosis Codes (F10-F19 series). Retrieved Nov 17, 2024 (earlier version), Link.

ICD10 (2026b) ICD-10-CM Diagnosis Codes (G89 series). Retrieved Oct 13, 2024 (earlier version), Link.

Johnson SP, Chung KC, Zhong L, Shauver MJ, Engelsbe MJ, and others (2016) Risk of prolonged opioid use among opioid-naïve patients following common hand surgery procedures. *The Journal of Hand Surgery.* 41(10):947–957.

Johnson DG, Ho VT, Hah JM, Humphreys K, Carroll I, and other (2022) Prescription quantity and duration predict progression from acute to chronic opioid use in opioid-naïve Medicaid patients. *PLOS Digital Health.* 1(8):e0000075.

Katz J, Melzack R (1999) Measurement of pain. *Surgical Clinics of North America.* 79(2):231–252.

Kendler KS, Lönn SL, Ektor-Andersen J, Sundquist J, Sundquist K (2023) Risk factors for the development of opioid use disorder after first opioid prescription: a Swedish national study. *Psychological Medicine.* 53(13):6223–6231.

Kondrup F, Jiralerspong T, Lau E, de Lara N, Shkrob J, and others (2023) Towards safe mechanical ventilation treatment using deep offline reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence.* 37:15696–15702.

Kumar A, Zhou A, Tucker G, Levine S (2020) Conservative Q-learning for offline reinforcement learning. *Advances in Neural Information Processing Systems.* 33:1179–1191.

Kurteva S, Abrahamowicz M, Gomes T, Tamblyn R (2021) Association of opioid consumption profiles after hospitalization with risk of adverse health care events. *JAMA Network Open.* 4(5):e218782.

Kostrikov I, Nair A, Levine S (2021) Offline reinforcement learning with implicit Q-learning. Retrieved Oct 15, 2025, Link.

Lee B, Yang KC, Kaminski P, Peng S, Odabas M, and others (2021) Substitution of nonpharmacologic therapy with opioid prescribing for pain during the COVID-19 pandemic. *JAMA Network Open.* 4(12):e2138453.

Levine S, Kumar A, Tucker G, Fu J (2020) Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems. Retrieved Feb 12, 2026, Link.

Lin S, Saghafian S, Lipschitz JM, Burdick KE (2025) A multiagent reinforcement learning algorithm for personalized recommendations in bipolar disorder. *PNAS Nexus.* 4(8): pgaf246.

Lo-Ciganic WH, Donohue JM, Yang Q, Huang JL, Chang CY, and others (2022) Developing and validating a machine-learning algorithm to predict opioid overdose in Medicaid beneficiaries in two US states: a prognostic modelling study. *The Lancet Digital Health.* 4(6):e455–e465.

Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes.* 151:90–103.

Luo F, Li M, Florence C (2021) State-level economic costs of opioid use disorder and fatal opioid overdose–United States, 2017. *Morbidity and Mortality Weekly Report.* 70(15):541–546.

Luo J, Stellato B (2024) Frontiers in operations: Equitable data-driven facility location and resource allocation to fight the opioid epidemic. *Manufacturing & Service Operations Management.* 26(4):1229–1244.

Mayhew M, DeBar LL, Deyo RA, Kerns RD, Goulet JL, and others (2019) Development and assessment of a crosswalk between ICD-9-CM and ICD-10-CM to identify patients with common pain conditions. *Journal of Pain.* 20(12):1429–1445.

McPhillips-Tangum CA, Cherkin DC, Rhodes LA, Markham C (1998) Reasons for repeated medical visits among patients with chronic back pain. Announcement. *Journal of General Internal Medicine.* 13(5):289–295.

Menendez ME, Neuhaus V, van Dijk CN, Ring D (2014) The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery. *Clinical Orthopaedics and Related Research.* 472(9):2878–2886.

Mišić VV (2020) Optimization of tree ensembles. *Operations Research.* 68(5):1605–1624.

NIDA: National Institute on Drug Abuse (2022) Overdose Death Rates. Retrieved Jul 07, 2022, Link.

NIH.gov (2016) Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Retrieved May 17, 2024, Link.

NPR: National Public Radio (2022) CDC weighs new opioid prescribing guidelines amid controversy over old ones. Retrieved Apr 21, 2022, Link.

Nuckols TK, Anderson L, Popescu I, Diamant AL, Doyle B, and others (2014) Opioid prescribing: a systematic review and critical appraisal of guidelines for chronic pain. *Annals of Internal Medicine.* 160(1):38–47.

Nursing (2017) Sickle cell disease. Retrieved Aug 02, 2023, Link.

Orfanoudaki A, Saghafian S, Song K, Chakkera HA, Cook C (2022) Algorithm, Human, or the Centaur: How to Enhance Clinical Care?. Retrieved Apr 17, 2024, Link.

Overton HN, Hanna MN, Bruhn WE, Hutfless S, Bicket MC, and others (2018) Opioid-prescribing guidelines for common surgical procedures: An expert panel consensus. *Journal of the American College of Surgeons.* 227(4):411–418.

Palumbo SA, Adamson KM, Krishnamurthy S, Manoharan S, Beiler D, and others (2020) Assessment of probable opioid use disorder using electronic health record documentation. *JAMA Network Open.* 3(9):e2015909–e2015909.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature.* 323:533—536.

Palliative Drugs (2009) Opioid conversions. Retrieved Jan 10, 2019, Link.

Phillips JK, Ford MA, Bonnie RJ (Eds.) (2017) Pain management and the opioid epidemic: balancing societal and individual benefits and risks of prescription opioid use. *National Academies Press.*

Pitt AL, Humphreys K, Brandeau ML (2018) Modeling health benefits and harms of public policy responses to the US opioid epidemic. *American Journal of Public Health.* 108(10):1394–1400.

Pletcher MJ, Kertesz SG, Kohn MA, Gonzales R (2008) Trends in opioid prescribing by race/ethnicity for patients seeking care in US emergency departments. *JAMA.* 299(1):70–78.

Powell WB (2011) *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd ed. (John Wiley & Sons, Hoboken, NJ).

Robins JM, Hernán MA, Brumback B (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology.* 11(5):550–560.

Saghafian S (2018) Ambiguous partially observable Markov decision processes: Structural results and applications. *Journal of Economic Theory.* 178:1–35.

Saghafian S, Murphy SA (2021) Innovative health care delivery: The scientific and regulatory challenges in designing mHealth interventions. *NAM Perspectives.*

Saghafian S, Idan L (2024) Effective Generative AI: The Human-Algorithm Centaur. Retrieved Apr 17, 2024, Link.

Saghafian S (2024) Ambiguous Dynamic Treatment Regimes: A Reinforcement Learning Approach. *Management Science.* 70(9): 5667–5690.

Sexton J (2018) Historical Tree Ensembles for Longitudinal Data. Retrieved Mar 15, 2020, Link.

Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. *Information Processing & Management.* 45(4):427–437.

Smith CA, Levett KM, Collins CT, Armour M, Dahlen HG, Suganuma M (2018) Relaxation techniques for pain management in labour. *Cochrane Database of Systematic Reviews.* Issue 3. Art. No.: CD009514. DOI: 10.1002/14651858.CD009514.pub2.

STAT (2022) As a pain specialist, I may have caused more harm by underprescribing opioids. Retrieved Aug 02, 2023, Link.

Teater D (2015) Evidence for the efficacy of pain medications. Retrieved Dec 10, 2019, Link.

University of Manitoba (2023) Concept: Elixhauser Comorbidity Index. Retrieved Apr 10, 2023, Link.

Vunikili R, Glicksberg BS, Johnson KW, Dudley J, Subramanian L, Khader S (2018) Predictive modeling of susceptibility to substance abuse, mortality and drug-drug interactions in opioid patients. Retrieved Mar 01, 2020, Link.

Wells N, Pasero C, McCaffery M (2008) Improving the quality of care through pain assessment and management. In *Patient safety and quality: An evidence-based handbook for nurses. Agency for Healthcare Research and Quality (US).*

Wilson JR, Lorenz KA (2015) Modeling Binary Correlated Responses using SAS, SPSS and R. *Springer.* 103–130.

Wu CL, King AB, Geiger TM, Grant MC, Grocott MP, and others (2019) American society for enhanced recovery and perioperative quality initiative joint consensus statement on perioperative opioid minimization in opioid-naïve patients. *Anesthesia and Analgesia.* 129(2):567.

Xiao H, Barber JP (2008) The effect of perceived health status on patient satisfaction. *Value in Health.* 11(4):719–725.

Xue L, Yin R, Cole ES, Lo-Ciganic WH, Gellad WF, and others (2025) Development and evaluation of a machine learning model to predict acute care for opioid use disorder among Medicaid enrollees engaged in a community-based treatment program. *Addiction.* 120(9):1780–1789.

Zgierska A, Rabago D, Miller MM (2014) Impact of patient satisfaction ratings on physicians and clinical care. *Patient Preference and Adherence.* 8:437–446.

Online Appendices to

"Human-Based Versus Algorithmic-Based Treatments: Evidence from the Opioid Epidemic"

A. Information related to characterization of variables

A.1. Surgical procedures

Admission type: 1 (surgical)

Procedure group/type: 2 (acne surgery), 8 (excision of breast tissue), 9 (other minor skin/breast surgery), 12 (other major skin surgery), 13 (other major breast surgery), 14 (other major musculoskeletal surgery), 15 (other minor musculoskeletal surgery), 47 (repair of inguinal hernia), 52 (transurethral surgery), 66 (decompression, carpal tunnel), 76 (cataract removal), 98 (other minor surgery procedures), 99 (other major surgery procedures), 158 (percutaneous angioplasty), 440 (cesarean section deliveries), 445 (vaginal deliveries), 470 (anesthesia services), 490–494 (dental operations)

Provider type: 5 (ambulatory surgery centers), 355 (plastic/maxillofacial surgery), 505 (surgical specialist), 510 (colon/rectal surgery), 520 (neurological surgery), 530 (orthopaedic surgery), 535 (abdominal surgery), 540 (cardiovascular surgery), 545 (dermatologic surgery), 550 (general vascular surgery), 555 (head/neck surgery), 560 (pediatric surgery), 565 (surgical critical care), 570 (transplant surgery), 575 (traumatic surgery), 580 (cardiothoracic surgery), 585 (thoracic surgery), 820 (midwife)

Place of service: 24 (ambulatory surgical center), 25 (birthing center)

Variables in bold are available from the MarketScan data.

A.2. Chronic pain conditions

M54.5: Lumbago (back); M54.3: Sciatica (back); M54.8, M54.9: Backache, unspecified; M48.02: Spinal stenosis in cervical region (neck); M54.2: Cervicalgia (neck); M53.0: Cervicocranial syndrome (neck); R60.0: Swelling of limb (limb/extremity); R25.2: Cramps in limb (limb/extremity); M79.89: Other disorders of soft tissue (limb/extremity); M19.90: Osteoarthrosis, unspecified whether generalized or localized (joint); M25.511, M25.512: Pain in joint, shoulder region (joint)

Sample ICD-10 diagnoses codes are provided here (Mayhew et al. 2019). For the full list of these codes, see link.

A.3. List of pharmacologic medications prescribed for pain treatment

(a) Before name decomposition (raw data)

Opioids (50 generic drug names)†: 'Acetaminophen/caffeine/dihydrocodeine Bitartrate' 'Acetaminophen/codeine Phosphate' 'Acetaminophen/hydrocodone Bitartrate' 'Acetaminophen/oxycodone Hydrochloride' 'Acetaminophen/propoxyphene Hydrochloride' 'Acetaminophen/propoxyphene Napsylate' 'Alfentanil Hydrochloride' 'Acetaminophen/butalbital/caff/codeine Phos' 'Apomorphine Hydrochloride' 'Aspirin/oxycodone Hcl/oxycodone Terephthalate' 'Aspirin/butalbital/caffeine/codeine Phosphate' 'Aspirin/caffeine/dihydrocodeine Bitartrate' 'Aspirin/caffeine/propoxyphene Hydrochloride'
Non-opioids (51 generic drug names)†: 'Acetaminophen' 'Acetaminophen/aspirin/caffeine' 'Acetaminophen/butalbital' 'Acetaminophen/butalbital/caffeine' 'Acetaminophen/caffeine/phenyltoloxamine Citrate' 'Acetaminophen/diphenhydramine Hydrochloride' 'Acetaminophen/pamabrom/pyrilamine Maleate' 'Acetaminophen/phenyltoloxamine' 'Acetaminophen/phenyltoloxamine Citrate'

† For brevity, samples of drugs' names are provided.

(b) After name decomposition (after step 1, see §3.2 in the main body)

Opioids (21 unique components): 'Alfentanil' 'Buprenorphine' 'Butorphanol' 'Codeine' 'Dihydrocodeine' 'Fentanyl' 'Hydrocodone' 'Hydromorphone' 'Levorphanol' 'Meperidine' 'Methadone' 'Morphine' 'morphine' 'Opium' 'Oxycodone' 'Oxymorphone' 'Pentazocine' 'Propoxyphene' 'Sufentanil' 'Tapentadol' 'Tramadol' 'Nalbuphine'

A.4. Non-pharmacologic pain treatments prescribed

Procedures CPT codes:

Chiropractic and other related services: 720x0 for x ∈{2,4,7}, 72170, 72190, 72200, 72220, 9701x for x ∈{0,2,4,8}, 97022, 97026, 9703x for x ∈{2,3,5,9}, 9711x for x ∈{0,2,3,6}, 97124, 97140, 97161, 97162, 97530, 97535, 97750, 9894x for x ∈{0,1,2,3}, 9920x for x ∈{2,3,4}, 9921x for x ∈{1,2,3,4}

Psychotherapy, cognitive behavioral therapy, etc.: 9083x for x ∈{2,3,4,6,7,8}, 90839-90840, 9084x for x ∈{5,6,7,9}, 90853

Physical medicine and rehabilitation: 97001-97002, 9701x for x ∈{0,2,4,6,8}, 9702x for x ∈{2,4,6,8}, 9703x for x ∈{2,3,4,5,6,9}, 9711x for x ∈{0,2,3,6}, 97124, 97129, 97130, 97139, 97140, 97150, 9715x for x ∈{0,1,2,3,4,5,6,7,8}, 9716x for x ∈{1,2,3,4,5,6,7,8,9}, 9717x for x ∈{0,1,2}, 9753x for x ∈{0,3,5,7}, 9754x for x ∈{2,5,6}, 97597-97598, 9760x for x ∈{2,5,6,7,8}, 97610, 97750, 97755, 9776x for x ∈{0,1,3}, 97799

Interventional pain management: 27096, 6228x for x ∈{0,1,2,7}, 6232x for x ∈{0,1,2,3,4,5,6,7}, 6235x for x ∈{0,1,5}, 6236x for x ∈{0,1,2,5}, 63650, 63655, 63663-63664, 63685, 63688, 64479, 64480, 64483-64484, 6449x for x ∈{0,1,2,3,4,5}, 64510, 64520, 6463x for x ∈{3,4,5,6}

Acupuncture: 9781x for x ∈{0,1,3,4}

Procedure groups: 181–189 (physical medicine), 191 and 195 (chiropractic/spinal manipulation)

Provider types: 120 (chiropractor), 140 (pain management), 350 (physical medicine and rehab), 865 (acupuncturist), 870 (spiritual healers)

Service categories: 3xy18 for x∈ $\{0,1\}$ and y∈ $\{1,2,3,4,5,6\}$ (behavioral health therapy)

Categories in bold are available from the MarketScan data.

B. Numerical experiments (supplement)

B.1. Summary of baseline parameters

Parameter	Description
$T = 12$	Time horizon (12 periods = 36 months), referred to as the therapeutic course
WTP = 20,000 (\$/QALY)	Willingness to pay
$q_{10} = q_{20} = 1/4 = 0.25$ yr[a]	Monthly quality-of-life (qol) score for not experiencing Events 1 or 2 (i.e., perfect health)
$q_{11} = 0.368/4 = 0.093$ yr[b]	Monthly qol score for experiencing Event 1 (see, e.g., De Maeyer et al. (2010))
$q_{21} = \begin{cases} 0.395/4 = 0.100 \text{ yr (acute)} \\ 0.470/4 = 0.108 \text{ yr (chronic)} \end{cases}$	Monthly qol score for experiencing Event 2 (for an acute and chronic pain pre-supply) (see, e.g., Katz (2002), Wu et al. (2003), Tüzün (2007), Ataoğlu et al. (2013), Taylor et al. (2013), Hadi et al. (2019) and references therein)
$C_1 = \$15{,}588.07$	Costs associated with an occurrence of Event 1[e]
$C_2 = \begin{cases} \$6,172.86 \text{ (acute)} \\ \$2,714.90 \text{ (chronic)} \end{cases}$	Costs associated with an occurrence of Event 2 (for an acute and chronic pain)[e]
$w_1 = w_2 = 0.5$	Weights assigned to Events 1-2

[a] Yearly qol scores are converted to their quarterly equivalents.
[b] We consider a same qol score for both genders (see, e.g., Giacomuzzi et al. (2005) and Domingo-Salvany et al. (2010)).

B.2. Operationalization of benchmark policies

We construct three benchmark policies against which we evaluate the cost-effectiveness of our algorithmic treatment policies (CQL and IQL). All benchmarks are derived from the same observational data (2017–2022) and evaluated under the common environment model (see §5.1 in the main body). Of note, the benchmarks differ exclusively in how opioid-related treatment decisions are constructed (i.e., strength and duration of supply); non-opioid and non-pharmacologic treatment modes are held fixed throughout at their observed values from our data.

Benchmark 1: Observed Clinical Practice. This benchmark uses the raw treatment decisions recorded in our data. For every patient i at visit t, the prescribed opioid dose a_{it}^{mme}, days of supply a_{it}^{days}, non-opioid dose, and non-pharmacologic treatment indicator are set exactly to their observed values. Benchmark 1, therefore, captures real-world prescribing behavior over the sample period.

Benchmark 2: CDC 2016 (Counterfactual Constrained Practice). The 2016 CDC guideline imposed structured limits on opioid prescribing: a maximum of 90 morphine milligram equivalents (MME) per day for opioid strength and a maximum of 7 (acute) or 90 (chronic) days of supply per

prescription (Dowell et al. 2016, CDC 2016, AZDHS 2017). Benchmark 2 operationalizes these recommendations by applying hard caps uniformly to the observed actions:

$$\tilde{a}_{it}^{\text{mme}} = \min\left(a_{it}^{\text{mme}}, 90\right), \qquad \tilde{a}_{it}^{\text{days}} = \max\left(0, \min\left(a_{it}^{\text{days}}, 90 - D_{it}\right)\right), \tag{EC.1}$$

where $D_{it} = \sum_{s=1}^{t-1} \tilde{a}_{is}^{\text{days}}$ denote the cumulative days of supply dispensed to patient i through period $t-1$, with $D_{i1} = 0$, so that the running total $D_{it} + \tilde{a}_{it}^{\text{days}}$ never exceeds 90 days. Once the cumulative budget is exhausted ($D_{it} \geq 90$), no further days of supply are dispensed for the remainder of the patient's observed history. Of note, all other treatment modes remain as observed. Benchmark 2 would represent a counterfactual in which the 2016 guideline recommendations were implemented *consistently* rather than *heterogeneously*.

Benchmark 3: CDC 2022 (Counterfactual Unconstrained Practice). The 2022 revision of the CDC guideline departed from its predecessor in one critical respect: it explicitly moved away from population-level hard thresholds toward patient-specific clinical judgment, acknowledging that rigid caps may be inappropriate for a subset of patients (CDC 2022).

Because our data predate widespread adoption of the 2022 guideline, we operationalize its intent as a structured counterfactual. The key identifying assumption is that patients whose observed treatment is bunched just at or below a CDC 2016 cap could be those whose treatment might have been constrained. These are the patients who would, under the 2022 guideline, be candidates for individualized upward adjustment. Based on this premise, we identify patient-visit (i, t) as constrained if

$$a_{it}^{\text{mme}} \in \left[(1-\delta)\cdot 90,\ 90\right] \quad \text{or} \quad a_{it}^{\text{days}} \in \left[(1-\delta)\cdot D_{it},\ D_{it}\right], \tag{EC.2}$$

In our baseline specification, we set $\delta = 0.15$ a moderate bunching bandwidth. In our robustness checks, we will use $\delta \in \{0.05, 0.25\}$ to evaluate less/more conservative thresholds; see Appendix C.5.

Of note, for the rest of patients (i.e., unidentified), Benchmark 3 mimics Benchmark 1: the observed action remains unchanged. For identified patient-visits, we model the observed treatment intensity (e.g., opioid dose or days of supply) as censored from above by the corresponding thresholds imposed by the CDC 2016 guideline and impute the latent treatment intensity using a censored (Tobit) regression (Tobin 1958, Duan et al. 1983):

$$a_{it}^{*} = \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathcal{N}(0, \sigma^2), \tag{EC.3}$$

where a_{it}^{*} is the latent (uncensored) treatment intensity, $\mathbf{x}_{it}$ is a vector of patient-level covariates (e.g., pain pathology, demographics, etc.), and the model is estimated on all observations with censoring at the applicable CDC 2016 cap.

Counterfactual doses for constrained patient-visits are then drawn from the right-truncated distribution $\mathcal{N}(\hat{\mu}_{it}, \hat{\sigma}^2)$ truncated to $[\bar{c}_{it}, \infty)$, where $\bar{c}_{it}$ is the relevant cap. By construction, every imputed value weakly exceeds the observed value, consistent with the notion of the 2022 guideline.

B.3. Validation of CDC counterfactuals

To verify the validity of our benchmark constructions prior to cost-effectiveness evaluation, we conduct five diagnostic checks, serving a dual purpose: they confirm that the counterfactual simulations behave as intended, and they surface any implementation errors before they propagate into the main results.

Plot 1: Bunching Identification. Figure EC.1 displays the histogram of opioid dose (MME/day) in the neighborhood of the CDC 2016 threshold (90 MME), separately for Benchmark 1 (observed practice) and Benchmark 3 (CDC 2022). Under Benchmark 1, a pronounced spike is visible just at or below the 90 MME threshold, consistent with clinicians actively managing prescriptions to the guideline limit. This bunching pattern provides direct empirical support for the identifying assumption underlying our Benchmark 3 construction: a non-trivial share of patients in the data were receiving doses that were artificially constrained by the 2016 cap rather than determined by unconstrained clinical judgment. Under Benchmark 3, this spike is more pronounced and the distribution extends past 90 MME, reflecting the upward imputation that would resemble less constrained opioid dosing. The distributions coincide below approximately 85 MME, confirming that unconstrained patient-visits are left unchanged. The residual mass at exactly 90 MME in Benchmark 3 is expected: for bunched patients whose Tobit latent mean lies close to the 90-MME cap, the truncated draw will naturally return values near the lower truncation point.

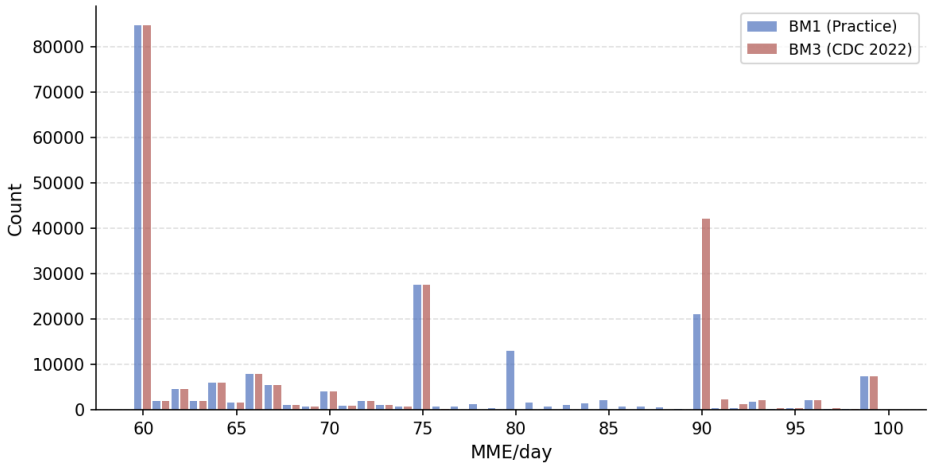


Figure EC.1 Dose distribution near the CDC 2016 threshold (90 MME)

Plot 2: Benchmark Comparison of Treatment Intensity. Figure EC.2 compares mean opioid dose (MME) across the three benchmarks, separately for the full patient population and for the bunched subgroup. We notice that the ordering BM2 $\leq$ BM1 $\leq$ BM3 in the mean opioid dose holds in both panels. For the full population, the differences are modest: Benchmark 2 reduces mean MME slightly relative to Benchmark 1, while Benchmark 3 essentially coincides with Benchmark 1. This is expected, given that the bunched subgroup represents a relatively small share of the overall population and the imputed dose increases are moderate in magnitude. For the bunched subgroup, the ordering is more pronounced and clinically meaningful: Benchmark 2 reduces mean MME substantially (the hard cap biting for patients who were previously prescribed at or above 90 MME), while Benchmark 3 restores and marginally exceeds the Benchmark 1 level. Furthermore, no violation of the monotonicity condition (BM3 $\geq$ BM1 for all patient-visits) was detected.

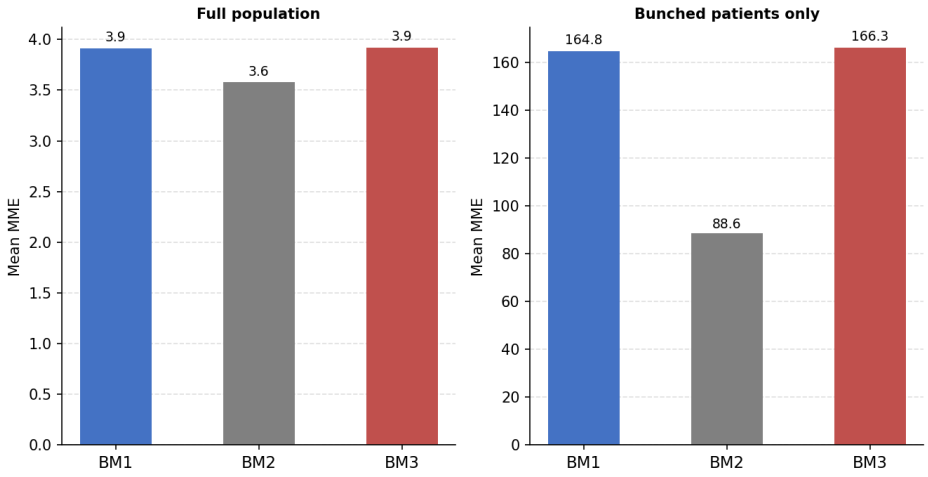


Figure EC.2 Mean opioid dose (MME) across benchmarks

Plot 3: Monotonicity Verification. Figure EC.3 displays the distribution of $\tilde{a}_{it}^{\text{mme}}(\text{BM3}) - a_{it}^{\text{mme}}(\text{BM1})$ for bunched patient-visits only. By construction of the truncated normal draw, every value must be weakly non-negative. The distribution is entirely supported on $[0, \infty)$, confirming that the monotonicity guarantee is satisfied without exception. The distribution is highly concentrated at zero, with a secondary mode around 10 MME. The spike at zero reflects the fact that for the majority of bunched patients, the Tobit latent mean lies close to or below the 90 MME cap, so the truncated draw returns a value near the lower truncation point. This pattern is statistically expected and does not indicate an implementation error; it does, however, imply that the practical magnitude of dose adjustment under Benchmark 3 is modest for most constrained patients.

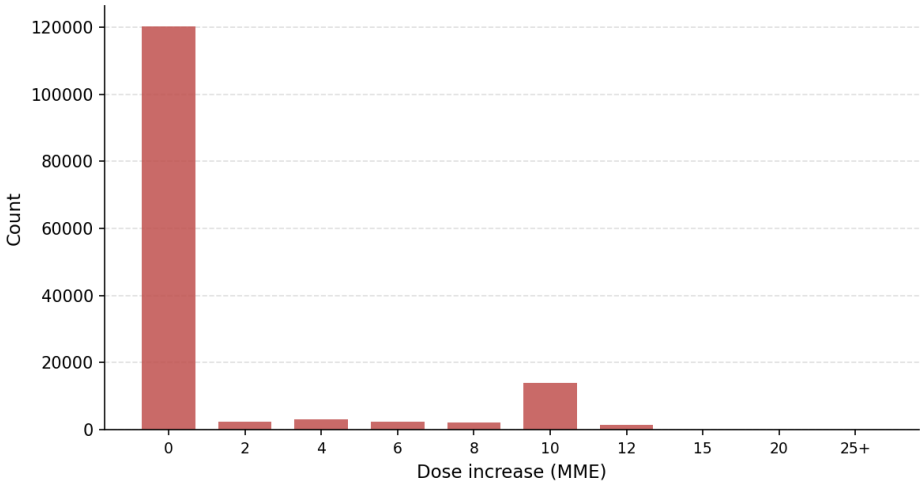


Figure EC.3 Distribution of BM3 − BM1 opioid dose for bunched patient-visits

Plot 4: Robustness Across Bandwidth Specifications. Figure EC.4 reports the share of constrained patient-visits and the mean Benchmark 3 dose for bunched patients, across the three bandwidth specifications $\delta \in \{0.05, 0.15, 0.25\}$. The bunching share increases monotonically with δ, as expected: a wider bandwidth classifies more patient-visits as constrained. The mean Benchmark 3 dose among bunched patients decreases modestly with δ. This pattern reflects a composition effect: as the bandwidth widens, newly included patient-visits correspond to patients further below the cap, whose Tobit latent means are lower on average, pulling down the group mean. Importantly, the qualitative cost-effectiveness findings are stable across all three specifications, supporting the robustness of our conclusions to the choice of δ (see Appendix C.5).

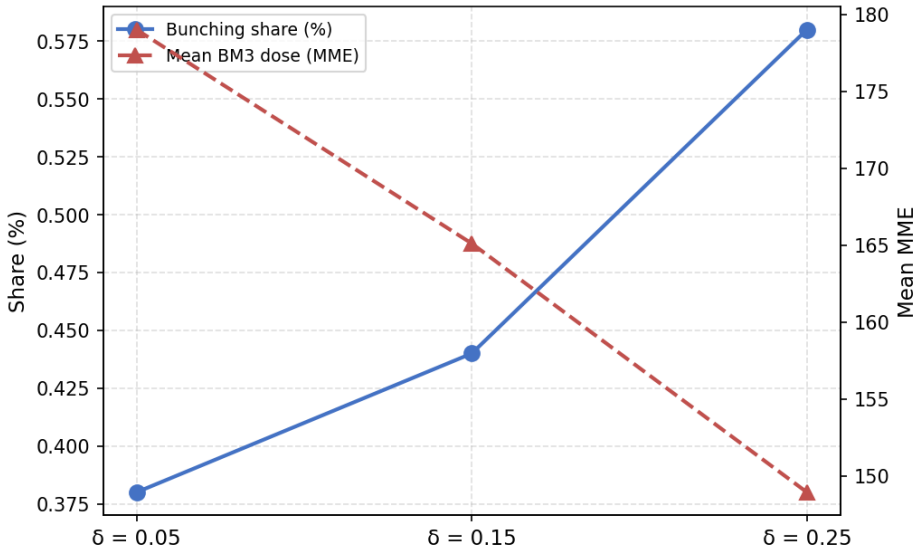


Figure EC.4 Bunching share and mean Benchmark 3 dose by bandwidth δ

Plot 5: Non-Opioid and Non-Pharmacologic Stability. Figure EC.5 verifies that non-opioid dose and non-pharmacologic treatment use rates are identical across all three benchmarks. By design, the benchmark constructions modify only opioid dose (MME) and days of supply; all other treatment modes are held fixed at their observed values. The plot confirms this: mean non-opioid dose and non-pharmacologic use rate are numerically identical across Benchmarks 1, 2, and 3. This rules out any inadvertent contamination of the non-opioid treatment modes during the benchmark construction process.

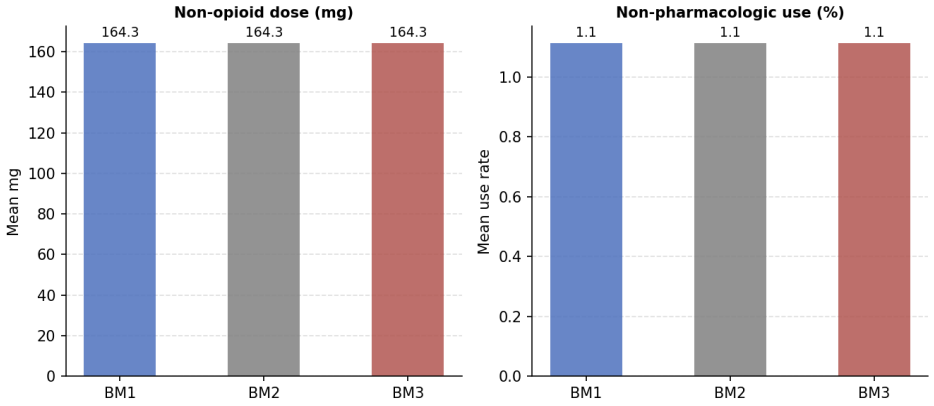
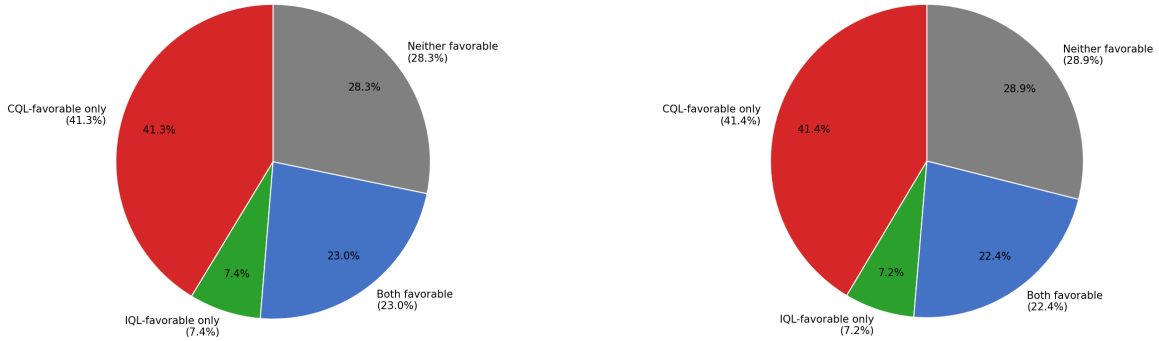


Figure EC.5 Non-opioid dose and non-pharmacologic use rate across benchmarks

B.4. Algorithmic performance across four subgroups

Our classification of four subgroups is robust to the choice of WTP threshold: an identical analysis at WTP = \$50K yields distributions within 0.2 percentage points of those reported under WTP = \$20K (in the main body), confirming that the patient-algorithm affinity may be an intrinsic property of the patient's clinical profile rather than an artifact of the threshold assumption.



(a) WTP = \$20K (b) WTP = \$50K

Figure EC.6 Patient classification by incremental NMB under CQL and IQL relative to observed practice

B.5. OUD and UTP Risk Trajectories

We examine how the policies translate into the two clinical risks: opioid use disorder (OUD) and undertreated pain (UTP). From the results in Figure EC.7, we make the following observation:

(i) Compared to the human-based practice, the CQL treatment policy results in a lower OUD risk but a comparable UTP risk. (ii) Under the observed practice of physicians, for very one person to ever experience OUD over the three-year therapeutic course, there are on average 2.25 persons to experience UTP over the same period. This number under our proposed CQL and IQL policies is 3.91 and 3.59, respectively.

Part (i) indicates that, under the CQL treatment policy, the risk of UTP is comparable to that observed from the practice. However, the risk of OUD resulted from the CQL policy is lower than that observed in the practice. To gain further insights into the risk trajectories, we also investigate the relative risks of experiencing UTP compared to OUD; i.e., the number of patients that would have to experience undertreated pain for one patient to experience opioid-related disorders at any time during the therapeutic course. This sheds light on a quantity known as "number needed to undertreat" (NNUT) introduced by the American Academy of Pain Medicine (Carr 2016). Part (ii) reveals that the relative risk under our policies is typically higher than that observed from the practice. Caution should be exercised when interpreting this observation, as it would appear to indicate that the human-based practice achieves a more favorable balance between the two events than either algorithmic policies. By definition, the relative risk is dominated by whichever risk is smaller. Holding UTP approximately fixed, the CQL policy that reduces the OUD rate would inflate the rate: the numerator changes little while the denominator falls. The higher rate under the algorithmic policies can therefore be a consequence of the algorithms' lower OUD rates rather than evidence of inferior pain management.

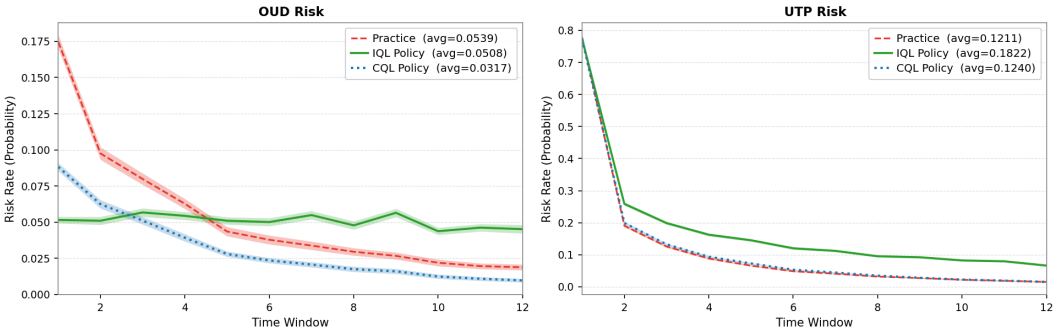


Figure EC.7 Likelihoods of incidence for OUD and UTP under different policies

$N = 2,738,118$ patients. Each time window = 3 months. Shades represent 95% confidence intervals. Under each policy, the average risk of each outcome is measured over the 3-year time horizon.

C. Robustness checks (supplement)

C.1. Quality-of-life scores, costs, and weights

Table EC.2 Average CE probability under various qol scores, costs, and weights scenarios

(a) qol scores

Algorithm: CQL

For Event 2 (q21)	For Event 1 (q11)						
	0.051		**0.093**		0.150		
	$20K	$50K	$20K	$50K	$20K	$50K	**Comparison**
(0.051,0.060)	64.4%	63.6%	64.4%	63.6%	64.4%	63.6%	vs. Practice
	60.9%	60.2%	60.9%	60.2%	60.9%	60.1%	vs. CDC 16
	64.4%	63.6%	64.4%	63.6%	64.4%	63.6%	vs. CDC 22
(0.100,0.108)	64.4%	63.6%	**64.4%**	**63.6%**	64.4%	63.6%	vs. Practice
	60.9%	60.2%	**60.9%**	**60.2%**	60.9%	60.2%	vs. CDC 16
	64.4%	63.7%	**64.4%**	**63.6%**	64.4%	63.6%	vs. CDC 22
(0.150,0.165)	64.4%	63.7%	64.4%	63.7%	64.4%	63.7%	vs. Practice
	60.9%	60.2%	60.9%	60.2%	60.9%	60.2%	vs. CDC 16
	64.4%	63.7%	64.4%	63.7%	64.4%	63.7%	vs. CDC 22

Algorithm: IQL

For Event 2 (q21)	For Event 1 (q11)						
	0.051		**0.093**		0.150		
	$20K	$50K	$20K	$50K	$20K	$50K	**Comparison**
(0.051,0.060)	29.9%	28.9%	29.9%	28.8%	29.9%	28.8%	vs. Practice
	28.4%	27.4%	28.4%	27.4%	28.4%	27.4%	vs. CDC 16
	30.0%	28.9%	29.9%	28.9%	29.9%	28.8%	vs. CDC 22
(0.100,0.108)	30.0%	28.9%	**30.0%**	**28.9%**	30.0%	28.9%	vs. Practice
	28.4%	27.4%	**28.4%**	**27.4%**	28.4%	27.4%	vs. CDC 16
	30.0%	28.9%	**30.0%**	**28.9%**	30.0%	28.9%	vs. CDC 22
(0.150,0.165)	30.0%	28.9%	30.0%	28.9%	30.0%	28.9%	vs. Practice
	28.5%	27.5%	28.5%	27.5%	28.4%	27.4%	vs. CDC 16
	30.0%	29.0%	30.0%	28.9%	30.0%	28.9%	vs. CDC 22

Notes. For qol scores, q11 and q21 represent the qol score for Events 1 and 2, respectively. Each scenario is evaluated under WTP ∈ {$20K, $50K}. The values in () represent the corresponding values for acute and chronic pain, respectively. Highlighted cells show results under baseline parameters.

(b) Costs

Algorithm: CQL

For Event 2 (C2)	For Event 1 (C1)						
	7,500		**15,588.07**		30,000		
	$20K	$50K	$20K	$50K	$20K	$50K	**Comparison**
(2000, 880)	65.2%	64.1%	67.3%	66.2%	68.7%	67.9%	vs. Practice
	61.6%	60.6%	63.4%	62.4%	64.6%	63.9%	vs. CDC 16
	65.2%	64.1%	67.3%	66.2%	68.7%	67.9%	vs. CDC 22
(6173, 2715)	62.3%	62.0%	**64.7%**	**64.2%**	66.7%	66.2%	vs. Practice
	59.1%	58.8%	**61.2%**	**60.7%**	62.9%	62.4%	vs. CDC 16
	62.3%	62.0%	**64.7%**	**64.2%**	66.7%	66.2%	vs. CDC 22
(20000, 8800)	58.4%	58.7%	60.9%	60.9%	63.3%	63.1%	vs. Practice
	55.8%	56.0%	58.0%	57.9%	60.0%	59.8%	vs. CDC 16
	58.4%	58.7%	60.9%	60.9%	63.3%	63.1%	vs. CDC 22

Algorithm: IQL

For Event 2 (C2)	For Event 1 (C1)						
	7,500		**15,588.07**		30,000		
	$20K	$50K	$20K	$50K	$20K	$50K	**Comparison**
(2000, 880)	31.3%	29.6%	35.1%	33.0%	38.2%	36.4%	vs. Practice
	29.7%	28.1%	33.5%	31.3%	36.5%	34.7%	vs. CDC 16
	31.3%	29.6%	35.2%	33.0%	38.2%	36.4%	vs. CDC 22
(6173, 2715)	26.9%	26.6%	**30.4%**	**29.6%**	34.0%	32.9%	vs. Practice
	25.6%	25.3%	**28.9%**	**28.1%**	32.3%	31.3%	vs. CDC 16
	26.9%	26.6%	**30.4%**	**29.6%**	34.0%	32.9%	vs. CDC 22
(20000, 8800)	22.1%	22.4%	25.0%	25.0%	28.2%	28.0%	vs. Practice
	21.2%	21.5%	23.9%	23.9%	26.8%	26.5%	vs. CDC 16
	22.1%	22.4%	25.0%	25.0%	28.2%	28.0%	vs. CDC 22

Notes. C1 and C2 represent the cost of Events 1 and 2, respectively. Each scenario is evaluated under WTP ∈ {$20K, $50K}. The values in () represent the corresponding values for acute and chronic pain, respectively. Highlighted cells show results under baseline parameters.

(c) Weights

Algorithm: CQL

(w1, w2)	WTP		Comparison
	$20K	**$50K**	
(1.00, 0.00)	70.7%	70.7%	vs. Practice
	66.4%	66.4%	vs. CDC 16
	70.7%	70.7%	vs. CDC 22
(0.75, 0.25)	68.0%	67.7%	vs. Practice
	64.0%	63.8%	vs. CDC 16
	68.0%	67.7%	vs. CDC 22
(0.50, 0.50)	**64.7%**	**64.2%**	vs. Practice
	61.2%	**60.7%**	vs. CDC 16
	64.7%	**64.2%**	vs. CDC 22
(0.25, 0.75)	60.7%	60.0%	vs. Practice
	57.7%	57.1%	vs. CDC 16
	60.7%	60.0%	vs. CDC 22
(0.00, 1.00)	47.5%	47.5%	vs. Practice
	46.2%	46.2%	vs. CDC 16
	47.5%	47.5%	vs. CDC 22

Algorithm: IQL

(w1, w2)	WTP		Comparison
	$20K	**$50K**	
(1.00, 0.00)	42.7%	42.7%	vs. Practice
	41.0%	41.0%	vs. CDC 16
	42.7%	42.7%	vs. CDC 22
(0.75, 0.25)	36.7%	36.0%	vs. Practice
	35.0%	34.3%	vs. CDC 16
	36.7%	36.0%	vs. CDC 22
(0.50, 0.50)	**30.4%**	**29.6%**	vs. Practice
	28.9%	**28.1%**	vs. CDC 16
	30.4%	**29.6%**	vs. CDC 22
(0.25, 0.75)	24.8%	24.1%	vs. Practice
	23.6%	23.0%	vs. CDC 16
	24.8%	24.1%	vs. CDC 22
(0.00, 1.00)	14.7%	14.7%	vs. Practice
	14.6%	14.6%	vs. CDC 16
	14.7%	14.7%	vs. CDC 22

Notes. w1 and w2 are the weights assigned to Events 1 and 2, respectively. Each scenario is evaluated under WTP ∈ {$20K, $50K}. Highlighted cells show results under baseline parameters.

C.2. Temporal validity: pre-COVID versus post-COVID

Table EC.3 Summary statistics

Variable	Pre-COVID $N = 1,068,788$ (39.03%) Mean (SD) / Num (%)	Post-COVID $N = 1,669,330$ (60.97%) Mean (SD) / Num (%)	p-value
Incidence of opioid-use disorder (Event 1) [a]	9,203 (0.9%)	6,396 (0.4%)	<0.001 *
Incidence of undertreated pain (Event 2) [a]	130,470 (12.2%)	183,400 (11.0%)	<0.001 *
Daily opioid strength (MME)	34.23 (40.22)	25.76 (25.26)	<0.001 *
Daily non-opioid strength (mg)	1281.09 (1013.97)	1176.37 (893.14)	<0.001 *
Days supply	14.97 (23.10)	13.01 (21.48)	<0.001 *
Use of nonpharm. treatments	0.03 (0.17)	0.02 (0.15)	<0.001 *
Age	47.47 (15.78)	45.27 (15.48)	<0.001 *
State annual opioid dispense rate (per 100 residents)	52.63 (14.71)	45.77 (12.27)	<0.001 *
Number of unique days a patient makes visits in a window	0.09 (0.46)	0.06 (0.36)	<0.001 *
Number of unique overlapping prescriptions in a window	3.69 (2.95)	3.85 (3.11)	<0.001 *
History of substance use (excluding opioid) [a]	273,483 (25.6%)	156,281 (9.4%)	<0.001 *
History of surgical procedures [a]	348,374 (32.6%)	417,172 (25.0%)	<0.001 *
Gender: Male	439,109 (41.1%)	715,454 (42.9%)	<0.001 *
Female	629,679 (58.9%)	953,876 (57.1%)	
Pain pathology: Surgery	288,511 (27.0%)	402,261 (24.1%)	<0.001 *
Acute	329,084 (30.8%)	459,317 (27.5%)	
Chronic	451,193 (42.2%)	807,752 (48.4%)	
Comorbidity (Elixhauser): <0	630,771 (59.0%)	1,105,883 (66.2%)	<0.001 *
1–4	208,693 (19.5%)	286,879 (17.2%)	
5+	229,324 (21.5%)	276,568 (16.6%)	
Insurance: Commercial	960,455 (89.9%)	1,560,974 (93.5%)	<0.001 *
Medicare	108,333 (10.1%)	108,356 (6.5%)	
Region: Northeast	119,797 (11.2%)	191,295 (11.5%)	<0.001 *
North Central	207,292 (19.4%)	376,694 (22.6%)	
South	587,303 (55.0%)	845,140 (50.6%)	
West	153,059 (14.3%)	253,188 (15.2%)	
Population size: Metro ≥1 million	456,478 (42.7%)	704,891 (42.2%)	<0.001 *
Metro 0.25-1 million	184,560 (17.3%)	262,228 (15.7%)	
Metro < 250K	65,683 (6.1%)	89,436 (5.4%)	
Non-metro <50K	132,029 (12.4%)	209,979 (12.6%)	
Unknown	230,038 (21.5%)	402,796 (24.1%)	

Pre- and post-COVID are defined based on whether the index opioid prescription was before or after March 2020, respectively (first wave of COVID-19 in the US).
Continuous variables: mean (SD). Categorical variables: number (%). Kruskal-Wallis and Chi-square tests for continuous and categorical variables, respectively.
[a] For this variable, the value reported is the number of patients who ever experienced an event or had a history of a medical condition.
* Estimates are statistically significant with p <0.001.

Table EC.4 Average CE probability

	WTP = $20K			WTP = $50K			
Algorithm	**Pre-COVID**	**Post-COVID**	**Whole sample**	**Pre-COVID**	**Post-COVID**	**Whole sample**	**Comparison**
CQL	68.4%	62.4%	64.4%	67.8%	61.9%	63.6%	vs. Practice
	63.4%	59.7%	60.9%	62.9%	59.3%	60.2%	vs. CDC 2016
	68.4%	62.4%	64.4%	67.9%	61.9%	63.6%	vs. CDC 2022
IQL	34.2%	28.0%	30.0%	33.4%	27.2%	28.9%	vs. Practice
	31.9%	27.0%	28.4%	31.1%	26.2%	27.4%	vs. CDC 2016
	34.2%	28.0%	30.0%	33.4%	27.2%	28.9%	vs. CDC 2022

C.3. Population generalizability: Medicaid population

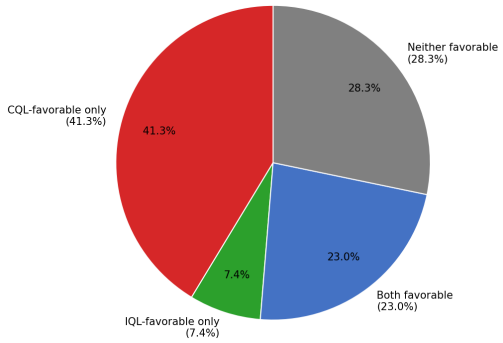
Table EC.5 Summary statistics

Variable	Comm + Mdcr $N = 2,738,118$ Mean (SD) / Num (%)	Medicaid $N = 991,529$ Mean (SD) / Num (%)	p-value
Incidence of opioid-use disorder (Event 1) [a]	15,599 (0.6%)	59,268 (6.0%)	<0.001 *
Incidence of undertreated pain (Event 2) [a]	313,870 (11.5%)	97,152 (9.8%)	<0.001 *
Daily opioid strength (MME)	29.07 (32.21)	42.17 (162.88)	<0.001 *
Daily non-opioid strength (mg)	1217.25 (943.54)	1594.20 (1201.36)	<0.001 *
Days supply	13.77 (22.15)	17.79 (23.36)	<0.001 *
Use of nonpharm. treatments	0.02 (0.16)	0.03 (0.16)	<0.001 *
Age	46.13 (15.64)	35.99 (14.56)	<0.001 *
State annual opioid dispense rate (per 100 residents)	50.21 (14.28)	—	N/A
Number of unique days a patient makes visits in a window	0.07 (0.40)	0.43 (2.37)	<0.001 *
Number of unique overlapping prescriptions in a window	3.79 (3.05)	4.66 (4.77)	0.537
History of substance use (excluding opioid) [a]	279,162 (10.2%)	461,408 (46.5%)	<0.001 *
History of surgical procedures [a]	765,546 (28.0%)	711,386 (71.7%)	<0.001 *
Gender: Male	1,154,563 (42.2%)	321,791 (32.5%)	<0.001 *
Female	1,583,555 (57.8%)	669,738 (67.5%)	
Pain pathology: Surgery	690,772 (25.2%)	655,932 (66.2%)	<0.001 *
Acute	788,401 (28.8%)	302,138 (30.5%)	
Chronic	1,258,945 (45.9%)	33,459 (3.4%)	
Comorbidity (Elixhauser): <0	1,736,654 (63.4%)	190,262 (19.2%)	<0.001 *
1–4	495,572 (18.1%)	355,470 (35.9%)	
5+	505,892 (18.5%)	445,797 (45.0%)	
Insurance: Commercial	2,521,429 (92.1%)	0 (0.0%)	N/A
Medicare	216,689 (7.9%)	0 (0.0%)	
Medicaid, dual eligible: No	0 (0.0%)	927,655 (93.6%)	
Medicaid, dual eligible: Yes	0 (0.0%)	63,874 (6.4%)	
Region: Northeast	311,092 (11.4%)	—	N/A
North Central	583,986 (21.3%)	—	
South	1,436,793 (52.5%)	—	
West	406,247 (14.8%)	—	
Population size: Metro ≥1 million	1,161,369 (42.4%)	—	N/A
Metro 0.25-1 million	446,788 (16.3%)	—	
Metro < 250K	155,119 (5.7%)	—	
Non-metro <50K	342,008 (12.5%)	—	
Unknown	632,834 (23.1%)	—	

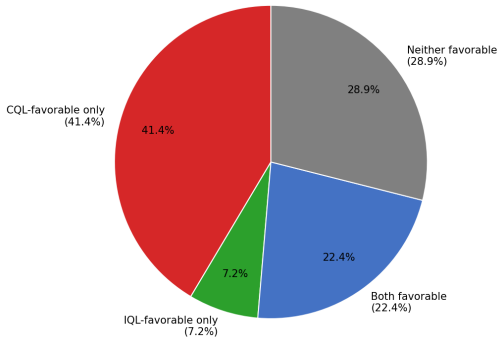
Our Medicaid population is from a multi-state data, with no geographical-related information (e.g., State annual opioid dispense rate, region, and population size).
Continuous variables: mean (SD). Categorical variables: number (%). Kruskal-Wallis and Chi-square tests for continuous and categorical variables, respectively.
[a] For this variable, the value reported is the number of patients who ever experienced an event or had a history of a medical condition.
* Estimates are statistically significant with $p < 0.001$.

Table EC.6 Average CE probability

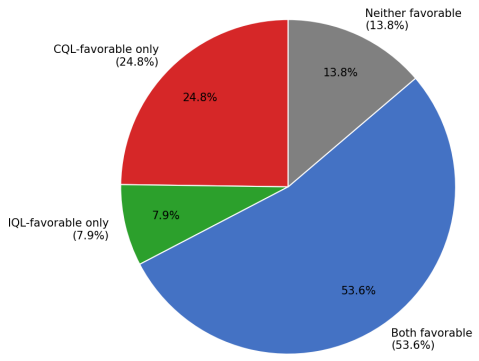
Algorithm	WTP = $20K		WTP = $50K		Comparison
	Sample: Medicaid	**Sample: baseline (Commercial + Medicare)**	**Sample: Medicaid**	**Sample: baseline (Commercial + Medicare)**	
CQL	78.4%	64.4%	78.1%	63.6%	vs. Practice
	75.3%	60.9%	75.0%	60.2%	vs. CDC 2016
	78.2%	64.4%	77.9%	63.6%	vs. CDC 2022
IQL	61.5%	30.0%	61.1%	28.9%	vs. Practice
	58.6%	28.4%	58.3%	27.4%	vs. CDC 2016
	61.4%	30.0%	61.0%	28.9%	vs. CDC 2022



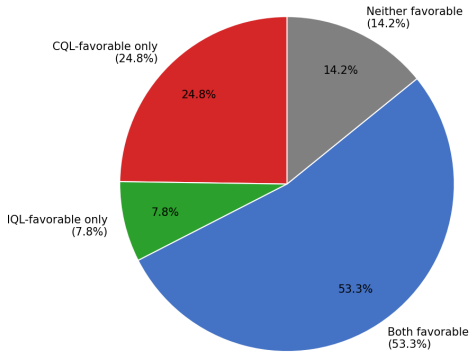
(a) Data (Comm + Mdcr), WTP = $20K



(b) Data (Comm + Mdcr), WTP = $50K



(c) Data (Medicaid), WTP = $20K



(d) Data (Medicaid), WTP = $50K

Figure EC.8 Patient classification by incremental NMB under CQL and IQL relative to observed practice

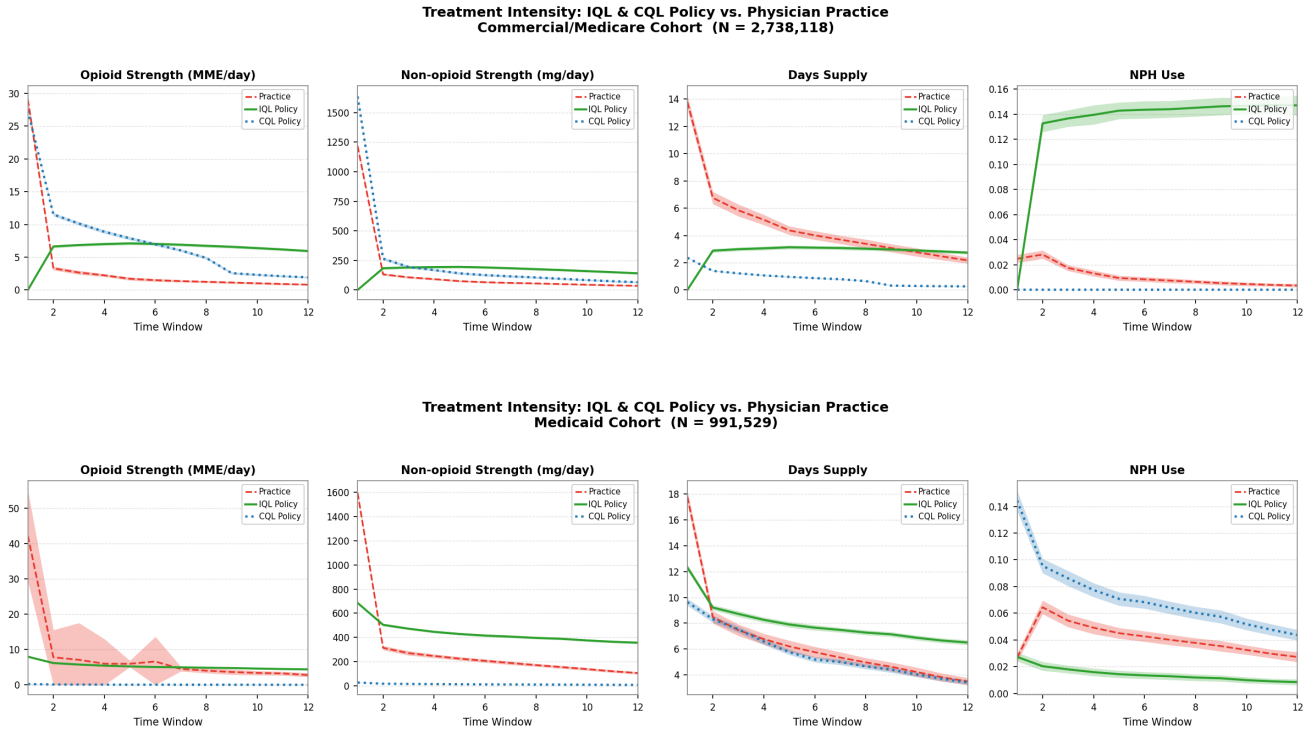


Figure EC.9 Multi-modal pain treatments under baseline (Commercial + Medicare) vs. Medicaid populations Results are obtained for WTP = \$20K. x-axes represent time window (each window = 3 months). Shades represent 95% confidence intervals. Supply is zero when both opioid/non-opioid strengths are zero (and vice versa). non-pharmacologic treatments: 1 (0) indicates use (no use).

C.4. Data preprocessing validity: SMOTE and Event 1 incidence rate

Table EC.7 Summary statistics

Variable	Original $N = 2,738,118$ Mean (SD) / Num (%)	SMOTE $N = 3,093,771$ Mean (SD) / Num (%)	p-value
Incidence of opioid-use disorder (Event 1) [a]	15,599 (0.6%)	371,252 (12.0%)	<0.001 *
Incidence of undertreated pain (Event 2) [a]	313,870 (11.5%)	330,579 (10.7%)	<0.001 *
Daily opioid strength (MME)	29.07 (32.21)	27.38 (31.62)	<0.001 *
Daily non-opioid strength (mg)	1217.25 (943.54)	1113.11 (943.67)	<0.001 *
Days supply	13.77 (22.15)	14.26 (21.94)	<0.001 *
Use of nonpharm. treatments	0.02 (0.16)	0.03 (0.15)	<0.001 *
Age	46.13 (15.64)	46.76 (15.75)	<0.001 *
State annual opioid dispense rate (per 100 residents)	50.21 (14.28)	25.01 (26.52)	<0.001 *
Number of unique days a patient makes visits in a window	0.07 (0.40)	0.09 (0.40)	<0.001 *
Number of unique overlapping prescriptions in a window	3.79 (3.05)	3.78 (3.04)	0.537
History of substance use (excluding opioid) [a]	279,162 (10.2%)	536,326 (17.3%)	<0.001 *
History of surgical procedures [a]	765,546 (28.0%)	1,034,410 (33.4%)	<0.001 *
Gender: Male	1,154,563 (42.2%)	1,311,906 (42.4%)	<0.001 *
Female	1,583,555 (57.8%)	1,781,865 (57.6%)	
Pain pathology: Surgery	690,772 (25.2%)	761,561 (24.6%)	<0.001 *
Acute	788,401 (28.8%)	913,427 (29.5%)	
Chronic	1,258,945 (45.9%)	1,418,783 (45.9%)	
Comorbidity (Elixhauser): <0	1,736,654 (63.4%)	1,822,973 (58.9%)	<0.001 *
1–4	495,572 (18.1%)	613,697 (19.8%)	
5+	505,892 (18.5%)	657,101 (21.2%)	
Insurance: Commercial	2,521,429 (92.1%)	2,793,164 (90.3%)	<0.001 *
Medicare	216,689 (7.9%)	300,607 (9.7%)	
Region: Northeast	311,092 (11.4%)	452,700 (14.6%)	<0.001 *
North Central	583,986 (21.3%)	688,286 (22.2%)	
South	1,436,793 (52.5%)	1,532,383 (49.5%)	
West	406,247 (14.8%)	420,402 (13.6%)	
Population size: Metro ≥1 million	1,161,369 (42.4%)	1,305,075 (42.2%)	<0.001 *
Metro 0.25-1 million	446,788 (16.3%)	486,455 (15.7%)	
Metro < 250K	155,119 (5.7%)	169,752 (5.5%)	
Non-metro <50K	342,008 (12.5%)	351,655 (11.4%)	
Unknown	632,834 (23.1%)	780,834 (25.2%)	

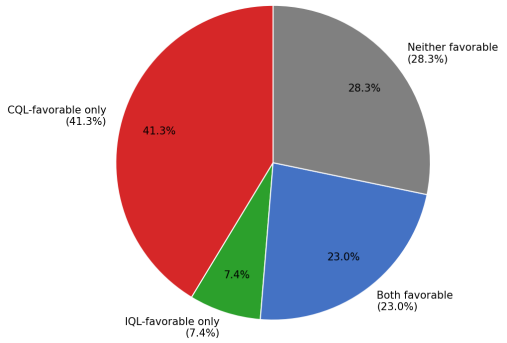
Continuous variables: mean (SD). Categorical variables: number (%). Kruskal-Wallis and Chi-square tests for continuous and categorical variables, respectively.

[a] For this variable, the value reported is the number of patients who ever experienced an event or had a history of a medical condition.

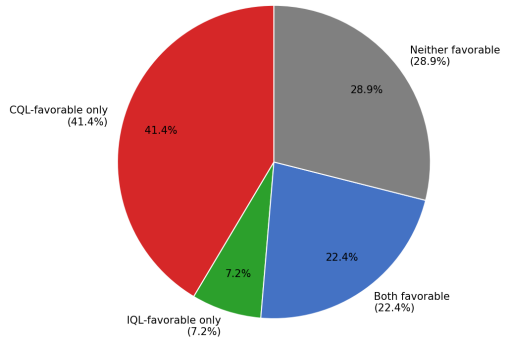
* Estimates are statistically significant with $p < 0.001$.

Table EC.8 Average CE probability

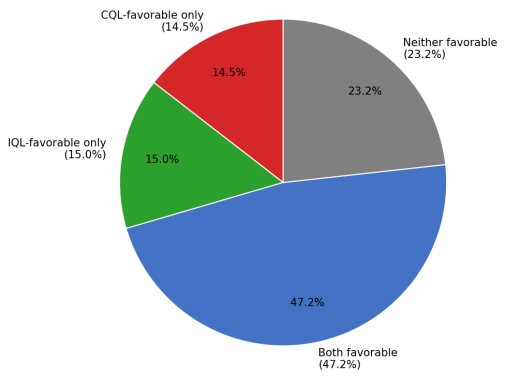
Algorithm	WTP = $20K		WTP = $50K		Comparison
	Sample: SMOTE + SMOTE Transformer	Sample: baseline	Sample: SMOTE + SMOTE Transformer	Sample: baseline	
CQL	61.7%	64.4%	61.4%	63.6%	vs. Practice
	58.5%	60.9%	58.3%	60.2%	vs. CDC 2016
	61.7%	64.4%	61.4%	63.6%	vs. CDC 2022
IQL	62.2%	30.0%	62.1%	28.9%	vs. Practice
	60.5%	28.4%	60.3%	27.4%	vs. CDC 2016
	62.3%	30.0%	62.1%	28.9%	vs. CDC 2022



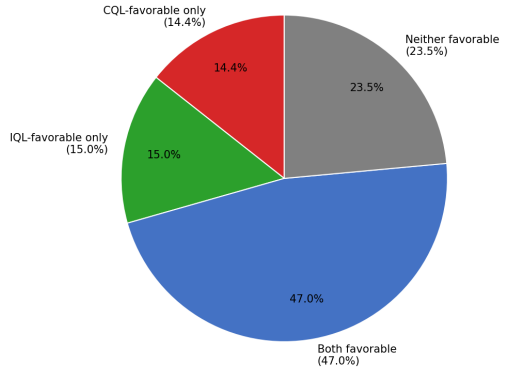
(a) Original data, WTP = $20K



(b) Original data, WTP = $50K



(c) Synthetic (SMOTE) data, WTP = $20K



(d) Synthetic (SMOTE) data, WTP = $50K

Figure EC.10 Patient classification by incremental NMB under CQL and IQL relative to observed practice

C.5. CDC 2022 (counterfactual unconstrained practice)

Table EC.9 Average CE probability

	WTP = $20K			
Algorithm	**δ = 0.05**	**δ = 0.15**	**δ = 0.25**	**Comparison**
CQL	63.9%	64.4%	64.0%	vs. CDC 2022
IQL	28.9%	30.0%	29.0%	vs. CDC 2022

Notes. δ is the bunching bandwidth for simulating the counterfactual for the CDC 2022 guidelines.

References

Ataoğlu E, Tiftik T, Kara M, Tunc H, Ersöz M, and others (2013) Effects of chronic pain on quality of life and depression in patients with spinal cord injury. *Spinal Cord.* 51(1):23–26.

Carr DB (2016) NPS versus CDC: Scylla, Charybdis and the "number needed to [under-] treat". *Pain Medicine.* 17(6):999–1000.

De Maeyer J, Vanderplasschen W, Broekaert E (2010) Quality of life among opiate-dependent individuals: A review of the literature. *International Journal of Drug Policy.* 21(5):364–380.

Domingo-Salvany A, Brugal MT, Barrio G, González-Saiz F, Bravo MJ, and others (2010) Gender differences in health related quality of life of young heroin users. *Health and Quality of Life Outcomes.* 8(1):1–10.

Duan N, Manning WG, Morris CN, Newhouse JP (1983) A comparison of alternative models for the demand for medical care. *Journal of Business & Economic Statistics.* 1(2):115–126.

Giacomuzzi SM, Riemer Y, Ertl M, Kemmler G, Rössler H, and others (2005) Gender differences in health-related quality of life on admission to a maintenance treatment program. *European Addiction Research.* 11(2):69–75.

Hadi MA, McHugh GA, Closs SJ (2016) Impact of chronic pain on patients' quality of life: a comparative mixed-methods study. *Journal of Patient Experience.* 6(2):133–141.

Katz N (2002) The impact of pain management on quality of life. *Journal of Pain and Symptom Management.* 24(1):S38–S47.

Taylor RS, Ullrich K, Regan S, Broussard C, Schwenkglenks M, and others (2013) The impact of early postoperative pain on health-related quality of life. *Pain Practice.* 13(7):515–523.

Tobin J (1958) Estimation of relationships for limited dependent variables. *Econometrica.* 26(1):24–36.

Tüzün E (2007) Quality of life in chronic musculoskeletal pain. *Best Practice & Research Clinical Rheumatology.* 21(3):567–579.

Wu CL, Naqibuddin M, Rowlingson AJ, Lietman SA, Jermyn RM, and others (2003) The effect of pain on health-related quality of life in the immediate postoperative period. *Anesthesia & Analgesia.* 97(4):1078–1085.