INSEAD The Business School for the World®

Working Paper

2021/14/TOM

(Revised version of 2020/30/TOM)

# Curbing the Opioid Epidemic at its Root: The Effect of Provider Discordance after Opioid Initiation

Katherine Bobroske
University of Cambridge, kab80@cam.ac.uk

Michael Freeman
INSEAD, michael.freeman@insead.edu

Lawrence Huan
University of Cambridge, Lawrence.Huan.MD@gmail.com

Anita Cattrell
Evolent Health, acattrell@evolenthealth.com

Stefan Scholtes
University of Cambridge, s.scholtes@jbs.cam.ac.uk

While medical research has addressed the clinical management of chronic opioid users, little is known about how operational interventions shortly after opioid initiation can impact a patient's likelihood of long-term opioid use. Using a nationwide US database of medical and pharmaceutical claims, we investigate the care delivery process at the most common entry point to opioid use: the primary care setting. For patients who return to primary care for a follow-up appointment within 30 days of opioid initiation, we ask: who should revisit (and potentially revise) the opioid-based treatment plan, the initial prescriber (provider concordance) or an alternate clinician (provider discordance)? First, using a fully controlled logistic model, we find that provider discordance reduces the likelihood of long-term opioid use 12 months after opioid initiation by 31% (95% CI: [18%, 43%]). Both an instrumental variable analysis and propensity-score matching (utilizing the minimum-bias estimator approach) that account for omitted variable bias indicate this is a conservative estimate of the true causal effect. Second, looking at patient activities immediately after the follow-up appointment, we find that this long-term reduction is at least partially explained by an immediate reduction in opioids prescribed after the follow-up appointment. Third, the data suggest that the benefit associated with provider discordance remains significant regardless of whether the patient's initial prescriber was their regular provider (versus another clinician). Overall, our analysis indicates that systematic, operational changes in the early stages of managing new opioid patients may offer a promising, and hitherto overlooked, opportunity to curb the opioid epidemic.

Keywords: Opioid Crisis; Healthcare Operations; Primary Care; Provider Discordance; Econometrics

# Curbing the Opioid Epidemic at its Root:
# The Effect of Provider Discordance after Opioid Initiation

**ABSTRACT.** While medical research has addressed the clinical management of chronic opioid users, little is known about how operational interventions shortly after opioid initiation can impact a patient's likelihood of long-term opioid use. Using a nationwide US database of medical and pharmaceutical claims, we investigate the care delivery process at the most common entry point to opioid use: the primary care setting. For patients who return to primary care for a follow-up appointment within 30 days of opioid initiation, we ask: who should revisit (and potentially revise) the opioid-based treatment plan, the initial prescriber (provider concordance) or an alternate clinician (provider discordance)? First, using a fully controlled logistic model, we find that provider discordance reduces the likelihood of long-term opioid use 12 months after opioid initiation by 31% (95% CI: [18%, 43%]). Both an instrumental variable analysis and propensity-score matching (utilizing the minimum-bias estimator approach) that account for omitted variable bias indicate this is a conservative estimate of the true causal effect. Second, looking at patient activities immediately after the follow-up appointment, we find that this long-term reduction is at least partially explained by an immediate reduction in opioids prescribed after the follow-up appointment. Third, the data suggest that the benefit associated with provider discordance remains significant regardless of whether the patient's initial prescriber was their regular provider (versus another clinician). Overall, our analysis indicates that systematic, operational changes in the early stages of managing new opioid patients may offer a promising, and hitherto overlooked, opportunity to curb the opioid epidemic.

## 1. INTRODUCTION

Over the past 30 years, the opioid epidemic has escalated to claim more than 130 lives each day in the United States. Opioid overdoses resulted in 47,600 deaths in 2017 alone, a 12% increase from 2016 (Scholl et al. 2019), rendering the epidemic one of the driving contributors to the nation's recent decrease in life expectancy (Dowell et al. 2017). To curb inappropriate opioid usage, the US has introduced interventions including prescription drug monitoring programs, increasing the availability of overdose-preventing drugs, and funding research into abuse-deterrent or tamper-resistant opioid formulations (Grosser et al. 2017, Pitt et al. 2018). However, as demonstrated by the surge in opioid-related overdoses and deaths during the recent Covid-19 crisis (American Medical Association 2020, Holland et al. 2021), the opioid epidemic is far from over.

One limitation of many current opioid-reducing initiatives is their focus on high-risk chronic opioid users – patients filling high dosage opioid prescriptions or heroin users. These high-risk patients can be difficult to treat, as they are often dependent on prescribed opioids or else obtain opioids illicitly and no longer rely on the healthcare system to provide opioid access (Pitt et al. 2018, Scholl et al. 2019). Moreover, while these patients are at the highest immediate risk of overdose, they are, in a sense, only the tip of the iceberg. With the majority of heroin users reporting that their journey to dependence began with a legal opioid prescription (Compton et al. 2016), more research is needed to identify early interventions that could effectively disrupt the pathway from first prescription to opioid dependence.

We address this gap in the literature by empirically examining the process of managing new opioid initiates in the primary care setting. Primary care clinicians are the largest group of opioid prescribers and often serve as the first encounter for patients with noncancer pain (Grosser et al. 2017, Levy et al. 2015). While some patients do not return to the primary care setting after their initial opioid prescription (either because their pain has subsided or because they receive subsequent care from a different specialty), other patients return for further diagnosis or to assess the progress of treatment. This follow-up appointment offers a chance to revisit and potentially revise the initial treatment plan based on health progression and any other new information.

For those patients who continue seeking treatment in the primary care setting after opioid initiation, we therefore ask: who should revisit the treatment plan with the patient, the original prescriber (i.e., provider concordance) or another clinician (i.e., provider discordance)?[1] While a different clinician can expose the patient to a "fresh perspective" and prevent anchoring to the original opioid treatment course, it may also lead to more fragmented care and reduce physician "ownership" of the long-term health of the patient (Ahuja et al. 2020, Ariely et al. 2003, Senot 2019). Although prescriber continuity is typically recommended for patients already dependent on opioids (Hallvik et al. 2018, Jena et al. 2014), the overall impact of exposing a patient to variation in providers in the initial stages of opioid use is not immediately clear.

We examine the care delivery process for new opioid initiates in the primary care setting by leveraging a nationwide claims database of more than 3.5 million patients. First, we find empirical evidence that incorporating provider discordance early in the care management process may reduce long-term opioid use rates by at least 31% (95% confidence interval [CI]: 18%, 43%). After identifying the main effect, we analyze a potential mechanism for the relationship: whether the patient fills a subsequent opioid prescription after the follow-up appointment. We also examine different discordant care pathways to determine whether the effect size differs depending on whether the patient was prescribed opioids by their regular (i.e., most frequently visited) primary care provider or by another clinician.

While continued opioid use is clinically justified and effectively reduces pain for some patients, the risks of inappropriate or long-term opioid use can be severe (Dowell et al. 2017, Glod 2017, Pitt et al. 2018). Overall, this research suggests that investigation into systematic changes following the initiation of opioids could be a promising and hitherto overlooked opportunity to reduce the influx of patients afflicted by the opioid epidemic.

## 2. THEORY AND HYPOTHESIS DEVELOPMENT

### 2.1 Related Literature

In the early stages of opioid use, it is not always clear whether the pain relief offered by opioid treatment will outweigh the possible risks of clinical harm and dependence (Deyo et al. 2017, Dowell et al. 2016). This uncertainty can lead to variation in care, such that a patient may receive a different treatment based on non-clinical factors, including the medical knowledge of the diagnosing clinician and location of care (Green 2012). This phenomenon has been acutely observed in the opioid context, where the number of opioid pills consumed

[1] Note we do not make any evaluation on the clinical efficacy of the opioid regimen. Instead, we explore a change to the management process (i.e. provider discordance or concordance after opioid initiation) that does not directly impact clinical judgement but may materially impact the patient's likelihood of long-term opioid use.

by patients after surgery varied markedly depending on, e.g., the default opioid pill count presented in the electronic medical record system and the quantity contained in a first prescription (Chiu et al. 2018, Farley et al. 2019).

Traditionally, the healthcare operations community has helped address care variation by developing decision support tools that, e.g., prioritize hospital discharge decisions (Chan et al. 2012), optimize cancer screening schedules (Ayer et al. 2012), inform hospital capacity decisions (Song et al. 2019), and balance fairness with efficiency when allocating organ transplant resources (Dickerson et al. 2019). In their respective clinical contexts, these tools effectively minimize unwarranted care variation and improve overall patient outcomes. However, in the context of opioids where clinicians lack an objective scale to measure the patient's level of pain (Morone and Weiner 2013), algorithms have limited utility in reducing variation of care.

Meanwhile, researchers have recently started to explore the potential *benefits* of variation when introduced within the care pathway of an individual patient (Ahuja et al. 2020, Chan et al. 2019, Kuntz et al. 2019). Especially for medical conditions where multiple viable treatment options exist, patients may benefit from exposure to a variety of providers and treatment options as they search for the best solution to their problems (Christensen et al. 2009). For pain management in particular, pharmacological therapy such as opioids represents only one of many potential treatment modalities; alternatives include physical rehabilitation, psychological therapy, and interventional treatments such as injections and surgical procedures (Owen et al. 2018).

In the early stages of pain management, the question of which provider is best positioned to revisit the treatment plan is akin to the notion of gatekeeping. Past research has investigated the initial provider, who decides whether to manage the patient's treatment alone or to refer the patient to a specialist (Freeman et al. 2017, 2020, Shumsky and Pinker 2003). In these settings, the additional cost of treatment by a specialist is weighed against the cost of failing to solve the patient's problem without specialist input. Our context also involves the patient potentially seeing two different clinicians. However, while the gatekeeping notion is predicated on referral to a specialist, we focus on alternative providers within the same care level (i.e., primary care). Furthermore, gatekeeping is an active, endogenous escalation decision by the initial clinician; in contrast, we study an exogenous process change (e.g., to patient scheduling) that is not necessarily taken by the initial prescriber.

This concept of routing during the early stages of opioid use is also related to studies on continuity of care. Such work typically focuses on the long-term management of patients who are already suffering from specific chronic conditions, e.g., heart failure (Senot 2019), diabetes (Ahuja et al. 2020), or opioid dependence (Hallvik et al. 2018). While most studies in these chronic settings find that patients benefit from repeated appointments with the same practitioner, we focus on patients *prior* to chronicity. Additionally, we highlight the importance of early-stage management as part of a strategy to *prevent* risk escalation in the first place. As outlined below, we hypothesize that the process of care delivery (namely provider concordance or discordance) during these critical initial stages may have a significant bearing on later outcomes.

### 2.2 The Effect of Provider Discordance

In medical settings, clinician discordance is generally recognized to confer two main advantages. First, combining the knowledge pools of two clinicians enlarges the information base used in decision-making (Ahuja et al. 2020,

Brooks et al. 2015). The new clinician may know of alternative treatments or could elicit new information from the patient that proves important in weighing the trade-offs between potential treatments. This new perspective may be particularly important for the early stages of opioid use given the array of clinically viable treatment options available for pain management (Owen et al. 2018).

Second, the sequential, independent decision-making process involved when seeing two different clinicians can help counteract cognitive biases such as confirmation bias and anchoring (Ariely et al. 2003, Gino and Pisano 2008). Following the initial decision to prescribe an opioid, a practitioner's desire to remain coherent over time may bias them against stopping the prescription. For example, Staats et al. (2018) show that providers with more experience may get caught in the status quo, rendering them less able to objectively update their beliefs in light of new information. A change in provider may thus increase the odds of a patient being transitioned to a new (i.e., non-opioid) course of treatment.

Despite potential benefits, clinician discordance is typically discouraged in medicine as it increases fragmentation of care. Receiving care from multiple clinicians can result in lapses in communication and coordination, exacerbated when providers do not share clinical data (Clark et al. 2013, Senot 2019). Both tacit and explicit information can be lost, reducing providers' opportunities to learn (e.g., by augmenting existing data with new information) and to adjust treatment over time (Reagans et al. 2005).

With respect to opioid-management specifically, the medical literature has examined the role of provider discordance in the *post-acute* phase of opioid-use. Since the past prescription is a known risk factor for long-term use (Deyo et al. 2017), a provider's awareness of the patient's opioid history may reduce their likelihood of offering additional opioids. Studies have also identified that involving fewer distinct clinicians in chronic opioid management is associated with a reduced number of high-risk prescriptions, opioid-related hospitalizations, and overdose (Hallvik et al. 2018, Jena et al. 2014).

Generally, in this *post-acute* phase of opioid use, it appears important for patients to avoid the potential disadvantages of care fragmentation. However, to our knowledge, the value of seeing multiple primary care clinicians during the *initial* phase of opioid use has not been evaluated. We believe that for these new opioid users, the advantages of clinician discordance (an enlarged information base and correction to confirmation bias and anchoring) are likely to outweigh the disadvantages. This is captured in Hypothesis 1:

***Hypothesis 1.*** *After an initial opioid prescription, if the patient's follow-up appointment is with an alternate clinician compared to their initial opioid prescriber (provider discordance), then the patient's likelihood of becoming a long-term opioid user is reduced.*

### 2.3 A Potential Mechanism

If a follow-up appointment with an alternate clinician reduces the likelihood of becoming a long-term opioid user (Hypothesis 1), then one would also expect a change in the patient's short-term opioid usage. A natural candidate for a mechanism in the relationship between provider discordance and long-term opioid use is the patient's opioid use shortly after the follow-up visit.

If this relationship exists, we expect to observe that patients who see a different doctor for their follow-up appointment will be less likely to fill an opioid prescription soon after the follow-up visit, following the arguments of Section 2.2. We also expect to observe an association between filling an opioid prescription soon after the follow-up appointment and the rate of long-term opioid use. These associations are tested in Hypotheses 2a-b:

***Hypothesis 2a.*** *Patients who experience provider discordance are less likely to fill an opioid prescription immediately after the follow-up appointment.*

***Hypothesis 2b.*** *Patients who fill an opioid prescription immediately after their follow-up appointment are more likely to become long-term opioid users.*

Note that our intention is not to establish a causal connection through these hypotheses. Instead, we aim to provide further evidence of the effect of provider discordance on long-term opioid use (Hypothesis 1) by investigating an observable change in patient activity shortly after the follow-up appointment.

**2.4 Provider Discordance Pathways**

Many patients have a regular primary care provider who the patient sees most frequently for their health needs (Atlas et al. 2009). Through repeated interactions, the patient's regular provider typically will have acquired knowledge of the patient's comorbidities, medical, and family history. Likewise, a patient is more likely to have familiarity and trust in their regular provider, creating an environment where they may be more willing to share personal information (Siemsen et al. 2009).

As a regular provider is more likely to take personal responsibility or ownership for the patient's long-term health (Jena et al. 2014, Senot 2019), we anticipate that they are well-positioned to balance short-term pain against long-term risks of continued opioid use. Therefore, we would expect a regular provider to be more likely than an alternate clinician to encourage conservative treatment plans. Likewise, greater familiarity and trust may increase a patients' willingness to explore alternative non-opioid treatment options proposed by a regular provider (Ahuja et al. 2020b).

Not only could seeing a regular provider have a net impact on opioid use, but patients who switch to and from their regular clinician in the early stages of opioid use may not benefit equally from provider discordance. For instance, we might expect that patients who were initially prescribed opioids by their regular provider will receive little benefit from discordance if the follow-up provider is less willing to adjust the regular provider's care plan. By comparison, a patient who initiates opioids with an alternate provider (perhaps due to acute onset of pain and unavailability of their regular provider) then switches to their regular provider for follow-up may benefit greatly from the fresh perspective and familiarity of their regular provider. This motivates Hypothesis 3:

***Hypothesis 3.*** *The relative reduction of long-term opioid use rates from provider discordance is stronger when patients change from an alternate clinician to their regular provider for the follow-up appointment, compared to patients who switch away from their regular provider to an alternate clinician.*

## 3. DATA SOURCES AND VARIABLE CONSTRUCTION

### 3.1 Data Description

For this retrospective cohort study, we used a multi-payer dataset that includes the insurance claim information of approximately 3.5 million beneficiaries in the US. While claims data are complex to analyze – often containing duplicated or missing information – they are also a rich source for observational studies (Hopp et al. 2018, Jena et al. 2014). They contain past utilization patterns, information about patient comorbidities, prescription fills, and details of the clinicians involved in the patient journey.

Patients in the sample were insured by a commercial or government (Medicare or Medicaid) plan and each has at least 18 months of continuous medical and pharmaceutical claims eligibility between September 2012 and December 2018. This continuous eligibility requirement ensured that we could observe the patient history both before and after their initial opioid prescription. The patient-level eligibility data also included the date of birth, gender, and location information for each unique patient-payer entry.

The medical claims data contain information including the unique patient and payer identifier, up to the first ten diagnosis codes (International Classification of Disease, ICD9 and ICD10), rendering provider identifier (linked to the National Provider Identifier [NPI]), and service location type (office, inpatient, emergency department, etc.). The NPI was linked to the National Plan and Provider Enumeration System (NPPES) database to obtain clinician information including credentials (MD, DO, PA, NP, etc.), specialty description, and office name and location(s).

The pharmaceutical claims data contain information including the unique patient identifier, prescribing physician identifier (linked to the NPI), generic product identifier (GPI) classification, drug base, drug strength, drug dosage form (e.g., tablet, patch), prescription quantity, days supplied, and filled date. The research methodology was reviewed and approved by the Ethics Committee at *[INTENTIONALLY LEFT BLANK].*

### 3.2 Sample Selection

To ensure clinical accuracy and interpretability, we closely followed related medical literature on opioid misuse when selecting the sample (Barnett et al. 2017, Dowell et al. 2016, Jena et al. 2014). New opioid users were identified using the two-digit GPI classification in the pharmaceutical claims. After a clean period of at least six months during which no opioid prescriptions were filled, we identified the start of a new opioid episode based on when the patient filled their index opioid prescription.

We then used the prescriber's unique identifier to subset the sample to new patient opioid use episodes that could be linked to a corresponding medical claim (considered to be the prescribing visit) in the 14 days prior to filling the prescription (Hoopes et al. 2018). The prescribing appointment must have taken place in the primary care office setting, which we identified using the Centers for Medicare and Medicaid Services (CMS) service location code and the prescriber's specialty as reported in NPPES.

To study the beginning phase of primary care management for new opioid initiates, we excluded all patients who did not return to the primary care setting within 30 days of filling the initial opioid prescription for the same

condition. Consistent with defining the study sample, any corresponding intervention should be designed only to target patients who were already returning to the primary care for follow-up (i.e., not the "one-and-done" patients). However, for completeness, models within Supplement Methods S14 account for this censoring to estimate the effect of provider discordance if the intervention were applied to the full population of opioid initiates.

We also excluded patients who were diagnosed with cancer or end-stage renal disease or who were under the age of 18 at the start of the episode. In this way, we identified 11,340 new opioid patients who had follow-up appointments within 30 days. Supplement Part I provides summary statistics of the patient sample and a detailed list of inclusion and exclusion criteria.

**3.3 Independent Variable**

For each opioid episode, we identified the patient's first follow-up appointment in the primary care setting occurring within 30 days after filling the initial opioid prescription. To increase confidence that the follow-up was related to the initial prescription, the prescribing and follow-up appointments were required to share at least one of the top three ICD diagnosis chapters. A binary independent variable captured whether the initial prescriber was different from the clinician seen for the follow-up appointment (i.e., provider discordance). Of the 11,340 episodes in the sample, 3,211 (28.3%) experienced provider discordance.

**3.4 Dependent Variable**

The primary dependent variable was long-term opioid use. This is typically defined in the medical literature as the patient having filled a minimum of 180 days supplied of opioid prescriptions within the first 360 days after opioid initiation, excluding any days supplied within the first 30 days of the initial prescription (Barnett et al. 2017). The 30-day exclusion period separates the exposure period from the outcome measure, eliminating the concern that the exposure and outcome are correlated by definition.

Because the days supplied field is often auto-populated in pharmaceutical claims, we confirmed robustness of the findings by computing an alternative definition of long-term opioid use that does not depend on the days supplied. Using the CMS standard baseline conversion to compute the full prescription dosage, we calculated patients' daily strength of opioids using the number of days between prescription filled dates (Centers for Medicare and Medicaid Services 2018). Patients were flagged for long-term opioid use if they had a daily dosage of 20 MME or higher 360 days after the initial prescription, a moderate threshold associated with an elevated risk of opioid-related complications and overdose (Dowell et al. 2016).

The two measures of long-term opioid use will be referred to as the *days supplied* or *daily strength* definitions throughout the paper. Whenever the measurement is not specified in the paper, the results correspond to the days supplied measure. In our sample, 10.0% of patients were considered long-term opioid users based on the days supplied definition and 5.9% using the daily strength definition. The rates of long-term opioid use were higher amongst patients who experienced provider concordance (11.5% and 6.9%, using the two long-term opioid use definitions, respectively) compared to those who experienced provider discordance (6.2% and 3.4%,

respectively). More information on the construction of the dependent variable can be found in Supplement Methods S2.

### 3.5 Controls

To isolate the impact of provider discordance on long-term opioid use, it is important to control for any observable factors which influence both provider discordance and a patient's propensity to continue opioid use. First, we gathered characteristics of the initial prescription which may be related to the patient's acuity (i.e., the speed of onset of the condition and need for timely evaluation) and severity (i.e., how bad the condition is). These included covariates for the drug base, average daily opioid strength, and prescription days supplied of the initial opioid prescription (Deyo et al. 2017, Levy et al. 2015).[2] We also included the specialty category of the patient's first prescriber (family medicine, internal medicine, nurse practitioner, or physician assistant), which may be related to the urgency under which the first appointment was scheduled (e.g., a more acute patient may schedule a last-minute appointment with a covering nurse practitioner or physician assistant).

Second, patient characteristics and clinical factors may have influenced both the patient's access to and preference for an alternate primary care clinician, as well as their likelihood of continuing long-term opioid use (Green 2012, Scholl et al. 2019). As such, we included the following patient characteristics: age, sex, payer type, disability status (defined as patients under 65 qualified for a Medicare plan). Condition-related controls included the ICD diagnosis chapter common to the initial and follow-up appointments of the episode, as well as a set of binary variables identifying chronic conditions within the patient's medical history: asthma, coronary artery disease, congestive heart failure, chronic obstructive pulmonary disease, depression, diabetes, and hypertension.

Last, we controlled for the context in which the initial opioid prescription was made. As the study included over five years of observations for patients located across the United States, macro-level changes to primary care scheduling practices or opioid management may have occurred during that time (Dowell et al. 2017, Grosser et al. 2017). For example, patients have become more aware of the risks associated with opioids throughout the study period, e.g., driven by media attention. As such, a time covariate that counts the months since the beginning of the study. We also linked the patient's address information to the annual state-level opioid prescription rate to account for the opioid prescribing culture in the patient's state at the start of their episode (Centers for Disease Control and Prevention 2018).

Within this set of context-related controls, we also adjusted for geographical disparities in opioid prescription and overdose rates by controlling for the patient's geographical census region as well as a nine-level classification of the degree of urbanization and proximity to a metropolitan area for the patient's home address (Cromartie 2013). In addition, we controlled for the opioid-prescribing practices of the initial primary care provider. For this covariate, we calculated the rate at which the prescribing clinician's *other* patients continued long-term opioid use 12 months after opioid initiation. Supplement Methods S1-S2 contain further details on definitions and calculations of the control variables.

[2] To estimate the average daily opioid strength, the total MME of the initial prescription was divided by the days supplied reported in the pharmaceutical claims (Shah et al. 2017). If multiple prescriptions were filled on the same day, we combined the total MME and divided by the longest period of days supplied.

### 3.6 Identifying the Regular Primary Care Provider

To investigate the different discordant care pathways, we used up to two years of a patient's medical history prior to opioid initiation to infer the patient's regular primary care provider. This regular provider was identified as the provider who saw the patient most frequently; in case of a tie, the regular provider was assigned based on which provider saw the patient most recently. As detailed in Supplement Methods S2 and S8, we could identify a regular provider for 10,465 of the 11,340 episodes. We then created an alternate independent variable with the following four levels: (1) provider concordance (N=8,129), (2) patient saw their regular provider for the initial prescribing appointment and had a follow-up with an alternate clinician (N=1,123), (3) patient saw an alternative clinician for their initial prescribing appointment and followed up with their regular provider (N=934), and (4) patient experienced provider discordance, but neither clinician involved was identified as the patient's regular clinician (N=1,154).

## 4. STATISTICAL ANALYSIS AND RESULTS

### 4.1 Testing Hypothesis 1: The Main Effect

We first tested Hypothesis 1 using a standard logistic model on the 11,340-episode sample. The full model output and a description of the estimation procedure are provided in Supplement Methods S4. As summarized in Table 1, the estimation results for the logistic models identify a large and statistically significant effect of provider discordance on long-term opioid use.

| **Control Structure** | **None** | **Initial Prescription ($O_i$)** | **+ Patient Character-istics ($P_i$)** | **+ Patient Conditions ($C_i$)** | **+ Context ($T_i$)** |
|---|---|---|---|---|---|
| Initial Prescription **($O_i$)** | No | Yes | Yes | Yes | Yes |
| Patient Characteristics **($P_i$)** | No | No | Yes | Yes | Yes |
| Patient Condition **($C_i$)** | No | No | No | Yes | Yes |
| Context **($T_i$)** | No | No | No | No | Yes |
| **Adjusted Odds Ratio of Provider Discordance on Long-Term Opioid Use: Days Supplied Measurement** | | | | | |
| Adjusted Odds Ratio | 0.51 | 0.66 | 0.72 | 0.70 | 0.69 |
| 95% Confidence Interval | (0.43-0.60) | (0.55-0.78) | (0.60-0.85) | (0.59-0.84) | (0.57-0.82) |
| P-Value | <0.001*** | <0.001*** | <0.001*** | <0.001*** | <0.001*** |
| AUC$^{+}$ | 0.56 | 0.78 | 0.82 | 0.83 | 0.84 |
| **Adjusted Odds Ratio of Provider Discordance on Long-Term Opioid Use: Daily Strength Measurement** | | | | | |
| Adjusted Odds Ratio | 0.47 | 0.61 | 0.66 | 0.65 | 0.63 |
| 95% Confidence Interval | (0.38-0.58) | (0.49-0.77) | (0.53-0.83) | (0.51-0.82) | (0.50-0.80) |
| P-Value | <0.001*** | <0.001*** | <0.001*** | <0.001*** | <0.001*** |
| AUC$^{+}$ | 0.56 | 0.80 | 0.84 | 0.85 | 0.86 |

* p<0.05, ** p<0.01, *** p<0.001; AUC$^{+}$ = Area under the receiver operating characteristics (ROC) curve

**Table 1. Logistic models.** *The models consistently estimate a significant effect of provider discordance on long-term opioid use at the 0.1% level, suggesting that the effect of provider discordance was not driven by model specification. The largest gain in fit (measured by AUC) comes from the initial prescription controls; additional patient characteristics, condition, and context controls make little difference to model fit or estimates.*

After accounting for all initial prescription, patient, condition, and context covariates in the episode-level logistic model, we found that patients who experienced provider discordance were less likely to become long-term opioid users compared to those who returned to the initial prescribing clinician for the follow-up appointment (adjusted odds ratio of 0.69 [95% CI: (0.57, 0.82)] using the days supplied measure and 0.63 [95% CI: (0.50, 0.80)] using the alternative daily strength outcome measure). These findings support Hypothesis 1.

### 4.2 Investigating Endogeneity Bias Within the Main Effect

Although results of the controlled logistic models above are in line with Hypothesis 1, the data are retrospective, and patients were not assigned to a control group (provider concordance) or treatment group (provider discordance) at random. Consequently, coefficients may be biased by the presence of unobserved confounders that simultaneously affected the likelihood of a patient experiencing provider discordance and their likelihood of becoming a long-term opioid user. We therefore introduced two additional analysis techniques to identify and correct for endogeneity biases: (1) an instrumental variable method and (2) a propensity-score matching approach called the minimum-bias estimator.

#### 4.2.1 Instrumental Variable Method

When correctly specified, the instrumental variable (IV) approach addresses concerns about self-selection bias in the model by effectively randomizing patient assignment into the exposure group to balance differences in both observed and unobserved confounders (Clarke and Windmeijer 2012, Heckman 1979). This technique makes it possible to accurately estimate causal treatment effects despite the lack of randomization in our study design (Wooldridge 2015 p. 594).

The IV approach typically requires the inclusion of an IV: a variable that should be predictive of the treatment (i.e., relevant) but have no direct impact on the outcome, nor be correlated with omitted variables that affect the outcome (i.e., valid). We chose the IV as the rate at which the prescribing clinician's *other* patients saw a different clinician for their follow-up appointment. Specifically, for each initial prescriber, we identified primary care appointments in the 12 months prior to the start of the episode where the patients in those appointments had a follow-up primary care visit within 30 days. The prescribing clinician's switching rate was defined as the proportion of appointments where their past patients saw a different clinician for their follow-up appointment.

If the prescribing clinician's other patients were frequently changing clinicians for their follow-up appointment (for any reason, e.g., scheduling availability of the provider), we expected the focal patient of the episode also to be more likely to experience provider discordance (i.e., relevant). However, because the IV is constructed using patients *other* than the focal patient, we have no reason to expect that the switching rate of these other patients should directly impact the focal patient's likelihood of becoming a long-term opioid user (i.e., valid).

That said, the IV may be invalidated if other, unobserved, factors (such as a patient's motivation to seek out non-opioid treatment) are correlated with both the IV and the likelihood of long-term opioid use. The extensive set of covariates – including the long-term opioid use rate for other opioid initiates at the same initial prescriber – reduces the likelihood that this will be the case. Further exploration of the IV relevance and validity is presented in Supplement Methods S5.

Using this IV, we estimated a patient's selection for treatment and outcome simultaneously in a recursive bivariate model:

$$\text{Treatment Equation: } x_i^* = \alpha + \beta_1 z_i' + b_1 V_i' + \gamma_1 O_i' + \delta_1 P_i' + \lambda_1 C_i' + \theta_1 T_i' + \varepsilon_{2i}, \quad \textbf{[1]}$$

$$\text{Outcome Equation: } y_i^* = \alpha + \beta_2 x_i + b_2 V_i' + \gamma_2 O_i' + \delta_2 P_i' + \lambda_2 C_i' + \theta_2 T_i' + \varepsilon_{1i}, \quad \textbf{[2]}$$

$$x_i = \mathbb{I}[x_i^* > 0], y_i = \mathbb{I}[y_i^* > 0],$$

where $x_i^*$ and $y_i^*$ are latent variables indicating provider discordance and long-term opioid usage, respectively. The vectors $O_i$, $P_i$, $C_i$, and $T_i$ contain the set of all prescription, patient, condition, and context covariates, respectively, $\mathbb{I}[\cdot]$ is the indicator function, $z_i'$ contains the IV and $V_i'$ contains an additional binary IV control that equals 1 when fewer than ten observations are available to calculate the IV and 0 otherwise. The error terms $\varepsilon_{1i}$ and $\varepsilon_{2i}$ are jointly distributed, with the estimated correlation capturing the aggregate effect of all omitted variables that affect selection into both the treatment and outcome.

To minimize misspecification error, we imposed no prior restriction on the distributional form of $(\varepsilon_{1i}, \varepsilon_{2i})$. We tested a variety of link functions (probit, logit, complementary log-log [cloglog]) and joint error distributions (modeled as copulas, such as Normal, Clayton, and Gumbel) to allow for either bivariate normal or non-normal dependencies between the treatment and outcome equations. The Supplement Methods S6 contains more details on the methodology and results from the full set of model specifications.

In the selection equation, the IV was highly predictive of provider discordance: a coefficient of 0.23 (95% CI: [0.22-0.25], p<0.001) was estimated for a 10% change in the IV. The bivariate models indicated an insignificant or small positive selection bias across a range of model specifications, as evidenced by the respective Kendall tau parameters. The best-fit recursive bivariate model with probit-logit marginals and a Clayton copula yielded an adjusted odds ratio of provider discordance of 0.46 (95% CI: [0.30, 0.70]) using the days supplied outcome measure and 0.42 (95% CI: [0.24, 0.72]) using the daily strength outcome measure. As such, the results from the logistic model appear conservative.

#### 4.2.2 Minimum Bias Estimator

Despite the range of sensitivity analyses we performed on the IV model (see Supplement Methods S5-S6) to validate the IV and corresponding results, there may still be concerns about the validity of the instrument. We therefore leveraged a second approach to address endogeneity concerns based on propensity-score matching methods. We matched patients in the control group (concordance) in a 1:1 ratio with patients in the treatment group (discordance) using nearest neighbor matching, with the condition that the closest propensity score can be no greater than 0.2 standard deviations away from the switchers' propensity score. This condition has been shown to reduce more than 90% of the bias due to observable differences (Gu and Rosenbaum 1993) as well alleviate some of the impact of unobserved bias (Rosenbaum 2005).

We then followed a technique detailed by Millimet and Tchernis (2013, p. 983) that uses the propensity scores from matching to estimate a minimum-biased estimator (MBE). The MBE aims to minimize the potential impact of omitted variable bias by restricting the sample to matched cases with propensity scores within a defined interval around 0.5, the propensity score which most closely mimics a coin-flip or random assignment to the

treatment and control groups. By estimating the logistic model on increasingly restricted subsamples (for example, limiting to propensity scores between 0.10 and 0.90, then 0.25 to 0.75, etc.), the method will allow us to detect the direction of endogeneity and reduce bias in the estimated effect size.

As shown in Table 2, the adjusted odds ratios get increasingly smaller as the range of propensity scores narrow. When we reach the narrowest propensity-score range ([0.33, 0.67], as recommended by Black and Smith (2004)) the estimated size of the effect of provider discordance on long-term opioid use increases to 0.55 (95% CI: [0.38-0.79]) and 0.45 (95% CI: [0.28-0.73]) for the days supplied and daily strength measures, respectively.[3] These estimates of the effect size using the MBE approach are similar to those observed using the IV approach.

| | **Days Supplied Measurement** | | **Daily Strength Measurement** | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Full sample (N=11,340) | 0.69 (0.57-0.82) | <0.001*** | 0.62 (0.50-0.80) | <0.001*** |
| Propensity matched sample (N=6,084) | 0.75 (0.62-0.91) | 0.003** | 0.62 (0.48-0.79) | <0.001*** |
| Propensity matched sample with scores 0.10-0.90 (N=5,966) | 0.72 (0.59-0.88) | 0.001** | 0.59 (0.46-0.76) | <0.001*** |
| Propensity matched sample with scores 0.25-0.75 (N=4,113) | 0.65 (0.49-0.87) | 0.003** | 0.45 (0.32-0.65) | <0.001*** |
| Propensity matched sample with scores 0.33-0.67 (N=2,745) | 0.55 (0.38-0.79) | 0.001** | 0.45 (0.28-0.73) | 0.001** |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

**Table 2. Estimated effect of provider discordance using the Minimum Bias Estimator method.** *The propensity-matched sample is reduced to observations with scores closer to 0.5 to correct for unobserved confounders. The estimates of the effect of provider discordance on long-term opioid use become more negative as endogeneity bias is increasingly addressed.*

In summary, the IV and MBE methods consistently suggest that unobserved factors, such as self-selection, rendered the group of patients who experienced discordance *more* likely to become long-term opioid users. In other words, those patients who chose not to change provider or who were unable to do so (i.e., from limited provider access in their region) would have benefitted even more from discordant care than those patients in our sample who already received the treatment. In conjunction with additional analyses presented in Supplement Part VI, all evidence indicates that the logistic model estimates underestimate the beneficial effect of provider discordance on long-term opioid use and support Hypothesis 1.

### 4.3 Testing Hypothesis 2: A Mechanism

Hypotheses 2a and 2b investigate a potential mechanism (a change in short-term opioid usage) that could explain the relationship between provider discordance and long-term opioid use. To test the hypotheses, we separately estimated the following equations:

$$m_i^* = \mu + \eta x_i + O_i'\nu + P_i'\xi + C_i'\varphi + T_i'\psi + \nu_i,\ m_i = \mathbb{I}[m_i^* > 0] \quad \textbf{[3]}$$

$$y_i^* = \alpha + \omega m_i + O_i'\gamma + P_i'\delta + C_i'\lambda + T_i'\theta + \varepsilon_i,\ y_i = \mathbb{I}[y_i^* > 0] \quad \textbf{[4]}$$

[3] Note that the p-values may be less significant as the effects are estimated on a smaller sample.

where $m_i^*$ captures whether the patient filled an opioid prescription after their follow-up appointment, and both $\nu$ and $\varepsilon$ are error terms following the standard logistic distribution. To remain consistent with the 14-day time window used when matching the initial prescription with the prescribing appointment, this subsequent opioid prescription must have been filled within 14 days of that follow-up appointment. The coefficient $\eta$ in Equation [3] captures the extent to which patients who change clinicians for their follow-up appointment are more likely to fill a subsequent opioid prescription (Hypothesis 2a). Meanwhile, $\omega$ in Equation [4] estimates the association between filling a subsequent prescription and becoming a long-term opioid user (Hypothesis 2b).

The model estimates support Hypotheses 2a and 2b. Patients who experienced provider discordance were less likely to fill an opioid prescription in the subsequent 14-day period (adjusted odds ratio 0.79 [95% CI: 0.71, 0.88]). Furthermore, patients who filled an opioid prescription within 14 days after the follow-up appointment were much more likely than those who did not to become long-term users (adjusted odds ratio 6.31 [95% CI: 5.41, 7.36] using the days supplied outcome measure, and 5.34 [95% CI: 4.40, 6.49] using the daily strength outcome measure). While we do not claim causal identification using these models, these correlational results identify a potential mechanism underlying the effect that provider discordance has on long-term opioid use.

**4.4 Testing Hypothesis 3: Provider Discordance Pathways**

To test the consistency of the provider discordance effect, we examined whether the strength of the effect differed depending on whether and when the patient saw their regular provider. In the fully controlled logistic models, we replaced $x_i$ with the categorical variable $X_i$ which has four levels corresponding to each of the pathways described in Section 3.6. We also added a binary variable $r_i$ that specified whether the patient saw their regular provider for either the first or second appointment ($r_i = 0$ if neither the first nor the second appointment was with the regular provider). This ensured that any advantage associated with a patient seeing their regular provider during the observed care episode was already accounted for in the models before testing for the potential additional benefit of discordant care.

$$y_i^* = \alpha + \beta X_i + \sigma r_i + \gamma O_i' + \delta P_i' + \lambda C_i' + \theta T_i' + \varepsilon_i, y_i = \mathbb{I}[y_i^* > 0]. \quad \textbf{[5]}$$

As displayed in Table 3 (and further detailed in Supplement Methods S8), the effect of provider discordance appears strongest for the set of patients who started with a non-regular provider for their first visit then switched to their regular PCP (adjusted odds ratio 0.58 [95% CI: (0.40-0.84)] and 0.51 [95% CI: (0.31-0.84)] for the days supplied and daily strength measurements, respectively). However, even for those patients who start with their regular provider and then switch to an alternative clinician, we find that discordance reduces their likelihood of becoming long-term opioid users.

We ran a series of hypothesis tests to determine whether the discordant care effect sizes are statistically different depending on the discordance pathway. The corresponding p-values (all greater than 0.31 using the chi-squared test) do not allow us to reject the hypothesis that the effect sizes are the same. While the sample size may limit the power to identify significant differences in effect sizes, the above results indicate that *all* types of patient pathways (even those where the patient is initially prescribed by their regular PCP) benefit from provider discordance at the beginning of the opioid episode. As such, we do not find evidence in support of Hypothesis 3.

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Effect on Long-Term Opioid Use** | **Adjusted Odds Ratio (95% CI)** | **P-value** | **Adjusted Odds Ratio (95% CI)** | **P-value** |
| Discordance pathway 1: regular provider to alternate provider | 0.73 (0.56-0.94) | 0.014* | 0.68 (0.49-0.95) | 0.024* |
| Discordance pathway 2: alternate provider to regular provider | 0.58 (0.40-0.84) | 0.004** | 0.51 (0.31-0.84) | 0.008** |
| Discordance pathway 3: alternate provider to alternate provider | 0.72 (0.53-0.97) | 0.034* | 0.65 (0.44-0.97) | 0.036* |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

**Table 3. Effect of provider discordance within different pathways.** *Pathways describe whether and when the patient was seen by their regular (most frequently visited) primary care provider, during the initial prescribing appointment or follow-up appointment.*

### 4.5 Alternative Explanations and Robustness Tests

One alternative explanation for our findings is that provider discordance is a proxy for the acuity of pain onset. High acuity patients may be more likely to receive an opioid script from a covering clinician (due to unavailability at short notice of their preferred provider) and then return to their regular provider for follow-up. If this is the case, these patients may be simultaneously more likely to experience discordance between the initial providers and less likely to become long-term opioid users (as their condition is more acute than chronic).

This potential omitted variable has been addressed in part by both the IV and MBE approaches, as they are designed to estimate the causal effect of provider discordance after adjusting for potential confounds such as acuity (see Supplement Methods S5-S7 for further discussion). In addition, the pathway analysis described in Section 4.4 indicates that all provider discordance pathways lead to reduced long-term opioid use rates, i.e., the effect we find does not appear to be explained exclusively by patients switching from a covering clinician *to* their regular clinician (the pathway expected to occur more likely amongst acute patients).

Despite these analyses, we acknowledge that unobserved factors that may still bias the results. For example, certain providers may be more available than others for more urgent appointments, and so may have a higher proportion of patients with acute pain onsets. To account for this and for other forms of provider-level heterogeneity, we included fixed and random effects at the initial prescriber-level in Supplement Methods S11. In Supplement Methods S12, we explored whether the observed effect of provider discordance differed depending on whether opioid initiates were being treated for a new (likely more acute pain) or a potentially pre-existing condition. In Supplement Methods S13, we included additional controls (available for 80.2% of patient episodes) that estimated the patient's socioeconomic status (differences which have been associated with variation in opioid use) and the estimated average patient knowledge of opioid risk, as measured by changes to the media attention and interest in the opioid epidemic across the study period. Other sensitivity analyses include relaxing the assumption that provider discordance must occur during the first two appointments of the patient's opioid journey (Supplement Methods S15), and measuring the impact of provider discordance on opioid use at different points throughout the first year, e.g., 6 months, 9 months, etc. (Supplement Methods S16). All estimates were consistent with the main results.

## 5. DISCUSSION AND PRACTICAL IMPLICATIONS

### 5.1 Discussion and Policy Recommendations

As opioid overdoses tend to occur with patients who are already in a pattern of chronic opioid use, many intervention strategies have aptly approached the opioid epidemic from the perspective of clinical management of ongoing dependence (Grosser et al. 2017, Hallvik et al. 2018, Pitt et al. 2018). However, our analysis suggests that operational interventions during the early phases of opioid use may also be extremely valuable. For opioid-naïve patients initially managed with an opioid treatment in the primary care setting, this research identified that clinician discordance can significantly disrupt the path to long-term opioid use.

A logistic regression model estimated that the likelihood of long-term opioid use at 360 days is 31% lower (95% CI: [18%, 43%]) for patients who followed up with an alternate primary care clinician, controlling for the strength and days supplied of the initial opioid prescription (known to be associated with risk of long-term opioid use and overdose) as well as a range of patient, condition, and context factors. Robustness tests including an IV analysis and propensity-score matching (utilizing the MBE approach) suggest that this estimate is, if anything, a conservative assessment of the true causal effect. While multiple potential mechanisms may explain the difference in long-term opioid use, one identified explanatory factor is that patients who experienced provider discordance were also 21% (95% CI: [12%, 29%]) less likely to fill a subsequent opioid prescription after the follow-up appointment. This decrease can have a large long-term impact, as filling an opioid prescription after the follow-up appointment was associated with a 6.31 [95% CI: 5.41, 7.36] times higher likelihood of continuing long-term opioid use.

We also performed a pathway analysis to assess whether the effects of provider discordance remained significant for patients whose initial prescriber was their regular provider. While there was evidence of a stronger effect on patients whose initial prescriber was *not* their non-regular clinician – with long-term opioid use at 360 days for such patients reduced by 42% (95% CI: [16%, 60%]) – the difference compared to other pathways was not statistically significant. This indicates that the benefits of provider discordance are pervasive and that, on average, patients benefit from seeing a different clinician regardless of who they saw for the first or follow-up appointments.

While it is critical for the results of this study to be corroborated through other data sets and study designs, there are multiple potential policy changes that could help facilitate provider discordance within the early stages of opioid use. As an example, approximately 85% of primary care physicians practice in an office with at least one other physician (Liaw et al. 2016). If a patient is flagged in the medical record system upon receiving a new opioid prescription, and the patient subsequently contacts the practice to schedule a follow-up appointment, then the office manager could schedule this appointment with a clinician other than the initial prescriber.[4] Without posing significant additional overhead or risking patient loss for practices involved, such a policy could impact a large portion of new opioid initiates in the primary care setting. As an operational process change, this intervention also circumvents the need for clinician education and other resource-intensive efforts to elicit

[4] As shown in Supplement S10, we did not observe any significant difference in the effect of provider discordance when the two clinicians shared an office versus when the two clinicians operated out of different offices.

longstanding behavior change for individual providers (Morris et al. 2011). In a new era of opioid management where tools such as prescription drug monitoring programs are becoming increasingly available for clinicians to see a broader picture of the patient's history and opioid-use patterns, the medical community is well poised to incorporate safeguarding measures focused on the early stages of opioid use.

### 5.2 Limitations

While these results provide promising evidence that the design of care management shortly after opioid initiation may significantly disrupt the pathway to long-term opioid use, the implications are not without limitations. First, it must be emphasized that this research does not claim that all opioid use is unnecessary or that conservative, alternative treatments are always more appropriate to manage ongoing pain; while aggressive opioid prescribing may induce significant clinical harm without providing meaningful pain improvement, some patients and conditions may respond well to pain management through opioids (Glod 2017). There is also a possibility that provider discordance shortly after opioid initiation could negatively affect other aspects of care quality.[5]

Second, this study is focused on new opioid initiates who return for a follow-up appointment in the primary care setting. As such, the findings cannot necessarily be extended to all new opioid users – for instance, to patients who require a one-off prescription or who have higher acuity pain that requires them to receive their first prescription in an emergency or urgent care setting. While our analyses show an aggregate net positive effect of introducing provider discordance within the primary care setting (see Supplement Methods S14), further analyses should investigate the impact of potential policy interventions on specific at-risk populations.

Third, although there are plausible explanations as to why provider discordance might be effective in this context (including an enlarged information base and correction to confirmation bias and anchoring), this study does not identify the precise behavioral mechanism behind the observed effect. Developing a clearer understanding of why this effect exists and whether the impact would be observed in other (i.e., non-opioid related) contexts could help researchers design and target potential interventions effectively.

Last, while we have taken various steps in our paper to address the issue of causality (e.g., by including an extensive set and combinations of controls in sensitivity analyses and employing IV and MBE methods), as is the case with all observational studies we cannot fully rule out the possibility of omitted variable bias. Although retrospective cohort studies are the norm in opioid research (Longhurst et al. 2014) and we recommend that similar analyses be replicated with other datasets, the only way to definitively confirm our findings is through a randomized control trial. The size and consistency of our findings across models suggest that such trials and corresponding investment into designing and testing clinical interventions may be warranted.

### 5.3 Conclusion

High-quality medical decision-making is at the heart of good medicine, yet it cannot be taken for granted. This study demonstrates that systematic operational changes in the primary care setting following opioid initiation may be a promising target to reduce the influx of patients afflicted by the opioid epidemic. We advocate for future

[5] In Supplement Methods S17, we investigate whether provider discordance shortly after opioid initiation is associated with an increase in emergency room usage (a quality measure that indicates poor pain management). While we recommend further analyses to explore this possibility of negative externalities, we did not identify any effect.

research to identify potential interventions early in the process of opioid management that can safely and effectively lower rates of long-term opioid use.

## REFERENCES


Ahuja V, Alvarez C, Staats BR (2020) Maintaining Continuity in Service: An Empirical Examination of Primary Care Physicians. *Manuf. Serv. Oper. Manag.*

American Medical Association (2020) *Issue brief: Reports of Increases in opioid-related overdose during COVID pandemic*

Ariely D, Loewenstein G, Prelec D (2003) "Coherent arbitrariness": Stable demand curves without stable preferences. *Q. J. Econ.* 118(1):73–105.

Atlas S J, Grant R W, Ferris T G, Chang Y, Barry M J, Atlas Steven J, Grant Richard W, Ferris Timothy G, Chang Yuchiao, Barry Michael J (2009) Patient-physician connectedness and quality of primary care. *Ann. Intern. Med.* 150(5):325–335.

Ayer T, Alagoz O, Stout NK (2012) OR forum-A POMDP approach to personalize mammography screening decisions. *Oper. Res.* 60(5):1019–1034.

Barnett M, Olenski A, Jena A (2017) Opioid-Prescribing Patterns of Emergency Physicians and Risk of Long-Term Use. *N. Engl. J. Med.* 376(7):663–673.

Black DA, Smith JA (2004) How robust is the evidence on the effects of college quality? Evidence from matching. *J. Econom.* 121(1–2):99–124.

Brooks AW, Gino F, Schweitzer ME (2015) Smart people ask for (My) advice: Seeking advice boosts perceptions of competence. *Manage. Sci.* 61(6):1421–1435.

Centers for Disease Control and Prevention (2018) U.S. Opioid Prescribing Rate Maps. *U.S. Dep. Heal. Hum. Serv. Opioid Overdose*. Retrieved (July 25, 2019), https://www.cdc.gov/drugoverdose/maps/rxrate-maps.html.

Centers for Medicare and Medicaid Services (2018) Opioid Oral Morphine Milligram Equivalent (MME) Conversion Factors. *Prescr. Drug Cover.*:1–2. Retrieved https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/.

Chan CW, Farias VF, Bambos N, Escobar GJ (2012) Optimizing Intensive Care Unit Discharge Decisions with Patient Readmissions. *Oper. Res.* 60(6):1323–1341.

Chan CW, Green L V., Lekwijit S, Lu L, Escobar G (2019) Assessing the impact of service level when customer needs are uncertain: An empirical investigation of hospital step-down units. *Manage. Sci.* 65(2):751–775.

Chiu AS, Jean RA, Hoag JR, Freedman-Weiss M, Healy JM, Pei KY (2018) Association of Lowering Default Pill Counts in Electronic Medical Record Systems with Postoperative Opioid Prescribing. *JAMA Surg.* 153(11):1012–1019.

Christensen CM, Grossman JH, Hwang J (2009) *The Innovator's Prescription: A Disruptive Solution for Health Care* (McGraw-Hill).

Clark JR, Huckman RS, Staats BR (2013) Learning from customers: Individual and organizational effects in outsourced radiological services. *Organ. Sci.* 24(5):1539–1557.

Clarke PS, Windmeijer F (2012) Instrumental variable estimators for binary outcomes. *J. Am. Stat. Assoc.*

107(500):1638–1652.

Compton WM, Jones CM, Baldwin GT (2016) Relationship between Nonmedical Prescription-Opioid Use and Heroin Use. *N. Engl. J. Med.* 374(2):154–163.

Cromartie J (2013) Rural-Urban Continuum Codes. *US Dep. Agric. Econ. Res. Serv.* Retrieved (November 19, 2019), https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.

Deyo RA, Hallvik SE, Hildebran C, Marino M, Dexter E, Irvine JM, O'Kane N, et al. (2017) Association Between Initial Opioid Prescribing Patterns and Subsequent Long-Term Use Among Opioid-Naïve Patients: A Statewide Retrospective Cohort Study. *J. Gen. Intern. Med.* 32(1):21–27.

Dickerson JP, Procaccia AD, Sandholm T (2019) Failure-aware kidney exchange. *Manage. Sci.* 65(4):1768–1791.

Dowell D, Arias E, Kochanek K, Anderson R, Guy GPJ, Losby J, Baldwin G (2017) Contribution of Opioid-Involved Poisoning to the Change in Life Expectancy in the United States, 2000-2015. *JAMA* 318(11):1065–1067.

Dowell D, Haegerich TM, Chou R (2016) CDC Guideline for Prescribing Opioids for Chronic Pain - United States, 2016. *JAMA* 315(15):1624–1645.

Farley KX, Anastasio AT, Kumar A, Premkumar A, Gottschalk MB, Xerogeanes J (2019) Association Between Quantity of Opioids Prescribed After Surgery or Preoperative Opioid Use Education With Opioid Consumption. *JAMA* 321(24):2465–2466.

Freeman M, Robinson S, Scholtes S (2020) Gatekeeping, Fast and Slow: An Empirical Study of Referral Errors in the Emergency Department. *Manage. Sci.* Forthcomin.

Freeman M, Savva N, Scholtesa S (2017) Gatekeepers at work: An empirical analysis of a maternity unit. *Manage. Sci.* 63(10):3147–3167.

Gino F, Pisano G (2008) Toward a theory of behavioral operations. *Manuf. Serv. Oper. Manag.* 10(4):676–691.

Glod SA (2017) The other victims of the opioid epidemic. *N. Engl. J. Med.* 376(22):2101–2102.

Green L V. (2012) The vital role of operations analysis in improving healthcare delivery. *Manuf. Serv. Oper. Manag.* 14(4):488–494.

Grosser BT, Woolf CJ, Fitzgerald GA (2017) Time for nonaddictive relief of pain. *Science (80-. ).* 355(6329):1026–1027.

Gu XS, Rosenbaum PR (1993) Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms. *J. Comput. Graph. Stat.* 2(4):405–420.

Hallvik S, Geissert P, Wakeland W, Hildebran C, Carson J, O'Kane N, Deyo R (2018) Opioid-Prescribing Continuity and Risky Opioid Prescriptions. *Ann. Fam. Med.* 16(5):440–442.

Heckman JJ (1979) Sample Selection Bias as a Specification Error. *Econometrica* 47(1):153–161.

Holland KM, Jones C, Vivolo-Kantor AM, Idaikkadar N, Zwald M, Hoots B, Yard E, et al. (2021) Trends in US Emergency Department Visits for Mental Health, Overdose, and Violence Outcomes Before and During the COVID-19 Pandemic. *JAMA Psychiatry* 30341:1–8.

Hoopes M, Angier H, Raynor LA, Suchocki A, Muench J, Marino M, Rivera P, Huguet N (2018) Development of an algorithm to link electronic health record prescriptions with pharmacy dispense claims. *J. Am. Med. Informatics Assoc.* 25(10):1322–1330.

Hopp WJ, Li J, Wang G (2018) Big Data and the Precision Medicine Revolution. *Prod. Oper. Manag.* 27(9):1647–1664.

Jena AB, Goldman D, Weaver L, Karaca-Mandic P (2014) Opioid prescribing by multiple providers in Medicare: Retrospective observational study of insurance claims. *BMJ* 348(February):1–12.

Kuntz L, Scholtes S, Sülz S (2019) Separate and concentrate: Accounting for patient complexity in general hospitals. *Manage. Sci.* 65(6):2482–2501.

Levy B, Paulozzi L, Mack KA, Jones CM (2015) Trends in Opioid Analgesic-Prescribing Rates by Specialty, U.S., 2007-2012. *Am. J. Prev. Med.* 49(3):409–413.

Liaw WR, Jetty A, Petterson SM, Peterson LE, Bazemore AW (2016) Solo and small practices: A vital, diverse part of primary care. *Ann. Fam. Med.* 14(1):8–15.

Longhurst CA, Harrington RA, Shah NH (2014) A "green button" for using aggregate patient data at the point of care. *Health Aff.* 33(7):1229–1235.

Millimet DL, Tchernis R (2013) Estimation of the treatment effects without an exclusion restriction: with an application to tha analysis of the school breakfast program. *J. Appl. Econom.* 28:982–1017.

Morone NE, Weiner DK (2013) Pain as the fifth vital sign: Exposing the vital need for pain education. *Clin. Ther.* 35(11):1728–1732.

Morris ZS, Wooding S, Grant J (2011) The answer is 17 years, what is the question: Understanding time lags in translational research. *J. R. Soc. Med.* 104(12):510–520.

Owen GT, Bruel BM, Schade CM, Eckmann MS, Hustak EC, Engle MP (2018) Evidence-based pain medicine for primary care physicians. *Baylor Univ. Med. Cent. Proc.* 31(1):37–47.

Pitt A, Humphreys K, Brandeau M (2018) Modeling Health Benefits and Harms of Public Policy Responses to the US Opioid Epidemic. *Am. J. Public Health* 108(10):1394–1400.

Reagans R, Argote L, Brooks D (2005) Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Manage. Sci.* 51(6):869–881.

Rosenbaum PR (2005) Heterogeneity and causality: Unit heterogeneity and design sensitivity in observational studies. *Am. Stat.* 59(2):147–152.

Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G, Release E (2019) Drug and Opioid-Involved Overdose Deaths — United States, 2013 – 2017. *Morb. Mortal. Wkly. Rep.* 67(51–52):1419–1427.

Senot C (2019) Continuity of Care and Risk of Readmission: An Investigation into the Healthcare Journey of Heart Failure Patients. *Prod. Oper. Manag.* 28(8):2008–2030.

Shumsky RA, Pinker EJ (2003) Gatekeepers and referrals in services. *Manage. Sci.* 49(7):839–856.

Siemsen E, Roth A V., Balasubramanian S, Anand G (2009) The influence of psychological safety and confidence in knowledge on employee knowledge sharing. *Manuf. Serv. Oper. Manag.* 11(3):429–447.

Song H, Tucker A, Graue R, Moravick S, Yang J (2019) Capacity Pooling in Hospitals: The Hidden Consequences of Off-Service Placement. *Manage. Sci.*:1–18.

Staats BR, Diwas SKC, Gino F (2018) Maintaining beliefs in the face of negative news: The moderating role of experience. *Manage. Sci.* 64(2):804–824.

Wooldridge JM (2015) *Introductory Econometrics: A Modern Approach* Sixth. (Australia: Cengage Learning).

**SUPPLEMENTARY MATERIAL**

for

# Curbing the Opioid Epidemic at its Root: The Effect of Provider Discordance after Opioid Initiation

**CONTENTS**

**SUPPLEMENT PART I. DATA SAMPLE**

### Methods S1. Medical and pharmaceutical claims dataset

The study observation period was set from February 2012 through April 2019 to allow approximately 90 days between April 2019 and when the data was analyzed to account for lag in the claims collection process. Patients included in the sample had both medical and pharmaceutical claims within the first 30 days of the start of their episode, increasing the likelihood that we observed the full patient journey via payers included in the dataset. Each claim contains a unique patient identifier and all patients were required to be continuously eligible with payers in the dataset for a minimum of 18 months. Patient episodes were identified from the claims data using the criteria in **Table S1**.

**Table S1. Criteria used to define new opioid-use episodes**

| Criterion | Patient episode count |
|---|---|
| Restricted to patient episodes with an opioid prescription other than buprenorphine or methadone[1] after an 'opioid-free' clean period of at least 180 days. Prescriptions must have non-zero and non-negative values for the days supplied and quantity fields. Excluded the following drug administration modes: intravenous, subcutaneous, intramuscular, intra-articular, and injection.[2] Required patients to have continuous medical and pharmaceutical eligibility with insurers in the dataset during the full analysis period: 180-day clean period prior to and 375 days following the new opioid initiation. | 543,238 |
| Restricted to episodes for which the initial opioid prescription could be linked (using the patient and prescriber's unique identifiers) to at least one corresponding medical claim in the 14 days prior to the prescription fill date. If the patient saw the prescribing clinician more than once prior to filling the prescription, the appointment closest to the opioid fill date was designated as the prescribing event. | 278,462 |
| Restricted to episodes in which the prescribing event is in the primary care office setting, identified using the CMS service location code (11, 50, 71, 72), clinician's credential (MD, DO, PA, PAC or NP) and specialty as reported in the National Plan and Provider Enumeration System (NPPES) data registry (see **Table S2** for breakdown of specialties). | 53,491 |
| Restricted to episodes where the patient returned to the primary care office setting (defined using the primary care location and specialty requirements as above) within 30 days of the initial prescribing visit. Required the two visits must have at least one matching diagnosis International Classification of Diseases (ICD) 9 or 10 chapter within the first 3 diagnosis codes to ensure that the appointments were related.[3] If the patient had multiple primary care appointments within the 30-day window, designated the first of these visits as the follow-up appointment. | 13,195 |
| Excluded episodes where the patient had any history of cancer diagnosis (ICD9 codes 140-165, 170-176, 179-208, 230-234 or ICD10 codes C01-C26, C30-C34, C37-C41, C43-C58, C60-C86, C88, C90-C96, C4A, C7A, C7B, D00-D09, D40-D49, and R97), end-stage renal disease (at least 2 visits with a diagnosis of ESRD and an Evaluation and Management procedure - inpatient, outpatient, or at the emergency department - at any point in their claims history), or was under the age of 18 at the start of the episode. | 11,340 |

[1] Often used for treating opioid dependence or addiction, and therefore potentially indicates that the patient is an opioid user, even if those opioids were not prescribed for the patient in a medical setting.

[2] As the analysis is focused on the primary care setting, prescriptions should be able to be administered by the patient at home (e.g. oral, patch).

[3] This logic was incorporated so as exclude episode in which, for example, the patient was prescribed opioids for back pain and then returned within 30 days to their primary care clinic with an unrelated complaint of bronchitis.

**Table S2. Included primary care specialties**

| Primary Care Category | Provider Taxonomy Code | Specialty Description | Initial Prescribers N = 11340, No. (%) | Follow-Up Clinicians N = 11340, No. (%) |
|---|---|---|---|---|
| Family Medicine | 207Q00000X | Family Medicine (FM) | 5249 (46%) | 5011 (44%) |
| | 207QA0000X | FM – Adolescent Medicine | 4 (0%) | 4 (0%) |
| | 207AG0300X | FM – Geriatric Medicine | 13 (0%) | 15 (0%) |
| | 208D00000X | General Practice | 90 (1%) | 81 (1%) |
| Internal Medicine | 207R00000X | Internal Medicine (IM) | 3508 (31%) | 3516 (31%) |
| | 207RA0000X | IM – Adolescent Medicine | 0 (0%) | 2 (0%) |
| Nurse Practitioner | 363L00000X | Nurse Practitioner (NP) | 477 (4%) | 535 (5%) |
| | 363LA2200X | NP – Adult Health | 83 (1%) | 111 (1%) |
| | 363LF0000X | NP – Family | 1094 (10%) | 1055 (9%) |
| | 363LP0200X | NP – Pediatrics | 2 (0%) | 1 (0%) |
| | 363LP2300X | NP – Primary Care | 17 (0%) | 19 (0%) |
| Physician Assistant | 363A00000X | Physician Assistant (PA) | 589 (5%) | 760 (7%) |
| | 363AM0700X | PA – Medical | 214 (2%) | 230 (2%) |

*Provider-specific taxonomy codes are obtained for each clinician from the National Plan and Provider Enumeration System (NPPES). All clinicians must also have corresponding credentials: MD, DO, PA, PAC, or NP.*

**Table S3. List of control variables**

| Variable | Type | Description |
|---|---|---|
| **Initial opioid prescription ($O_i$)** | | |
| Days supplied | Categorical (4) | Estimated days of opioids supplied to the patient; field pulled directly from insurance pharmaceutical claims |
| Drug base | Categorical (5) | Classification of opioids into bases (hydrocodone, oxycodone, etc.) using the Generic Product Identifier, GPI (Cooper 2012) |
| Daily strength | Categorical (4) | Total strength of the prescription (calculated based on the drug base, strength, unit of measure, quantity, and the CMS milligram of morphine equivalent, MME, conversion chart) divided by the prescription's days supplied (Centers for Medicare and Medicaid Services 2018) |
| Prescriber specialty | Categorical (4) | Primary care specialty category of the initial prescriber (family medicine, internal medicine, nurse practitioner, or physician assistant) identified using the NPPES database |
| **Patient characteristics ($P_i$)** | | |
| Gender | Binary | Patient's sex, male or female |
| Age | Categorical (6) | Patient's age at the start of the episode |
| Insurance type | Categorical (3) | Primary payer of the patient's medical claims (commercial, Medicare or Medicaid) |
| Disability | Binary | Flags patients covered by Medicare who are under 65 years |
| **Condition controls ($C_i$)** | | |

| | | |
|---|---|---|
| Asthma | Binary | Patient comorbidities flagged based on set of emergency/medicine procedure codes, diagnoses, medications, and encounters; logic uses the patient's full medical history prior to the start of the opioid episode (minimum 6 months of complete medical and pharmaceutical claims, per clean period requirements) |
| Coronary artery disease | Binary | |
| Congestive heart failure | Binary | |
| Coronary obstructive pulmonary disease | Binary | |
| Depression | Binary | |
| Diabetes | Binary | |
| Hypertension | Binary | |
| Diagnosis chapter | Categorical (14) | ICD diagnosis chapter common to the initial and follow-up appointments |
| **Context controls ($T_i$)** | | |
| Geographical census region | Categorical (5) | Based on census tract identifier for the patient home address at the start of the episode; linked to the state-region crosswalk from the US Census Bureau (U.S. Census Bureau 2019) |
| Rural-urban continuum | Categorical (10) | Based on the Federal Information Processing Standards (FIPS) county code for the patient home address; linked to the USDA's 2013 9-level metro, urban, rural continuum (Cromartie 2013) |
| State annual opioid prescription rate | Continuous | Based on state of the patient home address and the year at the start of the episode; linked to the CDC's opioids prescribed per 100 residents for the state each year (Centers for Disease Control and Prevention 2018) |
| Initial prescriber's long-term opioid use rate | Continuous | Percent of other patients (other than the focal patient in the episode) who initiated opioids with the same initial prescriber and were considered long-term opioid users 12-months after opioid initiation |
| Insufficient prescriber initiations | Binary | Indicates when there were insufficient other opioid initiations (<5) in the dataset over which to calculate a long-term opioid use rate for the initial prescriber |
| Months into study | Numeric | Designates number of months between February 2012 and the start of the episode |
| **Instrumental variable controls ($V_i$)** | | |
| Insufficient appointment history | Binary | Flags episodes where fewer than 10 follow-up appointment sets are observed for the prescribing clinician in the 12 months prior to episode start |
| Clean period appointment count | Categorical (5) | Number of appointments that the patient had related to the episode condition in the primary care setting with any clinician during the 6 month clean period prior to the episode start |

**Methods S2. Notes on variable calculations**

**i. Independent variable: provider discordance**

To increase the likelihood that the follow-up appointment is related to the initial prescribing appointment, we required that there to be a match within at least one of the first 3 ICD diagnosis chapters[4] between the two appointments. If the appointments have multiple distinct diagnosis chapters in common, we designated the chapter based on the primary order within the medical claims.

**ii. Dependent variable: long-term opioid use**

The long-term opioid use 'days supplied' measure only includes days supplied of opioids (as reported in the pharmaceutical claims) between day 30 and 360 after opioid initiation (Barnett et al. 2017). If multiple opioid prescriptions were filled on the same date, the maximum of the days supplied was included. This addressed both the concern of duplicate pharmaceutical entries, and does not double-count concurrent prescriptions (Shah et al. 2017). If a patient started an opioid prescription near the end of the observation period (360 days from opioid initiation), only the portion of the days supplied that occur prior to the end of the evaluation period are included.

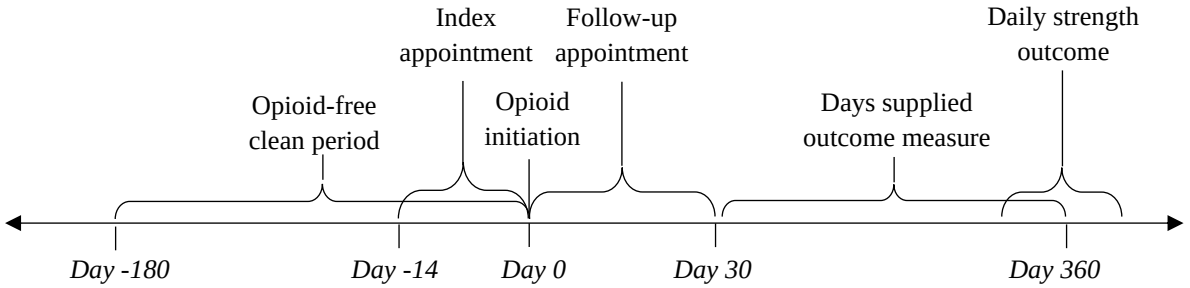
To calculate the 'daily strength' measure of long-term opioid use, each prescription's total morphine milligram equivalent (MME) was distributed by the number of days between prescription fills (or by days supplied on the longest prescription if the time gap between prescriptions was greater than 90 days). Following CMS protocol, we used the drug strength of the opioid component of the drug, unit of measure, quantity, and the morphine milligram conversion to calculate the total MME for each prescription (Centers for Medicare and Medicaid Services 2018). Then, to reduce noise stemming from prescription filled dates, we averaged the daily strengths of prescriptions filled in the time period [345, 375] days to estimate the daily dosage for the patient 360 days after opioid initiation.

In the case of methadone, the equivalent drug strength in MME does not scale linearly (the conversion factor is 4 for 1-20 mg/day, 8 for 21-20 mg/day, 10 for 41-60 mg/day and 12 for 60+ mg/day). We divided the total methadone drug strength by the days supplied in the prescription claim to obtain an estimated daily dosage and then multiplied the total prescription methadone strength by the appropriate MME conversion factor. If multiple opioid prescriptions are filled on the same day (possibly of different drug bases), we combined their dosages accordingly.

For both dependent variable measures, prescriptions administered through the following modes were excluded: intravenous, subcutaneous, intramuscular, intra-articular, and injection. This ensured that the opioids filled and administered by the patient or at home (i.e., not in an inpatient setting). A timeline summary of the independent and dependent variable calculations is provided in **Figure S1**.

[4] In a sensitivity analysis, we subset this sample to only include the 5,084 (44.8%) of episodes where the initial prescribing and follow-up appointments matched on their primary ICD diagnosis chapter. Despite having lower power in the analysis, the effects were nearly identical to those including the full sample: estimating on the fully controlled logistic model, the adjusted odds ratio of provider discordance on long-term opioid use was 0.58 [95% CI: 0.46-0.74] and 0.64 [95% CI: 0.47-0.86] using the days supplied measure and daily strength measure, respectively.

**Figure S1. Episode timeline and calculations for new opioid initiates**




*The dependent variable is calculated over days 31-360 after opioid initiation for the days supplied measurement and days 345-375 after opioid initiation for the daily strength measurement. All patients were required to have continuous eligibility in the dataset between 360 days prior to opioid initiation until 375 days after opioid initiation.*

### iii. Initial prescription controls

Characteristics of the initial prescription (specifically the opioid drug base, average daily opioid strength, and prescription days supplied) may be related to the patient's level of acuity or severity, which may influence the patient's likelihood of experiencing provider discordance in the early stages of opioid-use, as well as their likelihood of continuing opioid usage (Barnett et al. 2017, Deyo et al. 2017, Shah et al. 2018). We divided the total strength (in MME) of the index opioid prescription by the days supplied reported in the pharmaceutical claims to obtain an estimate of the daily strength of opioids prescribed at the beginning of the patient episode.

We followed the same approach as for the daily strength dependent variable measurement to convert methadone to MME. If more than one opioid was filled as part of the initial opioid prescription, the total MME's were added together to obtain a total opioid strength of the first prescription. In these cases, the total strength was divided by the longest of the days supplied of the prescriptions to obtain an estimated daily dose; the index days supplied corresponded to the longest of the days supplied and the drug base of the opioid with the largest total drug strength was designated as the index drug base. The quantities were categorized by common thresholds: 1-19, 20-49, 50-89 and 90+ MME/day for daily strength and 1-3, 4-7, 8-14, 15+ days supplied.

The initial prescription was linked to the prescribing appointment by matching the NPI of the prescription to the NPI's of the available medical claims. If the prescription was linked to more than one medical claim with the same NPI, we selected the medical claim with the closest date, no more than 14 days prior to the prescription (Hoopes et al. 2018). Over 90% of prescriptions were filled within 3 days of the designated prescribing appointment. Initial prescriber specialties were classified using the provider taxonomy codes shown in **Table S2**.

### iv. Patient characteristics controls

Age, sex, payer type, and location are defined based on the information current as of the day of the patient's initial opioid-prescribing appointment.

### v. Patient condition controls

Patient medical comorbidities were flagged using all available medical and pharmaceutical data (at least 18 months for each patient in the sample, and up to 6 years for those with continuous eligibility from the beginning of the dataset). Although it is possible that comorbidities are not observed for some patients within that timeframe, we expect that the active comorbidities (that may be affecting the patient's likelihood of continuing opioid prescriptions during the observation period) will be observed during the minimum 18-month time window.

The comorbidity definitions are based on emergency/medicine procedure codes, diagnoses, medications, and encounters and defined internally within the claims database, i.e., provided to, but not coded by the researchers. As these definitions may vary slightly from standard definitions (such as those Chronic Conditions Warehouse database), we ran additional sensitivity checks (see **Table 1** in the manuscript) with and without the comorbidity flags and confirmed that the comorbidity definitions are not driving our results.

### vi. Context controls

Patient-level location information was obtained based on the home or mailing address submitted to the insurance company. When available in the claims, the patient's state was cross-referenced with the state-region crosswalk from the US Census Bureau to assign each patient with a geographical census region (2019). The Federal Information Processing Standards (FIPS) county code associated with the patient's address was used to classify the patient's geography as one of nine categories in the rural-urban continuum (Cromartie 2013). The annual state opioid prescription rate (average opioids prescribed per 100 residents) was obtained from the Centers for Disease Control and Prevention (2018) and assigned to each patient episode, joining on the episode start year and the patient's location. As there is high level of regional variation in opioid prescribing habits, this control took on values between 25 and 125. Within each control, patients missing location information are flagged.

The initial prescriber long-term opioid rate variable captured the percent of other patients in our data who initiated opioids with the same initial prescriber as the focal patient and who subsequently became long-term opioid users. This covariate addresses the concern that certain patients (such as patients motivated to find a clinician who will prescribe opioids) are more likely to see certain clinicians for this first appointment, while simultaneously having a higher underlying likelihood of becoming a long-term opioid user.[5] An additional binary variable indicated when there was insufficient data (fewer than 5 appointments) to calculate a reliable estimate for the initial prescriber. In these cases of insufficient data, the control was set equal to the average long-term opioid use control across the other episodes in the dataset.[6]

[5] Specifically, suppose there are two types of provider, A and B. Assume that patients who see a type A provider are more likely to become long-term opioid users (e.g., this provider sees mainly chronic patients as their main PCP), while those who see a type B provider are less likely to become long-term opioid users (e.g., more acute patients visiting a covering PCP). The proposed control adjusts for the baseline risk profile of the initial prescriber, so that regardless of whether the patient saw a type A or type B provider for their first appointment, they have the same expected likelihood of becoming a long-term opioid user.

[6] The results in the table are identical when we set the control equal to 0 for the episodes with insufficient data to calculate a rate and including the binary flag to indicate an imputed value. To increase the sample size, the average long-term opioid use rate associated with the initial prescriber was calculated across all opioid initiates, not just the subset who returned to the primary care setting within 30 days. In total, there are 148,233 observations across the set of 4,894 unique initial prescribers.

### vii. Instrumental variables

The prescribing clinician's follow-up rate ($IV_1$) was calculated for each episode by first identifying all primary care appointments by the initial prescribing clinician in the 12 months prior to the start of the episode. We then subset the sample to appointments where a follow-up visit for a related condition (based on the first three ICD chapters) could be identified within 30 days in the primary care office setting. This was similar to how follow-up appointments were identified in the main analysis. The prescribing clinician's switching rate was defined the proportion of appointments where their past patients saw a different clinician for their follow-up visit. Note that this calculation includes *all* appointments conducted by the prescribing clinician, not just those related to opioid use.

As a robustness check (see **Methods S5.vi**), we included an additional variable $IV_2$ to improve predictive power in the selection equation. Unlike $IV_1$, which is calculated at the level of the initial prescriber, this second IV ($IV_2$) is calculated at the episode-level. For $IV_2$, we evaluated whether the patient had at least one appointment during the 6-month opioid-free clean period with their initial prescribing clinician, identifying past appointments based on the clinician's National Provider Identifier (NPI). To address the concern that patients who saw their prescribing clinician in the prior 6 months may be of higher acuity (compared, especially, to patients who did not see any primary care doctor in the past 6 months), we included a control for the count of primary care appointments during the clean period wherever $IV_2$ was used in the equation. To deal with outliers (as a small subsample of patients had frequent appointments with their primary care doctor, as often as every other week) the clean period appointment count was classified into categories: 0, 1, 2, 3, and 4+.

### viii. Regular provider

Each patient's regular primary care provider (PCP) was identified using up to two years of continuous medical history prior to the start of the episode. The regular PCP was identified as the provider who saw the patient most frequently over those two years; in the case of a tie, the regular PCP was assigned based on which provider saw the patient most recently.

See **Figure S2** for a distribution of the historical appointments counts by patient episode used to identify the regular PCP. **Figure S3** shows the proportion of appointments that were affiliated with the regular PCP versus other clinicians for each appointment set. An average of 75.7% of historic primary care appointments were held with the regular PCP, giving us high confidence in proper identification of the regular PCP.

Although this meant there were approximately 30 observations per prescriber, on average, 37% of initial prescribers had fewer than 5 appointments.

**Figure S2. Distribution of historical appointment counts used to identify regular primary care provider**

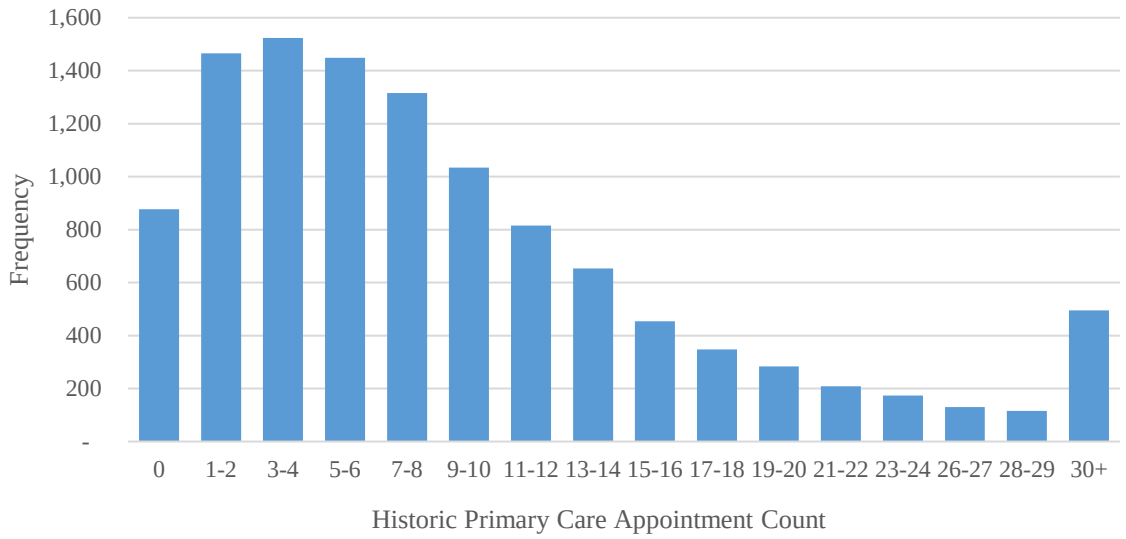
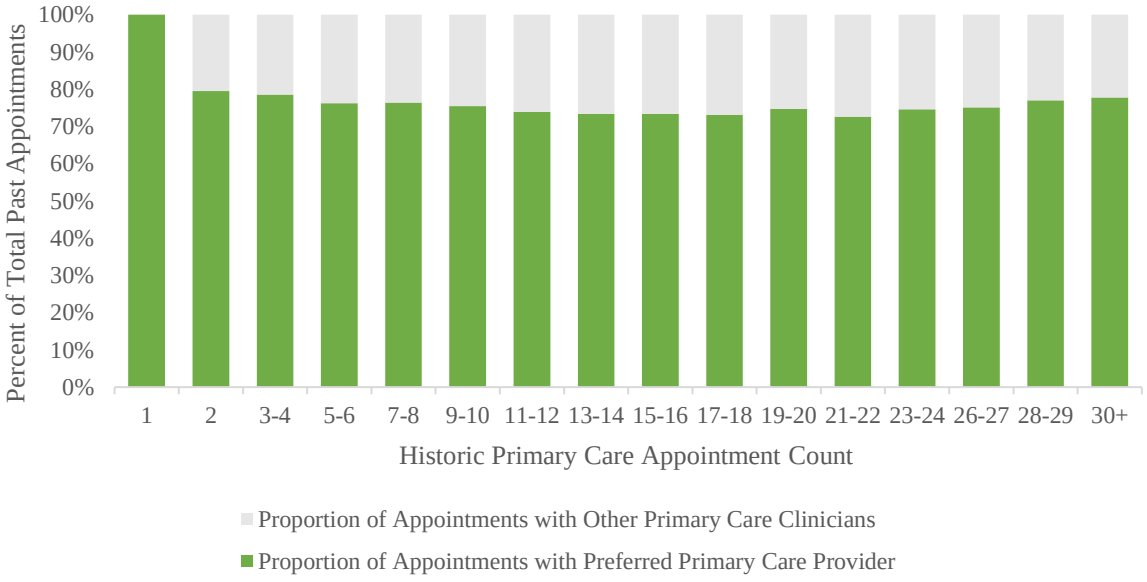



**Figure S3. Proportion of historic appointments with regular primary care provider**




Using this logic, regular we were able to identify regular PCPs for 10,465 of the 11,340 patient episodes. Approximately 61% (6,931) of the patients in the sample saw their regular PCP for the first appointment and 59% (6,742) saw their regular PCP for the follow-up appointment. Of the 8,129 episodes where patients experienced provider concordance, 5,808 (71.4%) of episodes were observed to have both appointments with the regular PCP versus 2,321 (28.6%) where the patient saw a PCP not identified as their regular provider for both appointments.

We then split the independent variable (provider discordance) into three categories (pathways) to test whether there was a difference in the effect of provider discordance between the sample of patients who:

- Pathway 1: Started with their regular PCP and switched to an alternative provider
- Pathway 2: Started with an alternative provider and then switched back to their regular PCP
- Pathway 3: Started with an alternative provider and then switched to another provider who was also not their regular PCP

**Table S4** shows the episode counts for patients who follow each pathway.

**Table S4. Pathway variations for patients experiencing provider discordance**

| | | Follow-up Appointment | |
|---|---|---|---|
| | | Regular PCP | Other PCP |
| First Appointment | Regular PCP | N/A | 1,123 (Pathway 1) |
| | Other PCP | 934 (Pathway 2) | 1,154* (Pathway 3) |

**A preferred primary care provider was not identified for 221 of these episodes.*

**Methods S3. Overview of Patient Opioid Episodes**

Of the 11,340 opioid episodes initiated by opioid-naïve patients, 8,129 (71.7%) of the patients saw the same initial prescribing clinician for their primary care follow-up appointment. The 3,211 patients who changed clinician for the follow-up appointment were given opioid prescriptions which averaged shorter in duration (median 7 days supplied [interquartile range (IQR): 5, 15] versus 10 days supplied [IQR: 7, 23]), but with similar daily dosages (median 20 MME/day [IQR: 15, 30] for both patient sets). Across the full sample, 2,887 (25.5%) of the patients filled an initial opioid prescription with a strength of at least 20 MME/day with over 7 days supplied. Patients who changed clinicians were slightly younger on average (median 62 years [IQR: 47, 74] versus 63 years [IQR: 50, 74]) and less likely to have coronary obstructive pulmonary disease, diabetes, or hypertension, making it particularly important to control for patient and condition characteristics. The condition for which opioids were prescribed were highly variable, with injuries and musculoskeletal diagnoses accounting for fewer than half of the opioid episodes (4,520 of 11,340). Complete summary statistics can be found in **Table S5**.

Of the 3,211 patients who changed clinicians for the follow-up appointment, 199 (6.2%) were considered long-term opioid users using the days supplied definition (filled 180 or more days of opioid prescriptions between days 30 and 360 after opioid initiation), compared to 935 (11.5%) of the 8,129 patients who saw the same initial prescribing clinician. Similarly, 108 (3.4%) of patients who saw an alternate clinician for the follow-up appointment continued to receive a daily dosage of 20 or more MME/day 12 months after the initial prescription, compared to 557 (6.9%) of patients who saw the same initial prescribing clinician.

**Table S5. Summary of patient episode characteristics**

| | Overall n = 11,340 | Provider Concordance n = 8,129 | Provider Discordance n = 3,211 | p-value |
|---|---|---|---|---|
| **Initial Daily Prescription Strength (% of total)** | | | | |
| 1-19 MME | 42.9% | 44.8% | 38.1% | <0.001*** |
| 20-49 MME | 50.0% | 48.3% | 54.4% | <0.001*** |
| 50-89 MME | 5.4% | 5.2% | 6.1% | 0.04* |
| 90+ MME | 1.7% | 1.8% | 1.4% | 0.13 |
| **Initial Prescription Days Supplied (% of total)** | | | | |
| 1-3 | 9.1% | 6.4% | 15.8% | <0.001*** |
| 4-7 | 31.7% | 29.3% | 37.7% | <0.001*** |
| 8-14 | 22.9% | 23.8% | 20.6% | <0.001*** |
| 15+ | 36.3% | 40.5% | 25.9% | <0.001*** |
| **Initial Prescription Drug Base (% of total)** | | | | |
| Hydrocodone | 45.1% | 44.3% | 47.2% | 0.006** |
| Tramadol | 36.8% | 37.4% | 35.3% | 0.04* |
| Codeine | 8.0% | 7.8% | 8.8% | 0.08^ |
| Oxycodone | 9.3% | 9.8% | 8.1% | 0.006** |
| Other | 0.7% | 0.7% | 0.7% | 0.63 |
| **Initial Prescriber Specialty (% of total)** | | | | |
| Family Medicine | 47.2% | 48.2% | 44.7% | <0.001*** |
| Internal Medicine | 30.9% | 35.2% | 20.1% | <0.001*** |
| Nurse Practitioner | 14.8% | 12.4% | 20.7% | <0.001*** |
| Physician Assistant | 7.1% | 4.2% | 14.5% | <0.001*** |
| **Gender (% prevalence)** | | | | |
| Female | 63.5% | 63.0% | 64.8% | 0.06^ |
| **Age (% of total)** | | | | |
| 18-34 | 8.0% | 7.2% | 10.2% | <0.001*** |
| 35-44 | 10.7% | 10.4% | 11.6% | 0.07^ |
| 45-54 | 15.8% | 16.1% | 15.1% | 0.21 |
| 55-64 | 18.0% | 18.2% | 17.7% | 0.51 |
| 65-74 | 23.5% | 24.1% | 22.1% | 0.02* |
| 75+ | 23.9% | 24.1% | 23.4% | 0.44 |
| **Insurance (% of total)** | | | | |
| Commercial | 30.5% | 27.2% | 38.7% | <0.001*** |
| Medicaid | 16.7% | 18.8% | 11.2% | <0.001*** |
| Medicare | 52.8% | 53.9% | 50.0% | <0.001*** |
| **Disability (% prevalence)** | | | | |
| Under 65 years and on Medicare | 6.6% | 6.9% | 5.7% | 0.002** |
| **Comorbidities (% prevalence)** | | | | |
| Asthma | 10.0% | 10.3% | 9.1% | 0.04* |
| Coronary Artery Disease | 16.4% | 16.5% | 15.9% | 0.42 |
| Congestive Heart Failure | 4.4% | 4.5% | 4.0% | 0.23 |
| Coronary Obstructive Pulmonary Disease | 12.1% | 13.1% | 9.7% | <0.001*** |
| Depression | 9.7% | 9.6% | 9.9% | 0.70 |
| Diabetes | 21.3% | 22.2% | 18.9% | <0.001*** |
| Hypertension | 28.6% | 29.3% | 26.6% | 0.004** |
| **Primary ICD Diagnosis Chapter (% of total)** | | | | |
| XIII. Musculoskeletal | 39.9% | 37.1% | 46.9% | <0.001*** |
| IX. Circulatory | 11.7% | 13.7% | 6.9% | <0.001*** |

| | | | | |
|---|---|---|---|---|
| XVIII. Abnormalities not classified | 10.7% | 10.8% | 10.6% | 0.84 |
| IV. Endocrine nutrition, metabolic | 8.3% | 9.8% | 4.6% | <0.001*** |
| XXI. Factors influencing health | 5.6% | 6.1% | 4.6% | 0.003** |
| XIX. Injury, poisoning | 5.3% | 4.1% | 8.3% | <0.001*** |
| X. Respiratory system | 3.7% | 3.9% | 3.2% | 0.08^ |
| XII. Skin and tissue | 2.9% | 2.7% | 3.4% | 0.06^ |
| VI. Nervous system | 2.6% | 2.8% | 2.3% | 0.20 |
| I. Infectious and parasitic diseases | 2.4% | 2.4% | 2.6% | 0.51 |
| XIV. Genitourinary | 2.3% | 2.2% | 2.5% | 0.38 |
| XI. Digestive system | 1.8% | 1.7% | 2.1% | 0.09^ |
| Other | 2.7% | 2.9% | 2.0% | 0.004** |
| **Geographical Census Region (% of total)** | | | | |
| South | 46.0% | 47.9% | 41.4% | <0.001*** |
| Midwest | 17.8% | 18.5% | 16.1% | 0.003** |
| West | 17.2% | 13.9% | 25.6% | <0.001*** |
| Northeast | 3.0% | 3.5% | 1.7% | <0.001*** |
| Unknown | 16.0% | 16.3% | 15.2% | 0.13 |
| **Rural-Urban Continuum (% of total)** | | | | |
| 1 - Large metropolitan area | 33.6% | 33.6% | 33.4% | 0.82 |
| 2 - Medium metropolitan area | 18.5% | 17.2% | 21.8% | <0.001*** |
| 3 - Small metropolitan area | 9.8% | 9.9% | 9.6% | 0.67 |
| 4 - Large urban, near metropolitan area | 2.5% | 2.6% | 2.1% | 0.14 |
| 5 - Large urban, not near metropolitan area | 5.2% | 5.0% | 5.7% | 0.14 |
| 6 - Small urban, near metropolitan area | 6.4% | 6.9% | 5.1% | <0.001*** |
| 7 - Small urban, not near metropolitan area | 5.4% | 5.7% | 4.6% | 0.02* |
| 8 - Rural, near metropolitan area | 1.7% | 1.8% | 1.6% | 0.38 |
| 9 - Rural, not near metropolitan area | 1.5% | 1.5% | 1.6% | 0.82 |
| Not known | 15.4% | 15.7% | 14.5% | 0.12 |
| **Initial Prescriber Long-term Opioid Use Rate** | | | | |
| Rate using days supplied measurement | 8.7% | 9.1% | 7.7% | <0.001*** |
| Rate using daily strength measurement | 3.6% | 3.8% | 3.3% | <0.001*** |
| **Insufficient Observations of Prescriber's Opioid Initiations** | | | | |
| Yes (<5 initiations) | 37.3% | 35.3% | 42.4% | <0.001*** |
| **State Opioid Prescription Rate** | | | | |
| Opioids prescribed annually per 100 residents | 80.0 | 80.0 | 79.8 | 0.56 |
| **Months into Study (since Feb 2012)** | | | | |
| Months to episode start | 50.2 | 50.1 | 50.4 | 0.40 |
| **Long-term Opioid Use Outcome Measure (% prevalence)** | | | | |
| ≥180 days supplied of opioids at 12 months | 10.0% | 11.5% | 6.2% | <0.001*** |
| Daily MME ≥20 at 12 months | 5.9% | 6.9% | 3.4% | <0.001*** |

^ *<0.10, *<0.05, **<0.01, ***<0.001*

## SUPPLEMENT PART II. LOGISTIC MODEL

### Methods S4. Overview of effects on long-term opioid use

The marginal effects of each control variable on the long-term opioid use outcome are estimated in the main model:

$$y_i^* = \alpha + \beta x_i + \gamma O_i' + \delta P_i' + \lambda C_i' + \theta T_i' + \varepsilon_i, \qquad y_i = \mathbb{I}[y_i^* > 0], \qquad \textit{[S1]}$$

where $y_i^*$ is an unobserved latent variable, $x_i$ is a binary variable that takes value 1 when the patient changes clinician for the follow-up appointment and zero otherwise, the vectors $O_i$, $P_i$, $C_i$, and $T_i$ contains the set of all initial prescription, patient characteristics, condition, and context covariates, respectively, and $y_i$ is the observed dichotomous variable that indicates whether the patient became a long-term opioid user, with $\mathbb{I}[\cdot]$ the indicator function. The main effect (provider discordance) is specified by $\beta$. The adjusted odds ratios of the effects of $\beta$ and all other control variables have on long-term opioid use are displayed in **Table S6**.

As expected, higher daily dosages and longer days supplied were significant risk factors of long-term opioid use (Barnett et al. 2017, Mundkur et al. 2017, Shah et al. 2018). Tramadol and codeine, drug bases commonly used for low-severity pain, are associated with lower rates of long-term opioid use compared to hydrocodone or oxycodone (Babalonis et al. 2013). Compared to the commercially insured population, Medicaid and Medicare patients are at higher risk of long-term opioid use. This observed effect in the Medicare population may be a function of the model set-up, however, as it is at least partially offset by how patients above 65 (who are qualified for Medicare) are observed to have lower risks of long-term opioid use. There were little observed differences between the types of geographies, perhaps due to the relatively small sample size.

Patients with a comorbidity of depression and coronary obstructive pulmonary disease have higher rates of continuing opioid use compared to patients without the diagnosis; however, patients having a comorbidity of asthma have relatively lower rates of continuing opioid use. Most patients are first prescribed opioids for musculoskeletal-related diagnoses (e.g., back pain); these diagnoses are also associated with one of the highest conversions of opioid initiates to long-term users. The only episodes with a higher conversion rate are those categorized with a nervous system-related diagnosis.

**Table S6. Complete logistic model adjusted odds ratios output**

| | Adjusted Odds Ratio (95% Confidence Interval) on Long-Term Opioid Use | |
|---|---|---|
| **Long-term Opioid Use Measure[a]** | **Days Supplied** | **Daily Strength** |
| **Provider discordance, Reference Category (RC)[b] = Follow-up with original prescribing clinician** | | |
| Follow-up with alternate clinician | 0.69 (0.57-0.82)*** | 0.63 (0.50-0.80)*** |
| **Initial Opioid Prescription Days Supplied, RC = 1-3 Days** | | |
| 4-7 Days | 1.29 (0.82-2.01) | 2.21 (1.23-3.99)** |
| 8-14 Days | 2.87 (1.85-4.46)*** | 4.11 (2.26-7.45)*** |
| 15+ Days | 11.35 (7.47-17.24)*** | 15.76 (8.94-27.78)*** |
| **Initial Opioid Prescription Drug Base, RC = Hydrocodone** | | |
| Tramadol | 0.55 (0.47-0.65)*** | 0.39 (0.31-0.49)*** |
| Codeine | 0.61 (0.45-0.85)** | 0.31 (0.18-0.55)*** |
| Oxycodone | 1.12 (0.87-1.43) | 1.01 (0.76-1.33) |
| Other | 1.78 (1.00-3.17)^ | 1.91 (1.03-3.54)* |
| **Initial Opioid Prescription Daily Strength in Morphine Milligram Equivalents, RC = 1-19 MME/day** | | |
| 20-49 MME/day | 1.29 (1.10-1.50)** | 2.49 (2.03-3.06)*** |
| 50-89 MME/day | 1.91 (1.36-2.66)*** | 4.64 (3.21-6.70)*** |
| 90+ MME/day | 3.21 (2.07-4.99)*** | 7.32 (4.56-11.76)*** |
| **Initial Prescriber Specialty, RC = Family Medicine** | | |
| Internal Medicine | 0.88 (0.75-1.04) | 0.75 (0.61-0.93)** |
| Nurse Practitioner | 1.04 (0.85-1.28) | 0.85 (0.65-1.10) |
| Physician Assistant | 1.09 (0.76-1.56) | 0.94 (0.60-1.47) |
| **Sex, RC = Female** | | |
| Male | 1.24 (1.07-1.43)** | 1.33 (1.11-1.59)** |
| **Age, RC = 18-34** | | |
| 35-44 | 1.09 (0.79-1.52) | 0.93 (0.62-1.39) |
| 45-54 | 1.18 (0.87-1.60) | 1.11 (0.76-1.61) |
| 55-64 | 0.89 (0.65-1.22) | 0.77 (0.52-1.14) |
| 65-74 | 0.44 (0.23-0.83)* | 0.62 (0.30-1.28) |
| 75+ | 0.30 (0.15-0.58)*** | 0.24 (0.11-0.52)*** |
| **Payer Type, RC = Commercial Insurance** | | |
| Medicare | 2.40 (1.30-4.42)** | 1.68 (0.84-3.38) |
| Medicaid | 2.74 (2.10-3.59)*** | 2.96 (2.10-4.17)*** |
| **Disability Status, RC = No** | | |
| Yes | 1.13 (0.61-2.10) | 1.67 (0.82-3.38) |
| **Comorbidities, RC = No** | | |
| Asthma | 0.94 (0.74-1.19) | 1.02 (0.76-1.38) |
| Coronary Artery Disease | 0.97 (0.78-1.21) | 1.13 (0.86-1.50) |
| Congestive Heart Failure | 0.95 (0.65-1.41) | 1.01 (0.62-1.64) |
| Coronary Obstructive Pulmonary Disease | 1.60 (1.30-1.97)*** | 1.63 (1.25-2.11)*** |
| Depression | 1.49 (1.21-1.85)*** | 1.41 (1.08-1.85)* |
| Diabetes | 0.81 (0.67-0.99)* | 0.77 (0.60-0.99)* |
| Hypertension | 0.94 (0.78-1.13) | 0.91 (0.72-1.15) |
| **ICD Diagnosis Chapter, RC = XIII. Musculoskeletal** | | |
| IX. Circulatory | 0.73 (0.58-0.91)** | 0.58 (0.43-0.79)*** |
| XVIII. Abnormalities | 0.76 (0.59-0.98)* | 0.63 (0.45-0.89)** |
| IV. Endocrine, nutrition | 0.71 (0.55-0.91)** | 0.73 (0.53-1.00)* |
| XXI. Health factors | 0.77 (0.58-1.03)^ | 0.63 (0.43-0.92)* |
| XIX. Injury, poisoning | 0.31 (0.18-0.55)*** | 0.29 (0.14-0.58)*** |
| X. Respiratory | 0.70 (0.47-1.04)^ | 0.47 (0.27-0.83)** |

| | | |
|---|---|---|
| XII. Skin and tissue | 0.07 (0.02-0.28)*** | N/A [c] |
| VI. Nervous system | 1.43 (1.01-2.03)* | 1.65 (1.11-2.47)* |
| I. Infectious diseases | 0.28 (0.13-0.63)** | 0.24 (0.08-0.67)** |
| XIV. Genitourinary | 0.47 (0.23-0.94)* | 0.34 (0.12-0.94)* |
| XI. Digestive system | 0.56 (0.31-0.99)* | 0.58 (0.29-1.18) |
| Other | 0.92 (0.63-1.34) | 0.82 (0.50-1.33) |
| **Geographical Census Region, RC = South** | | |
| Midwest | 0.81 (0.63-1.06) | 0.80 (0.56-1.12) |
| Northeast | 0.86 (0.53-1.40) | 0.74 (0.41-1.35) |
| West | 1.32 (1.00-1.73)^ | 1.57 (1.12-2.21)** |
| Unknown | N/A [c] | N/A [c] |
| **Rural-Urban Continuum, RC = 1 – Large metropolitan area** | | |
| 2 - Medium metropolitan area | 0.98 (0.77-1.26) | 0.73 (0.53-1.00)^ |
| 3 - Small metropolitan area | 0.90 (0.69-1.18) | 0.65 (0.46-0.93)* |
| 4 - Large urban, near metropolitan area | 1.04 (0.65-1.68) | 1.15 (0.64-2.08) |
| 5 - Large urban, not near metropolitan area | 1.22 (0.87-1.71) | 0.85 (0.56-1.31) |
| 6 - Small urban, near metropolitan area | 1.29 (0.96-1.72)^ | 0.99 (0.68-1.44) |
| 7 - Small urban, not near metropolitan area | 1.22 (0.91-1.65) | 0.83 (0.57-1.23) |
| 8 - Rural, near metropolitan area | 1.23 (0.74-2.04) | 0.78 (0.40-1.53) |
| 9 - Rural, not near metropolitan area | 1.43 (0.87-2.35) | 0.99 (0.52-1.86) |
| 0 – Unknown | N/A [c] | N/A [c] |
| **Initial Prescriber Long-term Opioid-use Rate, Increment = 10%** | | |
| Additional 10% | 1.22 (1.14-1.31)*** | 1.16 (1.03-1.30)* |
| **Insufficient Observations of Prescriber's Opioid Initiations, RC = No** | | |
| Yes | 1.46 (1.25 -1.70)*** | 1.43 (1.18-1.73)*** |
| **Annual State Opioid Prescription Rate, Increment = Additional 10 prescriptions per 100 residents** | | |
| Additional 10 opioids prescribed | 1.11 (1.05-1.18)*** | 1.07 (0.99-1.16)^ |
| **Years into Study, Increment = 1 additional year into study** | | |
| Years since February 2012 | 0.97 (0.90-1.05) | 0.90 (0.82-1.00)* |

*^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

*[a] Adjusted odds ratios corresponding to categories with small episode subsets should be interpreted with caution.*

*[b] Reference category designates the episode characteristic to which all other categories should be compared.*

*[c] Insufficient patient episodes to compute an estimate.*

## SUPPLEMENT PART III. INSTRUMENTAL VARIABLE MODEL

### Methods S5. Validity of the instrumental variables

#### i. Overview of the instrumental variable approach and validity concerns

In the paper we report results using an instrumental variable (IV) model to address concerns of endogeneity. Using the IV approach, the goal is to arrive at an unbiased estimate of the average effect of the treatment if it were applied to the population. The IV approach thus accounts for omitted variables (e.g., acuity) as well as other factors that might cause patients to self-select into the treatment or control group (discordance or concordance) in a non-random way (e.g., awareness of opioid risks, demographics).

The recursive bivariate modelling approach that we use estimates a selection equation (whether the patient experiences provider discordance) and an outcome equation (whether the patient becomes a long-term opioid user) simultaneously. Any unobserved factors that affect both selection and outcome are controlled for in the outcome estimate via the estimated correlation between the error terms of the two equations.

To improve reliability of this method, it is common (but not required) to include one or more IVs. These are variables that are included in the selection equation but that are excluded from the outcome equation. For these instruments to be appropriate, they must be relevant and not weak (i.e., have a significant impact on whether the patient will be in the treatment or control group). In addition, they must satisfy the validity assumption which requires that the instruments must (i) have no direct impact on the outcome after controlling for the control variables, and (ii) must not be correlated with other omitted variables that might impact the outcome.

As described in **Methods S2.vi** above, the main IV ($IV_1$) is the rate at which the prescribing clinician's other patients saw a different clinician for their follow-up appointment. Similar to identifying follow-up appointments in the main analysis, we then subset to appointments where the patient's follow-up visit (for a related condition, based on the first three ICD chapters) could be identified within 30 days in the primary care office setting. The prescribing clinician's switching rate was defined the proportion of appointments where their past patients saw a different clinician for their follow-up visit. In models that include IV1, we also incorporated a binary control that identified the 18% of episodes where we did not observe sufficient appointment history (fewer than 10 follow-up appointment sets) for the prescribing clinician to accurately estimate the IV.

If the prescribing clinician's other patients were frequently changing clinicians for their follow-up appointment (for any reason, including scheduling availability at the practice), the focal patient of the episode is also expected to be more likely to experience provider discordance. This relevance criteria (i.e., that the IV is predictive of the patient's likelihood of experiencing provider discordance) is confirmed by looking at the large and statistically significant effect of the IV on provider discordance in the selection equation: the coefficient 0.23 (95% CI: [0.22-0.25]) in the selection equation corresponding to a 10% increase in the initial provider discordance rate. However, given how the IV is measured using patients other than the focal patient, we have no reason to expect that the switching rate of these other patients for the follow-up appointment should directly impact on the focal patient's likelihood of becoming a long-term opioid user (i.e., should be valid).

However, we acknowledge that it is possible that other factors are correlated with both the IV and the likelihood of long-term opioid use – for example, if patients are well-informed about quality of care and appropriate prescribing practices, they may be more motivated to seek out a different opinion after being prescribed opioids. The main concern that we must address is therefore whether there exists any factor that affects (i) who a patient saw for their first appointment, (ii) the value of the IV, and (iii) a patient's likelihood of becoming a long-term opioid user, as this may bias the results.

A typical example for such an IV-confounder is the patient's demographics: patients living in lower-income neighborhoods may have a limited choice of provider, and therefore $IV_1$ may take a lower value (as it is less feasible for patients to switch providers during the follow-up appointment). If living in a poor neighborhood is also a risk factor for opioid dependence, then this would invalidate the IV. Alternatively, some providers may specialize in treating pain. If patients who opt to see these providers differ in their likelihood to subsequently switch provider but also differ in their likelihood of becoming a long-term opioid user, then this could also invalidate the IV. Finally, acuity may also have an impact: more acute patients may be more likely to see a covering clinician, and patients who see covering clinicians are more likely to switch to alternative providers. If acuity also affects a patient's likelihood of using opioids for the long-term, this could again invalidate the IV. We address each of these concerns in the sections below.

**ii. Controlling for "unobserved" factors**

One approach looks to improve validity of the IV by directly controlling for factors that might be biasing the results. Since the main issue with the instrument arises due to how it is measured at the level of the initial prescriber, the most direct approach is to control for heterogeneity in long-term opioid use rates across initial prescribers. Within the manuscript, we controlled for the long-term opioid use rate of the initial prescriber: the % of other patients in our data who saw the same initial provider for their first appointment as the focal patient and who subsequently became long-term opioid users.

This control will partially, if not perfectly, account for heterogeneity in average acuity across patients based on who they saw for their first appointment (e.g., their main PCP or a less familiar, potentially covering, provider). When we include this variable as a control in our IV model, we in effect remove any link between the initial prescriber and the outcome, i.e., it is equivalent to assuming that each patient saw the same clinician for their first appointment with respect to their baseline risk of becoming a long-term opioid user. This is the same approach used in Freeman et al. (2020), where their instrument was measured at the level of an ED physician.[7] Adopting the same approach and following the same logic, any omitted variable that is correlated with the initial provider that may affect a patient's long-term opioid use rate (e.g., the specialization of the provider, demographics of patients who see that provider, etc.) will be captured through this additional control.

[7] In their case, the treatment being studied was a patient being admitted to the clinical decisions unit (CDU), while the outcome studied was whether or not a patient was admitted unnecessarily or discharged wrongfully. They use as an instrument the rate at which that ED physician admits patients to the CDU. However, they were concerned that clinicians who may admit more frequently to the CDU may be more inexperienced and hence more likely to make admission or discharge errors. To account for this, they controlled also for the rate at which that ED physician admitted patients unnecessarily or discharged patients wrongfully.

The estimated effect of provider discordance when controlling for the initial prescriber's rate of long-term opioid use across other patients did not change when the model was estimated without this initial prescriber covariate. This observation provides some evidence that our instrumental variable analysis is not being affected by unobserved variables correlated to our IV.

**iii. Including the IV in the outcome equation**

As an alternative validity test, if there were an important omitted variable that was correlated with the IV and omitted from the outcome equation, then the IV would serve as a proxy for that variable. In this case, we would expect the IV to be significant in the outcome equation when we estimate the models based on functional form alone. However, when we estimate the best-fit recursive bivariate model and include the IV ($IV_1$) in both the selection and the outcome equation (i.e., estimate based on functional form), we find no evidence that the instrument has any effect on the outcome (see results in **Table S7**).

**Table S7. Recursive bivariate model estimated with IV in outcome equation**

| | **Days Supplied Measurement** | | **Daily Strength Measurement** | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)**[+] | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Provider discordance | 0.39 (0.23-0.67) | <0.001*** | 0.41 (0.21-0.79) | 0.008** |
| $IV_1$: 10% increase in initial prescriber's switching rate | 1.03 (0.97-1.08) | 0.342 | 1.01 (0.94-1.08) | 0.885 |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*
[+] *Estimated using the best fit recursive bivariate model with probit-logit marginals and a Clayton copula.*

**iv. Robustness check: recursive bivariate model estimated with two IVs**

As introduced in **Methods S2.vii**, a second proposed IV ($IV_2$) can be included in the model to improve predictive power. $IV_2$ measures patient familiarity with the prescribing clinician, evaluated as whether the patient had at least one appointment with their initial prescribing clinician during the 6-month opioid-free clean period. Unlike $IV_1$ which was measured at the level of the initial prescriber, this IV was measured at the level of the patient. In the selection equation, $IV_2$ is also confirmed to be highly relevant (with coefficient of 0.47 [95% CI: (0.42-0.52), p<0.001***]) and added further predictive power to the model.

As shown in **Table S8**, we observe little difference in the estimated effect of provider discordance on long-term opioid use when using both IVs. For simplicity, and because including this second IV made little difference in the overall estimates, we reserved this model with two IVs as a robustness check and did not include it in the manuscript.

**Table S8. Robustness of instrumental variable estimates with additional IV**

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval in Treatment Equation)[+]** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| **Two-IV model (standard + IV controls)** | | | | |
| Provider discordance | 0.47 (0.31-0.70) | <0.001*** | 0.47 (0.31-0.70) | 0.002** |
| **Two-IV model including both IVs in the outcome equation (standard + IV controls)** | | | | |
| Provider discordance | 0.37 (0.22-0.64) | <0.001*** | 0.38 (0.20-0.72) | 0.003** |
| $IV_1$: 10% increase in initial prescriber's switching rate | 1.31 (0.76-2.24) | 0.330 | 1.01 (0.94-1.08) | 0.811 |
| $IV_2$: Initial prescriber seen before | 0.98 (0.95-1.01) | 0.227 | 0.80 (0.52-1.22) | 0.296 |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

[+]*Estimated using the best fit recursive bivariate model with probit-logit marginals and a Clayton copula.*

**v. Recursive bivariate model estimated without IVs**

Finally, we estimated the recursive bivariate model without any IVs. Similar to the test in Methods S5.iii, this model specification estimates the effect of provider discordance using functional form alone. If there were an issue with the validity of the IV that was causing the estimates to overstate the true effect, we would expect the estimated effects of provider discordance to be smaller in the model. However, as shown in **Table S9** we observe the opposite, giving further confidence in our IV models presented in the manuscript.

**Table S9. Recursive bivariate model estimates without IVs**

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)[+]** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Provider discordance | 0.31 (0.17-0.57) | <0.001*** | 0.34 (0.16-0.69) | 0.003** |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

[+]*Estimated using the best fit recursive bivariate model with probit-logit marginals and a Clayton copula.*

**vi. Summary**

Between the base set of controls included in the model that adjust for heterogeneity across episodes and the three sensitivity analyses run on the instrumental variables, we find support that our instrumental variables are reliable and that the results of the analyses are not meaningfully impacted by correlations between our IV and unobserved confounders. In addition, we test different model specifications (**Methods S6**) and carry out a propensity-score matching with bias-minimization methodology (**Methods S7**). Overall, therefore, we find consistent support that the main effect presented in the paper is likely a conservative estimate of the main impact that provider discordance has on long-term opioid use.

**Methods S6. Copula and joint error specification**

As introduced in the manuscript, although the recursive bivariate model is structured to address confounding, in practice, a recursive bivariate probit model with normal error dependencies is likely to have biased estimates of the coefficients when underlying assumptions are not perfectly satisfied (Clarke and Windmeijer 2012). We use the GJRM package in the statistical software R to test a variety of link functions (probit, logit, complementary log-log (cloglog)) and joint error distributions (modelled as copulas, including Clayton, Joe, and Gumbel) to allow for either bivariate normal or non-normal dependencies between the treatment and outcome equations (Marra et al. 2017, Winkelmann 2011).

If there is no confounding by unobserved variables or the confounding is small, then the simultaneous equations approach is less efficient than the single estimation of the outcome equation, leading to larger confidence intervals. As shown in **Table S10**, while the IV models yield large confidence intervals (some overlapping with 1), the estimated association between provider discordance and long-term opioid use are consistently negative across specifications. This increases the confidence in our main results.

Using the Akaike Information Criteria (AIC) as the comparison metric, the models specified with probit-logit marginal and a Clayton copula outperformed all other model specifications (including those with non-logit outcome marginal) for the days supplied measure of long-term opioid use. The same model structure also performed well for the daily strength measure (AIC = 15,298). The only model specifications with a lower AIC for the daily strength measure had a cloglog outcome marginals (slight decrease in AIC to 15,293); however, the Vuong and Clarke tests were either (a) unable to differentiate between the models with cloglog outcome marginals and the probit-logit Clayton model, or (b) the probit-logit Clayton model was observed to be a better fit. Especially because it is simpler to interpret coefficients using a logit marginal in the outcome equation, we therefore use the probit-logit Clayton model specification across both measures of long-term opioid use.

**Table S10** displays the sensitivity results from models estimated with the logit outcome marginal as these models yield estimated adjusted odds ratios. For each set of covariates, the average tau ($\tau$) value that indicates the direction of the confounding is positive and, in many cases, statistically significant at the 95% level. As such, these recursive bivariate models suggest that the basic logistic regression is underestimating the impact that provider discordance has on long-term opioid use. While the size of the effect and associated confidence intervals vary by specification, the models consistently report a negative and significant effect of provider discordance on long-term opioid use.

**Table S11** and **Table S12** display the complete list of coefficient estimations (presented in the form of adjusted odds ratios, for simplicity) of both the treatment and outcome equations for the best fit recursive bivariate models.

**Table S10. Alternate specifications of recursive bivariate instrumental variable model**

| Long-term Opioid Use Outcome Measure | Exposure Marginal[a] | Copula[b] | AIC[c] | Direction of Endogeneity, Tau (95% CI) | Adjusted Odds Ratio of Provider Discordance (95% CI) | p-value |
|---|---|---|---|---|---|---|
| Days Supplied | probit | C0 | 17,178 | 0.05 (0.02-0.14) | 0.46 (0.30-0.70) | <0.001*** |
| Days Supplied | probit | N | 17,181 | 0.07 (-0.01-0.17) | 0.48 (0.29-0.80) | 0.005** |
| Days Supplied | probit | G0 | 17,182 | 0.09 (0.01-0.36) | 0.51 (0.29-0.91) | 0.022* |
| Days Supplied | probit | HO | 17,182 | 0.09 (0.02-0.33) | 0.51 (0.29-0.91) | 0.022* |
| Days Supplied | probit | PL | 17,182 | 0.08 (-0.11-0.24) | 0.50 (0.25-1.00) | 0.050^ |
| Days Supplied | probit | F | 17,183 | 0.06 (-0.08-0.19) | 0.54 (0.29-0.99) | 0.048* |
| Days Supplied | probit | FGM | 17,183 | 0.04 (-0.06-0.12) | 0.58 (0.35-0.97) | 0.038* |
| Days Supplied | probit | C180 | 17,183 | 0.00 (0.00-0.98) | 0.69 (0.57-0.82) | <0.001*** |
| Days Supplied | probit | J0 | 17,183 | 0.00 (0.00-0.96) | 0.69 (0.57-0.82) | <0.001*** |
| Days Supplied | logit | C0 | 17,184 | 0.05 (0.02-0.13) | 0.45 (0.29-0.69) | <0.001*** |
| Days Supplied | logit | N | 17,187 | 0.08 (-0.02-0.17) | 0.47 (0.28-0.80) | 0.005** |
| Days Supplied | logit | HO | 17,188 | 0.09 (0.01-0.39) | 0.50 (0.28-0.90) | 0.020* |
| Days Supplied | logit | G0 | 17,188 | 0.09 (0.02-0.43) | 0.50 (0.28-0.90) | 0.020* |
| Days Supplied | logit | PL | 17,188 | 0.08 (-0.08-0.23) | 0.49 (0.24-1.00) | 0.049* |
| Days Supplied | logit | F | 17,189 | 0.06 (-0.08-0.17) | 0.53 (0.29-0.99) | 0.046* |
| Days Supplied | logit | FGM | 17,189 | 0.04 (-0.07-0.13) | 0.58 (0.35-0.97) | 0.037* |
| Days Supplied | logit | J0 | 17,189 | 0.00 (0.00-0.96) | 0.69 (0.57-0.82) | <0.001*** |
| Daily Strength | probit | C0 | 15,298 | 0.04 (0.01-0.13) | 0.42 (0.24-0.72) | 0.002** |
| Daily Strength | probit | N | 15,299 | 0.09 (-0.04-0.22) | 0.40 (0.21-0.78) | 0.007** |
| Daily Strength | probit | PL | 15,299 | 0.18 (-0.06-0.41) | 0.30 (0.11-0.81) | 0.018* |
| Daily Strength | probit | F | 15,299 | 0.14 (-0.12-0.35) | 0.36 (0.15-0.86) | 0.022* |
| Daily Strength | probit | HO | 15,299 | 0.13 (0.03-0.55) | 0.40 (0.18-0.87) | 0.021* |
| Daily Strength | probit | G0 | 15,299 | 0.13 (0.03-0.52) | 0.40 (0.18-0.87) | 0.021* |
| Daily Strength | probit | FGM | 15,300 | 0.09 (-0.08-0.18) | 0.45 (0.23-0.87) | 0.018* |
| Daily Strength | probit | J0 | 15,301 | 0.00 (0.00-0.96) | 0.63 (0.50-0.80) | <0.001*** |
| Daily Strength | logit | C0 | 15,304 | 0.05 (0.02-0.11) | 0.41 (0.24-0.72) | 0.002** |
| Daily Strength | logit | N | 15,305 | 0.09 (-0.04-0.23) | 0.39 (0.20-0.77) | 0.006** |
| Daily Strength | logit | PL | 15,305 | 0.20 (-0.09-0.48) | 0.28 (0.10-0.79) | 0.016* |
| Daily Strength | logit | F | 15,306 | 0.14 (-0.09-0.34) | 0.35 (0.14-0.85) | 0.021* |
| Daily Strength | logit | HO | 15,306 | 0.13 (0.03-0.39) | 0.39 (0.18-0.85) | 0.018* |
| Daily Strength | logit | G0 | 15,306 | 0.13 (0.03-0.47) | 0.39 (0.18-0.85) | 0.018* |
| Daily Strength | logit | FGM | 15,306 | 0.09 (-0.09-0.19) | 0.45 (0.23-0.87) | 0.017* |
| Daily Strength | logit | J0 | 15,307 | 0.00 (0.00-0.96) | 0.63 (0.50-0.80) | <0.001*** |

*^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

*[a] All estimations use logit outcome marginal for ease of interpretation.*

*[b] Joint terror distributions (copula) tested and displayed unless the model failed to converge: Normal (N), Clayton (C0), Clayton rotated 180[o] (C180), Joe (J0), and Gumbel (G0), Frank (F), Farlie-Gumbel-Morgentsern (FGM), and Plackett (PL).*

*[c] Lower values of the Akaike Information Criteria (AIC) suggest better fit*

**Table S11. Best fit instrumental variable model for days supplied outcome measure**

| | Coefficients (95% Confidence Interval)[a] | |
|---|---|---|
| **Recursive Bivariate Equation** | **Effect on Treatment (Provider Discordance)** | **Effect on Outcome (Long-term Opioid Use)** |
| **Provider Discordance, Reference Category (RC)[b] = Follow-up with original prescribing clinician,** | | |
| Follow-up with alternate clinician | - | 0.50 (0.33-0.75)*** |
| **$IV_1$ = Prescribing Clinician Follow-up Rate, Increment = 10%** | | |
| Increase in follow-up rate | 0.23 (0.22-0.25)*** | - |
| **$IV_1$ Control: Insufficient Appointment History for Prescribing Clinician, RC = No** | | |
| Yes | 0.32 (0.24-0.39)*** | 0.05 (-0.15-0.26) |
| **Initial Opioid Prescription Days Supplied, RC = 1-3 Days** | | |
| 4-7 Days | -0.28 (-0.38--0.18)*** | 0.21 (-0.23-0.65) |
| 8-14 Days | -0.42 (-0.53--0.31)*** | 0.97 (0.53-1.41)*** |
| 15+ Days | -0.47 (-0.58--0.36)*** | 2.33 (1.90-2.75)*** |
| **Initial Opioid Prescription Drug Base, RC = Hydrocodone** | | |
| Tramadol | 0.02 (-0.04-0.09) | -0.58 (-0.75--0.42)*** |
| Codeine | 0.03 (-0.07-0.14) | -0.48 (-0.79--0.16)** |
| Oxycodone | -0.10 (-0.21-0.00)^ | 0.10 (-0.14-0.35) |
| Other | 0.09 (-0.23-0.40) | 0.58 (0.01-1.16)* |
| **Initial Opioid Prescription Daily Strength in Morphine Milligram Equivalents, RC = 1-19 MME/day** | | |
| 20-49 MME/day | -0.02 (-0.08-0.04) | 0.25 (0.09-0.40)** |
| 50-89 MME/day | -0.11 (-0.24-0.02)^ | 0.63 (0.30-0.96)*** |
| 90+ MME/day | -0.17 (-0.40-0.06) | 1.13 (0.69-1.57)*** |
| **Initial Prescriber Specialty, RC = Family Medicine** | | |
| Internal Medicine | -0.13 (-0.19--0.06)*** | -0.14 (-0.30-0.03)^ |
| Nurse Practitioner | 0.11 (0.04-0.19)** | 0.06 (-0.14-0.27) |
| Physician Assistant | 0.22 (0.11-0.33)*** | 0.15 (-0.21-0.51) |
| **Sex, RC = Female** | | |
| Male | -0.03 (-0.09-0.02) | 0.22 (0.07-0.36)** |
| **Age, RC = 18-34** | | |
| 35-44 | -0.09 (-0.22-0.03) | 0.07 (-0.25-0.40) |
| 45-54 | -0.08 (-0.19-0.04) | 0.16 (-0.14-0.47) |
| 55-64 | -0.03 (-0.15-0.08) | -0.11 (-0.42-0.20) |
| 65-74 | -0.04 (-0.27-0.19) | -0.79 (-1.42--0.17)* |
| 75+ | 0.04 (-0.20-0.28) | -1.17 (-1.82--0.52)*** |
| **Payer Type, RC = Commercial Insurance** | | |
| Medicare | -0.08 (-0.30-0.14) | 0.83 (0.23-1.44)** |
| Medicaid | -0.21 (-0.32--0.10)*** | 0.98 (0.71-1.25)*** |
| **Disability Status, RC = No** | | |
| Yes | 0.00 (-0.24-0.24) | 0.14 (-0.47-0.76) |
| **Comorbidities, RC = No** | | |
| Asthma | 0.01 (-0.09-0.10) | -0.06 (-0.30-0.18) |
| Coronary Artery Disease | 0.11 (0.03-0.19)* | -0.01 (-0.23-0.20) |
| Congestive Heart Failure | 0.04 (-0.11-0.18) | -0.03 (-0.42-0.35) |
| Coronary Obstructive Pulmonary Disease | -0.06 (-0.15-0.04) | 0.47 (0.26-0.67)*** |
| Depression | 0.04 (-0.05-0.14) | 0.40 (0.19-0.61)*** |
| Diabetes | 0.00 (-0.08-0.07) | -0.21 (-0.41--0.02)* |
| Hypertension | -0.03 (-0.10-0.04) | -0.06 (-0.25-0.12) |
| **ICD Diagnosis Chapter, RC = XIII. Musculoskeletal** | | |
| IX. Circulatory | -0.33 (-0.43--0.23)*** | -0.35 (-0.58--0.13)** |

| | | |
|---|---|---|
| XVIII. Abnormalities | -0.15 (-0.24--0.06)** | -0.28 (-0.54--0.03)* |
| IV. Endocrine, nutrition | -0.36 (-0.48--0.25)*** | -0.38 (-0.64--0.13)** |
| XXI. Health factors | -0.22 (-0.34--0.10)*** | -0.28 (-0.56-0.01)^ |
| XIX. Injury, poisoning | -0.03 (-0.14-0.09) | -1.13 (-1.68--0.58)*** |
| X. Respiratory | -0.24 (-0.39--0.09)** | -0.40 (-0.80-0.00)* |
| XII. Skin and tissue | -0.22 (-0.38--0.06)** | -2.71 (-4.13--1.29)*** |
| VI. Nervous system | -0.16 (-0.33-0.01)^ | 0.32 (-0.03-0.67)^ |
| I. Infectious diseases | -0.17 (-0.35-0.01)^ | -1.26 (-2.04--0.47)** |
| XIV. Genitourinary | -0.05 (-0.22-0.13) | -0.75 (-1.43--0.07)* |
| XI. Digestive system | 0.07 (-0.13-0.26) | -0.59 (-1.16--0.01)* |
| Other | -0.35 (-0.53--0.17)*** | -0.11 (-0.49-0.26) |
| **Geographical Census Region, RC = South** | | |
| Midwest | -0.08 (-0.16-0.00)^ | -0.20 (-0.46-0.05) |
| Northeast | -0.09 (-0.30-0.12) | -0.15 (-0.63-0.33) |
| West | 0.03 (-0.06-0.12) | 0.28 (0.01-0.56)* |
| Unknown | N/A [c] | N/A [c] |
| **Rural-Urban Continuum, RC = 1 – Large metropolitan area** | | |
| 2 - Medium metropolitan area | -0.08 (-0.16-0.00)^ | -0.03 (-0.28-0.21) |
| 3 - Small metropolitan area | -0.16 (-0.26--0.06)** | -0.13 (-0.40-0.14) |
| 4 - Large urban, near metropolitan area | -0.05 (-0.22-0.13) | 0.03 (-0.44-0.51) |
| 5 - Large urban, not near metropolitan area | -0.15 (-0.28--0.02)* | 0.17 (-0.16-0.51) |
| 6 - Small urban, near metropolitan area | -0.18 (-0.30--0.06)** | 0.22 (-0.07-0.51) |
| 7 - Small urban, not near metropolitan area | -0.22 (-0.35--0.09)*** | 0.16 (-0.14-0.46) |
| 8 - Rural, near metropolitan area | -0.13 (-0.34-0.08) | 0.19 (-0.31-0.69) |
| 9 - Rural, not near metropolitan area | -0.12 (-0.34-0.10) | 0.32 (-0.17-0.82) |
| 0 - Unknown | N/A [c] | N/A [c] |
| **Initial Prescriber Long-term Opioid-use Rate, Increment = 10%** | | |
| Additional 10% | -0.01 (-0.04-0.03) | 0.20 (0.14-0.27)*** |
| **Insufficient Observations of Prescriber's Opioid Initiations, RC = No** | | |
| Yes | -0.02 (-0.08-0.04) | 0.38 (0.22-0.54)*** |
| **Annual State Opioid Prescription Rate, Increment = Additional 10 prescriptions per 100 residents** | | |
| Additional 10 opioids prescribed | 0.03 (0.01-0.05)* | 0.11 (0.05-0.17)*** |
| **Years into Study, Increment = 1 additional year into study** | | |
| Years since February 2012 | 0.05 (0.01-0.08)** | -0.02 (-0.10-0.06) |

*^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

[a] *The adjusted odds ratio corresponds to a one unit change in the increment (e.g. 10%) or compared to the reference category (e.g. provider concordance). Adjusted odds ratios corresponding to categories with small episode subsets should be interpreted with caution. Refer to summary statistics in* ***Table S5*** *for category breakdown percentages.*

[b] *Reference category designates the episode characteristic to which all other categories should be compared.*

[c] *Insufficient patient episodes to compute an estimate.*

Instrumental variable estimates are modeled with probit-logit marginal and a Clayton copula. The instrumental variable is highly significant in the treatment equation, as would be expected for a variable that has high predictive power of the independent variable. Compared to the logistic model that does not account for endogeneity, provider discordance in the instrumental variable model has a larger effect on long-term opioid use. This suggests that if unobserved factors are confounding our estimates, they are making the logistic model estimate of provider discordance more conservative.

**Table S12. Best fit instrumental variable model for daily strength outcome measure**

| Recursive Bivariate Equation | Coefficients (95% Confidence Interval)[a] | |
|---|---|---|
| | **Effect on Treatment (Provider Discordance)** | **Effect on Treatment (Provider Discordance)** |
| **Follow-up Appointment, Reference Category (RC)[b] = Seen by original prescribing clinician,** | | |
| Provider discordance | - | -0.87 (-1.41--0.32)** |
| **$IV_1$ = Prescribing Clinician Follow-up Rate, Increment = 10%** | | |
| 10% Increase in follow-up rate | 0.23 (0.22-0.25)*** | - |
| **$IV_1$ Control: Insufficient Appointment History for Prescribing Clinician, RC = No** | | |
| Yes | 0.32 (0.24-0.39)*** | 0.08 (-0.18-0.33) |
| **Initial Opioid Prescription Days Supplied, RC = 1-3 Days** | | |
| 4-7 Days | -0.28 (-0.38--0.18)*** | 0.74 (0.16-1.33)* |
| 8-14 Days | -0.41 (-0.52--0.31)*** | 1.34 (0.74-1.93)*** |
| 15+ Days | -0.46 (-0.57--0.36)*** | 2.66 (2.08-3.23)*** |
| **Initial Opioid Prescription Drug Base, RC = Hydrocodone** | | |
| Tramadol | 0.02 (-0.04-0.08) | -0.93 (-1.16--0.70)*** |
| Codeine | 0.03 (-0.08-0.13) | -1.14 (-1.69--0.59)*** |
| Oxycodone | -0.10 (-0.20-0.00)^ | 0.00 (-0.27-0.28) |
| Other | 0.08 (-0.23-0.40) | 0.63 (0.02-1.25)* |
| **Initial Opioid Prescription Daily Strength in Morphine Milligram Equivalents, RC = 1-19 MME/day** | | |
| 20-49 MME/day | -0.02 (-0.08-0.04) | 0.90 (0.70-1.11)*** |
| 50-89 MME/day | -0.11 (-0.24-0.02)^ | 1.51 (1.14-1.88)*** |
| 90+ MME/day | -0.17 (-0.40-0.06) | 1.95 (1.48-2.43)*** |
| **Initial Prescriber Specialty, RC = Family Medicine** | | |
| Internal Medicine | -0.13 (-0.20--0.06)*** | -0.30 (-0.51--0.09)** |
| Nurse Practitioner | 0.11 (0.03-0.19)** | -0.14 (-0.40-0.12) |
| Physician Assistant | 0.22 (0.11-0.33)*** | -0.01 (-0.46-0.44) |
| **Sex, RC = Female** | | |
| Male | -0.03 (-0.09-0.02) | 0.28 (0.10-0.46)** |
| **Age, RC = 18-34** | | |
| 35-44 | -0.09 (-0.21-0.03) | -0.08 (-0.48-0.32) |
| 45-54 | -0.08 (-0.19-0.04) | 0.11 (-0.27-0.48) |
| 55-64 | -0.03 (-0.14-0.09) | -0.25 (-0.64-0.13) |
| 65-74 | -0.03 (-0.26-0.20) | -0.46 (-1.17-0.25) |
| 75+ | 0.04 (-0.20-0.28) | -1.41 (-2.18--0.63)*** |
| **Payer Type, RC = Commercial Insurance** | | |
| Medicare | -0.08 (-0.30-0.14) | 0.51 (-0.18-1.19) |
| Medicaid | -0.20 (-0.31--0.09)*** | 1.06 (0.71-1.40)*** |
| **Disability Status, RC = No** | | |
| Yes | 0.00 (-0.24-0.25) | 0.51 (-0.19-1.21) |
| **Comorbidities, RC = No** | | |
| Asthma | 0.01 (-0.09-0.10) | 0.03 (-0.26-0.33) |
| Coronary Artery Disease | 0.11 (0.03-0.20)** | 0.14 (-0.14-0.41) |
| Congestive Heart Failure | 0.04 (-0.10-0.19) | 0.02 (-0.46-0.51) |
| Coronary Obstructive Pulmonary Disease | -0.06 (-0.16-0.03) | 0.47 (0.21-0.73)*** |
| Depression | 0.04 (-0.05-0.14) | 0.34 (0.08-0.61)* |
| Diabetes | 0.00 (-0.08-0.07) | -0.26 (-0.51--0.01)* |
| Hypertension | -0.04 (-0.10-0.03) | -0.09 (-0.33-0.14) |
| **ICD Diagnosis Chapter, RC = XIII. Musculoskeletal** | | |

| | | |
|---|---|---|
| IX. Circulatory | -0.33 (-0.43--0.24)*** | -0.58 (-0.89--0.28)*** |
| XVIII. Abnormalities | -0.15 (-0.24--0.06)** | -0.48 (-0.82--0.13)** |
| IV. Endocrine, nutrition | -0.36 (-0.48--0.25)*** | -0.35 (-0.66--0.03)* |
| XXI. Health factors | -0.22 (-0.34--0.10)*** | -0.48 (-0.85--0.10)* |
| XIX. Injury, poisoning | -0.03 (-0.14-0.09) | -1.22 (-1.91--0.52)*** |
| X. Respiratory | -0.23 (-0.38--0.09)** | -0.75 (-1.31--0.20)** |
| XII. Skin and tissue | -0.22 (-0.38--0.06)** | N/A [c] |
| VI. Nervous system | -0.16 (-0.33-0.01)^ | 0.46 (0.06-0.87)* |
| I. Infectious diseases | -0.17 (-0.35-0.01)^ | -1.44 (-2.47--0.40)** |
| XIV. Genitourinary | -0.04 (-0.22-0.13) | -1.06 (-2.05--0.06)* |
| XI. Digestive system | 0.08 (-0.12-0.27) | -0.51 (-1.21-0.18) |
| Other | -0.35 (-0.53--0.17)*** | -0.23 (-0.71-0.25) |
| **Geographical Census Region, RC = South** | | |
| Midwest | -0.08 (-0.16-0.00)^ | -0.23 (-0.57-0.11) |
| Northeast | -0.10 (-0.31-0.11) | -0.32 (-0.92-0.28) |
| West | 0.03 (-0.06-0.12) | 0.46 (0.12-0.79)** |
| Unknown | N/A [c] | N/A [c] |
| **Rural-Urban Continuum, RC = 1 – Large metropolitan area** | | |
| 2 - Medium metropolitan area | -0.08 (-0.16-0.00)^ | -0.32 (-0.63--0.01)* |
| 3 - Small metropolitan area | -0.16 (-0.26--0.06)** | -0.44 (-0.79--0.09)* |
| 4 - Large urban, near metropolitan area | -0.06 (-0.23-0.12) | 0.13 (-0.45-0.71) |
| 5 - Large urban, not near metropolitan area | -0.15 (-0.28--0.02)* | -0.19 (-0.62-0.24) |
| 6 - Small urban, near metropolitan area | -0.18 (-0.30--0.06)** | -0.04 (-0.42-0.33) |
| 7 - Small urban, not near metropolitan area | -0.23 (-0.36--0.10)*** | -0.23 (-0.62-0.16) |
| 8 - Rural, near metropolitan area | -0.14 (-0.35-0.07) | -0.29 (-0.96-0.38) |
| 9 - Rural, not near metropolitan area | -0.12 (-0.33-0.10) | -0.02 (-0.65-0.60) |
| 0 - Unknown | N/A [c] | N/A [c] |
| **Initial Prescriber Long-term Opioid-use Rate, Increment = 10%** | | |
| Additional 10% | -0.04 (-0.09-0.02) | 0.15 (0.03-0.27)* |
| **Insufficient Observations of Prescriber's Opioid Initiations, RC = No** | | |
| Yes | -0.03 (-0.09-0.04) | 0.35 (0.15-0.55)*** |
| **Annual State Opioid Prescription Rate, Increment = Additional 10 prescriptions per 100 residents** | | |
| Additional 10 opioids prescribed | 0.03 (0.01-0.05)* | 0.08 (0.00-0.15)^ |
| **Years into Study, Increment = 1 Year** | | |
| Years since February 2012 | 0.05 (0.02-0.08)** | -0.08 (-0.18-0.02)^ |

*^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

*[a] The adjusted odds ratio corresponds to a one unit change in the increment (e.g. 10%) or compared to the reference category (e.g. provider concordance). Adjusted odds ratios corresponding to categories with small episode subsets should be interpreted with caution. Refer to summary statistics in **Table S5** for category breakdown percentages.*

*[b] Reference category designates the episode characteristic to which all other categories should be compared.*

*[c] Insufficient patient episodes to compute an estimate.*

Instrumental variable estimates are modeled with probit-logit marginal and a Clayton copula. In addition to $IV_1$, we note that several other covariates (including specialty of the initial prescriber and the days supplied of the initial opioid prescription) are significant in the selection equation. This gives us some insight into the conditions under which patients are more (or less) likely to be exposed to provider discordance. The instrumental variable is highly significant in the treatment equation, as would be expected for a variable that has high predictive power of the independent variable.

Compared to the logistic model that does not account for endogeneity, provider discordance in the instrumental variable model has a larger effect on long-term opioid use. Consistent with the other models, this observation suggests that if unobserved factors are confounding our estimates, they are making the logistic model estimate of provider discordance more conservative.

## SUPPLEMENT PART IV. PROPENSITY-SCORE MATCHING

### Methods S7. Matching with the Minimum Bias Estimator Technique

#### i. Propensity-score matching

As discussed above, factors that affect a patient's likelihood of changing clinicians for the follow-up appointment (provider discordance) may also affect the patient's likelihood of long-term opioid use, thus confounding the logistic model estimates. Propensity score matching is one common approach used to account for confounding. While matching is typically used to adjust for differences in the underlying characteristics of patients in the control (provider concordance) and treatment (provider discordance) groups based on *observed* factors, importantly, Rosenbaum (2005) demonstrated that matching can alleviate the impact of *unobserved* bias as well. In particular, Rosenbaum notes that "reducing heterogeneity reduces both sampling variability and sensitivity to unobserved bias – with less heterogeneity, larger biases would need to be present to explain away the same effect."

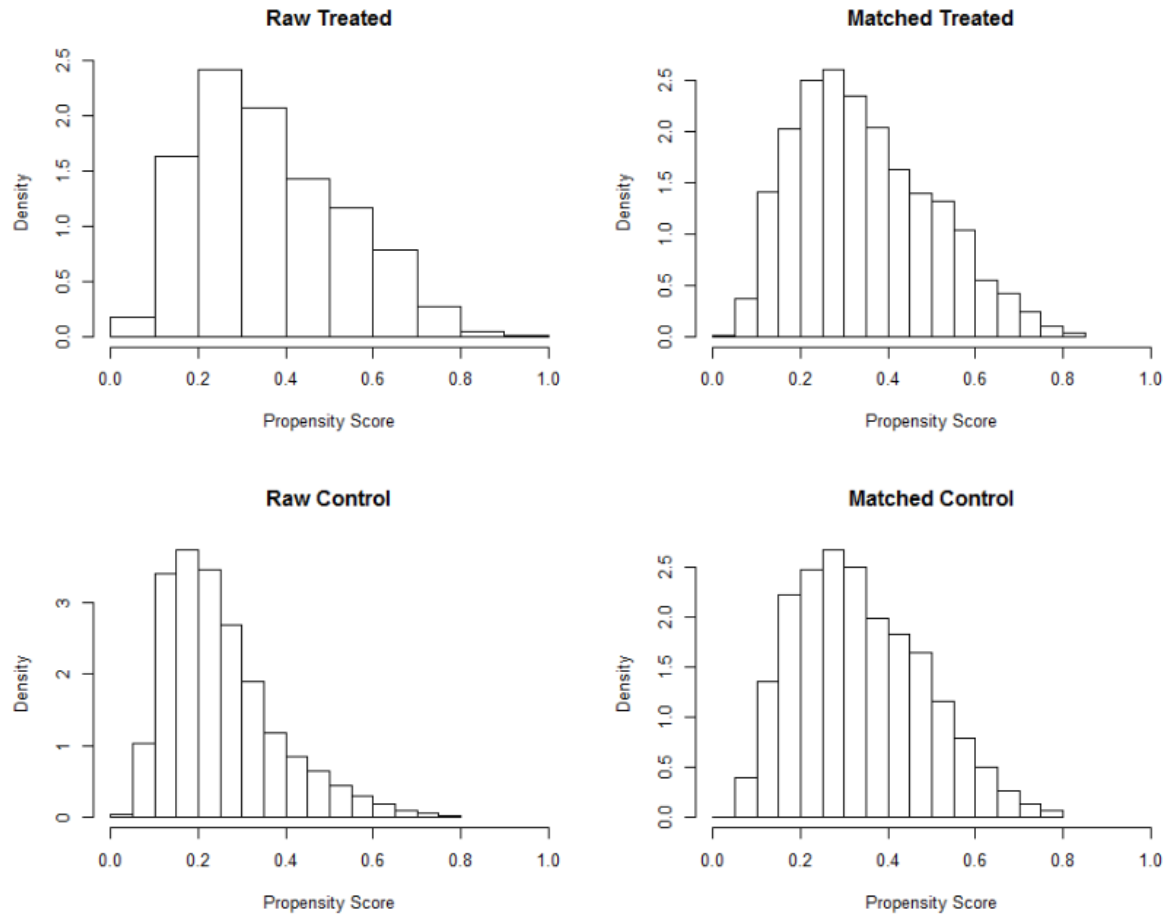
Therefore, one robustness test that can be performed to assess the extent to which findings are biased by unobserved factors is to re-run the estimations on a matched sample. If results are the same or very similar on the matched sample, then this provides more confidence in the findings and helps to rule out omitted variable bias as a major concern. Propensity score matching works to retain all (or most) of the treated observations in the data set and to select a matching set of non-treated observations that are similar. Especially when there are a large number of control variables over which the matching occurs, the goal of matching methods is to balance the covariate distributions across the two groups (treated and untreated).

We matched patients who change clinician (i.e., received the treatment) are matched in a 1:1 ratio to those patients with the closest propensity score who did not change clinician using nearest neighbor matching, with the condition that the closest propensity score can be no greater than 0.2 standard deviations away from the switchers' propensity score. As explained in the manuscript, this condition is a conservative distance that has been shown to reduce more than 90% of the bias due to observable differences between treatment and control groups (Gu and Rosenbaum 1993). Applying this condition means that we are unable to find a match for 169 of the patients in the treatment group. This leaves us a matched sample consisting of 6,084 patients (3,042 in each of the treatment and the control groups).[8]

The matching is performed using the MatchIt package in R (Ho et al. 2011). **Figure S4** shows the distribution of propensity scores before (left column) and after (right column) matching. This shows significantly improved balance after matching.

[8] There are 6,084 episodes in the sample when the long-term opioid use rate of the initial prescriber is measured using the days supplied measure. When the same covariate is measured using the daily strength definition, there are 6,086 episodes in the sample (i.e., 165 episodes in the treatment group are left unmatched).

**Figure S4. Propensity score matching: raw vs matched sample**




One of the main benefits of matching is that it can increase efficiency by removing observations outside of the region in which the model can reasonably extrapolate. Turning to the estimated coefficients, we report results when re-estimating the logistic models from the paper on the matched sample in **Table S13** below.

**Table S13. Estimated effect of provider discordance using propensity-matched sample**

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Full sample (N=11,340) | 0.69 (0.57-0.82) | <0.001*** | 0.62 (0.50-0.80) | <0.001*** |
| Full matched sample (N=6,084 days supplied; N=6,086 daily strength) | 0.75 (0.62-0.91) | 0.003** | 0.62 (0.48-0.79) | <0.001*** |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

The observed effects of provider discordance on long-term opioid use are similar when estimated on the full sample and matched samples. The consistency confirms that our findings are unlikely to be biased by observations outside of the region in which the model can reasonably extrapolate.

**ii. Bias minimization method**

Within the context of propensity score matching, any unobserved bias will have more of an impact on the estimate of the treatment effect for matched cases that are in the tails of the propensity score distribution (i.e., those with predicted probabilities of being in the treatment group close to 0 or close to 1). Bias is therefore minimized for matched observations with propensity scores closest to 0.5. Based on that observation, Millimet and Tchernis (2013, p. 983) detail a method that they refer to as the minimum-biased estimator (MBE). This method aims to minimize the potential impact of omitted variable bias by restricting the sample to matched cases with propensity scores within a defined interval around 0.5. Standard treatment coefficients can then be compared with coefficients estimated on this restricted matched sample to determine the robustness of the results and the direction of any omitted variable bias, if any.

As to the degree to which potential bias is reduced is greater in intervals centered around 0.5, Black and Smith (2004) propose that the reduced matched sample should be formed by taking observations with propensity scores between of 0.33 and 0.67. A limitation of this narrow window is that it reduces bias at the expensive of an increase in variability, since the results are estimated on a narrower sample. This makes it harder to find an effect even if one truly exists. Moreover, because the analysis focuses in on a smaller subset of patients, it becomes harder to generalize to the full population and to extrapolate outside of the common support.

As a result of the above limitations, rather than directly interpret the effects from the MBE approach, instead a common application of this approach is to test for the direction of any potential bias. Peel (2018) suggests the following four potential impactions of MBE treatment estimates:

> *"(a) MBE treatment effects are similar and significant, implying unobserved bias is not a significant threat and offering support for the CIA;*
> *(b) MBE treatment effects are higher and significant, implying that standard treatment estimates are subject to downward bias (under-estimated) due to the presence of an unobserved correlated variable;*
> *(c) MBE treatment effects are lower and significant, implying that standard treatment estimates are subject to upward bias (over- estimated) due to the presence of an unobserved correlated variable; and*
> *(d) MBE treatment effects are statistically insignificant and close to zero, implying that standard significant treatment estimates arise as a consequence of omitted variable bias."*

Our instrumental variable (IV) approach suggested that the results from our paper fall under condition (b), i.e., we underestimate the potential effect of the treatment due to omitted variable bias. In particular, our standard logistic model suggested a reduction in long-term opioid use of 31% (95% CI: 18%-43%) when using the days supplied measure, while the IV models indicated that this was biased downwards and instead the effect was closer to 54% (95% CI: 30%-70%). For the daily strength measure, the standard treatment estimates using the logistic model were estimated at 37% (95% CI: 20%-50%) and increased to 58% (95% CI: 28%-76%) when using the IV models.

We can thus apply the MBE approach to determine whether this method is consistent with the IV approach, i.e., also finds that the standard treatment estimates using the logistic model under-estimate the true treatment effect.

We estimate the MBE models on increasingly restricted subsamples (note that as the propensity score range narrows, we should expect the bias to reduce, as noted earlier):

- **Propensity matched case I**: Patients whose propensity score takes a value between the range of 0.10 and 0.90 (5,966 patients when matched with days supplied measurement and 6,012 when matched with daily strength measurement).
- **Propensity matched case II**: Patients whose propensity score takes a value between the range of 0.25 and 0.75 (4,113 patients when matched with days supplied measurement and 4,127 when matched with daily strength measurement).
- **Propensity matched case III**: Patients whose propensity score takes a value between the range of 0.33 and 0.67 (2,745 patients when matched with days supplied measurement and 2,809 when matched with daily strength measurement).

Consistent with the findings from the IV models, the MBE models suggest that, if anything, the logistic results presented in the paper under-estimate the true effect (see **Table 2** in the main manuscript for full results). We find that as we narrow the propensity score range for the matched subsamples further, the coefficient estimates become increasingly negative. When we reach the range recommended by Black and Smith (2005) – i.e., $0.33 \leq$ propensity score $\leq 0.67$ in Case III – the estimated size of the effect of provider discordance on long-term opioid use increases to 0.55 (95% CI: 0.38-0.79) and 0.45 (95% CI: 0.28-0.73) for the days supplied and daily strength measures, respectively.[9] Observe also that these effect sizes are very similar to those estimated using the IV approach (0.46 [95% CI: 0.30-0.70] and 0.42 [95% CI: 0.24-0.72], respectively). Therefore, the MBE approach implies that standard treatment estimates are subject to downward bias (under-estimated) due to the presence of an unobserved correlated variable.

The observation that both the IV approach and the MBE approach – one which relies on valid instrumental variables and the other that is estimated without instruments – provide consistent findings is further confirmation of the results in the paper. All this evidence points to the fact that the headline result presented in the paper (a 31% [95% CI: 18%-43%] reduction in long-term opioid use for patients who are exposed to the treatment, provider discordance) is, if anything, conservative.

[9] Note that the p-values also decrease but this is entirely expected due to the variance-bias trade-off noted earlier: that minimizing since we are estimating these effects on a significantly smaller sample.

## SUPPLEMENT PART V. PROVIDER DISCORDANCE PATHWAYS

### Methods S8. Pathways as a proxy for patient acuity

An important aspect of provider discordance is associated with the relationship between the patient and the providers involved in their care management. For example, it could be that a patient has an acute onset of pain and does not have time to wait for their regular primary care provider to become available. If this is the case, then the patient may be more likely to see an alternate, or "covering" primary care provider for the first appointment and then schedule a follow-up appointment with their regular provider for the follow-up appointment.

In this scenario, the patient may be more likely to experience provider discordance and also more likely to receive a short-term initial opioid script and not continue long-term opioid use. Because we cannot control for unobserved variables like patient acuity directly (as this is an unobserved variable that is not recorded in our data set), this pathway analysis helps us determine whether the provider discordance differs across the potential pathways (and thus, across differing levels of acuity of the patients).

As alluded to in the example above, we would expect patients with an acute pain onset to be ***less*** likely to see their main (regular) primary care provider at their initial appointment, as the acute onset of pain requires them to quickly see whichever provider has availability. If acuity were an important omitted variable, we would expect that patients who see their regular primary care provider at their first appointment benefit less from a change in provider than those who see an alternate (potentially "covering") primary care provider.

In the logistic regression we replace the binary independent variable $x_i$ (provider discordance) with categorical variable $X_i$ where $X_i$ takes one of four levels:

- Base level: Provider concordance
- Pathway 1: Started with their regular primary care provider and switched to an alternative provider
- Pathway 2: Started with an alternative provider and then switched back to their regular provider
- Pathway 3: Started with an alternative provider and then switched to another provider who was also not their regular primary care provider

In the model, we also include a dummy variable that specifies whether or not the patient saw their regular PCP for either appointment, as this adjusts for differences in long-term opioid rates between patients who see their regular provider at any point during the initial stages of the opioid-use episode. Without this control, it would not be possible to determine whether the effect of provider discordance when separated into different patient pathways is due to the discordance effect or, rather, whether the effect of a patient being treated by their regular primary care provider.[10]

Results reported in **Table 3** of the manuscript suggest that the discordance effect estimates are significant for all patient pathways, regardless of whether a patient started with or switched to their regular primary care provider or did neither. Note that because we have fewer observations associated with each of the three pathways compared to the overall "provider discordance" category, the significance levels for each pathway are lower than when

[10] Note that if we erroneously omit this control variable, the effects of the 3 pathways reported in **Table 3** of the manuscript become even stronger.

patients are combined into a single "discordance" category. However, all discordance pathway effects remain statistically significant at the 5% level.

The effects are strongest for the set of patients who started with a non-regular provider for their first visit then switched back to their regular PCP – i.e., pathway 2 (adjusted odds ratio 0.58 [95% CI: 0.40-0.84] and 0.51 [95% CI: 0.31-0.84], respectively, for the days supplied and daily strength measurements). However, even for those patients who start with their regular provider and then switch to an alternate provider (pathway 1), we find that discordance reduces their likelihood of becoming long-term opioid users (adjusted odds ratio 0.73 [95% CI: 0.56-0.94] and 0.68 [95% CI: 0.49-0.95], respectively, for the days supplied and daily strength measurements). The effect of provider discordance was also significant for those who start with an alternate provider and then switch to a different provider who is still not the patients' regular provider (pathway 3).

We ran a series of hypothesis tests to determine whether the effect size of provider discordance differs depending on which of the three pathways a patient follow. The p-values (all greater than 0.30 using the chi-squared test) do not provide evidence to reject the hypothesis that the effect sizes are the same. While the power to identify differences in effect sizes may be limited by our sample size, the above results indicate that all types of patient pathways (to and from the patient's identified 'regular' primary care provider) are positively impacted by provider discordance at the beginning of the opioid episode. If acuity is in fact correlated with whether a patient sees their preferred provider for the first or second appointment (as has been argued above), then the above results indicate that both more acute and less acute patients benefit from provider discordance. The data does not provide evidence that the sizes of the effects are statistically different.

In addition, we observe that seeing a regular provider at all during the initial stage of opioid use (for the initial prescribing appointment, follow-up appointment, or both appointments – as captured in $r_i$), is associated with a lower likelihood of continuing long-term opioid use. The adjusted odds ratio corresponding to this binary control covariate's effect on long-term opioid use is 0.80 (95% CI: [0.67-0.95], p=0.01) using the days supplied measurement and 0.83 (95% CI: [0.67-1.02], p = 0.08). While not a main point of our study, this observation corroborates other continuity-of-care studies that establishing a relationship with a regular provider and seeing this regular provider around the time of opioid initiation makes it less likely for patients to continue with the (potentially high-risk) pathway of long-term opioid use. Because this control was incorporated in the model, the observed effects of provider discordance reported above capture solely the impact beyond this regular provider effect.

**Methods S9. Regular provider subgroup analyses**

In the analyses presented in the main manuscript as well as the pathway analyses estimated in **Methods S8**, the base level of the independent variable, *provider concordance*, was a single category made up of both patients who saw their regular primary care provider during opioid initiation as well as those who did not. To further explore the importance of an established relationship with this initial prescriber, we now conduct separate subgroup analyses: one analysis on the cohort of patients who saw their regular primary care clinician for their initial opioid prescription (N=6,931) and a second analysis on the cohort of patients who saw a non-regular clinician (N=4,409).

We estimated the base logistic model from the main manuscript twice, once for each patient cohort:

$$y_i^* = \alpha + \beta x_i + \gamma O_i' + \delta P_i' + \lambda C_i' + \theta T_i' + \varepsilon_i, \qquad y_i = \mathbb{I}[y_i^* > 0] \qquad \textit{[S2]}$$

where $x_i$ took on the value 0 under provider concordance and 1 under provider discordance.[11] The estimated effects of the association between provider discordance and long-term opioid use can be found in **Table S14**.

**Table S14. Sub-sample analyses based on initial prescriber**

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Effect on Long-Term Opioid Use (95% Confidence Interval in Treatment Equation)**[+] | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| **Sub-sample 1 - regular initial prescriber (N=6,931)** | | | | |
| Provider discordance | 0.70 (0.54-0.91) | 0.007** | 0.67 (0.48-0.94) | 0.020* |
| **Sub-sample 2 - non-regular initial prescriber (N=4,409)** | | | | |
| Provider discordance | 0.60 (0.46-0.78) | <0.001*** | 0.54 (0.38-0.77) | <0.001*** |

Similar to the pathway analysis described in **Methods S8**, we observe that the effect of provider discordance appears stronger in the cohort of patients who initiated their opioid journey with a clinician *other than* their regular primary care provider. However, we still detect a statistically significant and large effect within the cohort of patients whose initial prescriber was their regular primary care clinician. While further analysis should be pursued prior to implementing an intervention, our data consistently suggest that provider discordance is negatively associated with long-term opioid use, irrespective of the relationship that a patient has with their initial prescribing clinician.

[11] In the non-regular initial prescriber cohort, $x_i$=1 includes both patients who transitioned from their non-regular primary care clinician to another non-regular clinician, as well as patients who transitions from their non-regular clinician to their regular primary care clinician. We repeated the analysis by separating the independent variable into 3 categories (the two provider discordance categories described above). Although there were only a small number of patients who switched between two non-regular clinicians (N=934), we observed the following effects on long-term opioid use: 0.73 (95% CI: [0.53-1.00*]) for patients who switched between two non-regular clinicians and 0.44 (95% CI: [0.30-0.65]) for patients who switched to their regular clinician using the days supplied measure; using the daily strength measure, the adjusted odds ratios were 0.64 (95% CI: [0.42-0.97*]) and 0.41 (95% CI: [0.24-0.69***]), respectively.

**Methods S10. Impact of provider discordance within shared provider offices**

We next investigated the possibility that communication between the initial prescribing and follow-up providers, in particular whether two providers shared an office, has a tangible impact on long-term opioid use. We hypothesize that if clinicians share an office, they would be more likely to have access to any electronic medical records pertaining to the initial prescribing appointment. We investigated whether this potential for information-sharing or the relationship between the two clinicians may impact the propensity that the follow-up clinician continues the patient on the opioid journey (Senot 2019).

As the claims data did not contain rendering service location information, we used provider-level sources such as the National Plan and Provider Enumeration System (NPPES) to determine whether the initial prescriber and follow-up clinician had at least one office location in common (Bindman 2013). We then estimated the following logistic model:

$$y_i^* = \alpha + \pi f_i + \gamma O_i' + \delta P_i' + \lambda C_i' + \theta T_i' + \varepsilon_i, \qquad y_i = \mathbb{I}[y_i^* > 0] \qquad \textit{[S3]}$$

where $f_i$ is a 4-level categorical variable that classifies the nature of the follow-up appointment:

(1) the patient returned to the same prescribing clinician for the follow-up, N = 8,129
(2) the follow-up clinician was different from the initial prescriber and the two clinicians shared an office, N =1,842
(3) the follow-up clinician was different from the initial prescriber and the two clinicians did not share an office, N = 1,364
(4) the office address of either the prescribing or follow-up clinician was unknown, N = 5

The base level of $f_i$ is that the patient returned to the prescribing clinician for a follow-up appointment. The adjusted odds ratio all non-base levels of the categorical variable $f_i$ are shown in **Table S15**.

**Table S15. Estimate of provider discordance effect in shared versus different offices**

| | **Days Supplied Measure** | | **Daily Strength Measure** | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| (2) Shared office | 0.69 (0.55-0.87) | <0.001*** | 0.67 (0.49-0.90) | 0.008** |
| (3) Different office | 0.68 (0.53-0.88) | 0.003** | 0.58 (0.42-0.81) | 0.002** |
| (4) Unknown address(es) | 1.15 (0.09-14.90) | 0.916 | 1.54 (0.10-24.23) | 0.758 |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Provider discordance has a significant effect on long-term opioid use both when the providers share an office, and when the providers are in different offices. The Wald chi-squared test shows no statistical difference between the two cohorts ($\chi^2$ = 0.005 and p-value = 0.98 for the days supplied outcome measure; $\chi^2$ = 0.378 and p-value = 0.54 for the daily strength outcome measure). Therefore, while the results should be corroborated with larger or additional data sets, we do not find evidence that the effect of provider discordance differs depending on whether the providers share the same office space.

## SUPPLEMENT PART VI. OTHER APPROACHES TO ADDRESS ENDOGENEITY

### Methods S11. Incorporating initial prescriber fixed and random effects

It is possible that there is heterogeneity across primary care providers (PCPs) in their likelihood to see patients with an underlying propensity to continue long-term opioid use. This variation could be attributed to the demographics of their patients, how their practices are organized, or how providers certain professional qualification (e.g., nurse practitioners) may be more likely to see high-acuity patients. We therefore want to decompose the unobserved variation between patients in the sample into (a) variation between patients seen by a given clinician, i.e., variation *within* clinicians and (b) variation in the proportion of acute patients *between* clinicians.

If provider-level variation such as differences in acuity are an important omitted variable, then we argue that provide an alternative explanation for our findings, then both variations (a) and (b) should affect the results. The control structure, sensitivity tests, and instrumental variable approach presented in Methods S5-S7 address bias resulting from case (a). We now take two related approaches to directly address case (b):

1. Add a fixed effect (FE) for the first PCP.
2. Add a random effect (RE) for the first PCP.

We discuss each of these, with their limitations below, and report results in Table 1. In summary, controlling with either fixed or random effects does not weaken the estimated effects.

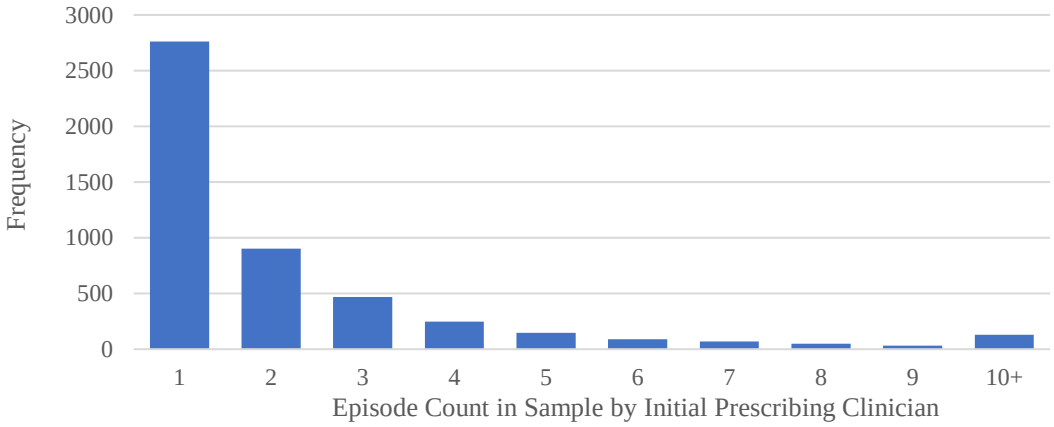
#### i. Add a fixed effect for the initial prescriber

Adding a fixed effect (an indicator variable) for the initial provider controls for the difference between the two providers and in effect removes all heterogeneity across the PCPs in both the treatment (rate of discordance vs. concordance) and outcome (rate of long-term opioid use). For example, suppose there were only two (types of) providers, A and B. Patients who see provider A are more likely to become long-term opioid users (e.g., less acute patients visiting their main PCP), while those who see provider B are less likely to become long-term opioid users (e.g., more acute patients visiting a covering PCP). The provider-level fixed effects adjust for the average levels of acuity or long-term opioid use in patients seen by each initial prescriber.

One issue with this approach is that our dataset contains 4,894 unique initial PCP's. As our data are sourced from insurance claims, we have complete medical and pharmaceutical information for each patient as they seek care across multiple clinicians. However, we may only observe a fraction of the patients seen by each of their clinicians (i.e., we only observe the other patients also covered by the same insurance provider). It is therefore uncommon within the dataset to find instances where a single clinician serves as the initial opioid prescriber for multiple patients satisfying all medical eligibility and clinical criteria (see **Figure S5**).[12]

[12] This sparsity of patients per prescriber is why all analyses for the main manuscript were conducted at the episode level without nesting at the level of the clinician or primary care practice.

**Figure S5. Distribution of episode count in sample by initial prescriber**




Introducing too many fixed effects into the model (especially when the number of observations per FE is low) results in overfitting to the data and unreliable coefficient estimates. While this will lead to unbiased estimates of the effect of interest (the impact of discordance on the rate of long-term opioid use), the estimate can be subject to high sample-to-sample variability. We have this problem in our data set, since we have 11,340 patients who see nearly 4,894 different PCPs for their first visit.

To address this concern, we "bucketed" all PCPs with fewer than 5 observations into an "Other" category. We then estimated the model using the "Other" level as well as the fixed effects for each clinician with 5 or more observations. In total, this leaves to estimate 516 fixed effects: a large, but feasible number of additional variables to estimate given our sample size of 11,340 patient episodes.

**ii. Add a random effect for the initial prescriber**

As an alternative to a fixed effects model, we also estimate a random effects model. The main advantage of the random effects estimator is that when the number of observations per physician is low, the estimated random effects term is closer to the overall group mean; meanwhile, when the number of observations per physician is high, the random effects term converges to the random effects estimator. Unfortunately, the random effects model will not completely eliminate bias in the estimate of the effect of interest (discordance versus concordance). However, it reduces this bias significantly and can also greatly constrain the variance of the estimate and thereby lead to an estimate that is closer, on average, to the true value in the population.

**iii. Model results estimated with provider-level effects**

Overall, the fixed effects approach has higher variance but lower bias, and the random effects approach has lower variance but higher bias.[13] If when estimating with both fixed effects and random effects we find consistent results, then this provides greater confidence in our findings with respect to low bias and relatively good out-of-sample generalizability. We report results from **i-ii** below in comparison to the original model output in **Table S16**.

[13] For more on this see: Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods*, *3*(2), 399-408.

**Table S16. Incorporating initial prescriber fixed and random effects**

| | Days Supplied Measurement | | Daily Strength Measurement | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Base logistic models (manuscript) | 0.69 (0.58-0.83) | <0.001*** | 0.63 (0.50-0.80) | <0.001*** |
| Initial prescriber fixed effects (S11.1) | 0.68 (0.56-0.82) | <0.001*** | 0.63 (0.49-0.82) | <0.001*** |
| Initial prescriber random effects (S11.2) | 0.69 (0.57-0.83) | <0.001*** | 0.64 (0.50-0.81) | <0.001*** |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Across all models, there was no evidence that characteristics of the initial prescribing clinician (potentially associated with patient acuity levels) were important omitted variables. The coefficient estimates are nearly identical to the effects estimated in the base logistic model. This suggests that unobserved confounders at the initial prescriber level are unlikely to be biasing the estimates presented in the main manuscript.

**Methods S12. Interacting provider discordance with patient condition history**

We next examine whether patients who have presented with the same condition before the opioid initiation (e.g., potentially more chronic patients) differ in how they are affected by provider discordance compared to patients who have no initial primary care visit for the condition prior to the opioid initiation (e.g., potentially more acute patients).

To capture this, we create a binary variable that indicates whether this is the first time that a patient has been seen for the condition in the primary care setting over the past 6 months. Such patients are assumed to be more likely to have presented with an acute condition compared to those for who have visited primary care for the same condition over the past 6 months. Approximately 39% (4,420) of patients in the sample fit into the "new condition (more likely acute)" category. To ensure the binary variable does not introduce additional bias into the estimates, we also control for the number of primary care appointments (for any condition) that the patient had in the 6 months prior to the prescribing appointment.

As shown in **Table S17**, interacting the "new condition (more likely acute)" indicator variable with the exposure yields no statistical difference in the size of the effect of the exposure. In other words, the effect of provider discordance is not significantly different whether this was the first time the patient was seen for that condition (more acute) over the past 6 months or whether they had been seen before (more chronic). Specifically, p-values for the interaction term equal 0.538 and 0.292 for the days supplied and daily strength measurements, respectively.

**Table S17. Incorporating new condition interaction**

| | **Days Supplied Measure** | | **Daily Strength Measure** | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Provider discordance | 0.71 (0.57-0.88) | 0.002** | 0.68 (0.52-0.90) | 0.007** |
| New condition (more likely acute) | 1.00 (0.81-1.24) | 0.965 | 0.90 (0.68-1.17) | 0.420 |
| Provider discordance x new condition | 0.92 (0.63-1.34) | 0.648 | 0.77 (0.46-1.28) | 0.317 |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

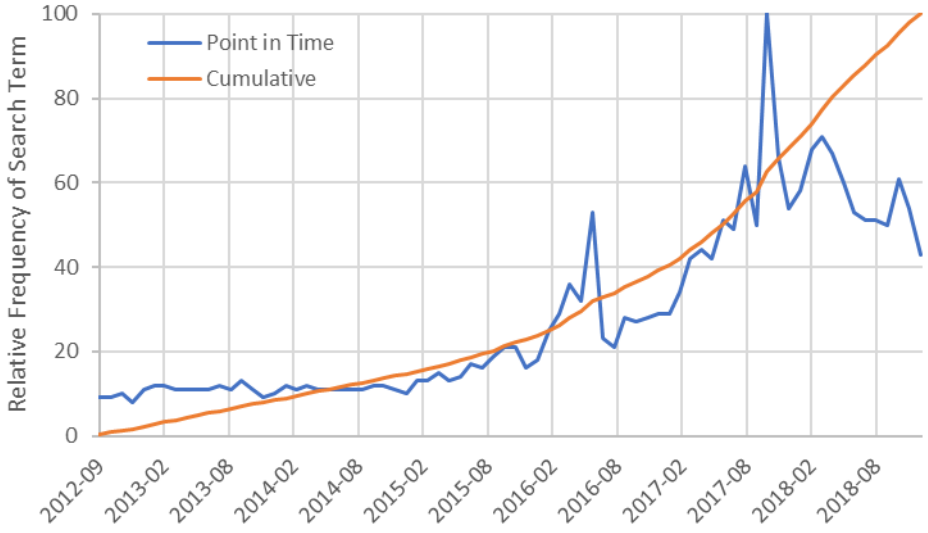
Again, if acuity is correlated with whether a patient was treated for that condition before (as has been argued above), then the above results indicate that both acute and chronic patients benefit from provider discordance. The data does not provide evidence that the effect differs between the two groups.

**Methods S13. Controlling for socio-economic factors and patient knowledge of opioid risk**

It is reasonable to hypothesize that a patient's understanding of the risks related to opioids, or that their ability to change clinicians could affect the likelihood of provider discordance. For example, the education level of a patient may be associated with their understanding of the risks of opioid use, thereby affecting the rate at which they seek out an alternative clinician after opioid initialization. Likewise, lower-income patients or those on restrictive insurance plans such as Medicaid may have less flexibility in scheduling appointments with an alternative clinician, even if they would have preferred to see a different clinician.

We additionally acknowledge that throughout the duration of our study, patients were exposed to more media coverage related to the opioid crisis. A graph of the relative frequency that "opioid" was typed using Google's search engine between September 2012 (the earliest episode start date in the sample) and December 2019 can be found in **Figure S6**.

**Figure S6. Google searches for "opioid" throughout the study observation period**




We adjusted for many of these concerns using the base set of control variables. For instance, we included a linear control for time ("months into study") which approximately follows the cumulative distribution of the frequency of Googling of the term "opioid," a potential indicator of the accumulated knowledge that an average patient would have about the opioid crisis, throughout the study period. We controlled for the patient's insurance type, region, and the type of county (rural, urban, metropolitan, etc.) in which they live as these factors may be associated with the patient's ability to seek out alternative care. Additionally, we controlled for the opioid prescription rate per 100 patients at the state level – updated annually – to adjust for geographical differences in media coverage or the cultural acceptability of opioids within the patient's living context.

To confirm whether other socio-economic or knowledge-based factors are important omitted variables in the model, we proposed two extra sets of controls. We had access to location information for 80.2% of the patients in the sample. The locations could be linked to economic and education data recorded by the US Census and available at the census block group level (areas of 600-3000 residents) (U.S. Census Bureau 2010). Though these controls are not specific to the individual patient, these data give detailed insight into neighborhood characteristics,

including (for residents of at least 18 years of age): the percent of residents who have not completed high school, the percent of residents who have a college degree, the percentage of unemployment, the percent below the poverty line, and the average annual income.

Second, we control for the relative frequency at which people across the US were Google searching for "opioids" over time as a proxy for the general population-wide knowledge about opioids and their risks. Since the cumulative trend was already generally captured in our "months into study" control, we added in the relative frequency of googled "opioid" searches during the month of the patient's opioid initiation (as seen in the blue point-in-time line in **Figure S6**).

If any of the education level, economic, or knowledge-based factors gathered for this study were important omitted variables, we would expect to observe the coefficient of provider discordance to become smaller, possibly less significant. Recall that the estimated effect of provider discordance in the main manuscript is 0.69 (95% CI: 0.58-0.83) for the days supplied measurement of long-term opioid use and 0.63 (95% CI: 0.50-0.80) using the daily strength measurement. The results of the logistic model with the additional controls are presented in **Table S18**.

**Table S18. Incorporating economic and opioid risk knowledge controls ledge: logistic model**

| | **Days Supplied Measurement** | | **Daily Strength Measurement** | |
|---|---|---|---|---|
| **Adjusted Odds Ratio on Long-Term Opioid Use (95% Confidence Interval)** | **Adjusted Odds Ratio** | **P-value** | **Adjusted Odds Ratio** | **P-value** |
| Provider discordance | 0.69 (0.57-0.82) | <0.001*** | 0.63 (0.50-0.80) | <0.001*** |
| No high school degree[+] | 0.61 (0.19-1.99) | 0.414 | 0.78 (0.20-3.09) | 0.723 |
| Completed college degree[+] | 0.84 (0.57-1.23) | 0.366 | 0.58 (0.33-1.01) | 0.056^ |
| Unemployed[+] | 4.04 (0.88-18.66) | 0.0734 | 5.52 (1.10-27.63) | 0.037* |
| Below poverty level[+] | 1.01 (0.62-1.63) | 0.977 | 0.97 (0.55-1.70) | 0.909 |
| Income (in $1000's) [+] | 1.00 (1.00-1.00) | 0.764 | 1.00 (0.99-1.00) | 0.192 |
| 10% increase in frequency of "opioid" Google searches in initiation month | 1.07 (1.01-1.14) | 0.022* | 1.03 (0.95-1.11) | 0.447 |
| Years into study[++] | 0.88 (0.78-0.98) | 0.026* | 0.86 (0.75-1.00) | 0.050* |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

[+]*Estimated per patient using census track data at the census group block level. While the coefficients are not all displayed, these models include all base controls including insurance type, region, state opioid rate, and county classification.*

[++]*For precision, this was estimated as a 'months into study' variable and converted to years into study for the display.*

We additionally estimated the recursive bivariate model with the additional controls, including the controls in both the first and second equations. The estimates for clinician discordance can be found in **Table S19**.

**Table S19. Incorporating economic and opioid risk knowledge controls: recursive bivariate model**

| | Selection Equation | | Outcome Equation | |
|---|---|---|---|---|
| **Long-Term Opioid Use Measure (95% Confidence Interval)** | **Coefficient** | **P-value** | **Coefficient** | **P-value** |
| **Days supplied outcome measure** | | | | |
| Provider discordance | -- | -- | -0.76 (-1.18--0.33) | <0.001*** |
| No high school degree[+] | 0.04 (-0.40-0.49) | 0.851 | -0.47 (-1.66-0.71) | 0.432 |
| Completed college degree[+] | 0.08 (-0.04-0.19) | 0.187 | -0.17 (-0.55-0.21) | 0.391 |
| Unemployed[+] | 0.59 (-0.21-1.39) | 0.148 | 1.40 (-0.14-2.94) | 0.075^ |
| Below poverty level[+] | -0.03 (-0.25-0.19) | 0.785 | 0.00 (-0.48-0.48) | 0.996 |
| Income (in $1000's) [+] | 0.00 (0.00-0.00) | 0.522 | 0.00 (0.00-0.00) | 0.753 |
| 10% increase in frequency of "opioid" Google searches in initiation month | 0.01 (-0.01-0.03) | 0.347 | 0.07 (0.01-0.13) | 0.019* |
| Years into study[++] | 0.03 (-0.01-0.08) | 0.188 | -0.12 (-0.24--0.01) | 0.039* |
| **Daily strength outcome measure** | | | | |
| Provider discordance | -- | -- | -0.86 (-1.40--0.32) | 0.002** |
| No high school degree[+] | 0.04 (-0.41-0.49) | 0.861 | -0.24 (-1.61-1.14) | 0.738 |
| Completed college degree[+] | 0.08 (-0.04-0.19) | 0.176 | -0.52 (-1.08-0.03) | 0.063^ |
| Unemployed[+] | 0.64 (-0.15-1.43) | 0.114 | 1.77 (0.16-3.38) | 0.031* |
| Below poverty level[+] | -0.03 (-0.25-0.19) | 0.786 | -0.05 (-0.61-0.52) | 0.867 |
| Income (in $1000's) [+] | 0.00 (0.00-0.00) | 0.556 | 0.00 (-0.01-0.00) | 0.193 |
| 10% increase in frequency of "opioid" Google searches in initiation month | 0.01 (-0.01-0.03) | 0.328 | 0.03 (-0.05-0.11) | 0.421 |
| Years into study[++] | 0.03 (-0.01-0.08) | 0.171 | -0.13 (-0.28-0.02) | 0.082^ |

^ *p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

[+]*Estimated per patient using census track data at the census group block level. While the coefficients are not all displayed, these models include all base controls in both the selection and outcome equations including insurance type, region, state opioid rate, and county classification.*

[++]*For precision, this was estimated as a months into study term and converted to years into study for the display.*

The models show that some of these additional controls are associated with the patient's likelihood of long-term opioid use. For example, unemployment is significantly associated with an increase in long-term opioid use across all four models. Increased overall knowledge of opioid risk, proxied by years into study, is also associated with a decrease in long-term opioid use. Note that because we captured the opioid awareness trend in two ways (through the point-in-time and cumulative changes – through the time into study variable – in opioid use over time), it is not unsurprising that the point-in-time frequency in Google searches of "opioid" in the initiation month do not also have a negative effect on long-term opioid use.

Most importantly, however, even with the inclusion of this range of additional controls, the point estimate of the effect of provider discordance on long-term opioid varies by less than 0.01 compared to what is presented in the main manuscript: when the estimates from **Table S19** are converted to adjusted odds ratios, the corresponding to the effect of provider discordance on long-term opioid use are 0.47 (95% CI: [0.31-0.72]) using the days supplied measure and 0.42 (95% CI: [0.25-0.73]) using the daily strength measure. We therefore do not find any evidence that suggests patient education, resources, or knowledge about opioid risks impacts the main model estimates in a way that was not already accounted for by the initial set of control variables.

## SUPPLEMENT PART VII. ROBUSTNESS AND SENSITIVITY ANALYSES

### Methods S14. Addressing concerns of sample selection

One concern with encouraging provider discordance within the primary care for new opioid initiates is self-selection –that unobserved factors both affect the likelihood of a patient returning for a repeat visit, their likelihood of seeing the same clinician during that visit *and* their likelihood of becoming a long-term opioid user. Practice-level workload is a great example of this: patients may be less likely to return to a busy practice, may be less likely to see the same provider, and may also differ in their long-term opioid use propensity. It is important to address self-selection if the findings are ever extrapolated to the question of "*what would have been the impact of provider discordance on those patients who both did and did not return for a follow-up visit, supposing that all (or more) patients returned for a follow-up visit?*"

**Figure S7. Patient sample censoring of new opioid initiates in the primary care setting.**

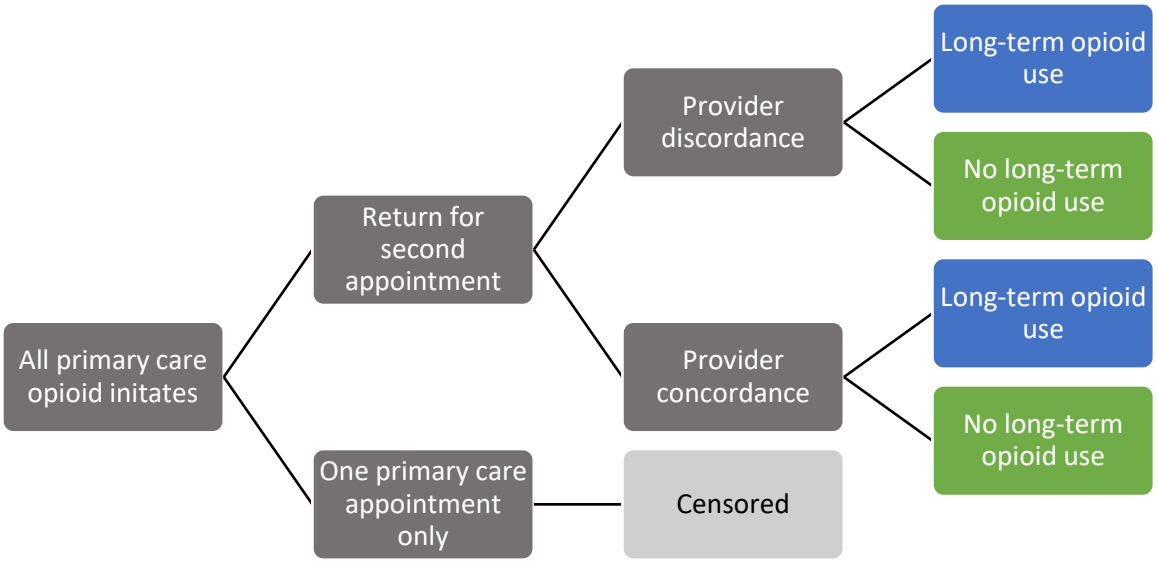



**Figure S7** provides a visual summary of the sample set-up and the censoring that occurs for the new opioid initiates. Self-selection may cause issues if interventions are designed to encourage provider discordance beyond the sample of those who naturally were expected to return for a second appointment. Consider two contrasting interventions:

1. [Proposed in the manuscript]: If a patient called back into the primary care office to request an appointment after opioid initiation, the administrators should schedule the patient with a different clinician, wherever possible.
2. [Alternative intervention]: Whenever an opioid-naive patient is prescribed an opioid in the primary care setting, they are automatically scheduled for a follow-up appointment with a different clinician. Under this alternative intervention, it is of critical importance to consider the full population of opioid initiates because some patients that would not have self-initiated a follow-up appointment may now be swayed to come back to the primary care setting – thus forcing an extrapolation of findings.

By design, all patients included in intervention 1 (the intervention proposed in the manuscript) were naturally returning for a follow-up in the primary care setting anyway; therefore, it is less important to understand how provider discordance affects the full set of opioid initiates. However, intervention 2 (the alternate intervention) may impact the population returning to a follow-up: some patients who may not have naturally called back into the office to schedule a follow-up appointment may return for a follow-up appointment as the appointment has been pre-scheduled. This type of alternate intervention extrapolates the findings beyond the original study sample and requires further investigation.

To explore the impact of self-selection in the context of opioid initiation, we identified a sample of patients who initiated opioids but did not return for their follow-up appointment in the claims data. Of the 53,491 patient episodes that met this criteria, 45,215 episodes met the same exclusion criteria as the episodes in the main sample (e.g., no history of cancer, end-stage renal disease, or <18 years of age). In **Figure S7** above, the 45,215 episodes represents pre-censored sample "all primary care opioid initiates"; of those, 11,340 were contained in the "return for second appointment" bucket and 33,875 were censored within the main sample.

As shown in **Table S20**, there exist significant observable differences between the population that did and did not return to the primary care for a second appointment, including age (patients with follow-up were older) and comorbidities (patients with follow-up had higher rates of comorbidities).

**Table S20. No follow-up appointment censored episode comparison**

| | **Overall**<br>**n = 45,215** | **No follow-up Appointment**<br>**n = 33,875** | **Follow-up Appointment**<br>**n = 11,340** | **p-value** |
|---|---|---|---|---|
| **Initial Daily Prescription Strength (% of total)** | | | | |
| 1-19 MME | 42.8% | 42.7% | 42.9% | 0.734 |
| 20-49 MME | 50.2% | 50.2% | 50.0% | 0.681 |
| 50-89 MME | 5.5% | 5.5% | 5.4% | 0.730 |
| 90+ MME | 1.6% | 1.5% | 1.7% | 0.350 |
| **Initial Prescription Days Supplied (% of total)** | | | | |
| 1-3 | 9.4% | 9.5% | 9.1% | 0.211 |
| 4-7 | 30.6% | 30.2% | 31.7% | 0.003** |
| 8-14 | 21.7% | 21.3% | 22.9% | <0.001*** |
| 15+ | 38.3% | 39.0% | 36.3% | <0.001*** |
| **Initial Prescription Drug Base (% of total)** | | | | |
| Hydrocodone | 45.1% | 45.1% | 45.1% | 0.955 |
| Tramadol | 37.4% | 37.7% | 36.8% | 0.105 |
| Codeine | 8.3% | 8.3% | 8.0% | 0.334 |
| Oxycodone | 8.6% | 8.3% | 9.3% | <0.001*** |
| Other | 0.6% | 0.6% | 0.7% | 0.086^ |
| **Initial Prescriber Specialty (% of total)** | | | | |
| Family Medicine | 48.9% | 49.5% | 47.2% | <0.001*** |
| Internal Medicine | 30.7% | 30.6% | 30.9% | 0.559 |
| Nurse Practitioner | 13.6% | 13.2% | 14.8% | <0.001*** |
| Physician Assistant | 6.8% | 6.7% | 7.1% | 0.164 |
| **Gender (% prevalence)** | | | | |
| Female | 64.4% | 64.7% | 63.5% | 0.017* |
| **Age (% of total)** | | | | |

| | | | | |
|---|---|---|---|---|
| 18-34 | 9.7% | 10.2% | 8.0% | <0.001*** |
| 35-44 | 12.2% | 12.7% | 10.7% | <0.001*** |
| 45-54 | 16.8% | 17.1% | 15.8% | 0.001** |
| 55-64 | 19.5% | 19.9% | 18.0% | <0.001*** |
| 65-74 | 22.1% | 21.6% | 23.5% | <0.001*** |
| 75+ | 19.7% | 18.4% | 23.9% | <0.001*** |
| **Insurance (% of total)** | | | | |
| Commercial | 39.4% | 42.4% | 30.5% | <0.001*** |
| Medicaid | 14.6% | 13.9% | 16.7% | <0.001*** |
| Medicare | 46.0% | 43.7% | 52.8% | <0.001*** |
| **Disability (% prevalence)** | | | | |
| Under 65 years and on Medicare | 6.1% | 5.8% | 7.1% | <0.001*** |
| **Comorbidities (% prevalence)** | | | | |
| Asthma | 6.3% | 5.0% | 10.0% | <0.001*** |
| Coronary Artery Disease | 8.3% | 5.5% | 16.4% | <0.001*** |
| Congestive Heart Failure | 3.1% | 2.7% | 4.4% | <0.001*** |
| Coronary Obstructive Pulmonary Disease | 6.1% | 4.1% | 12.1% | <0.001*** |
| Depression | 5.8% | 4.6% | 9.7% | <0.001*** |
| Diabetes | 12.0% | 8.9% | 21.3% | <0.001*** |
| Hypertension | 22.3% | 20.2% | 28.6% | <0.001*** |
| **Primary ICD Diagnosis Chapter (% of total)** | | | | |
| XIII. Musculoskeletal | 36.0% | 34.7% | 39.9% | <0.001*** |
| IX. Circulatory | 10.5% | 10.0% | 11.7% | <0.001*** |
| XVIII. Abnormalities not classified | 10.4% | 10.3% | 10.7% | 0.195 |
| IV. Endocrine nutrition, metabolic | 8.7% | 8.9% | 8.3% | 0.058^ |
| XXI. Factors influencing health | 7.7% | 8.4% | 5.6% | <0.001*** |
| XIX. Injury, poisoning | 6.8% | 7.4% | 5.3% | <0.001*** |
| X. Respiratory system | 4.4% | 4.6% | 3.7% | <0.001*** |
| XII. Skin and tissue | 2.3% | 2.1% | 2.9% | <0.001*** |
| VI. Nervous system | 3.1% | 3.2% | 2.6% | 0.001** |
| I. Infectious and parasitic diseases | 2.3% | 2.2% | 2.4% | 0.151 |
| XIV. Genitourinary | 2.6% | 2.8% | 2.3% | 0.005** |
| XI. Digestive system | 2.3% | 2.4% | 1.8% | <0.001*** |
| Other | 2.8% | 2.9% | 2.7% | 0.251 |
| **Geographical Census Region (% of total)** | | | | |
| South | 45.2% | 44.7% | 46.6% | <0.001*** |
| Midwest | 21.9% | 23.3% | 17.8% | <0.001*** |
| West | 18.8% | 19.4% | 17.2% | <0.001*** |
| Northeast | 2.9% | 2.9% | 3.0% | 0.432 |
| Unknown | 11.2% | 9.8% | 15.4% | <0.001*** |
| **Rural-Urban Continuum (% of total)** | | | | |
| 1 - Large metropolitan area | 35.4% | 36.0% | 33.6% | <0.001*** |
| 2 - Medium metropolitan area | 18.7% | 18.8% | 18.5% | 0.508 |
| 3 - Small metropolitan area | 9.0% | 8.7% | 9.8% | <0.001*** |
| 4 - Large urban, near metropolitan area | 2.7% | 2.7% | 2.5% | 0.225 |
| 5 - Large urban, not near metropolitan area | 3.1% | 2.4% | 5.2% | <0.001*** |
| 6 - Small urban, near metropolitan area | 8.1% | 8.7% | 6.4% | <0.001*** |
| 7 - Small urban, not near metropolitan area | 5.6% | 5.6% | 5.4% | 0.507 |
| 8 - Rural, near metropolitan area | 3.7% | 4.4% | 1.7% | <0.001*** |
| 9 - Rural, not near metropolitan area | 2.5% | 2.8% | 1.5% | <0.001*** |
| Not known | 11.3% | 9.9% | 15.4% | <0.001*** |
| **State Opioid Prescription Rate** | | | | |

| | | | | |
|---|---|---|---|---|
| Opioids prescribed annually per 100 residents | 80.1 | 80.1 | 80.0 | 0.82 |
| **Months into Study** | | | | |
| Months elapsed from Feb 2012 to episode start | 48.7 | 48.1 | 50.2 | <0.001*** |
| **Long-term Opioid Use Outcome Measure (% prevalence)** | | | | |
| ≥180 days supplied of opioids at 12 months | 7.3% | 6.3% | 10.0% | <0.001*** |
| Daily MME ≥20 at 12 months | 3.9% | 3.3% | 5.9% | <0.001*** |

^ <0.10, *<0.05, **<0.01, ***<0.001

A thoughtfully designed econometrics model can help us investigate the impact of self-selection within this full opioid initiate sample. Self-selection occurs in economics when estimating the wage equation for women, whereby a woman's wage is only observed if she makes the decision to enter the workplace and is unobservable if she does not. Heckman's (1976) paper that introduced the Heckman Selection model worked on this very problem. The paper looks to estimate the effect of various factors (e.g., hours worked, education, experience) on a woman's wage. However, a bias arises because the effects can only be estimated for women who choose to participate in the labor market. This means that any findings are not generalizable to all women. Heckman's selection model tries to address this.

To thoroughly capture and account for this sample selection problem, we formulated a trivariate model with endogeneity and non-random sample selection (Filippou et al. 2017, 2019). In this model, **Equation S4** estimates whether patients return for a follow-up visit. **Equation S5** estimates whether patients see the same (concordance) or a different (discordance) provider, conditional on returning for a follow-appointment, and censoring the observation if the patient did not return. **Equation S6** estimates the effect of provider discordance on long-term opioid use. The model formulation is:

$$f_i^* = \alpha + \beta_1 F_i + b_1 V_i' + \gamma_1 O_i' + \delta_1 P_i' + \lambda_1 C_i' + \theta_1 T_i' + \varepsilon_{3i}, \quad \textbf{[S4]}$$

$$\widehat{x_i^*} = \alpha + \beta_2 z_i' + b_2 V_i' + \gamma_2 O_i' + \delta_2 P_i' + \lambda_2 C_i' + \theta_2 T_i' + \varepsilon_{2i}, \quad \textbf{[S5]}$$

$$y_i^* = \alpha + \beta_3 \widehat{x_i} + b_3 V_i' + \gamma_3 O_i' + \delta_3 P_i' + \lambda_3 C_i' + \theta_3 T_i' + \varepsilon_{1i}, \quad \textbf{[S6]}$$

$$f_i = \mathbb{I}[f_i^* > 0], x_i = \mathbb{I}[x_i^* > 0], y_i = \mathbb{I}[y_i^* > 0],$$

where $F_i$ is a new instrument that corrects for endogeneity in returns for a follow-up appointment (which is similar to the instrument in the main paper): the % of patients who saw the same first provider as the focal patient and who returned for a second appointment;[14] $f_i$ indicates whether the patient returns for a follow-up appointment; $\widehat{x_i}$ is a modified independent variable that indicates provider discordance, and is censored in the case that the patient does not return for a follow-up appointment. All other variables including the IV $z_i'$ and coefficients take the same definitions as in the main paper.

This trivariate model corrects both the endogeneity associated with whether a patient returns for a follow-up visit as well as endogeneity associated with whether the non-censored patients have their follow-up appointment with

[14] When many other patients return to see their initial provider, so too might the focal patient (e.g., as the provider may be more available, may encourage patients to return, etc.). This new instrument and the original instrument from the paper are included in **Equation S4**, only the original instrument is included in **Equation S5**, and no instruments are included in **Equation S6**.

the same initial prescribing clinician. The term $\beta_3$ can thus be converted to an adjusted odds ratio to estimate the causal impact of provider discordance on long-term opioid use while adjusting for censoring: 0.51 (0.31-0.88)* using the days supplied measure of long-term opioid use and 0.41 (0.21-0.80)** using the daily strength measure.[15] Both estimates show that, if anything, the results reported in the main paper are conservative.

Therefore, while we recommend rigorous evaluation of any intervention prior to implementation, this analysis suggests that self-selection surrounding the decision to have a follow-up appointment in the primary care setting does not invalidate the main results.

[15] While the findings were robust across model specifications, the figures reported correspond to logit-probit-logit marginals, similar to the probit-logit specifications used in the IV model reported in the manuscript. The probit-probit-logit marginals resulted in the following adjusted odds ratios: 0.51 (0.31-0.88)* using the days supplied measure of long-term opioid use and 0.42 (0.21-0.80)** using the daily strength measure.

**Methods S15. Sensitivity to independent variable definition**

With evidence that provider discordance is associated with a change in long-term opioid use, we investigated the extent to which the definition of the independent variable affects the results. We relaxed the assumption that the patient must be exposed to provider discordance between the initial prescribing and follow-up appointment for the independent variable to be set to 1. Instead, we identified patients who were exposed to a primary care provider other than the initial prescriber at any point within the first X days of opioid initiation. In other words, even if a patient returned to the same initial prescriber several times before seeing an alternate clinician on day X, this new outcome variable would be set to 1 on day X as well as on all days where $x > X$, to indicate that the patient has been exposed to provider discordance in the primary care setting (regardless of their pathway or number of appointments prior to seeing this alternate primary care provider).

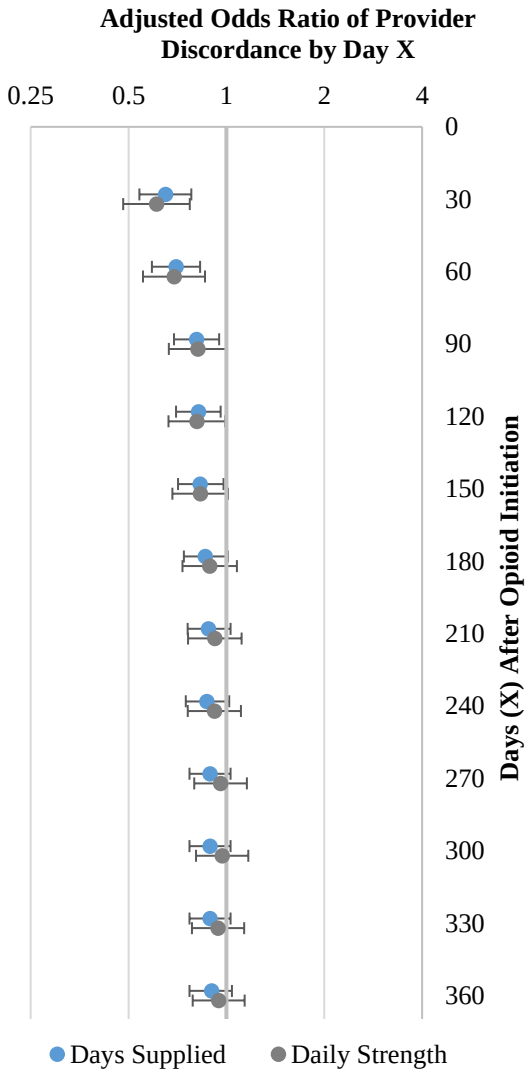
We estimated a series of fully controlled logistic models to identify the impact of provider discordance by day X (set to 30 days, 60 days, 90 days, etc.) on the patient's likelihood of long-term opioid use. As shown in **Figure S8**, within 30, 60 and 360 days of the initial opioid prescription, 3,243 (29%), 3,524 (31%) and 5,021 (44%) of patients in the sample, respectively, had experience provider discordance in the primary care setting. Seeing an alternate clinician within the first 30 days of opioid initiation was negatively associated with long-term opioid use, irrespective of whether the patient returned to their initial prescriber in the interim (adjusted odds ratio 0.66 [95% CI: 0.55, 0.79] using the days supplied measure, 0.61 [95% CI: 0.48, 0.78] using the daily strength measure). Exposure to provider discordance beyond 60 days was not consistently associated with lower rates of long-term opioid use, evaluated at the 5% significance level for both measures of long-term opioid use.

Overall, these findings suggest that exposure to provider discordance soon after opioid initiation can help curb long-term opioid use – even if this exposure occurs beyond the first follow-up appointment. To incorporate the intervention into healthcare policy, future research should further examine potential cut-off dates and investigate whether provider discordance that crosses outside of primary care (for example, to specialists or physical therapists) yield a similar effect.

**Figure S8. Sensitivity to the timing of the provider discordance exposure**

| | | Adjusted Odds Ratio of Provider Discordance by Day X on Long-Term Opioid Use (95% CI) | |
|---|---|---|---|
| Days (X) into Opioid-Use Journey | Patients with Provider Discordance to Date, No. (%) | Days Supplied Outcome Measure | Daily Strength Outcome Measure |
| 30 | 3243 (29%) | 0.65 (0.54-0.78), p=<0.001*** | 0.61 (0.48-0.77), p=<0.001*** |
| 60 | 3524 (31%) | 0.70 (0.59-0.83), p=<0.001*** | 0.69 (0.55-0.86), p=<0.001*** |
| 90 | 3751 (33%) | 0.81 (0.69-0.95), p=0.010** | 0.82 (0.67-1.00), p=0.054^ |
| 120 | 3961 (35%) | 0.82 (0.70-0.96), p=0.012* | 0.81 (0.66-0.99), p=0.041* |
| 150 | 4115 (36%) | 0.83 (0.71-0.98), p=0.023* | 0.83 (0.68-1.01), p=0.065^ |
| 180 | 4254 (38%) | 0.86 (0.74-1.01), p=0.062^ | 0.89 (0.73-1.08), p=0.229 |
| 210 | 4396 (39%) | 0.88 (0.76-1.03), p=0.104 | 0.92 (0.76-1.11), p=0.397 |
| 240 | 4516 (40%) | 0.87 (0.75-1.02), p=0.078^ | 0.92 (0.76-1.11), p=0.374 |
| 270 | 4639 (41%) | 0.89 (0.77-1.03), p=0.124 | 0.96 (0.80-1.16), p=0.664 |
| 300 | 4784 (42%) | 0.89 (0.77-1.03), p=0.133 | 0.97 (0.81-1.17), p=0.751 |
| 330 | 4902 (43%) | 0.89 (0.77-1.03), p=0.131 | 0.94 (0.78-1.13), p=0.532 |
| 360 | 5021 (44%) | 0.90 (0.77-1.04), p=0.139 | 0.95 (0.79-1.14), p=0.560 |

^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

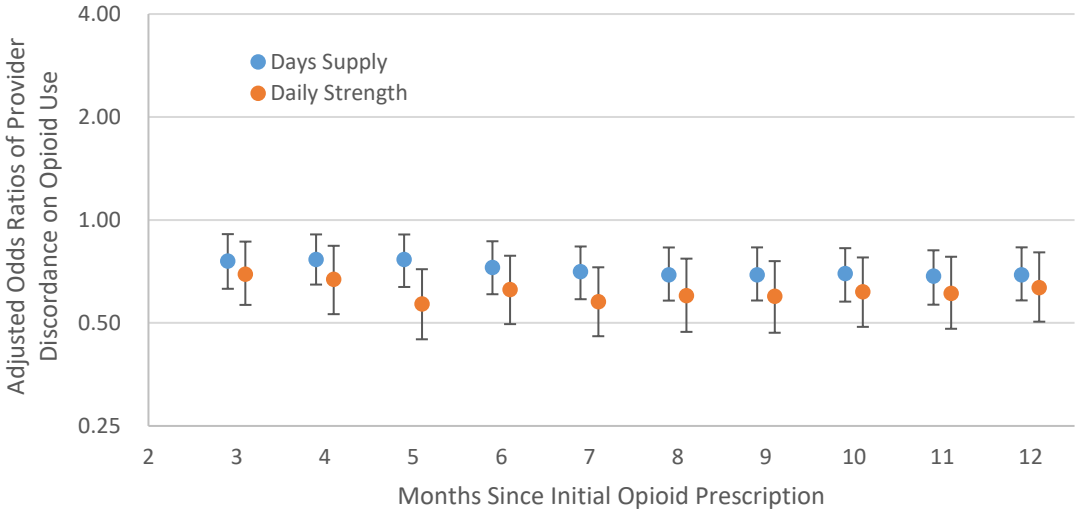



Patients who experience provider discordance in the primary care setting within the first 60 days are less likely to continue long-term opioid use, significant at the 0.1% level for both measures of long-term opioid use.

**Methods S16. Sensitivity to dependent variable definition**

In the main analysis, patients are flagged for long-term opioid usage when they are prescribed at least 180 days of opioid prescription within the first 360 days after the index prescription. For the days supplied outcome measure, the threshold that defines a patient as an opioid user is 0.5*X where X is the number of days since the index prescription. For example, at 180 days (approximately 6 months), the patient is required to have at least 90 days supplied of opioids to be considered a long-term user). The daily strength measure uses the 20 MME/day cut-off threshold to define opioid use at each month. Results for the two sets of logistic models are displayed in **Figure S9**.

**Figure S9. Sensitivity of long-term opioid use to various time horizons**




The estimated adjusted odds ratios shown that the estimates are not sensitive to the time horizon over which long-term opioid-use is defined, especially between months 6 and 12 (commonly benchmarked times for long-term opioid use). Between months 6 and 12 the point estimates of the adjusted odds ratio ranged between 0.69 and 0.73 for the days supplied measure, and between 0.58 and 0.63 for the daily strength measure. In all cases, 95% confidence bounds of the adjusted odds ratio were strictly less than 1.

**Methods S17. Alternative outcome analysis**

Even if provider discordance reduces the likelihood of long-term opioid use for new opioid initiates, provider discordance may have externalities, or ancillary negative consequences on a patient's care. For instance, if long-term opioid use is reduced because patients are prevented from having opioids that were appropriate for their condition, then they may continue to experience high levels of pain or discomfort. This pain and discomfort may have been avoided had the patient returned to the initial prescribing doctor for their follow-up appointment where they had a higher likelihood of being prescribed an additional opioid.

One way to detect negative consequences may be detected is by analysing the relationship between provider discordance and unplanned emergency room visits. Visits to the emergency room suggest that the patient's pain is not being managed by other preventative measures (including pain medicine, therapy, etc.) that could have been initiated by the follow-up primary care clinician. Other measures may also be good proxies of detecting negative consequences, but we leave these for future analyses.

To test the impact on subsequent emergency room visits, we created a binary variable $e_i$ that indicated whether the patient had at least one emergency room encounter in days 30 to 360 after the index opioid prescription (after the follow-up appointment) related to the opioid diagnosis.[16] Because some patients have higher underlying likelihoods of visiting the emergency room (often known as 'frequent fliers'), we also included an additional binary control variable $E_i$ that indicated whether the patient visited the emergency room at least once during the 12-month clean period. The logistic model formulation is as follows, where all other variable definitions are consistent with the main model presented in the paper:

$$e_i^* = \alpha + \beta x_i + \varphi E_i + \gamma O_i' + \delta P_i' + \lambda C_i' + \theta T_i' + \varepsilon_i, \qquad e_i = \mathbb{I}[e_i^* > 0],$$

If negative consequences resulted from provider discordance (as indicated emergency room visits after the follow-up appointment), we would expect the coefficient $\beta$ to be positive and significant. However, as shown in **Table S21**, we observe that provider discordance is not a significant predictor of emergency room visits (adjusted odds ratio 1.01 [95% CI: 0.89-1.15] when model controls are calculated using both the days supplied and the daily strength measurement). Therefore, while future studies should corroborate this finding with additional analyses prior to implementing an intervention, we do not find any evidence to suggest that provider discordance shortly after opioid initiation has a significant negative impact on other aspects of patient pain management.

**Table S21. Association Between Provider Discordance and Subsequent Emergency Room Usage**

| | **Adjusted Odds Ratio (95% Confidence Interval) on Emergency Room Visits** | |
|---|---|---|
| **Metric for Measuring Opioid Use** [a] | **Days Supplied** | **Daily Strength** |
| **Provider discordance, Reference Category (RC)**[b] **= Follow-up with original prescribing clinician** | | |
| Follow-up with alternate clinician | 1.01 (0.89-1.15) | 1.01 (0.89-1.15) |
| **1+ Related Emergency Visit in the Past 12 Months ($E_i$) RC = No** | | |

[16] For both variables $e_i$ and $E_i$, the emergency room visit was considered related to the opioid prescription when any of the diagnosis chapters recorded within the emergency visit claims matched the diagnosis chapter of the episode. The mean values of $e_i$ and $E_i$ across the patient sample were 0.17 and 0.19, respectively. For robustness, the analysis was repeated without requiring the emergency visit to be related to the opioid prescription (mean values of $\widehat{e_i}$ and $\widehat{E_i}$ were 0.35 and 0.35, respectively). Similar to the model shown in **Table XX**, the coefficient $\beta$ was small and not significant at the p=0.10 level.

| | | |
|---|---|---|
| Yes | 3.04 (2.71-3.42)*** | 3.04 (2.71-3.42)*** |
| **Initial Opioid Prescription Days Supplied, RC = 1-3 Days** | | |
| 4-7 Days | 0.84 (0.68-1.05) | 0.84 (0.68-1.05) |
| 8-14 Days | 0.90 (0.71-1.13) | 0.90 (0.71-1.13) |
| 15+ Days | 1.02 (0.82-1.29) | 1.03 (0.82-1.29) |
| **Initial Opioid Prescription Drug Base, RC = Hydrocodone** | | |
| Tramadol | 0.92 (0.82-1.05) | 0.92 (0.81-1.05) |
| Codeine | 1.08 (0.88-1.33) | 1.08 (0.88-1.33) |
| Oxycodone | 1.13 (0.93-1.39) | 1.13 (0.93-1.39) |
| Other | 1.18 (0.64-2.16) | 1.18 (0.64-2.16) |
| **Initial Opioid Prescription Daily Strength in Morphine Milligram Equivalents, RC = 1-19 MME/day** | | |
| 20-49 MME/day | 0.99 (0.87-1.12) | 0.99 (0.88-1.12) |
| 50-89 MME/day | 0.82 (0.62-1.10) | 0.82 (0.62-1.10) |
| 90+ MME/day | 1.03 (0.66-1.61) | 1.03 (0.66-1.61) |
| **Initial Prescriber Specialty, RC = Family Medicine** | | |
| Internal Medicine | 1.01 (0.89-1.14) | 1.01 (0.89-1.14) |
| Nurse Practitioner | 1.00 (0.84-1.18) | 1.00 (0.84-1.18) |
| Physician Assistant | 0.73 (0.55-0.96)* | 0.73 (0.55-0.96)* |
| **Sex, RC = Female** | | |
| Male | 0.90 (0.80-1.01)^ | 0.90 (0.80-1.01)^ |
| **Age, RC = 18-34** | | |
| 35-44 | 0.72 (0.56-0.92)** | 0.72 (0.56-0.92)** |
| 45-54 | 0.67 (0.53-0.84)*** | 0.67 (0.53-0.84)*** |
| 55-64 | 0.58 (0.46-0.73)*** | 0.58 (0.46-0.73)*** |
| 65-74 | 0.53 (0.32-0.86)** | 0.53 (0.32-0.86)** |
| 75+ | 0.70 (0.42-1.15) | 0.70 (0.42-1.15) |
| **Payer Type, RC = Commercial Insurance** | | |
| Medicare | 1.71 (1.07-2.74)* | 1.71 (1.07-2.74)* |
| Medicaid | 2.72 (2.19-3.37)*** | 2.73 (2.20-3.38)*** |
| **Disability Status, RC = No** | | |
| Yes | 1.40 (0.85-2.29) | 1.40 (0.86-2.30) |
| **Comorbidities, RC = No** | | |
| Asthma | 1.06 (0.89-1.27) | 1.06 (0.89-1.27) |
| Coronary Artery Disease | 1.34 (1.15-1.57)*** | 1.34 (1.15-1.57)*** |
| Congestive Heart Failure | 1.47 (1.14-1.90)** | 1.47 (1.14-1.90)** |
| Coronary Obstructive Pulmonary Disease | 1.39 (1.18-1.64)*** | 1.39 (1.18-1.64)*** |
| Depression | 1.16 (0.98-1.38)^ | 1.16 (0.98-1.38)^ |
| Diabetes | 1.14 (0.99-1.32)^ | 1.14 (0.99-1.31)^ |
| Hypertension | 0.73 (0.63-0.84)*** | 0.73 (0.63-0.84)*** |
| **ICD Diagnosis Chapter, RC = XIII. Musculoskeletal** | | |
| IX. Circulatory | 2.27 (1.93-2.66)*** | 2.27 (1.93-2.66)*** |
| XVIII. Abnormalities | 2.21 (1.87-2.60)*** | 2.20 (1.87-2.60)*** |
| IV. Endocrine, nutrition | 1.41 (1.16-1.71)*** | 1.41 (1.16-1.71)*** |
| XXI. Health factors | 0.80 (0.61-1.04)^ | 0.80 (0.61-1.04)^ |
| XIX. Injury, poisoning | 0.80 (0.60-1.07) | 0.80 (0.60-1.07) |
| X. Respiratory | 1.27 (0.96-1.67)^ | 1.27 (0.96-1.67)^ |
| XII. Skin and tissue | 0.42 (0.26-0.68)*** | 0.42 (0.26-0.68)*** |
| VI. Nervous system | 0.76 (0.52-1.11) | 0.76 (0.52-1.11) |
| I. Infectious diseases | 0.12 (0.05-0.34)*** | 0.12 (0.05-0.34)*** |
| XIV. Genitourinary | 0.79 (0.52-1.19) | 0.79 (0.52-1.19) |
| XI. Digestive system | 0.89 (0.58-1.35) | 0.89 (0.58-1.35) |

| | | |
|---|---|---|
| Other | 0.70 (0.48-1.03)^ | 0.70 (0.48-1.03)^ |
| **Geographical Census Region, RC = South** | | |
| Midwest | 0.89 (0.74-1.06) | 0.89 (0.74-1.06) |
| Northeast | 0.61 (0.42-0.87)** | 0.61 (0.42-0.87)** |
| West | 1.00 (0.80-1.23) | 1.00 (0.80-1.23) |
| Unknown | N/A [c] | N/A [c] |
| **Rural-Urban Continuum, RC = 1 – Large metropolitan area** | | |
| 2 - Medium metropolitan area | 1.30 (1.09-1.54)** | 1.30 (1.09-1.54)** |
| 3 - Small metropolitan area | 1.04 (0.85-1.27) | 1.04 (0.85-1.27) |
| 4 - Large urban, near metropolitan area | 1.45 (1.05-2.02)* | 1.45 (1.05-2.02)* |
| 5 - Large urban, not near metropolitan area | 0.95 (0.72-1.26) | 0.95 (0.72-1.26) |
| 6 - Small urban, near metropolitan area | 0.92 (0.72-1.18) | 0.92 (0.72-1.18) |
| 7 - Small urban, not near metropolitan area | 1.07 (0.84-1.38) | 1.07 (0.83-1.38) |
| 8 - Rural, near metropolitan area | 0.88 (0.58-1.35) | 0.88 (0.58-1.35) |
| 9 - Rural, not near metropolitan area | 1.05 (0.68-1.63) | 1.05 (0.68-1.62) |
| 0 – Unknown | N/A [c] | N/A [c] |
| **Initial Prescriber Long-term Opioid-use Rate, Increment = 10%** | | |
| Additional 10% | 1.00 (0.95-1.07) | 1.00 (0.91-1.09) |
| **Insufficient Observations of Prescriber's Opioid Initiations, RC = No** | | |
| Yes | 1.08 (0.96-1.21) | 1.08 (0.96-1.21) |
| **Annual State Opioid Prescription Rate, Increment = Additional 10 prescriptions per 100 residents** | | |
| Additional 10 opioids prescribed | 0.98 (0.94-1.02) | 0.98 (0.94-1.03) |
| **Years into Study, Increment = 1 additional year into study** | | |
| Years since February 2012 | 0.96 (0.90-1.01) | 0.96 (0.90-1.01) |

*^ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

[a] *Adjusted odds ratios corresponding to categories with small episode subsets should be interpreted with caution.*

[b] *Reference category designates the episode characteristic to which all other categories should be compared.*

[c] *Insufficient patient episodes to compute an estimate.*

## REFERENCES


Babalonis S, Lofwall MR, Nuzzo PA, Siegel AJ, Walsh SL (2013) Abuse liability and reinforcing efficacy of oral tramadol in humans. *Drug Alcohol Depend.* 129(1–2):116–124.

Barnett M, Olenski A, Jena A (2017) Opioid-Prescribing Patterns of Emergency Physicians and Risk of Long-Term Use. *N. Engl. J. Med.* 376(7):663–673.

Bindman AB (2013) Using the National Provider Identifier for Health Care Workforce Evaluation. *Medicare Medicaid Res. Rev.* 3(3):E1–E10.

Black DA, Smith JA (2004) How robust is the evidence on the effects of college quality? Evidence from matching. *J. Econom.* 121(1–2):99–124.

Centers for Disease Control and Prevention (2018) U.S. Opioid Prescribing Rate Maps. *U.S. Dep. Heal. Hum. Serv. Opioid Overdose*. Retrieved (July 25, 2019), https://www.cdc.gov/drugoverdose/maps/rxrate-maps.html.

Centers for Medicare and Medicaid Services (2018) Opioid Oral Morphine Milligram Equivalent (MME) Conversion Factors. *Prescr. Drug Cover.*:1–2. Retrieved https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/.

Clarke PS, Windmeijer F (2012) Instrumental variable estimators for binary outcomes. *J. Am. Stat. Assoc.* 107(500):1638–1652.

Cooper L (2012) GPI Classification Report. :1–265. Retrieved http://centricityusers.com/wp-content/forum-file-uploads/mitch/Drug-Classification-Report.pdf.

Cromartie J (2013) Rural-Urban Continuum Codes. *US Dep. Agric. Econ. Res. Serv.* Retrieved (November 19, 2019), https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.

Deyo RA, Hallvik SE, Hildebran C, Marino M, Dexter E, Irvine JM, O'Kane N, et al. (2017) Association Between Initial Opioid Prescribing Patterns and Subsequent Long-Term Use Among Opioid-Naïve Patients: A Statewide Retrospective Cohort Study. *J. Gen. Intern. Med.* 32(1):21–27.

Filippou P, Kneib T, Marra G, Radice R (2019) A trivariate additive regression model with arbitrary link functions and varying correlation matrix. *J. Stat. Plan. Inference* 199:236–248.

Filippou P, Marra G, Radice R (2017) *Penalized Likelihood Estimation of Trivariate Additive Binary Models*

Freeman M, Robinson S, Scholtes S (2020) Gatekeeping, Fast and Slow: An Empirical Study of Referral Errors in the Emergency Department. *Manage. Sci.* Forthcomin.

Gu XS, Rosenbaum PR (1993) Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms. *J. Comput. Graph. Stat.* 2(4):405–420.

Ho D, Imai K, King G, Stuart E (2011) MatchIt : Nonparametric Preprocessing for Parametric Casual Inference.

*J. Stat. Softw.* 42(8):1–28.

Hoopes M, Angier H, Raynor LA, Suchocki A, Muench J, Marino M, Rivera P, Huguet N (2018) Development of an algorithm to link electronic health record prescriptions with pharmacy dispense claims. *J. Am. Med. Informatics Assoc.* 25(10):1322–1330.

Marra G, Radice R, Access O (2017) Joint regression modeling framework for analyzing bivariate binary data in R. *Depend. Model.* 5(1):268–294.

Millimet DL, Tchernis R (2013) Estimation of the treatment effects without an exclusion restriction: with an application to tha analysis of the school breakfast program. *J. Appl. Econom.* 28:982–1017.

Mundkur ML, Rough K, Huybrechts KF, Levin R, Gagne JJ, Desai RJ, Patorno E, Choudhry NK, Bateman BT (2017) Patterns of opioid initiation at first visits for pain in United States primary care settings. *Pharmacoepidemiol. Drug Saf.* 27(5):495–503.

Peel MJ (2018) Addressing Unobserved Selection Bias in Accounting Studies: The Bias Minimization Method. *Eur. Account. Rev.* 27(1):173–183.

Rosenbaum PR (2005) Heterogeneity and causality: Unit heterogeneity and design sensitivity in observational studies. *Am. Stat.* 59(2):147–152.

Senot C (2019) Continuity of Care and Risk of Readmission: An Investigation into the Healthcare Journey of Heart Failure Patients. *Prod. Oper. Manag.* 28(8):2008–2030.

Shah A, Hayes C, Martin B (2017) Characteristics of Initial Prescription Episodes and Likelihood of Long-Term Opioid Use-United States, 2006-2015. *Morb. Mortal. Wkly. Rep.* 66(10):265–269.

Shah A, Hayes C, Martin B (2018) Factors influencing long-term opioid use among opioid naive patients: An examination of initial prescription characteristics and pain etiologies. *J. Pain* 18(11):1374–1383.

U.S. Census Bureau (2010) American Census Block Group Data. Retrieved (September 20, 2012), https://data.census.gov/cedsci/.

U.S. Census Bureau (2019) Census Regions and Divisions of the United States. Retrieved (April 1, 2019), https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf.

Winkelmann R (2011) Copula Bivariate Probit Models: With an Application to Medical Expenditures. *Health Econ.*