Misuse of prescription opioids is a leading cause of premature death in the United States. We use state government administrative data and machine learning methods to examine whether the risk of future opioid dependence, abuse, or poisoning can be predicted in advance of an initial opioid prescription. Our models accurately predict these outcomes and identify particular prior nonopioid prescriptions, medical history, incarceration, and demographics as strong predictors. Using our estimates, we simulate a hypothetical policy which restricts new opioid prescriptions to only those with low predicted risk. The policy’s potential benefits likely outweigh costs across demographic subgroups, even for lenient definitions of “high risk.” Our findings suggest new avenues for prevention using state administrative data, which could aid providers in making better, data-informed decisions when weighing the medical benefits of opioid therapy against the risks.

opiates | evidence-based policy | predictive modeling | machine learning | administrative data

Prescription opioids rank among the highest in terms of potential for dependence, abuse, and poisoning. In 2016, more Americans under the age of 50 y died from drug overdoses than from car crashes or gun violence, a trend driven by increases in opioid overdoses (1).

However, opioids may also be an important therapy for those who suffer from chronic pain. The majority of those prescribed opioids do not experience adverse outcomes; a survey of studies of opioid use found that rates of misuse, abuse, and addiction averaged between 8% and 12% (2). This rate is, however, higher than an early (and widely cited) claim that less than 1% of hospitalized patients receiving narcotics developed an addiction (3).

Moreover, many of those suffering from adverse outcomes were introduced to opioids through a legitimate opioid prescription. One study of 6 y of medical and pharmacy claims found that 79.9% of opioid abusers had a prescription prior to their first abuse diagnosis (4). Of the opioid abusers who did not themselves have a prior prescription, 50.8% had a family member with a prior prescription.

Given the risks and long-term consequences of adverse outcomes following legitimate opioid prescriptions, many providers now report a lack of confidence in managing their patients’ chronic pain through opioid therapy (5). Providers could benefit from better information on the risks of initiating a patient on opioid therapy, especially when that patient has never received an opioid prescription before.

Prior studies have identified risk factors for opioid abuse and dependence through descriptive analysis and statistical modeling of both medical claims and electronic health records (6–10), and two studies have also evaluated the predictive performance of such models (11, 12). However, these studies focus on individuals already persistently receiving opioid therapy and describe patterns of opioid use which are indicative of dependency and misuse within this subpopulation. Previous research has not yet developed a predictive model that is applicable to the larger population of recipients of opioid therapy using data on individuals known only prior to a prescription being given.

In this study, we use integrated administrative data to estimate models of adverse opioid-related outcomes for Medicaid enrollees in Rhode Island and conduct policy simulations of restricting opioid prescriptions to only those with low predicted risk. By some estimates, the opioid epidemic created $5.5 billion in additional health care costs to the Medicaid program nationally in 2013 (13). Estimating our model on state administrative data provides an avenue for state policymakers to predict the risk associated with prescribing opioids to Medicaid enrollees, which could be used to inform providers’ treatment decisions.

Materials and Methods

We use deidentified administrative records from a research data lake we helped build for the State of Rhode Island to support science- and data-driven policy (14). The data lake is housed in a secure enclave, and personally identifiable information has been removed and replaced with anonymous identifiers so that researchers with approved access can join and analyze records associated with the same individual across data sources while preserving anonymity (15). Because this study does not involve data that are
both identifiable and private, Brown University’s Institutional Review Board does not classify it as research with human subjects. The database includes Medicaid records from 2005 to 2017 and data on major social benefit and insurance programs, employment, incarceration, and criminal history.

We construct a panel dataset of 80,768 individuals who received an opioid prescription or injection according to the Medicaid claims records between 2006 and 2012 (16). There are 400,024 distinct Medicaid enrollees in this period. Further details and descriptive statistics are in SI Appendix, section 2 and Table S1.

We define an adverse opioid-related outcome as receiving a diagnosis of opioid dependence, abuse, or poisoning∗ or receiving treatment for an opioid use disorder in the 5 y following initial prescription. SI Appendix, Fig. S1 shows the cumulative frequency of adverse outcomes from the time of initial prescription, which peaks at 5.7% by year 5.

We construct variables from observations in the 12 mo prior to when an individual receives an opioid prescription. These include 84 variables for demographics, incarceration, citations, arrests, car crashes, wages, unemployment rates, household composition, and payments received from social benefit and insurance programs.

We construct 327 variables from Medicaid claims and enrollment records, including summary counts of the number of distinct diseases, chronic conditions, and procedures. The pharmacy claims data include 39,805 distinct drug product codes, and we use a pharmacological classification to consolidate opioid prescriptions into 262 distinct classes. There are 8,494 distinct diagnosis codes and 6,507 distinct procedure codes observed in the claims data. We intend to simultaneously reduce the dimensionality of these variables and estimate the underlying latent structure of the occurrence of the codes. One approach is to use the preexisting hierarchical structure of the codes: to use, for instance, the fact that all ICD-9 codes at a given level of codification are a subcollection of codes at a higher level. However, this constrains the model to nest codes in ways that may or may not be helpful for our predictive modeling purposes. For example, is it the case that codes 305.0 (nondependent alcohol abuse) and 305.2 (nondependent cannabis abuse) are together more likely to predict our outcome, as a combined measure of nondependent substance abuse? Or is 305.0 together with 303.0 (acute alcoholic intoxication) and 303.9 (other and unspecified alcohol dependence) a broader measure of alcohol use? Because we do not know a priori how to optimally nest the codes, we instead use natural language-processing topic-modeling techniques to consolidate the codes into 50 topics, based on the text descriptions and frequencies of the codes. For example, the 10 most frequent words in topic no. 39 are “hand sprain” lateral closed fracture foot minimum examination ankle views.” The variable for topic no. 39 measures how strongly this combination of diagnoses and procedures for hand, foot, and ankle injuries is represented in each individual’s medical history. Details on the topic modeling implementation appear in SI Appendix, section 3.

Finally, we construct 890 interaction terms from the 84 non-Medicaid and 327 Medicaid variables, for a total of 3,101 variables. We consider interactions among demographics, between the Medicaid summary counts and all non-Medicaid variables and between payments received from social benefit and insurance programs and all non-Medicaid variables.

We estimate predictive models using machine-learning algorithms that search over variables and functions of those variables to maximize out-of-sample predictive fit. We fit three kinds of models: a regularized regression, an ensemble, and a neural network. These models vary in complexity (17).

For example, the prediction function from a regularized regression is a linear combination of explanatory variables whose weightings are algorithmically selected from a set of variables and functions of those variables predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approximate functions predetermined by the researcher. Neural networks can approxim...
the BOLASSO model selected as the strongest, consistent predictors from the full set of 1,301 variables across the 100 bootstrap replicates. BOLASSO helps to identify consistent covariates, avoiding arbitrary choices among highly correlated pairs. While the coefficients on the selected variables do not necessarily have a causal interpretation, they pick up factors which are strong predictors among observables. For example, observed claims for routine preventative health (e.g., Fig. 1, topics 4 and 10) may themselves lower risk through increased or more frequent interactions with medical professionals, or they may proxy for attention to personal health or responsibility which is the true unobserved underlying factor that reduces risk. The primary purpose of our post-BOLASSO regression is to identify the strongest predictors which may point us in the direction of potential underlying mechanisms for further study.

The two variables with the largest odds ratios (indicating increased risk) are related to crime: release from prison and an indicator for an arrest. Individuals released from prison in the prior year are estimated as 119% more likely to develop an adverse outcome if given an initial prescription (odds ratio of 1.76). The next three variables with the smallest odds ratios (indicating decreased risk) were age 65+ y (0.13, indicating an almost complete, 87% reduction in risk), Hispanic ethnicity (0.41), age 55 to 64 y (0.43), and opiate agonists (1.36). Opioid agonists, such as cough syrups and mild painkillers, may have small dosages of an opioid ingre-
dients from the full set of 1,301 variables across the 100 bootstrap replicates. BOLASSO helps to identify consistent covariates, avoiding arbitrary choices among highly correlated pairs. While the coefficients on the selected variables do not necessarily have a causal interpretation, they pick up factors which are strong predictors among observables. For example, observed claims for routine preventative health (e.g., Fig. 1, topics 4 and 10) may themselves lower risk through increased or more frequent interactions with medical professionals, or they may proxy for attention to personal health or responsibility which is the true unobserved underlying factor that reduces risk. The primary purpose of our post-BOLASSO regression is to identify the strongest predictors which may point us in the direction of potential underlying mechanisms for further study.

The two variables with the largest odds ratios (indicating increased risk) are related to crime: release from prison and an indicator for an arrest. Individuals released from prison in the prior year are estimated as 119% more likely to develop an adverse outcome if given an initial prescription (odds ratio of 1.76), all else equal, and those with an arrest in the prior year are 76% more likely to do so (odds ratio of 1.76). The next three variables with the smallest odds ratios are prior prescriptions for benzodiazepines (1.51), centrally acting muscle relaxants (1.39), and opiate agonists (1.36). Opioid agonists, such as cough syrups and mild painkillers, may have small dosages of an opioid ingredient (SI Appendix, Table S3), but are not considered strong enough for chronic opioid therapy and therefore not classified as or considered to be opioids. Benzodiazepines are relaxants used to treat, for example, alcohol withdrawal, anxiety, and panic disorders.

Variables with the smallest odds ratios (indicating decreased risk) were age 65+ y (0.13, indicating an almost complete, 87% reduction in risk), Hispanic ethnicity (0.41), age 55 to 64 y (0.43), African-American race (0.50), and missing marital status (0.51). Because we use modal marital status across all administrative sources, the missing indicator is likely a proxy for individuals who are enrolled only in Medicaid and not in other programs where marital status is reported.

Twenty-four of the strongest predictors are derived from Medicaid records. These include enrollment in managed care, number of unique Medicaid IDs, summary counts of distinct procedures and CCS diseases, total pharmacy payments, and three indicators for prior prescriptions. The remainder of the Medicaid predictors are diagnosis/procedure topics. Some of the significant themes among the selected topics with positive coefficients are drug/alcohol screening, back pain and injury, sprains and strains, contusions, psychotherapy, and depression and anxiety; and those with negative coefficients are asthma/allergies, breast cancer, gynecological examination, cholesterol screening, eyeglasses, dental evaluations, and intellectual disability.

Cost-Benefit Analysis. Whether the prescription diversion policy delivers benefits overall depends on whether it delivers benefits for those denied prescriptions. This in turn depends on how the parameters $\alpha$, $C_D$, $C_A$, and $C_i$ covary with $TP_i$. Assume for simplicity that $\alpha$, $C_D$, and $C_A$ do not vary across individuals and that $\alpha = 1$. Fig. 2 shows the break-even cost ratio $C_D/C_A$ at which the policy is cost neutral using predictive risk from the neural network model, with the green line assuming a diversion rate $\alpha = 1$ and homogeneous diversion and adverse outcome costs ($C_{D,i} = C_D$, $C_A = C_A$). In the top risk decile, the break-even ratio is 0.223: It is net beneficial to recommend against opioid prescriptions for individuals in the top decile if $C_D$ is less than 23.3% of $C_A$. It is net beneficial to intervene with the entire population if $C_D$ is less than 5.7% of $C_A$.

The existing literature provides guidance on reasonable estimates for $C_D$ and $C_A$, and we detail the calculation for an estimate of $C_A \approx 450,000$ (2010 dollars) in SI Appendix, section 5 and Table S4. Diversion costs are more difficult to quantify. They may include lost productivity due to chronic pain after receiving an alternative therapy, or they may include lost time due to requirements for more frequent monitoring of high-risk individuals by prescribing physicians. The economic cost of pain in the United States is conservatively estimated at $560$ to $635 billion, with a value of lost productivity from $299$ to $335 billion (23). Treating pain compassionately is a moral imperative for physicians, who must balance protecting those experiencing...
chronic pain with the significant risk of harm that opioids can cause individuals, their families, and their communities (24). However, recent research suggests that opioid therapy may not be more effective at pain relief than nonopioid therapy in both the short and the long term. A randomized trial comparing opioid therapy to nonopioid therapy for acute short-term pain found similar levels of pain relief between the two treatments (25), and observational studies also show no advantage for opioid treatment in terms of pain relief, with some patients on higher-potency opioids reporting more psychological impairment than those on lower-potency opioids (26, 27).

Estimates of costs of time are often calculated and utilized in the transportation literature. The value of time (VOT) has been estimated using stated-preference surveys as well as using revealed preference methodologies (28, 29). Typical VOT estimates are on the order of $30/h (30). Using a 2,000-h work year, the VOT estimate would correspond to a $60,000 annual loss in productivity if diversion costs resulted in loss of 1 y of full-time VOT.

This suggests that \( C_D \) is likely lower than the $104,400 break-even cost (23.2% of $450,000) for the top risk decile predicted by our model at \( \alpha = 1; \) $104,400 is above the 86th percentile of the annual earnings distribution in the United States in 2017 (31). Thus, a low risk threshold that maximizes true positives at the cost of increased false positives could be optimal. These findings support a growing belief among some within the medical community that the risks of opioid prescription outweigh the benefits in many cases of prescription outside of cancer or palliative care (32).

A benefit of structuring our cost–benefit analysis in terms of the cost ratio is that a risk threshold can be reevaluated as better data on these costs become available or as knowledge about opioid dependency improves. For example, the cost–benefit analysis represented by the green line in Fig. 2 assumes perfect prevention of dependency for predicted high-risk individuals as a result of the policy (\( \alpha = 1 \)). Individuals may still get access to an opioid through prescriptions given to friends and family. Approximately 10.7% of dependents (50.8% of 21.1% who did not themselves have a prescription) claim friends and family as the source of their first opioid (4), and diversion may still fail if those who do not receive a prescription subsequently borrow pills from others. An \( \alpha \) of 0.893 (the red line in Fig. 2) would assume that 10.7% of people go on to seek opioids from a friend or family, and true positives would then develop a dependency. In this case, the break-even costs for the top decile would be $95,400 (21.2% of $450,000).

Furthermore, high-risk individuals who are diverted to alternative therapies could have a higher rate of seeking and obtaining alternative opioid sources (e.g., \( \alpha_i \) and \( TP_i \) are negatively correlated). This may occur, for example, if opioid addiction is rational. Rational addiction models (33) predict that those seeking doctor prescriptions for opioids may be rationally seeking them prior to their first prescription to form an addiction as a fully informed, forward-looking, rational decision. Therefore, while restricting opioids may raise the cost of acquiring them and decrease the total number of prescribed opioids, diversion effectiveness may still be imperfect if those seeking prescriptions are making a rational choice and are therefore more likely to obtain opioids and develop a dependency even without a prescription.

To explore whether rational addiction may drive first-time prescriptions for opioids, we examine data on adverse outcomes as a function of patients’ degree of knowledge that they are receiving an opioid. We use the fact that patients may receive opioids through epidural or intravenous injections during inpatient procedures. Under the assumption that these opioid recipients were less likely to be informed they were receiving an opioid than those receiving and filling a prescription from a physician, we would expect fewer adverse outcomes from opioids received through inpatient procedures than through prescriptions in a rational addiction framework. We find that, when used as an explanatory variable for dependency while controlling flexibly for observable characteristics, an indicator for opioid injection is not significantly different from zero (SI Appendix, section 6 of 7)

---

Fig. 2. The break-even cost ratio for three values of the efficacy rate \( \alpha \). The break-even cost ratio is the point at which the hypothetical policy becomes cost neutral. If the diversion cost is less than this ratio times the adverse outcome cost (estimated at $450,000), then the policy will be net beneficial. Lower diversion costs are required to make the policy net beneficial among lower risk scores. Error bars indicate the 95% confidence interval calculated from 100 bootstrap replicates.
calculation shows that for the top decile we are powered to detect FDRs in the top-risk decile is 3.2% and insignificant. A power decile of our model. The difference between white and minor-
Hispanic) have a higher point estimate for FDR in the top-risk not a selected predictor of adverse outcomes.
There is no significant difference by disability status, and this was incarceration is a strong positive predictor of adverse outcomes. The previously divert cost accruing to individuals from marginalized public (35).

In general, trialing a policy and evaluating outcomes would allow policymakers and scientists to uncover individual-level parameter distributions by estimating heterogeneous treatment effects. This could allow policy to improve dynamically over time and eventually predict prescription restriction efficacy for diverting adverse outcomes.

**Fairness.** In addition to evaluating the overall cost–benefit trade-off of a prescription restriction policy, our framework can help policymakers examine measures of “fairness” by quantifying the extent to which policy costs versus benefits accrue disproportionately to marginalized groups. The predictive model’s false discovery rate (FDR) is defined as the fraction of false positives among all individuals who are predicted to have an adverse outcome. Differences in FDR across subgroups can occur when the model predictions $Y$ are not independent of subgroup membership conditional on the true outcomes $Y$, which is a construct for evaluating fairness that is well cited in the literature (36, 37). Here, we focus on FDR because it represents a notion of unfairness arising from a dispropor-
tionate diversion cost accruing to individuals from marginalized groups.

Fig. 3 shows the FDR by risk decile and by minority status, incarceration history, and disability status. The previously incarcerated have a significantly lower FDR, as release from incarceration is a strong positive predictor of adverse outcomes. There is no significant difference by disability status, and this was not a selected predictor of adverse outcomes.

Minority status is a negative predictor of adverse outcomes, all else equal. Members of minority groups (African-American, Hispanic) have a higher point estimate for FDR in the top-risk decile of our model. The difference between white and minority FDRs in the top-risk decile is 3.2% and insignificant. A power calculation shows that for the top decile we are powered to detect an 8.2 percentage point difference given our sample size (SI Appendix, Table S7). For the lower-risk deciles, the FDR difference becomes significant as a fraction of minorities in the subsample increases. The break-even diversion costs for whites and minorities in the top decile are $107,100 and $92,700 (SI Appendix, Fig. S4), which are above the 86th and 76th percentiles of the annual earnings distribution.

Thus, while the FDR is higher for minorities, restricting opioid prescriptions to those with high predicted risk may generate net benefits in minority and nonminority communities alike. It could be that diversion costs are higher for minorities than for nonminorities. In our data, minorities receiving an opioid prescription have roughly the same number of provider visits in the 30 d prior to an initial prescription, but live on average closer to providers, suggesting that diversion costs may not be substantially different across minority groups to negate overall predicted benefits from prevention policies (SI Appendix, section 8 and Table S8). Our predictive modeling and cost–benefit approach allows policy makers to quantify and weigh benefits and costs within and across subpopulations when designing a data-driven preventative policy.

**Discussion**

Prevention and treatment policies can be complementary approaches to opioid use disorders. Treatment can help the many individuals already suffering from adverse outcomes, while prevention can stem the growth of new cases of opioid dependence, abuse, or poisoning.

The proven standard treatment for opioid use disorder is medication-assisted treatment (MAT) (38–40). However, it faces two significant hurdles. First, MAT is not widely available to those with opioid use disorders; only 36% of substance abuse treatment facilities offer one of three different kinds of medication treatment (41). Second, even when those suffering from opioid use disorders can be connected to treatment, the costs associated with treatment are high and recovery from an opioid use disorder is challenging. The probability of recovery after a year of MAT is estimated at 50% (42).

Prevention strategies can help prevent further cases of opioid use disorder. Current strategies are primarily designed around reducing the quantity or potency of opioid prescriptions to curb misuse and prevent poisoning among those with existing opioid use disorders. They are especially complementary to a treatment approach. A recent study suggests that limiting

---

1For example, a major health insurer’s effort to reduce extended-release oxycodone pre-
scription by requiring prior authorization led to an increase in the rate of short-acting opioid prescriptions and no overall change in the total morphine milligram equivalents prescribed (43).
opioid availability for those with an existing disorder may increase the use of illicit drugs such as heroin.\(^6\)

The most widespread approach to preventing misuse by those with a disorder has been the deployment of prescription drug monitoring programs (PDMPs). These electronic data systems present data on the prescription history of controlled drugs to providers and are now in use in almost every state (40). They have been shown to reduce prescription rates of opioids and increase provider comfort in prescribing opioids, as providers can be reassured that they are not enabling risky opioid-dependency-related behaviors such as doctor shopping or receiving multiple overlapping prescriptions (45, 46).

These strategies are reactive rather than proactive; they target individuals who have already begun opioid treatment and have likely developed dependency. Our models complement these policies by providing an opportunity to predict high-risk prescriptions among the larger population of patients based on their characteristics and health histories before an opioid prescription is given for the first time. The models can be applied to the broader population of Medicaid enrollees, alerting physicians to possible risk when an opioid prescription is being considered, along with, for example, risk indicators of existing dependency from prior opioid prescription patterns.

Our models and hypothetical policy aim to prevent dependency before it occurs. This is complementary to existing efforts and could make use of the infrastructure already in place, such as the PDMPs. For example, a PDMP could implement our modeling approach to show providers a risk categorization for all patients (e.g., a red, yellow, or green indicator for predicted risk). This could increase information available to providers, expand the population covered by the PDMP, and help providers consider the benefits and risks of initiating opioid therapy with a new patient.

The information policy could be implemented without disclosing particular and potentially sensitive information about the individual not known to the physician. By determining a threshold based on rough high-risk/low-risk categories, it may be possible to both protect privacy and communicate valuable information to support health care professionals in determining the best course of treatment for their patients. For example, the mean rate of prior incarceration in the top two risk deciles is 9.7%, implying that being in the highest-risk deciles does not imply an individual is highly likely to have a prior criminal record.

Moreover, diversion costs may be small and effectiveness relatively high as the number of opioid prescriptions will be reduced, reducing the probability of unintended dependency. Once dependency occurs, MAT typically costs $6,552 to $14,112 annually (47) and is estimated to be effective 50% of the time (42). This means that for 1,000 individuals, it would cost $5.7 to $12.3 million over 3 y to bring 88% into remission (\textit{SI Appendix, Table S9}). Prevention is lower cost than treatment and can reduce treatment costs going forward by decreasing dependency rates. Given that PDMP platforms have been deployed in most states, distribution channels exist for converting the government’s own data into actionable intelligence accessible by physicians.

A limitation of our models is that they are trained on data from individuals to whom a physician decided to give an opioid prescription. We do not observe the cases where a patient requested an opioid prescription or had a condition that was treatable by opioid therapy, but the physician decided not to give an opioid prescription. In this sense, our models face a “selective labels problem” (48, 49), in which the data that can be observed are determined by prior human decisions whose decision rules are not known and may respond to the policy once implemented. For example, if, given the publicity of the opioid crisis, some physicians decreased opioid prescriptions, having a risk indicator could lead them to increase overall prescribing if they now feel more confident to prescribe given a low-risk indicator. Any implementation of a prescription restriction policy based on a predictive model should be accompanied by a causal analysis of impact. For example, assume the information policy is rolled out through the PDMP to a treatment of group of physicians or providers, but not to a control group. The causal impact could then be estimated for high- and low-risk patients, allowing inference on heterogeneous changes in prescribing behavior across types of physicians for patients with high-predicted versus low-predicted baseline risk to uncover how physician decision rules adapt to information. This could then support further improvements to the predictive model, for example predicting diversion success incorporating physician responses.

Our definition of adverse outcomes is limited by the accuracy of diagnosis codes in the Medicaid records. Prior studies have found that opioid-related diagnoses can be underreported because of their potential stigma. Although it is unknown precisely what fraction of opioid use disorder go undiagnosed, Carrell et al. (50) found that diagnosis codes were missing for as many as one-quarter of patients for whom their providers were aware of opioid abuse. Similarly, a study by Barocas et al. (51) estimated that only 44% of individuals with opioid use disorder were identified as such in claims and administrative records. To address this limitation, we added an adverse outcome based on procedure codes for the treatment of opioid use disorder, which could indicate an adverse outcome even in the absence of a diagnosis.

Including treatment as an indicator of adverse outcomes is also a limitation. As noted in prior work, receiving treatment for an opioid use disorder is a positive outcome conditional on already having a disorder (51, 52). However, the goal of this study is to suggest opportunities for prevention by examining whether individuals at a high risk of developing an adverse outcome can be identified with confidence before they are given a prescription using administrative data. This complements important research being done on successfully treating opioid use disorders after they have occurred (53).

Rhode Island has a research data lake that enables predictive modeling using cross-agency data. While any state or county could develop a similar research data lake (14, 15), restricting our predictive model to use only Medicaid claims and enrollment data yields nearly the same accuracy as models using integrated, cross-agency data. This is because, in the case of opioid dependence, Medicaid claims data contain many variables correlated with key predictors found in non-Medicaid data. For example, Medicaid enrollment data contain information on prior incarceration through payer codes related to receipt of health services while incarcerated, indicating an incarceration in the base period. They also contain data on demographics, family structure, and income from the application process. \textit{SI Appendix, Figs. S5–S7} replicate Figs. 1–3 using only data from Medicaid in the predictive model, with minimal changes in the results.

That being said, all models achieve an AUC near 0.800, indicating they have strong predictive power but could still be improved. While the Rhode Island data lake is uniquely rich in the connected and anonymized administrative records it holds, it contains only medical claims records from Medicaid. Those receiving a first prescription outside of Medicaid and developing a dependency diagnosed in Medicaid records, or vice versa, will cause decreased predictive accuracy in our model. Expanding the data to include, for example, state-wide electronic health records to examine impact on predictive fit, false positive rates,\(^6\) Abuse-deterrent reformulations of prescription opioids were developed to make it more difficult to crush or dissolve pills to release the drug more quickly. Unfortunately, recent evidence suggests that the introduction of abuse-deterrent prescription opioids into the market caused opioid abusers to substitute away from prescription opioids to heroin, with differential increases in fatal heroin poisonings (44).

\(^6\)\text{Abuse-deterrent reformulations of prescription opioids were developed to make it more difficult to crush or dissolve pills to release the drug more quickly. Unfortunately, recent evidence suggests that the introduction of abuse-deterrent prescription opioids into the market caused opioid abusers to substitute away from prescription opioids to heroin, with differential increases in fatal heroin poisonings (44).}
fairness, and cost–benefit analysis of diverting opioid prescriptions from those predicted to have high dependency risk is an important topic for future research.

Conclusion

The opioid epidemic is a complex public health challenge that requires policy solutions spanning prevention to treatment and recovery. Our results demonstrate the feasibility of an approach to prevent abuse based on intervening with high-risk initial prescriptions through predictive modeling. Our data-driven, machine-learning approach to modeling adverse outcome risk provides insights into the benefits, costs, and fairness of policies limiting opioid prescriptions. Intervening at the earliest stage, before an individual receives an initial opioid prescription, has the potential to prevent future treatment costs and recovery challenges and, ultimately, the life-long consequences of opioid use disorders.

ACKNOWLEDGMENTS. We thank the Smith Richardson Foundation and the Laura and John Arnold Foundation for financial support. We thank Miraj Shah for contributions to the project; Tom Corderre, Brandon Marshall, Susan Atthey, and participants at the National Bureau of Economic Research Conference on Machine Learning in Healthcare for helpful comments; and the Office of the Governor of Rhode Island and the Rhode Island Executive Office of Health and Human Services for supporting research to improve fact-based policymaking.