Predicting premature discontinuation of medication for opioid use disorder from electronic medical records

Ivan Lopez[1*], Sajjad Fouladvand[1*], Scott Kollins[2], Chwen-Yuen Angie Chen[3], Jeremiah Bertz[4], Tina Hernandez-Boussard[1], Anna Lembke[5], Keith Humphreys[5, 6], Adam S Miner[5], Jonathan H Chen[1$]

[1] Department of Medicine, Biomedical Informatics Research, Stanford Medicine, Stanford University, CA

[2] Holmusk, NY

[3] Primary Care and Population Health, Stanford Medicine, Stanford University, CA

[4] Center for the Clinical Trials Network, National Institute on Drug Abuse, MD

[5] Department of Psychiatry and Behavioral Sciences, Stanford Medicine, Stanford University, CA

[6] Veterans Affairs Palo Alto Health Care System, Palo Alto, CA

[*] Co-first authors.

[$] Corresponding author. Address: 3180 Porter Drive, Palo Alto, CA 94304. Email: jonc101@stanford.edu

Abstract

Medications such as buprenorphine-naloxone are among the most effective treatments for opioid use disorder, but limited retention in treatment limits long-term outcomes. In this study, we assess the feasibility of a machine learning model to predict retention vs. attrition in medication for opioid use disorder (MOUD) treatment using electronic medical record data including concepts extracted from clinical notes. A logistic regression classifier was trained on 374 MOUD treatments with 68% resulting in potential attrition. On a held-out test set of 157 events, the full model achieved an area under the receiver operating characteristic curve (AUROC) of 0.77 (95% CI: 0.64-0.90) and AUROC of 0.74 (95% CI: 0.62-0.87) with a limited model using only structured EMR data. Risk prediction for opioid MOUD retention vs. attrition is feasible given electronic medical record data, even without necessarily incorporating concepts extracted from clinical notes.

Introduction

Opioid use disorder (OUD) is a major public health concern in the United States, with approximately 2.4 million Americans suffering from the condition.[1] The consequences of OUD can be severe if left untreated, with morbidity and mortality often resulting from an opioid overdose.[2] In order to address this issue and improve outcomes for those with OUD, it is important to identify and address the factors that contribute to treatment dropout, specifically for medication for opioid use disorder (MOUD). MOUD is a widely accepted and evidence-based treatment approach for OUD and is recommended by the World Health Organization and approved by the US Food and Drug Administration.[3] One such MOUD is buprenorphine-naloxone, also known as Suboxone, which is a partial opioid receptor agonist fused with an opioid antagonist.[4] Buprenorphine-naloxone is approved for the maintenance treatment of opioid dependence and has been associated with favorable clinical outcomes, including reduced illicit drug use, lower risk of overdose, reduced crime, and higher treatment retention.[5] Despite the effectiveness of buprenorphine-naloxone and other MOUD medications, premature treatment discontinuation remains a major obstacle for many patients with OUD.[6] This can be due to a variety of factors, including social and environmental barriers, psychological and emotional challenges, and issues related to access and availability of treatment. As a result, treatment retention is often poor, and patients may struggle to maintain long-term recovery from OUD.

Previous research has demonstrated the efficacy of machine learning and deep learning models in accurately predicting OUD by leveraging electronic health record (EHR) data.[7–11] These models hold

promise for enhancing clinical decision-making by enabling early intervention and prevention of OUD. Nonetheless, there remains a gap in our capacity to determine which patients on MOUD are at the greatest risk of being lost to follow-up. To address this issue, there have been previous efforts to examine the attrition patterns of OUD patients who were treated with buprenorphine-naloxone and assess the association between clinical, sociodemographic, and medication dosing features with attrition.[12] However, predicting premature discontinuation of MOUD is poorly characterized in the literature emphasizing the importance of providing clinicians with this knowledge to facilitate informed treatment decisions and personalized care plans tailored to the specific risk profiles of different patient populations seeking OUD treatment. In order to address this gap, we use both structured and unstructured data within electronic medical records of patients receiving buprenorphine-naloxone treatment to predict which patients are likely to prematurely discontinue MOUD. Early identification of patients at risk of treatment discontinuation may aid early intervention efforts, thus improving treatment adherence and outcomes. Our objective is to determine whether retrospective EMR data is predictive of which MOUD-treated patients will be retained in ongoing therapy. Additionally, we evaluate whether natural language processing of unstructured clinical notes improves identifying patients at risk for treatment attrition.

Implementation of such tools would have significant public health implications, including risk-adjusted outcome assessments of different treatment centers in collaboration with organizations such as the National Institute on Drug Abuse Clinical Trials Network (NIDA CTN), or organizations tasked with analyzing the state of the opioid crisis and proposing solutions to it domestically while attempting to stop its spread internationally.[13,14] These advancements may significantly enhance our ability to identify and address barriers to treatment retention for patients with OUD and aid development of targeted, personalized treatment plans, improving treatment retention and ultimately leading to better outcomes.

Materials and Methods

Data Source

The present study employed a retrospective analysis of EHR data obtained from an academic medical center. The data encompassed patients who were treated for opioid use disorder at an academic medical center between 2009 and 2022. All identifiable information was removed or obscured from the structured data, utilizing the Safe Harbor approach as per NIST guidelines. The unstructured data, comprising clinical text, was de-identified via TiDE, a text de-identification algorithm.[15] The study included all individuals aged 18-89 who were prescribed buprenorphine-naloxone medication for a duration exceeding one day at the institution. Approval for the study was obtained from the Institutional Review Board (IRB) at our academic medical center.

Predictors

Labels for the data were created using the length of a “drug era” for buprenorphine-naloxone treatment. A drug era is a period of time during which a patient is continuously exposed to buprenorphine-naloxone. A drug era can be defined based on the start and end dates of drug exposure for a patient. During a drug era, a patient may stop taking the drug temporarily and then start again. If this happens, the original drug era will be extended to include the time during which the patient was not taking the drug if the drug hiatus is no less than 30 days in length. If a patient starts taking a different drug or set of drugs, a new drug era will begin with a new start date. Furthermore, the categorization of drug eras was bifurcated through two distinct approaches: (1) A stringent demarcation based on their Observational Medical Outcomes Partnership (OMOP) drug ID, wherein comparable drugs could retain discrete identities within different drug eras if variations existed in their dosage or administration route; alternatively, (2) Drug eras had the flexibility for consolidation if they shared analogous active compounds, irrespective of differences in dosage or administration route. Our experimental emphasis centered on the utilization of approach (1); however, outcomes for approach (2) were also reported.

Features for a drug era were selected by taking all diagnosis, procedure, drug prescription, and note keyword data recorded 90 days prior to the start of a drug era. The resulting feature matrix for our 531 drug eras contained 2,791,773 features composed of 17,961 diagnosis features, 17,271 procedure

features, 47,476 drug features, 2,709,065 note keyword features, and 10 demographic features. Dimensionality reduction was performed to reduce the feature matrix space. Our dimensionality reduction method was based on a statistically significant cutoff between feature differences in the attrition and retention groups using Fisher's exact test. Features that show a statistically significant difference in their distribution between the two groups are retained, while those that do not show a significant difference are discarded. The final feature matrix contained 323 features composed of 27 diagnosis features, 37 procedure features, 57 drug features, 192 note keyword features, and 10 demographic features.

Outcome and Settings

To label a drug era as an attrition or retention event, we used a 180-day cut-off where drug eras less than recommended 180 days were considered MOUD attrition events. Research showed that patients that discontinued MOUD within the first 180 days after starting it, are at higher risk of relapse than those who discontinued treatment after 180 days.[16–18] The 180-day cut-off has been used in prior clinical trials and is a way to identify individuals who may be at a particularly high risk of relapse and may benefit from additional support or re-initiation of MOUD.

The model was trained on data from 12/30/2009 to 11/07/2020 (training set, 374 drug eras), and tested retrospectively on data from 11/08/2020 to 07/03/2022 (test set, 157 drug eras). Table 1 shows descriptive analyses of our train and test sets. Note, training and testing datasets have no patients in common to avoid data leakage. L1-regularized logistic regression (Lasso) was used for feature selection and prediction. Lasso's hyperparameter (lambda) was tuned using 10-fold cross-validation on 70% of the training set to maximize the area under the ROC curve (AUC).

Table 1: Patient characteristics and clinical characteristics for the training set, and retrospective test set.

Metric	Train set	Test set
Event-level attributes		
Total N distinct events	374	157
N attrition events (%)	68%	92%
Median N events per patient (95% CI)	1 (CI: 1-1)	1 (CI: 1-1)
Median age (years) at drug era start date	48 (CI: 45-51)	54 (46-60)
Age (years) at drug era start date		
18-25	24	17
26-45	140	46
46-60	112	27
>60	98	67
Patient drug eras		
Median drug era length (95% CI)	80 (CI: 67-91)	38 (CI: 28-42)
Patient diagnosis		

Total diagnosis features	29	
Median N distinct diagnosis (95% CI)	0 (CI: 0-1)	1 (CI: 1-2)
Patient drug prescriptions		
Total drug features	59	
Median N distinct drug prescriptions (95% CI)	1 (CI: 1-1)	4 (CI: 2-7)
Patient procedures		
Total procedure features	39	
Median N distinct procedures (95% CI)	1 (CI: 0-1)	2 (CI: 0-8)
Patient note features		
Total note features	194	
Median N distinct note features (95% CI)	9 (CI: 7-13)	25 (CI: 17-41)
Patient-level attributes		
Total N distinct patients	297	113
Gender		
Female	143	58
Male	154	55
Race		
American Indian or Alaska Native	1	2
Asian	9	10
Black or African American	20	7
Native Hawaiian or other Pacific Islander	1	1
White	234	78
Not Reported	32	15
Ethnicity		
Hispanic or Latino	25	19
Not Hispanic or Latino	266	91

Not Reported	6	3

Results

Area under the receiver operator curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were used to evaluate model discrimination. 95% confidence intervals were calculated using Clopper-Pearson for sensitivity, specificity, PPV, and NPV, and using the DeLong test for AUC.[19] Calibration in the mean and measures of weak and moderate calibration (calibration slope, calibration intercept, and calibration curves) were calculated for the retrospective test set. Table 2 shows the model's prediction performance results. On the retrospective test set, the model had an AUC of 0.77 (95% CI: 0.64-0.90), sensitivity of 0.91 (95% CI: 0.85-0.95), specificity of 0.29 (95% CI: 0.08-0.58), PPV of 0.93 (95% CI: 0.87-0.97), and NPV of 0.24 (95% CI: 0.07-0.5). The calibration slope was 1.10 and the calibration intercept was 1.34 before recalibration. After Platt scaling, the model had an AUC of 0.77 (CI: 0.64-0.90), sensitivity of 0.99 (CI: 0.96-1), specificity of 0.14 (CI: 0.02-0.43), PPV of 0.92 (CI: 0.87-0.96), and NPV of 0.67 (CI: 0.09-0.99), calibration slope of 1 and an intercept of 0. Drug era method (2) yields an AUC of 0.67 (CI: 0.55-0.76) before and after Platt scaling.

Further, we conducted an ablation study to explore the effect of using unstructured data on the model's predictive power. We trained and tested our model using only structured EMR data (excluding features extracted from patients' unstructured clinical notes) and the model's prediction performance (AUC) was 0.74 (95% CI: 0.62-0.87).

Table 2: Model performance for the retrospective test set (n = 157) before and after Platt scaling.

Metric	Retrospective test set	Calibrated retrospective test set
AUC	0.77 (CI: 0.64-0.90)	0.77 (CI: 0.64-0.90)
Threshold for binary classification	0.56	0.56
Sensitivity	0.91 (CI: 0.85-0.95)	0.99 (CI: 0.96-1)
Specificity	0.29 (CI: 0.08-0.58)	0.14 (CI: 0.018-0.43)
PPV	0.93 (CI: 0.87-0.97)	0.92 (CI: 0.87-0.96)
NPV	0.24 (CI: 0.07-0.50)	0.67 (CI: 0.094-0.99)
Calibration slope	1.10	1
Calibration intercept	1.34	0
AUC (Drug Era Method 2)	0.67 (CI: 0.55-0.76)	0.67 (CI: 0.55-0.76)

Predictive features

Table 3 shows feature coefficients computed by our model. The presence of a diagnosis related to sedative, hypnotic, and/or anxiolytic use disorder was the most predictive feature for attrition in our model, followed by the presence of a muscle pain diagnosis, the administration of the medications cyclobenzaprine hydrochloride, pregabalin, and mentions of 'very difficult' and 'restlessness' in clinical notes. Furthermore, the presence of a peripheral nerve disease diagnosis, a disorder of the sacrococcygeal

spine diagnosis, and the administration of the medications prazosin, cetirizine hydrochloride, mirtazapine, and rheumatoid arthritis diagnosis were the most predictive features for MOUD retention.

Table 3: Top 6 most predictive features of attrition and retention for MOUD.

	EHR concept predictive for attrition	Group	Coefficient
1	Sedative, hypnotic AND/OR anxiolytic-related disorder	diagnosis	1.0343
2	Muscle pain	diagnosis	0.8916
3	cyclobenzaprine hydrochloride 10 MG Oral Tablet	drug	0.8716
4	pregabalin 50 MG Oral Capsule	drug	0.8595
5	Very difficult	note_nlp	0.5657
6	Restlessness	note_nlp	0.5572
	EHR concept predictive for retention	**Group**	**Coefficient**
1	mirtazapine 30 MG Oral Tablet	drug	2.7958
2	cetirizine hydrochloride 10 MG Oral Tablet	drug	2.7743
3	Disorder of sacrococcygeal spine	diagnosis	2.0967
4	Peripheral nerve disease	diagnosis	2.0909
5	Rheumatoid arthritis	diagnosis	1.8261
6	prazosin 2 MG Oral Capsule	drug	1.7390

Discussion

MOUD combines non-pharmacological therapies, such as counseling or cognitive behavioral therapy, with FDA-approved medications to manage OUD as a chronic disorder. Consistent with the management of other chronic conditions, such as diabetes, treatment plans for OUD are personalized for each patient and formulated in collaboration with the patient, the prescriber, and other members of the healthcare team.[20] Many patients will benefit from staying on MOUD for the rest of their lives whereas others will eventually choose to slowly taper off in a closely monitored fashion.[20] Sudden treatment discontinuation has a poor prognosis and is difficult to manage given the lack of resources available to clinicians to identify patients who are most vulnerable to prematurely discontinuing therapy.

In this study, we proposed an automated prediction model that can anticipate when a patient is at risk of discontinuing MOUD. By identifying patients who are at high risk of loss to follow-up, clinicians can proactively intervene with personalized care and support, such as increasing the frequency of appointments or providing additional resources to address barriers to treatment adherence. This approach allows for a more individualized and patient-centered approach to care, where providers can address the unique needs and challenges of each patient. Additionally, the use of predictive models can reduce the incidence of treatment discontinuation, improving treatment outcomes and reducing the risk of relapse and other adverse events, such as overdose death, associated with untreated OUD.[21–25]

We believe our prediction model is a step in the right direction toward giving clinicians the ability to accurately predict which patients are going to discontinue MOUD. More specifically, our model reports

a sensitivity of 0.99, suggesting that we can correctly identify 99% of patients who are at risk of discontinuing MOUD. Additionally, our PPV of 0.92 suggests that out of all the patients who are predicted to discontinue therapy by the model, 92% of them will actually be lost to follow-up. Both of these measures are crucial in identifying patients who may require additional support or interventions. However, sensitivity and PPV alone may not provide a complete picture of the model's performance. Additionally, the specific context and consequences of false positives and false negatives should also be taken into account when evaluating the model's performance. Nonetheless, high sensitivity and PPV are promising results and suggest that the model is effective at identifying patients who may require additional support to remain engaged in life-saving MOUD.

Further, our experimental results suggest that individuals with a sedative, hypnotic, and/or anxiolytic-related disorder diagnosis may be at higher risk of attrition in MOUD, as indicated by the large positive coefficient associated with this feature. Additionally, the presence of physical pain or discomfort, as indicated by the other predictive features, may also be associated with an increased risk of attrition. On the other hand, individuals with a peripheral nerve disease diagnosis, a disorder of the sacrococcygeal spine diagnosis, and those who were prescribed prazosin, cetirizine hydrochloride, mirtazapine, or had a rheumatoid arthritis diagnosis had large negative coefficients, suggesting that they may be more likely to be retained in MOUD. Features selected by LASSO may be associated with the outcome, but this association does not imply causality.

Despite the rigorous methodology, our study comes with its limitations. Our study lacks sufficient post-COVID era training and testing data. The COVID-19 pandemic has brought about significant changes in clinical workflows and patient life. Patients with OUD have been severely affected, experiencing more complications than non-OUD patients and reduced access to OUD treatment.[26,27] COVID-19 patients with OUD had higher hospitalization odds, longer maximum lengths of stay, and greater odds of invasive ventilator dependence than those without OUD.[26] In addition, US counties with higher OUD mortality rates also had significantly higher rates of COVID-19 mortality.[28] As the pandemic subsides, the impact on MOUD and OUD patients could have changed over time. This data set shift can have significant implications for our machine learning model, where our training data is on pre-COVID and COVID-19 era data. Thus, proper application of model monitoring would be necessary to evaluate the model's performance in real-world clinical settings. Furthermore, our model's NPV was recorded at 0.67. This suggests the model is more likely to generate false negatives, i.e., predict that a patient will not be lost to follow-up when they actually will. In the context of a life-saving MOUD, patients who are at high risk of discontinuing therapy may not receive the necessary intervention or support. Therefore, it is important for future iterations of our model to utilize methods such as utility-based probability thresholding to properly describe the highest costs of false negatives relative to false positives for our prediction task. This will improve NPV to ensure that high-risk patients are identified and provided with appropriate support and interventions.Another limitation includes our inability to classify different types of attrition from MOUD therapy due to the costliness of manually labeled data. Attrition can result from a multitude of factors, ranging from patient preference to external circumstances. Our definition of attrition could include instances of attrition from MOUD therapy other than discontinuation due to a patient's choice. Lastly, while our cohort was centered on drug exposure events related to buprenorphine-naloxone, it is worth noting that there are several other medications available as MOUD for opioid use disorder. Future versions of our model should include a more extensive list of these medications to ensure that our results have a broader reach while also avoiding erroneous classifications of lost-to-follow-up cases in our training and testing data. In reality, patients may have transitioned from buprenorphine-naloxone to a different MOUD, and a more comprehensive medication list would enable us to capture these cases more accurately. However, our choice to focus on buprenorphine-naloxone exposures improves the likelihood our patients are on MOUD.

Our model's AUC of 0.77 indicates that the machine learning binary classifier has discriminatory power. Our ultimate aim is to develop a robust machine learning model that can identify patients at risk of premature MOUD discontinuation with precision while ensuring that the model's performance remains consistent across diverse patient populations and healthcare institutions. An intriguing avenue for future

research is to explore the potential of unstructured clinical text data and large language models in enhancing the model's predictive power. In this paper, we included unstructured clinical text data in the form of Note-NLP from OMOP's data structure. However, Note-NLP features can lead to redundancy across all information features. As a result, the addition of these data did not improve our model's performance. To accurately evaluate whether unstructured clinical text can enhance our ability to predict a loss to follow-up, comprehensive comparison using a state-of-the-art language model is merited. This will enable us to more effectively determine the impact of including unstructured clinical text in our feature matrix and whether it can provide valuable insights that may not be captured by structured data alone. Lastly, the patient population of our academic medical center might not be representative of the broader population due to factors such as geographic location, socioeconomic status, and healthcare access. Consequently, the model's predictions might be skewed towards the characteristics of this specific patient population. In the future, we plan on conducting multi-site training and prospective testing of the model, which includes federated studies on social determinants of health outcomes. Collaborating with multiple institutions to explore the influence of social determinants on health outcomes would provide a richer and more comprehensive understanding of the complexities involved in attrition from MOUD with the added benefit of assessing our model's robustness, generalizability, and performance.

Conclusions

Our findings show that retrospective EMR data can effectively predict opioid MOUD attrition and retention and that additional unstructured information in clinical notes does not significantly improve predictive performance.

Acknowledgment

This publication was made possible by Grant Number UG1 DA 012345 from the National Institute on Drug Abuse (NIDA). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Disclosures

Work supported by NIH/National Institute on Drug Abuse Clinical Trials Network (UG1DA015815 - CTN-0136)

JHC reports additional support from
Stanford Artificial Intelligence in Medicine and Imaging - Human-Centered Artificial Intelligence (AIMI-HAI) Partnership Grant
Doris Duke Charitable Foundation - Covid-19 Fund to Retain Clinical Scientists (20211260)
Google, Inc. Research collaboration Co-I to leverage EHR data to predict a range of clinical outcomes.
American Heart Association - Strategically Focused Research Network - Diversity in Clinical Trials

This research used data or services provided by STARR, STAnford medicine Research Data Repository," a clinical data warehouse containing live Epic data from Stanford Health Care (SHC), the University Healthcare Alliance (UHA) and Packard Children's Health Alliance (PCHA) clinics and other auxiliary data from Hospital applications such as radiology PACS. The STARR platform is developed and operated by the Stanford Medicine Research IT team and is made possible by the Stanford School of Medicine Research Office.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or Stanford Healthcare.

Jonathan H. Chen reports being co-founder of Reaction Explorer LLC which develops and licenses organic chemistry education software and having been paid consulting fees from Sutton Pierce and Younker Hyde MacFarlane PLLC as a medical expert witness.

References

1. Degenhardt L, Charlson F, Ferrari A, et al. The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry* 2018;5:987–1012.
2. Hser Y-I, Hoffman V, Grella CE, et al. A 33-Year Follow-up of Narcotics Addicts. *Arch Gen Psychiatry* 2001;58:503.
3. Research C for DE and. Information about Medication-Assisted Treatment (MAT). *FDA*. Available at https://www.fda.gov/drugs/information-drug-class/information-about-medication-assisted-treatment-mat. 2019, Accessed February 28, 2023.
4. Connery HS. Medication-Assisted Treatment of Opioid Use Disorder: Review of the Evidence and Future Directions. *Harvard Review of Psychiatry* 2015;23:63–75.
5. Simoens S, Matheson C, Bond C, et al. The effectiveness of community maintenance with methadone or buprenorphine for treating opiate dependence. *Br J Gen Pract* 2005;55:139–46.
6. Brorson HH, Ajo Arnevik E, Rand-Hendriksen K, et al. Drop-out from addiction treatment: A systematic review of risk factors. *Clinical Psychology Review* 2013;33:1010–24.
7. Dong X, Deng J, Rashidian S, et al. Identifying risk of opioid use disorder for patients taking opioid medications with deep learning. *Journal of the American Medical Informatics Association* 2021;28:1683–93.
8. Fouladvand S, Talbert J, Dwoskin LP, et al. Identifying Opioid Use Disorder from Longitudinal Healthcare Data using a Multi-stream Transformer. *AMIA Annu Symp Proc* 2022;2021:476–85.
9. Cochran BN, Flentje A, Heck NC, et al. Factors predicting development of opioid use disorders among individuals who receive an initial opioid prescription: Mathematical modeling using a database of commercially-insured individuals. *Drug and Alcohol Dependence* 2014;138:202–8.
10. Reps JM, Cepeda MS, Ryan PB. Wisdom of the CROUD: Development and validation of a patient-level prediction model for opioid use disorder using population-level claims data. *PLoS ONE* 2020;15:e0228632.
11. Ripperger M, Lotspeich SC, Wilimitis D, et al. Ensemble learning to predict opioid-related overdose using statewide prescription drug monitoring program and hospital discharge data in the state of Tennessee. *Journal of the American Medical Informatics Association* 2021;ocab218.
12. Ker S, Hsu J, Balani A, et al. Factors That Affect Patient Attrition in Buprenorphine Treatment for Opioid Use Disorder: A Retrospective Real-World Study Using Electronic Health Records. *Neuropsychiatr Dis Treat* 2021;17:3229–44.
13. Tai B, Dobbins R, Blackeney Q, et al. The NIDA clinical trials network: evolving, expanding, and addressing the opioid epidemic. *Addiction Science & Clinical Practice* 2021;16:28.
14. Humphreys K, Shover CL, Andrews CM, et al. Responding to the opioid crisis in North America and beyond: recommendations of the Stanford–Lancet Commission. *The Lancet* 2022;399:555–604.
15. Datta S, Posada J, Olson G, et al. A new paradigm for accelerating clinical data science at Stanford Medicine. Epub ahead of print March 17, 2020. DOI: 10.48550/arXiv.2003.10534.
16. Pharmacotherapy for Opioid Use Disorder. *NCQA*. Available at https://www.ncqa.org/hedis/measures/pharmacotherapy-for-opioid-use-disorder/. Accessed April 24, 2022.
17. Williams AR, Nunes EV, Bisaga A, et al. Development of a Cascade of Care for responding to the opioid epidemic. *The American Journal of Drug and Alcohol Abuse* 2019;45:1–10.
18. Jones CM, Shoff C, Hodges K, et al. Receipt of Telehealth Services, Receipt and Retention of Medications for Opioid Use Disorder, and Medically Treated Overdose Among Medicare Beneficiaries Before and During the COVID-19 Pandemic. *JAMA Psychiatry* 2022;79:981–92.
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
20. Medications for Substance Use Disorders. Available at https://www.samhsa.gov/medications-substance-use-disorders. Accessed February 23, 2023.

21. Zhang Z, Friedmann PD, Gerstein DR. Does retention matter? Treatment duration and improvement in drug use. *Addiction* 2003;98:673–84.
22. Mintzer IL, Eisenberg M, Terra M, et al. Treating opioid addiction with buprenorphine-naloxone in community-based primary care settings. *Ann Fam Med* 2007;5:146–50.
23. Armstrong G, Kermode M, Sharma C, et al. Opioid substitution therapy in manipur and nagaland, north-east india: operational research in action. *Harm Reduct J* 2010;7:29.
24. Kakko J, Svanborg KD, Kreek MJ, et al. 1-year retention and social function after buprenorphine-assisted relapse prevention treatment for heroin dependence in Sweden: a randomised, placebo-controlled trial. *Lancet* 2003;361:662–8.
25. Davoli M, Perucci CA, Forastiere F, et al. Risk factors for overdose mortality: a case-control study within a cohort of intravenous drug users. *Int J Epidemiol* 1993;22:273–7.
26. Qeadan F, Tingey B, Bern R, et al. Opioid use disorder and health service utilization among COVID-19 patients in the US: A nationwide cohort from the Cerner Real-World Data. *eClinicalMedicine*;37. Epub ahead of print July 1, 2021. DOI: 10.1016/j.eclinm.2021.100938.
27. Currie JM, Schnell MK, Schwandt H, et al. Prescribing of Opioid Analgesics and Buprenorphine for Opioid Use Disorder During the COVID-19 Pandemic. *JAMA Network Open* 2021;4:e216147.
28. Qeadan F, Mensah NA, Tingey B, et al. The association between opioids, environmental, demographic, and socioeconomic indicators and COVID-19 mortality rates in the United States: an ecological study at the county level. *Archives of Public Health* 2021;79:101.