# Quantifying the opioid use disorder crisis: PULSNAR finds nearly 3/4 undiagnosed

**Praveen Kumar[1], Fariha Moomtaheen[1], Scott A. Malec[1], Jeremy J. Yang[1], Cristian G. Bologa[1], Kristan A Schneider[1], Yiliang Zhu[1], Mauricio Tohen[2], Gerardo Villarreal[2,3], Douglas J. Perkins[1], Elliot M. Fielstein[4], Sharon E. Davis[4], Michael E. Matheny[4,5], Christophe G. Lambert[1]**

**[1]University of New Mexico, Department of Internal Medicine, Albuquerque, NM, USA**
**[2]University of New Mexico, Department of Psychiatry & Behavioral Sciences, Albuquerque, NM, USA**
**[3]VA New Mexico Healthcare System, Albuquerque, NM, USA**
**[4]Vanderbilt University Medical Center, Department of Biomedical Informatics, Nashville, TN, USA**
**[5]Tennessee Valley Healthcare System VA, Nashville, TN, USA**

## Background

The opioid crisis continues to be a significant global public health challenge.[1] In the US, 107,941 drug overdose deaths occurred in 2022, with opioids contributing to 81,806 (75.8%) of these fatalities.[2] The economic burden associated with opioid use disorder (OUD) and fatal opioid overdoses in the US was estimated at $1.02 trillion in 2017 (5.25% of the GDP),[3] escalating to nearly $1.5 trillion in 2020 (7.12% of the GDP).[4]

Accurate estimation and diagnosis of OUD is essential for identifying individuals at risk, assessing treatment needs, monitoring prevention and intervention efforts, and recruiting treatment-naive participants for clinical trials. However, OUD is substantially underdiagnosed and undercoded in electronic health records (EHRs) and claims data.[5] This poses a significant challenge in estimating the prevalence of OUD, and in applying cutting-edge machine learning (ML) techniques to model patient outcomes.

To address the issues of underdiagnosis and undercoding of OUD, our study employs a novel Positive and Unlabeled (PU) machine learning (ML) approach, termed "*Positive Unlabeled Learning Selected Not At Random (PULSNAR)",*[6] to estimate the proportion of OUD among undetected individuals. Furthermore, we utilized SHapley Additive exPlanations (SHAP)[7] values to analyze the relationships between important features and outcomes to understand the underlying risk factors and potential predictors of OUD. To the best of our knowledge, this is the first study to apply PU learning to opioid-related data to estimate the prevalence of undercoding and predict OUD.

## Materials and Methods

This study used the US MarketScan Commercial Claims and Encounters (CCAE, 2017-2021) database for cohort and feature selection mapped to the OMOP common data model.[8] The study cohort comprised individuals with a minimum of two years of observation after January 1, 2017, and exposure to at least one of the 37 opioid medications throughout their enrollment period (e.g., morphine, oxycodone). The OUD phenotype was defined using the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes F11.1*, F11.2*, and F11.93. Individuals with these codes were labeled positive cases (class 1), while those without were considered unlabeled (class 0). A total of 94,668 unique covariates were identified for the ML model, including age groups, sex, and two feature classes: Conditions (ICD-10-CM and ancestors) and Drug exposures (RxNorm) at an ingredient level. For positive cases, covariates were selected from the period preceding the first OUD coding date, while for unlabeled cases, all conditions and drugs during the observation period were included. This process resulted in 45,019 positive and 3,297,025 unlabeled examples.

The dataset had a significant class imbalance, with the ratio of unlabeled to positive examples being approximately 73 to 1. To address this, we created 73 balanced datasets. Each balanced dataset included all the positive (labeled) examples and a similar number of unlabeled examples, which were sampled without replacement from the original unlabeled set. We then employed the PULSNAR algorithm to estimate the proportion ($\alpha$) of OUD among uncoded examples and calibrate predictions for the unlabeled examples in each balanced dataset. The unlabeled examples were then sorted in descending order of the calibrated predictions, and the top $\alpha$ × (# unlabeled in the balanced set) examples were selected as probable positives. For the PULSNAR method, we used XGBoost[9] as a classifier to estimate the predictions. Figure 1 illustrates the complete process for calculating $\alpha$ using the PULSNAR method and generating calibrated predictions for uncoded individuals where the estimates are expected to accurately reflect the true likelihood of OUD.

We examined the SHAP plot for the top 15 features identified by XGBoost, selected based on their average gain score across 73 balanced dataset models. The SHAP plot provided insights into the relationships between the covariates and the predicted outcomes, thereby enhancing the model's interpretability.

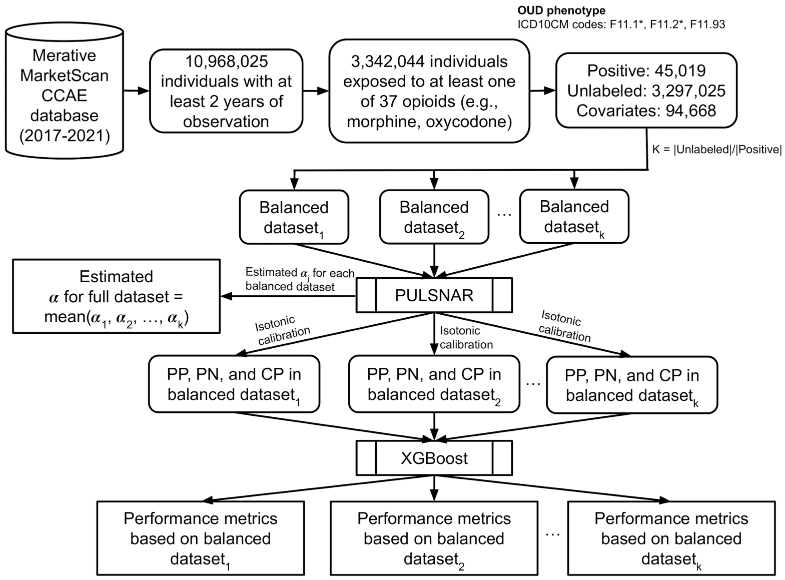
OUD phenotype
ICD10CM codes: F11.1*, F11.2*, F11.93
Merative MarketScan CCAE database (2017-2021)
10,968,025 individuals with at least 2 years of observation
3,342,044 individuals exposed to at least one of 37 opioids (e.g., morphine, oxycodone)
Positive: 45,019
Unlabeled: 3,297,025
Covariates: 94,668
K = |Unlabeled|/|Positive|
Balanced dataset$_1$
Balanced dataset$_2$
...
Balanced dataset$_k$
Estimated $\alpha_i$ for each balanced dataset
PULSNAR
Estimated $\alpha$ for full dataset = mean($\alpha_1$, $\alpha_2$, ..., $\alpha_k$)
Isotonic calibration
Isotonic calibration
Isotonic calibration
PP, PN, and CP in balanced dataset$_1$
PP, PN, and CP in balanced dataset$_2$
...
PP, PN, and CP in balanced dataset$_k$
XGBoost
Performance metrics based on balanced dataset$_1$
Performance metrics based on balanced dataset$_2$
...
Performance metrics based on balanced dataset$_k$

**Figure. 1. Steps to estimate the proportion of uncoded OUD using PULSNAR.**
PP: Probable positives identified by PULSNAR; PN: Probable negatives identified by PULSNAR; CP: coded positives

## Results

In the study cohort of 3,342,044 individuals exposed to opioid medication, only 45,019 cases (1.35%) were coded or diagnosed with OUD. However, applying the PULSNAR method identified an estimated 124,723 additional cases of undiagnosed OUD, representing 3.78% of individuals who were not initially coded for the disorder (imputed OUD, 95% CI: [3.76%, 3.80%]). Consequently, the overall cumulative prevalence of OUD among patients who received opioid medication over an average of 3.39 years of observation, was 5.08% across all age groups and sexes. This estimate combines both diagnosed and imputed undiagnosed cases, with 74.3% of the cases being imputed. The average age of patients with coded OUD was 42 years (standard deviation: 13), whereas the average age for those without coded OUD was 38 years (standard deviation: 15). Figure 2 presents the mean $\alpha$ estimates from the PULSNAR method for each iteration across 73 balanced datasets, along with the 95% confidence intervals (CIs).

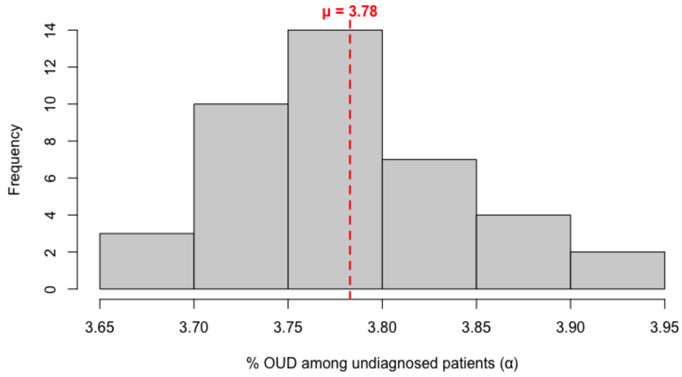

**Figure. 2. Distribution of α estimates by PULSNAR method.** Each iteration had 73 α estimates, each corresponding to one of the 73 balanced datasets. Red line: mean α value

The mean gain scores of the top 15 important covariates returned by the XGBoost model across all 73 balanced datasets are shown in Figure 3. These top covariates contributed most to the prediction (positive class or negative class) by the XGBoost model.

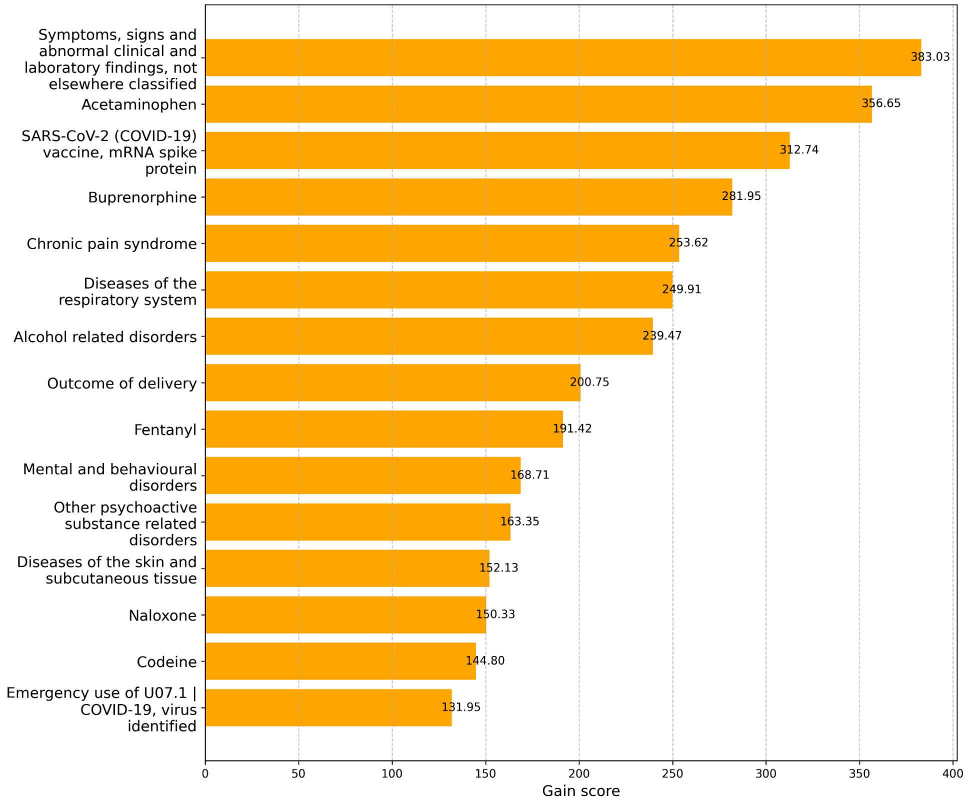

**Figure. 3. Gain scores for the top 15 features distinguishing patients with OUD**. The gain score represents the mean gain across 73 balanced datasets and 40 iterations. Higher gain scores mean the feature produced more separation between child nodes across the models.

Figure 4 displays the SHAP plot for the top 15 features identified by the XGBoost model. In this plot, red indicates high feature values, while blue indicates low feature values. The plot shows that individuals with high values for certain features, such as acetaminophen, buprenorphine, chronic pain syndrome, other psychoactive substance-related disorders, and naloxone, are likely to be predicted as having OUD by the model. The wide spread of a color (red/blue) for a feature highlights the varying effects of the feature on model predictions across different individuals. Conversely, a dense clustering of a color (red/blue) suggests that the feature consistently affects model predictions similarly for different individuals.

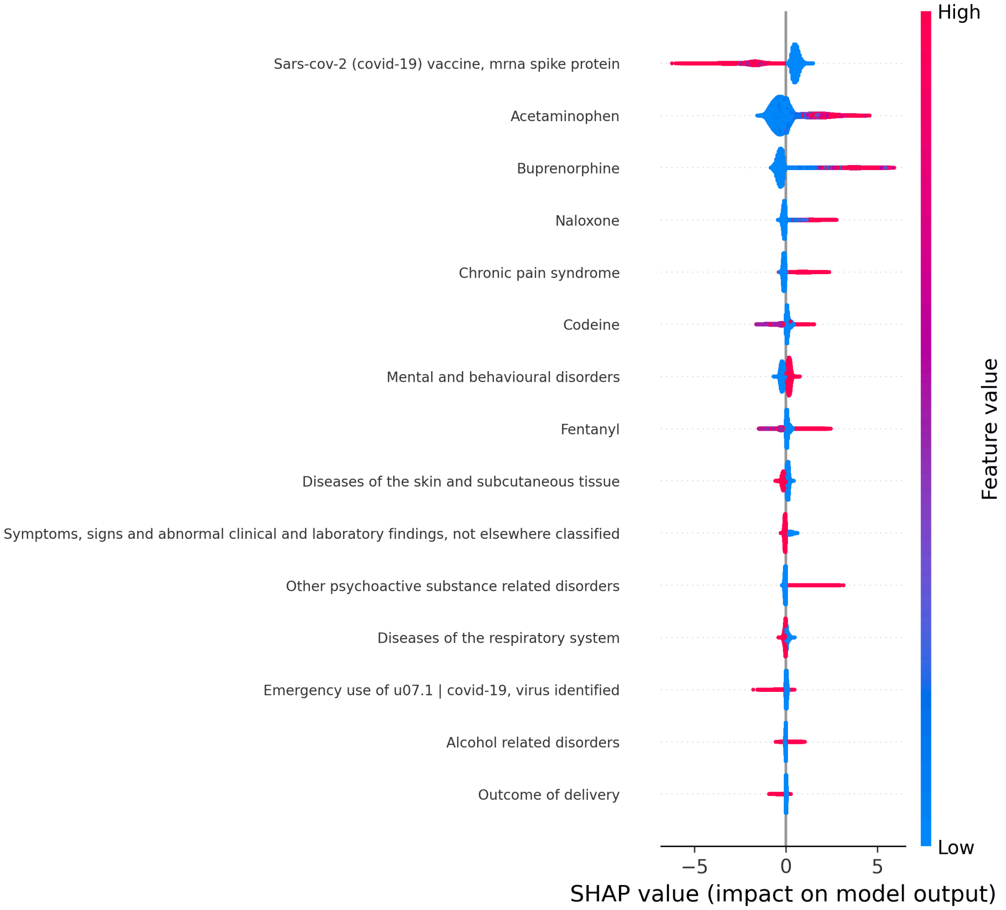

**Figure. 4. SHAP plot for the top 15 features identified by the XGBoost model across all 73 balanced datasets.** The plot shows the effects of these features on XGBoost's OUD prediction for individuals.

## Discussion and Conclusion

Accurate detection of OUD is crucial for several key reasons, including identifying at-risk individuals, improving responses to the opioid crisis, expanding access to treatment, guiding public health strategies, enhancing health outcomes, addressing co-occurring conditions, and ultimately saving lives. This study demonstrated the appropriateness of the PU learning algorithm, PULSNAR, in identifying undercoded or undiagnosed OUD. In our study cohort, merely 1 in 73 individuals were coded for OUD. However, by applying the PULSNAR method, our estimates suggest that approximately 1 in 20 individuals exposed to opioids have OUD, leading to an overall estimated cumulative prevalence of 5.08% across all age and sex demographics. This estimation is consistent with the prevalence ranges reported in other studies, justifying the applicability of PULSNAR.[10,11] The gain scores obtained from the XGBoost model provide valuable insights into potential risk factors and predictors associated with OUD. Unsurprisingly, treatments for OUD, such as buprenorphine and naloxone were highly predictive of OUD, as well as the presence of chronic pain and treatments for pain (e.g., acetaminophen). The lower incidence of OUD among patients who received a COVID-19 vaccine, as well as those with a positive COVID-19 test, may represent a temporal bias that warrants further investigation. The predictive power of other substance use disorders, including alcoholism, suggests that common causes may underlie multiple substance use conditions. Through this novel application of PULSNAR, we have not only identified a significant hidden burden of OUD, but also set the stage for improved diagnostic practices and interventions that are vital to intervening in the opioid crisis. Importantly each patient has a calibrated probability that can be used for screening as well as for probabilistic phenotyping.

## Acknowledgment

This research was supported by the National Institute of Mental Health of the National Institutes of Health under award numbers R01MH129764 and R56MH120826.

## References

1. Degenhardt L, Grebely J, Stone J, Hickman M, Vickerman P, Marshall BDL, et al. Global patterns of opioid use and dependence: harms to populations, interventions, and future action. Lancet. 2019;394(10208):1560–79.
2. Drug Overdose Death Rates https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates Accessed 5/28/2024
3. Florence C, Luo F, Rice K. The economic burden of opioid use disorder and fatal opioid overdose in the United States, 2017. Drug Alcohol Depend. 2021 Jan 1;218:108350. doi: 10.1016/j.drugalcdep.2020.108350. Epub 2020 Oct 27. PMID: 33121867; PMCID: PMC8091480.
4. The Economic Toll of the Opioid Crisis Reached Nearly $1.5 Trillion in 2020 https://www.jec.senate.gov/public/_cache/files/67bced7f-4232-40ea-9263-f033d280c567/jec-cost-of-opioids-issue-brief.pdf Accessed 5/28/2024
5. Haight SC, Ko JY, Tong VT, Bohm MK, Callaghan WM. Opioid use disorder documented at delivery hospitalization—United States, 1999–2014. Morbidity and Mortality Weekly Report. 2018 Aug 8;67(31):845.
6. Kumar P, Lambert CG. PULSNAR--Positive unlabeled learning selected not at random: class proportion estimation when the SCAR assumption does not hold. arXiv preprint arXiv:2303.08269. 2023 Mar 14.
7. An introduction to explainable AI with Shapley values https://shap.readthedocs.io/en/latest/example_notebooks/overviews/An%20introduction%20to%20explainable%20AI%20with%20Shapley%20values.html Accessed 6/20/2024
8. Voss EA , Makadia R, Matcho A, et al. . Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. J Am Med Inform Assoc 2015; 22 (3): 553–64.
9. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. InProceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining 2016 (pp. 785-794).
10. Keyes KM, Rutherford C, Hamilton A, Barocas JA, Gelberg KH, Mueller PP, Feaster DJ, El-Bassel N, Cerdá M. What is the prevalence of and trend in opioid use disorder in the United States from 2010 to 2019? Using multiplier approaches to estimate prevalence for an unknown population size. Drug and alcohol dependence reports. 2022 Jun 1;3:100052.
11. Lindner SR, Hart K, Manibusan B, McCarty D, McConnell KJ. State-and county-level geographic variation in opioid use disorder, medication treatment, and opioid-related overdose among medicaid enrollees. In JAMA Health Forum 2023 Jun 2 (Vol. 4, No. 6, pp. e231574-e231574). American Medical Association.