

Key features associated with opioid misuse in chronic pain: A machine learning cross-sectional study

Dokyoung S. You, PhD^{*,1}

Yiyu Wang, PhD^{*,2}

Samsuk Kim, PhD²

Mark A. Lumley, PhD³

Beth D. Darnall, PhD²

Sean C. Mackey, MD, PhD²

Maisa S. Ziadni, PhD²

*co-first authors

¹ Department of Family and Community Medicine, Health Promotion Research Center, Stephenson Cancer Center, University of Oklahoma-Tulsa

² Department Anesthesiology, Perioperative and Pain medicine, Stanford University School of Medicine, 1070 Arastradero Road, Suite 200, MC 5596, Palo Alto, CA 94304

³ Department of Psychology, Wayne State University, 5057 Woodward, Suite 7908 Detroit, Michigan 48202

Word counts

Abstract (limit 350): 333

Manuscript (limit 3,000) 2,994

Corresponding author: Dokyoung Sophia You, PhD.

Contact information: Department of Family and Community Medicine, Health Promotion Research Center, Stephenson Cancer Center, University of Oklahoma-Tulsa. 4502 E. 41st Street, Room 3G09, Tulsa, OK 74135, O: 918-660-3923, Email: dokyoung-you@ou.edu

Keywords: Prescription opioid misuse, Chronic pain, Anger, Ambivalence over emotion expression, Machine learning,

Abstract

Importance: Opioid misuse remains a critical public health concern, associated with increased risk of overdose, psychiatric comorbidity, and societal costs. While machine learning (ML) analyses have been used to identify features associated with opioid misuse across various clinical settings, few studies have examined both modifiable and non-modifiable features. Understanding modifiable features may inform targeted prevention and intervention strategies to reduce opioid misuse.

Objective: To identify key features associated with opioid misuse severity in patients with chronic pain taking long-term prescription opioids.

Design, Setting, and Participants: This analysis used data from clinical trials investigating a skills-based pain management intervention. Participants were 314 community-dwelling adults with chronic pain who were taking daily opioid medications (≥ 10 morphine-equivalent daily dose) for at least 3 months. Data was extracted from baseline assessments.

Outcomes and Measures: Opioid misuse severity was assessed using the Current Opioid Misuse Measure (COMM). Thirty-six demographic and clinical features were evaluated, including PROMIS symptom domains (e.g., pain rating, pain interference, physical function, fatigue, sleep disturbance, depression, anxiety, and anger) as well as emotional ambivalence, pain catastrophizing, trauma exposure, and substance use.

Results: Among the seven ML algorithms (Random Forest, XGBoost, Support Vector Regression, LASSO regression, Ridge Regression, Elastic Net, and Multilayer Perceptron), the Elastic Net model demonstrated the strongest performance, yielding the highest correlation with COMM scores (mean $r = 0.61$, 95% CI [0.40, 0.73]) and the lowest root mean square error (mean RMSE = 4.16, 95% CI [3.41, 4.80]). Feature ablation analysis identified anger ($\Delta r = 0.053$), emotional ambivalence ($\Delta r = 0.022$), pain catastrophizing ($\Delta r = 0.008$), and fatigue ($\Delta r = 0.002$) as the most influential features associated with the COMM scores. Shapley Additive exPlanations (SHAP) analysis confirmed that higher levels of these key features were associated with higher COMM scores.

Conclusions and Relevance: Emotional factors, particularly anger, emerged as key features associated with the severity of opioid misuse. These findings suggest that interventions targeting emotion regulation, especially anger management, may reduce opioid misuse among individuals receiving long-term opioid therapy for chronic pain.

Introduction

Chronic pain is a debilitating condition that affects 50.2 million (20.5%) adults in the United States (U.S.).¹ Musculoskeletal disorders such as chronic back pain and arthritis are the leading causes of disability in the U.S.^{2,3} To manage chronic pain, 29.3% of affected individuals use prescription opioid.⁴ Opioids are particularly utilized by those with worse health status, including higher levels of pain, pain interference, fatigue, sleep disturbance, and depression.⁵ Consistent with this pattern, its utilization is higher among those with activity-limiting pain,⁶ with 41.2% reporting prescription opioid use.⁴

Opioid medications are generally used as prescribed; however, they carry a risk of misuse. The prevalence of opioid misuse is estimated to be between 21% and 29% among individuals using opioids for pain management.⁷ Misuse refers to taking opioids in a manner other than prescribed, such as using higher doses, more frequently, or for a longer duration than directed.⁸ It also includes using opioids for non-medical reasons, taking someone else's prescription, continuing use despite negative consequences, or combining opioids with central nervous system depressants (e.g., benzodiazepines, alcohol).⁹ Opioid misuse is a significant health concern, as it increases the risk of overdose and is associated with greater emotional distress.^{5,10} Furthermore, it imposes a substantial financial burden. Individuals with opioid misuse spend \$15,000 more on healthcare annually compared to those without it, and the social cost is estimated at \$50 billion per year.¹¹

To address opioid misuse, machine learning (ML) has been used to identify risk factors across various healthcare settings. ML models can capture complex, non-linear interactions without strict parametric assumptions. Due to this flexibility, ML models are better suited in detecting risk factors for multifaceted conditions like opioid misuse. One ML study analyzing electronic health records (EHRs) from intensive care units (ICU) identified predictors of opioid misuse, including: a) ICD-9 codes related to drug withdrawal, chronic hepatitis, posttraumatic stress disorder (PTSD), tobacco use disorder, dysthymic disorder, unspecified drug abuse, cocaine abuse, and other chronic pain; and b) certain prescription medications such as methadone, nicotine patch, diazepam, zolpidem, quetiapine, and clonazepam.¹² Another ML study analyzing EHRs of surgical patients identified pre-operative opioid use, depression, and antidepressant use as predictors of prolonged opioid misuse after surgery.¹³ However, no study has applied ML to community-dwelling adults with chronic pain on long-term opioids. Moreover, prior ML studies have rarely incorporated modifiable psychological risk factors, which may inform behavioral treatment targets.

Traditional regression analyses have identified psychological risk factors for opioid misuse in chronic pain populations, such as pain catastrophizing^{10,14-16} and emotion dysregulation.^{17,18} Pain catastrophizing refers to a negative cognitive and emotional response to pain.¹⁹ Emotion dysregulation refers to the inability to regulate emotional reactions in adaptive ways.²⁰ Specifically, it involves difficulties with awareness, understanding, and accepting emotions, as well as engaging in a goal-directed behavior or inhibit inappropriate or impulsive behaviors when experiencing negative emotions.²⁰ Recently, ambivalence over emotion expression (hereafter emotional ambivalence), a form of emotion dysregulation, has been identified as a factor associated with opioid misuse in individuals with chronic pain.²¹

The current cross-sectional study applied an ML approach to examine both modifiable and non-modifiable factors associated with opioid misuse severity among adults with chronic pain receiving long-term opioid therapy. The primary aim was to identify the most influential features, with particular interest in pain catastrophizing and emotional ambivalence.

Methods

Study Design

Stanford's Institutional Review Board approved this study procedure. Participants were recruited for three clinical trials, investigating the effect of a single-session, pain self-management skills class. Recruitment began in 2019 and ended in 2024. The current study analyzed the baseline data from

different phases of clinical trials, which included: participants enrolled in the in-person version of the trial prior to COVID (N=46), single-arm Zoom-delivered version during COVID (N=55), and the final Zoom-delivered RCT (N=213). Inclusion and exclusion criteria were consistent across all three phases. Participants were eligible if they: a) were at least 18 years old, b) had chronic non-cancer pain (at least 6-month duration), c) had been using prescription opioids (≥ 10 morphine-equivalent daily dose: MEDD) for ≥ 3 months, d) were fluent in English, and e) had access to an electronic device to complete web-based surveys and attend a zoom-delivered class. Exclusion criteria included: a) ongoing litigation or disability claim, b) concurrent participation in another pain self-management skills class, c) inability to complete any of the study required tasks (e.g., cognitive impairment), and d) active suicidality. Eligible participants received the electronic informed consent, which was administered via the REDCap, Health Insurance Portability and Accountability Act (HIPAA)-compliant platform.²² Once signed the informed consent, baseline surveys were administered through REDCap.

Measures

Descriptive demographic and clinical variables

Demographic information, including age, gender, race/ethnicity, income, marital status, and employment status, was collected. Additionally, data on pain diagnoses, current disability status, mental health diagnoses, and history of psychological treatment were gathered. For pain and mental health diagnoses, participants responded to yes/no questions and, if endorsed, they were prompted to specify their conditions in a free-text format.

Outcome variable: Opioid misuse

The severity of opioid misuse was assessed using the 17-item Current Opioid Misuse Measure (COMM), which has demonstrated excellent internal consistency ($\alpha = 0.86^{23}$ and 0.83^{24}) and reliability (ICC = 0.86).²⁴ The cutoff score of 9 has been validated to identify problematic opioid misuse.²⁴ As expected, our sample's COMM scores were rightly skewed (Figure S1).

Independent variables

Thirty-six demographic and clinical variables were examined (Table 1.1 and 1.2). The correlation matrix illustrates the relationships among demographic and clinical variables (Figure S1). Income and education were dummy coded such that higher values would reflect greater income and higher levels of education (see Table 1.1 note for coding details). To improve model efficiency and reduce overfitting, employment status, marital status, race, and disability were recoded as binary variables: “employed vs unemployed,” “married vs non-married,” “white vs non-white,” and “disability vs no disability.” Binary coding reduced the number of predictors and addressed sparse category representation, improving model convergence.

The following measures were used to assess clinical variables. Alcohol-related problems were assessed using the 10-item **Alcohol Use Disorders Identification Test** (AUDIT; total score range 0 to 40), with scores of ≥ 8 indicating hazardous or harmful drinking. Drug-related problems were measured with the 10-item **Drug Abuse Screening Test** (DAST-10; total score range 0-10), where higher total scores indicate more problems with drug use. We also administered the **DSM-5 OUD symptom checklist**²⁵ to evaluate OUD severity; however, this variable was excluded from the ML analysis due to predictor-outcome overlap. Additionally, the **Childhood Trauma Questionnaire (CTQ)** was used to assess childhood maltreatment across five domains. The CTQ consists of 28 items that evaluate the frequency of specific traumatic experiences during childhood. Total scores range from 25 to 125, with higher scores indicating greater levels of reported childhood trauma. The CTQ includes five subscales: emotional abuse, physical abuse, sexual abuse, emotional neglect, and physical neglect. Each subscale contains five items rated on a 5-point Likert scale, with total scores for each domain ranging from 5 to 25 and higher scores indicating more severe maltreatment. The average CTQ total score among adults is 32

(SD = 9 and 11 for male and female, respectively).²⁶ The **Patient-Reported Outcomes Measurement Information System** (PROMIS) short-forms were also used to assess global physical and mental health,²⁷ as well as specific health domains, including pain intensity, pain interference (8a),²⁸ physical function (8a), fatigue (8a), sleep disturbance (8a), depression (8a), anxiety (8a), and anger (8a) over the past seven days. T-scores were calculated using the web-based scoring system (www.assessmentcenter.net/), with a mean of 50 (SD = 10). Higher scores indicate better global health or greater symptom severity or function, depending on the domain. The 13-item **Pain Catastrophizing Scale** (PCS; total scores 0-52) was used to assess the degree of pain catastrophizing, consisting of perceived helplessness about, ruminating on, and enhanced threat appraisal of pain. Higher scores indicate more catastrophic thinking about pain.¹⁹ Helplessness (6 items), rumination (4 items), and magnification (3 items) subscale scores range from 0 to 24, 0 to 16, and 0 to 12, respectively. The 14-item **Ambivalence over Emotion Expression Questionnaire (AEQ)** was used to measure degree of discomfort and uncertainty when expressing their negative emotions, reflecting their internal conflict between wanting to express and suppress these emotions.²¹ Each item is rated on a 1-5 scale and total summed scores range from 14 to 70, with higher scores indicating greater ambivalence about expressing emotions.

Missing values

Most missing data were found in the DAST, CTQ and AEQ such as 89 (28%), 86 (27%), and 25 (8%) cases, respectively. The other scores had less than 2% missing data. The higher missing rates in the DAST and CTQ may reflect the sensitive nature of assessing substance use problems and childhood trauma. All missing values were replaced using multivariate imputation to preserve statistical power and minimize potential bias.

Analysis: hyperparameters, model training, cross validation, and interpretation

For model training, we conducted a suite of ML analyses to examine features associated with COMM scores. Seven ML models were evaluated: Random Forest, XGBoost (XGB), Support Vector Regression (SVR), LASSO regression, Ridge Regression, Elastic Net, and Multilayer Perceptron (MLP; Figure 1). Hyperparameters for each model were optimized using a grid search procedure with five-fold cross validation, implemented via Scikit-learn's GridSearchCV.²⁹ The best hyperparameters were determined based on the highest average cross-validated r^2 value.

After hyperparameter selection, model performance was evaluated on the testing dataset to evaluate how well each trained model generalized to unseen data in predicting COMM scores. We used a 100-iteration of Monte Carlo cross-validation.^{30,31} For each iteration, the data were randomly split into a training set (80%) and a testing set (20%). The best performance model was selected based on the highest cross validated mean correlation coefficient value (r) (Figure 1a) and the lowest cross validated mean root mean squared error (RMSE; Figure 1b).

Following model selection, we conducted a feature ablation analysis to evaluate the relative contribution of each feature to the model performance. This involved systematically replacing the values of each independent variable with zero (i.e., "ablation") and reapplying the model with the same 100-iteration Monte Carlo cross-validation. We then calculated the change in the model's average correlation coefficient (r) compared to the original model containing all predictors. This change, referred to as Δr , reflects the reduction in model performance caused by removing a specific feature. A larger Δr indicates a greater decrease in model performance, suggesting that the removed feature has a greater impact on opioid misuse severity.

Next, we applied SHapley Additive exPlanations (SHAP) to the best-performing model to examine the contribution of individual features. Specifically, we examined a) the most influential features identified in a feature ablation and b) the direction of their relationship with the outcome.³² SHAP values quantify each feature's contribution to the model's prediction. A positive SHAP value associated with higher feature values indicates a positive relationship between the feature and the outcome. A wider

spread of SHAP values for a given feature suggests a stronger influence on the model's predicted outcome scores.

Sample Size Justification

Learning curve analysis³³ (Supplementary Figure S2) has shown that the validation score reaches plateau when training sample size reaches around 125, suggesting that the model performance is unlikely to improve by adding more data beyond this point. Therefore, our sample size of 314 would have sufficient power to detect meaningful patterns in the data.

Study Sample

A total of 314 participants completed the baseline survey. The demographic characteristics are summarized in Table 1.1. Regarding clinical characteristics (Table 1.2), the majority (92.0%) reported that their pain duration was more than 5 years. Among those who reported their pain diagnoses (73.6%), the mean number of pain diagnosis was 2.3 (SD = 1.6, range 1 – 8) and the three most common diagnoses were fibromyalgia (19.0%), arthritis (17.7%), and chronic pain syndrome (13.0%, Supplement Table 1). The mean opioid duration was 12 years, 66 participants (27.7%) were screened positive for problematic misuse (i.e., COMM score of ≥ 9), and 51 participants (16.2%) met the diagnostic criteria for OUD. Notably, all participants were taking opioids, but moderate to severe OUD was rare (2.8%). Similarly, alcohol use was common (47.5%), but hazardous or harmful drinking was rare (0.6%). Drug use for non-medical reasons was reported by 12.1% and the most common drug was cannabis (11.5%, Supplement Table 2). Of 176 participants with reported mental health diagnoses, the average number of diagnoses was 1.7 (SD = 0.9). Depression was most common (62.5%), followed by anxiety (42.1%) and PTSD (15.9%; Supplement Table 3). About 47.1% reported receiving psychological treatments.

Results

Among the seven models, the Elastic Net model demonstrated the strongest performance, yielding the highest mean correlation coefficient r of 0.61 (95% CI: 0.40, 0.73, Figure 1a) and the lowest mean RMSE of 4.12 (95% CI: 3.44, 4.85, Figure 1b).

Feature Interpretation

We performed two analyses to interpret model features. First, we conducted a feature ablation analysis to assess the relative contribution of each feature to the Elastic Net model. Because the Elastic Net uses regularization that assigns zero weights to less important features, we only conducted the feature ablation analysis on non-zero weighted features (Figure 2a). Top four features that made the most impact to the model were the PROMIS Anger T-scores ($\Delta r = 0.053$), AEQ total scores ($\Delta r = 0.022$), PCS total scores ($\Delta r = 0.008$), and the PROMIS Fatigue T-scores ($\Delta r = 0.002$). These results suggest that emotional factors such as anger and emotional ambivalence are important features associated with opioid misuse severity.

Secondly, the SHAP model, which identifies the additive impacts of each feature, provided complementary support to the feature ablation analysis. Features including the PROMIS-Anger T-scores, AEQ scores, PCS total scores, PROMIS-Fatigue T-scores show significant impact on predicted COMM total scores (Figure 2b). Higher values of anger, AEQ, PCS, and fatigue scores (in red, Figure 2b) were associated with more positive SHAP values, indicating that greater values of these features were associated with greater severity of opioid misuse. Both feature analyses highlight the important role of emotion factors in opioid misuse.

Discussion

The current study used an ML approach to analyze baseline data from community-dwelling adults with chronic pain who were receiving long-term opioid therapy and enrolled in clinical trials of a pain self-management skills class. Our ML model identified that anger, emotional ambivalence, pain catastrophizing, and fatigue were the most important features associated with opioid misuse severity.

Our ML model identified anger and emotional ambivalence as key features associated with opioid misuse severity. These findings are consistent with prior ML studies conducted in ICU and surgical populations, which also highlighted emotional factors as important predictors of opioid misuse.^{12,13} However, the specific emotional features associated with opioid misuse differed between the current study and previous studies. Specifically, previous ML studies based on EHRs identified depression and PTSD as primary predictors^{12,13} whereas our model identified anger as a primary factor. Notably, a prior stepwise regression analysis using data from a tertiary pain clinic sample found anxiety, anger, and depression, in that order, to be significant predictors of opioid misuse.⁵ One explanation for this discrepancy may be that anger is not routinely assessed or documented (ICD R45.4) in many clinical settings and there are no FDA-approved medications for anger. As a result, anger often goes undocumented in EHRs, limiting its detection in EHR-based research. To address this limitation, our learning healthcare system (<https://choir.stanford.edu/>) routinely assesses symptoms of depression, anxiety, and anger in all patients seen in a tertiary pain clinic. This comprehensive symptom tracking, in combination with an ML approach, may improve the identification of key emotional risk factors for opioid misuse in chronic pain populations.⁵

Pain catastrophizing and emotional ambivalence were of a particular interest in this current study because cumulative studies have found pain catastrophizing as a significant predictor for opioid misuse^{14,15,34-36} and a recent study identified emotional ambivalence as an additional predictor even when controlling for the known predictors like emotional distress and pain catastrophizing.²¹ Our findings extend this literature by demonstrating, through a ML approach, that both emotional ambivalence and pain catastrophizing are among the key features associated with opioid misuse severity in the chronic pain population. Besides emotional factors, fatigue was a significant feature. Fatigue is a known side effect of opioids³⁷ and is positively correlated with opioid misuse severity in individuals with chronic pain.^{5,10,38,39} Future study should examine whether reducing anger, emotional ambivalence, pain catastrophizing, and fatigue would reduce opioid misuse risk.

Limitations

Several limitations of the current study should be acknowledged. First, the cross-sectional design precludes any conclusions about causal relationship between emotional factors and opioid misuse. Second, to enhance generalizability and reduce the risk of overfitting, we employed multiple ML models and reported the final performance conservatively, based on the mean outcome from a rigorous 100-fold cross validation procedure. Future work should test the model in independent samples to evaluate its generalizability. Third, the high rates of missing data on the DAST and CTQ may have limited the model's ability to identify substance use problems and childhood trauma as important features. Therefore, these potentially relevant factors may have been underrepresented in the final model. Finally, the sample size may have limited our ability to test more complex predictive algorithms that might enhance performance. To balance model complexity and interpretability, we prioritized model interpretability that allowed for the identification of modifiable psychological factors, thereby informing potential treatment targets for opioid misuse.

Conclusions

The current study is the first to apply ML approach to identify key features associated with the severity of opioid misuse among community-dwelling U.S. adults with chronic pain receiving long-term opioid therapy. Consistent with previous studies, our findings underscore negative emotions as key features associated with opioid misuse in this population. Among negative emotions, anger and emotional ambivalence emerged as particularly relevant. Intervention targeting anger regulation and adaptive

emotional expression may reduce the risk of opioid misuse. Building on these findings, future large-scale studies are needed to develop and validate ML-based models that can identify individuals with problematic opioid misuse. Such models could support accurate identification of opioid misuse, enable targeted intervention, and promote safer opioid prescribing practices in clinical settings.

Table 1.1 Demographic characteristics (n = 314)

Variables		M	SD	Range
Age	(years)	54.4	11.9	19 - 86
		N	(%)	
Gender	Female	241	76.8	
	Male	71	22.6	
	Other (including decline to answer)	2	0.6	
Ethnicity	Non-Hispanic	297	94.6	
	Hispanic	17	5.4	
Race	White/Caucasian	282	89.8	
	Black/African	9	2.9	
	Asian	5	1.6	
	Native Hawaiian or Pacific Islander	1	0.3	
	American Indian or Alaska Native	3	1.0	
	More than one	10	3.2	
	Other or unknown	4	1.3	
Marital Status	Married	168	53.5	
	Partnered and living together	22	7.0	
	In a relationship but not living together	10	3.2	
	Divorced	59	18.8	
	Separated	11	3.5	
	Widowed	13	4.1	
	Never married	31	9.9	
Employment status*	Employed	69	22.0	
	Not currently employed	18	5.7	
	Temporarily laid off/sick leave or maternity leave	3	1.0	
	Disabled or not working due to pain	178	54.8	
	Retired	74	23.6	
	Other	15	4.8	
		M	SD	Range
Income		4.6	2.2	1 - 8
Education		6.6	1.4	3 - 10

Note: Income brackets were 1, Less than \$10,000; 2, <\$25,000; 3, \$25,000-\$44,999; 4, \$45,000-\$64,999; 5, \$65,000-\$84,999; 6, \$85,000-\$104,999; 7, \$105,000-\$124,999; 8, ≥\$125,000. Education brackets were: 1, Grade School/Elementary School (kindergarten through 6th grade); 2, Middle School (through 8th grade); 3, Some High School; 4, High School Diploma or GED; 5, Some College, No Degree; 6, Associate's Degree or Vocational Certificate; 7, Bachelor's Degree (e.g., BA, AB, BS, BBA); 8, Master's Degree (e.g., MA, MS, MEng, MEd, MBA); 9, Professional School Degree (e.g., MD, DDS, DVM, JD); 10, Doctoral Degree (e.g., PhD, EdD). *The total exceeds 314 because multiple selections were allowed. Variables shown in bold were included as binary variables in the ML model.

Table 1.2 Clinical characteristics (n = 314) *Descriptives are calculated before imputation

Variables	N	(%)	
Duration of pain			
6 – less than 12 months	1	0.3	
1 – 5 years	24	7.6	
> 5 years	289	92.0	
Received diagnosis for chronic pain? (Yes)	231	73.6	
Currently on disability? (Yes)	149	47.5	
COMM positive (total scores of ≥ 9)	66	27.7	
DSM 5 OUD symptom positive (total score of ≥ 2)	51	16.2	
Drinking alcohol? (Yes)	149	47.5	
AUDIT positive (total scores of ≥ 8)	2	0.6	
Use drug for non-medical reasons? (Yes)	38	12.1	
Received mental health diagnosis? (Yes)	176	56.1	
Receive psychological treatment? (Yes)	148	47.1	
	M	SD	Range
PROMIS-Pain Intensity (0-10 scale): Worst	8.6	1.3	3 – 10
Average	5.3	1.6	2 – 10
Current	5.5	1.9	0 – 10
Duration of opioid medication (years)	12.5	8.1	0 – 55
COMM total scores	6.8	5.2	0 – 33
AUDIT total scores	1.0	1.6	0 – 16
DAST total scores	0.3	0.3	0 – 4
CTQ total scores	47.8	17.7	25 – 119
Emotional abuse	11.4	5.2	5 – 25
Physical Abuse	8.3	4.5	5 – 25
Emotional Neglect	11.9	4.6	5 – 25
Physical Neglect	7.9	3.3	5 – 25
Sexual Abuse	7.7	4.7	5 – 25
PROMIS (T-scores)			
Global Physical Health	36.0	6.7	19.7 – 60.3
Global Mental Health	41.3	8.7	21.3 – 67.6
Pain Interference	64.4	6.8	40.7 – 77.0
Physical Function	35.9	6.6	20.9 – 59.7
Fatigue	60.4	9.3	33.1 – 77.7
Sleep Disturbance	57.3	8.7	30.5 – 77.5
Depression	53.4	9.6	38.2 – 81.1
Anxiety	53.8	9.6	37.1 – 82.8
Anger	49.2	8.5	32.9 – 79.6
PCS total scores	16.9	11.5	0 – 52
Helplessness	7.7	5.5	0 – 24
Rumination	6.1	4.4	0 – 16
Magnification	3.1	2.6	0 – 12
AEQ total scores	31.7	11.2	14 – 69

Note: Variables shown in bold were included in the ML model.

Data Sharing Statement

The data sets generated and analyzed during this study are available upon study team approval of written requests with analytic plan included in the request.

REFERENCE

1. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. *Pain*. 2022;163(2):e328-e332.
2. Courtney-Long EA, Carroll DD, Zhang QC, et al. Prevalence of disability and disability type among adults—United States, 2013. *MMWR Morb Mortal Wkly Rep*. 2015;64(29):777-783.
3. Theis KA, Roblin DW, Helmick CG, Luo R. Prevalence and causes of work disability among working-age US adults, 2011–2013, NHIS. *Disability and health journal*. 2018;11(1):108-115.
4. Zajacova A, Grol-Prokopczyk H, Limani M, Schwarz C, Gilron I. Prevalence and correlates of prescription opioid use among US adults, 2019–2020. *PLoS One*. 2023;18(3):e0282536.
5. Gilam G, Sturgeon JA, You DS, Wasan AD, Darnall BD, Mackey SC. Negative affect–related factors have the strongest association with prescription opioid misuse in a cross-sectional cohort of patients with chronic pain. *Pain medicine*. 2020;21(2):e127-e138.
6. Von Korff M, Scher AI, Helmick C, et al. United States national pain strategy for population research: concepts, definitions, and pilot data. *The Journal of Pain*. 2016;17(10):1068-1080.
7. Jantarada C, Silva C, Guimarães-Pereira L. Prevalence of problematic use of opioids in patients with chronic noncancer pain: a systematic review with meta-analysis. *Pain Practice*. 2021;21(6):715-729.
8. Key substance use and mental health indicators in the United States: results from the 2021 National Survey on Drug Use and Health (2022).
9. Abuse S, Administration MHS. 2015 National survey on drug use and health. 2016;
10. You DS, Hah JM, Collins S, et al. Evaluation of the preliminary validity of misuse of prescription pain medication items from the Patient-Reported Outcomes Measurement Information System (PROMIS)®. *Pain Medicine*. 2019;20(10):1925-1933.
11. Oderda GM, Lake J, Rüdell K, Roland CL, Masters ET. Economic burden of prescription opioid misuse and abuse: a systematic review. *Journal of pain & palliative care pharmacotherapy*. 2015;29(4):388-400.
12. Vega L, Conneen W, Veronin MA, Schumaker RP. A neural network approach to predict opioid misuse among previously hospitalized patients using electronic health records. *Plos one*. 2024;19(8):e0309424.
13. Emam OS, Eldaly AS, Avila FR, et al. Machine learning algorithms predict long-term postoperative opioid misuse: a systematic review. *The American Surgeon™*. 2024;90(1):140-151.
14. Arteta J, Cobos B, Hu Y, Jordan K, Howard K. Evaluation of how depression and anxiety mediate the relationship between pain catastrophizing and prescription opioid misuse in a chronic pain population. *Pain Medicine*. 2016;17(2):295-303.
15. Baranoff JA, Clubb B, Coates JM, Elphinston RA, Loveday W, Connor JP. The contribution of pain catastrophizing, depression and anxiety symptoms among patients with persistent pain and opioid misuse behaviours. *Journal of Behavioral Medicine*. 2024;47(2):342-347.
16. López-Martínez AE, Sainero-Tirado G, Esteve R, et al. Does pain catastrophizing and distress intolerance mediate the relationship between PTSD and prescribed opioid misuse in people with chronic noncancer pain? *Psychological Trauma: Theory, Research, Practice, and Policy*. 2023;15(3):394.
17. Lutz J, Gross RT, Vargovich AM. Difficulties in emotion regulation and chronic pain-related disability and opioid misuse. *Addictive behaviors*. 2018;87:200-205.
18. Rogers AH, Garey L, Bakhshaie J, Viana AG, Ditre JW, Zvolensky MJ. Anxiety, depression, and opioid misuse among adults with chronic pain: the role of anxiety sensitivity. *The Clinical Journal of Pain*. 2020;36(11):862-867.
19. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychological assessment*. 1995;7(4):524.

20. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of psychopathology and behavioral assessment*. 2004;26:41-54.
21. Oberleitner LM, Lumley MA, Grekin ER, et al. Problematic prescription opioid use in a chronic pain treatment facility: the role of emotional processes. *Substance use & misuse*. 2019;54(3):495-505.
22. Patridge EF, Bardyn TP. Research electronic data capture (REDCap). *Journal of the Medical Library Association: JMLA*. 2018;106(1):142.
23. Butler SF, Budman SH, Fernandez KC, et al. Development and validation of the current opioid misuse measure. *Pain*. 2007;130(1-2):144-156.
24. Butler SF, Budman SH, Fanciullo GJ, Jamison RN. Cross validation of the current opioid misuse measure to monitor chronic pain patients on opioid therapy. *The Clinical journal of pain*. 2010;26(9):770-776.
25. You DS, Mardian AS, Darnall BD, et al. A brief screening tool for opioid use disorder: EMPOWER study expert consensus protocol. *Frontiers in Medicine*. 2021;8:591201.
26. Scher CD, Stein MB, Asmundson GJ, McCreary DR, Forde DR. The childhood trauma questionnaire in a community sample: psychometric properties and normative data. *Journal of traumatic stress*. 2001;14:843-857.
27. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of life Research*. 2009;18:873-880.
28. Askew RL, Cook KF, Revicki DA, Cella D, Amtmann D. Evidence from diverse clinical populations supported clinical validity of PROMIS pain interference and pain behavior. *Journal of clinical epidemiology*. 2016;73:103-111.
29. Pedregosa F, Varoquaux Ge, Gramfort A, et al. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-2830.
30. Shao J. Linear model selection by cross-validation. *Journal of the American statistical Association*. 1993;88(422):486-494.
31. Xu QS, Liang YZ, Du YP. Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 2004;18(2):112-120.
32. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. 2017;30
33. Anzanello MJ, Fogliatto FS. Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*. 2011;41(5):573-583.
34. Martel M, Wasan A, Jamison R, Edwards R. Catastrophic thinking and increased risk for prescription opioid misuse in patients with chronic pain. *Drug and alcohol dependence*. 2013;132(1-2):335-341.
35. Martinez-Calderon J, Flores-Cortes M, Morales-Asencio JM, Luque-Suarez A. Pain catastrophizing, opioid misuse, opioid use, and opioid dose in people with chronic musculoskeletal pain: a systematic review. *The Journal of Pain*. 2021;22(8):879-891.
36. Lazaridou A, Franceschelli O, Buliteanu A, Cornelius M, Edwards RR, Jamison RN. Influence of catastrophizing on pain intensity, disability, side effects, and opioid misuse among pain patients in primary care. *Journal of applied biobehavioral research*. 2017;22(1):e12081.
37. Harris J-D. Management of expected and unexpected opioid-related side effects. *The Clinical journal of pain*. 2008;24:S8-S13.
38. Manning K, Rogers AH, Garey L, Zvolensky MJ. The moderating role of fatigue sensitivity in the relation between depression and alcohol and opioid misuse among adults with chronic low back pain. *Fatigue: biomedicine, health & behavior*. 2020;8(3):130-143.

39. Manning K, Rogers AH, Shepherd JM, et al. Fatigue severity among adults with chronic low back pain: relations to opioid use and hazardous alcohol use. *Fatigue: biomedicine, health & behavior*. 2021;9(3):148-158.

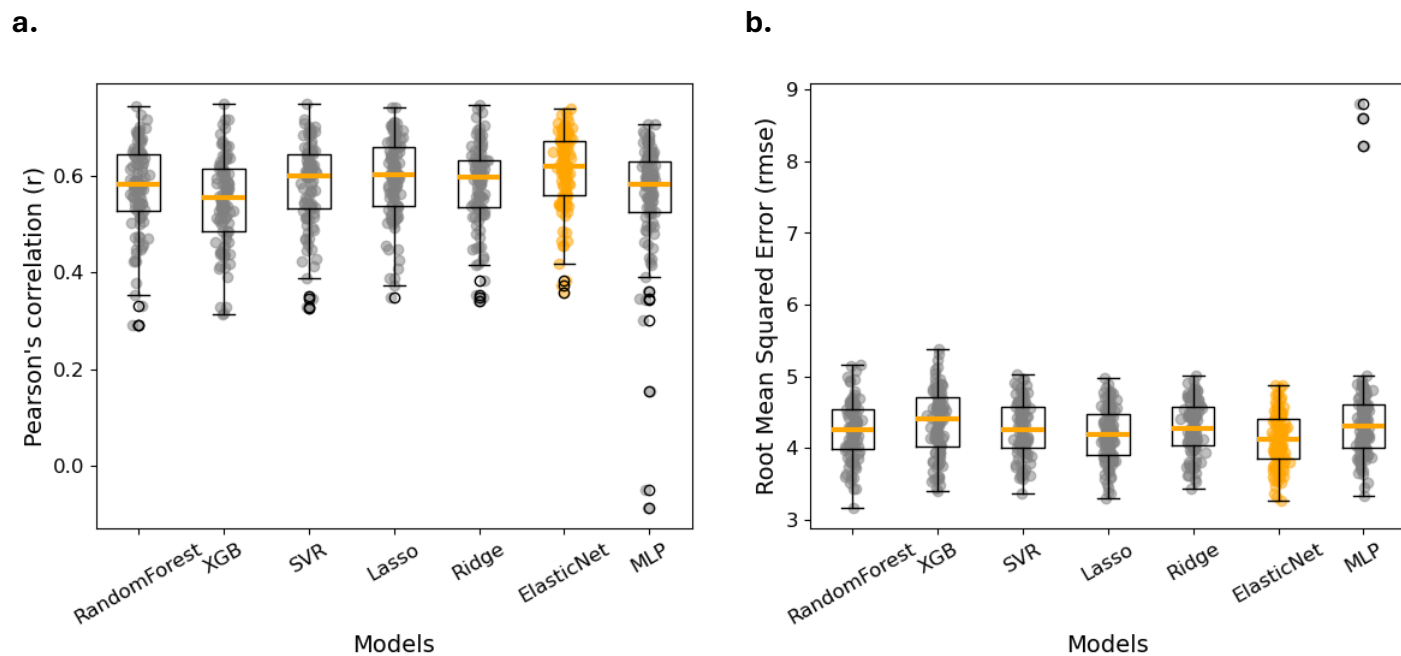


Figure 1. Model performance comparison. a) The distribution of the correlation coefficient (r) from the 100-fold cross validation. The Elastic Net Regression (orange) has the highest median r score ($r = 0.61$), and b) the lowest mean RMSE (RMSE = 4.12).

