Stanford MEDICINE | Center for Biomedical Informatics Research

Development and External Validation of ML Models for Identifying Patients at Risk of Postoperative Prolonged Opioid Use (PORPOISE)

A Network Study on OMOP Databases

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

OHDSI Community Call August 2nd, 2022

Behzad Naderalvojoud, PhD
Postdoctoral Research Fellow
behzadn@stanford.edu

Opioid

- Potent analgesic used to manage pain, often prescribed after surgery
- Highly addictive, even when prescribed appropriately and taken as directed
- Serious complications (e.g., dependence and abuse) result in significant morbidity and mortality
- Postoperative opioid exposure is a major risk factor for prolonged opioid use and abuse

Project objectives

- Improve **pain management** following surgery

- Identify **patients at risk** for prolonged opioid use **prior to prescribing** pain management regimes

- Develop and validate ML models **in a diverse**, **multisite cohort** by evaluating their **generalizability**, **discrimination** and **calibration** abilities

Cohort selection

Target Cohort

Patients who underwent one of the surgery in a CPT/ICD codes pre-defined list (keep only the first surgery if multiple) (n=121,404)

Surgery occurred between 2008-01-01 and 2020-01-01 (n=96,420)

Admission and Discharge dates could be retrieved for each patient (n=83,179)

At least one opioid prescription (among a predefined list of RxNorms) was found (n=82,016)

At least one of the opioid prescription periods overlaps with the time range +/- 30 days around surgery (n=81,399)

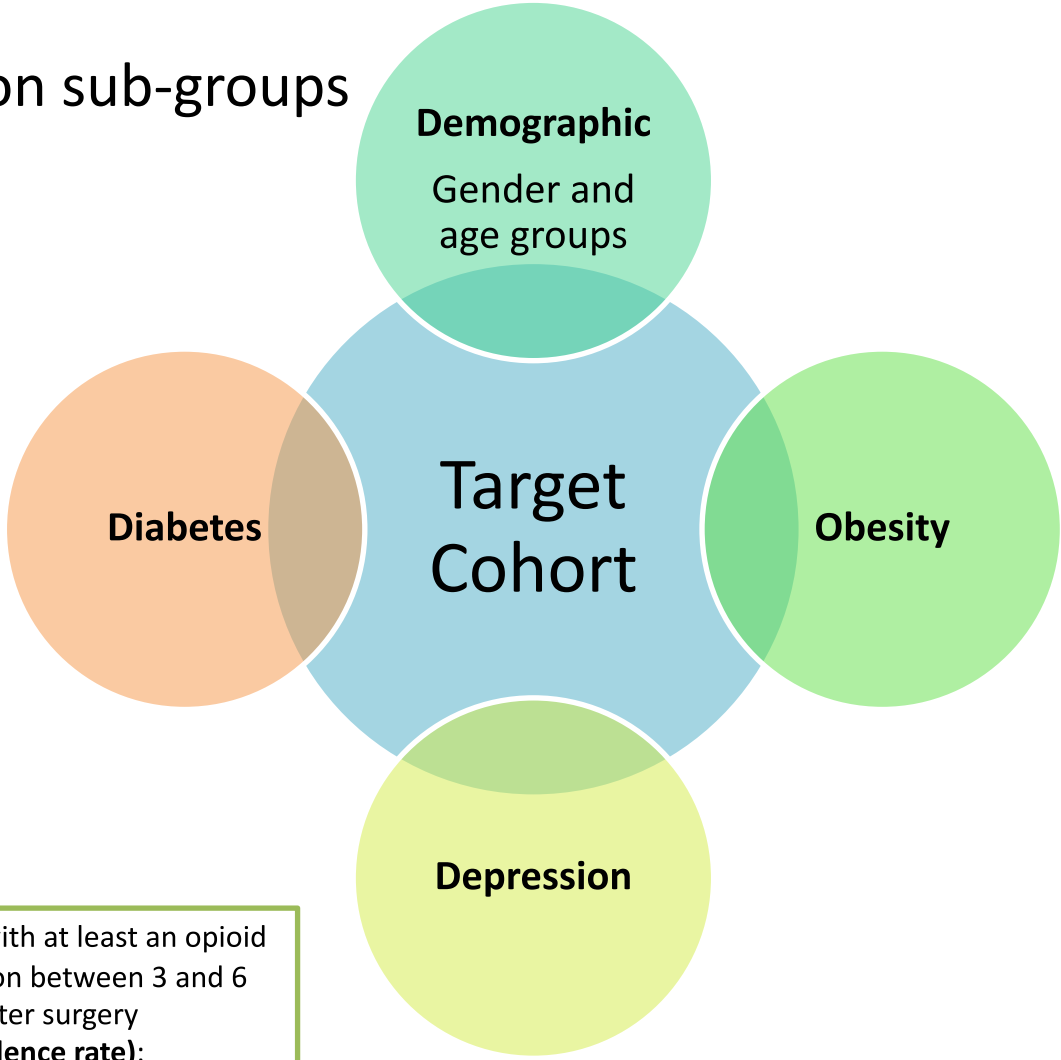
Patients aged between 18 and 90 (n=74,257)

At least 2 encounters in the window [2 years before ; 30 days before] surgery and at least 2 encounters in the window [30 days after; 2 years after] surgery (n=44,751)

At least 1 encounters and still alive more than 2 years after surgery (n=37,163)

No other surgery was conducted between 2 and 7 months after the first one **(n=35,406)**

Evaluation sub-groups



Outcome Cohort

Patients with at least an opioid prescription between 3 and 6 months after surgery **(15% incidence rate)**:

Cohort concept sets

Opioid prescription

RxNorm	Drugs Ingredients
5489	HYDROcodone
4337	fentaNYL
2670	Codeine
3423	HYDROmorphone
6754	Meperidine
6813	Methadone
7052	Morphine
7804	oxyCODONE
10689	Tramadol

- Drug concept set: 46,236 OMOP concepts
- Surgery concept set: 5182 OMOP concepts

Surgery categories (ICD/CPT Codes):

- Laminectomy, excision intervertebral disc
- Spinal fusion
- Cholecystectomy and common duct exploration
- Partial excision bone
- Hysterectomy, abdominal and vaginal
- Colorectal resection
- Excision, lysis peritoneal adhesions
- Appendectomy
- Treatment, fracture or dislocation of hip and femur
- Oophorectomy, unilateral and bilateral
- Coronary artery bypass graft (CABG)
- Inguinal hernia repair
- Distal radial Fracture
- Thoracotomy
- Mastectomy
- Knee Replacement
- Treatment, fracture or dislocation of lower extremity (other than hip or femur)

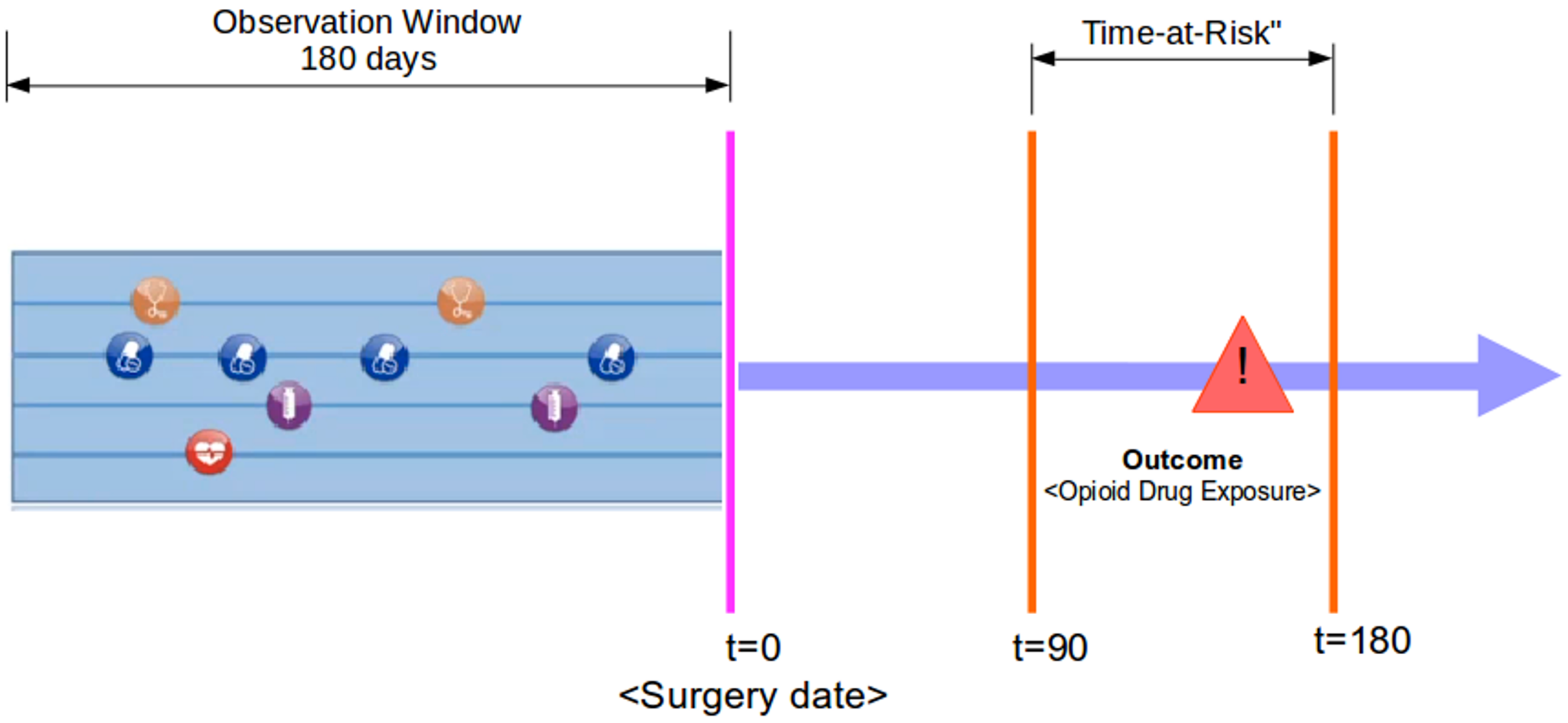
Cohort characterization

Domain	Covariate	Opioid Users	Prolonged Opioid Users	Not prolonged Opioid Users	Incidence (%)	Std Mean Diff
Gender	Female	17,805 (56.54%)	2,882 (59.11%)	14,923 (56.07%)	16.19%	0.0283
	Male	13,685 (43.46%)	1,994 (40.89%)	11,691 (43.93%)	14.57%	-0.0329
Race	White	20,227 (64.23%)	3,022 (61.98%)	17,205 (64.65%)	14.94%	-0.0237
	Black or African American	1,137 (3.61%)	261 (5.35%)	876 (3.29%)	22.96%	0.0701
	Asian	4,121 (13.09%)	583 (11.96%)	3,538 (13.29%)	14.15%	-0.0266
	Native Hawaiian	306 (0.97%)	70 (1.44%)	236 (0.89%)	22.88%	0.0360
	Other	5,103 (16.21%)	879 (18.03%)	4,224 (15.87%)	17.23%	0.0370
Ethnicity	Hispanic or Latino	4,039 (12.83%)	733 (15.03%)	3,306 (12.42%)	18.15%	0.0498
	Not Hispanic or Latino	26,954 (85.6%)	4,101 (84.11%)	22,853 (85.87%)	15.21%	-0.0135
Age Group	15-19	339 (1.08%)	37 (0.76%)	302 (1.13%)	10.91%	-0.0273
	20-24	816 (2.59%)	124 (2.54%)	692 (2.6%)	15.20%	-0.0025
	25-29	863 (2.74%)	145 (2.97%)	718 (2.7%)	16.80%	0.0116
	30-34	1,185 (3.76%)	208 (4.27%)	977 (3.67%)	17.55%	0.0211
	35-39	1,491 (4.73%)	247 (5.07%)	1,244 (4.67%)	16.57%	0.0125
	40-44	1,894 (6.01%)	320 (6.56%)	1,574 (5.91%)	16.90%	0.0184
	45-49	2,558 (8.12%)	433 (8.88%)	2,125 (7.98%)	16.93%	0.0218
	50-54	3,193 (10.14%)	553 (11.34%)	2,640 (9.92%)	17.32%	0.0308
	55-59	3,539 (11.24%)	613 (12.57%)	2,926 (10.99%)	17.32%	0.0325
	60-64	3,719 (11.81%)	642 (13.17%)	3,077 (11.56%)	17.26%	0.0323
	65-69	4,113 (13.06%)	593 (12.16%)	3,520 (13.23%)	14.42%	-0.0211
	70-74	3,432 (10.90%)	405 (8.31%)	3,027 (11.37%)	11.80%	-0.0691
	75-79	2,411 (7.66%)	330 (6.77%)	2,081 (7.82%)	13.69%	-0.0275
	80-84	1,470 (4.67%)	173 (3.55%)	1,297 (4.87%)	11.77%	-0.0457
	85-89	467 (1.48%)	53 (1.09%)	414 (1.56%)	11.35%	-0.0288
Cancer	Cancer Diagnosis	10,683 (33.40%)	2,174 (43.91%)	8,509 (31.47%)	20.42%	-0.2588

Cohort characterization

Domain	Top Covariates (P value < 0.0001)	Prolonged Opioid Users	Not prolonged Opioid Users	Incidence (%)	Std Mean Diff
Cancer	Primary malignant neoplasm of female breast	630 (12.72%)	1574 (5.82%)	1.97%	0.1603
	Carcinoma in situ of breast	252 (5.09%)	516 (1.91%)	0.79%	0.1202
	Primary malignant neoplasm of upper outer quadrant of female breast	154 (3.11%)	330 (1.22%)	0.48%	0.0908
	Primary malignant neoplasm of rectum	188 (3.80%)	477 (1.76%)	0.59%	0.0862
	Primary malignant neoplasm of liver	82 (1.66%)	148 (0.55%)	0.26%	0.0747
	Primary malignant neoplasm of prostate	103 (2.08%)	1251 (4.63%)	0.32%	-0.0983
	Secondary malignant neoplastic disease	157 (3.17%)	420 (1.55%)	0.49%	0.0744
	Secondary malignant neoplasm of bone	138 (2.79%)	352 (1.30%)	0.43%	0.0734
	Primary malignant neoplasm of rectosigmoid junction	113 (2.28%)	291 (1.08%)	0.35%	0.0658
	Neoplasm of uncertain behavior of breast	42 (0.85%)	71 (0.26%)	0.13%	0.0556
Condition	Transplanted lung present	222 (4.48%)	80 (0.30%)	0.69%	0.1915
	Dyspnea	550 (11.11%)	1155 (4.27%)	1.72%	0.1743
	Hypoxemia	329 (6.65%)	524 (1.94%)	1.03%	0.1606
	Complication of transplanted lung	95 (1.92%)	33 (0.12%)	0.30%	0.1258
	Chronic respiratory failure	102 (2.06%)	43 (0.16%)	0.32%	0.1276
	Primary malignant neoplasm of female breast	630 (12.72%)	1574 (5.82%)	1.97%	0.1603
	Pleural effusion	507 (10.24%)	1167 (4.32%)	1.59%	0.1552
	Disorder of lung	445 (8.99%)	1005 (3.72%)	1.39%	0.1479
	Fever	383 (7.74%)	812 (3.00%)	1.20%	0.1444
	Insomnia	1240 (4.59%)	501 (10.12%)	3.88%	0.1443
Procedure	Anesthesia for heart/lung transplant	227 (4.58%)	76 (0.28%)	0.71%	0.1951
	Breast reconstruction, with tissue expander	347 (7.01%)	345 (1.28%)	1.08%	0.1991
	Backbench standard preparation of cadaver donor lung allograft prior to transplantation	154 (3.11%)	32 (0.12%)	0.48%	0.1665
	Lung transplant with cardiopulmonary bypass	156 (3.15%)	38 (0.14%)	0.49%	0.1659
	Pulmonary stress testing	157 (3.17%)	65 (0.24%)	0.49%	0.1587
	Mastectomy, simple, complete	491 (9.92%)	843 (3.12%)	1.54%	0.1883
	Reconstructive procedures on breast	347 (7.01%)	484 (1.79%)	1.08%	0.1759
	Biologic implant for soft tissue reinforcement (ie, breast, trunk)	224 (4.52%)	206 (0.76%)	0.70%	0.1636
	Double lung transplant	95 (1.92%)	26 (0.10%)	0.30%	0.1284
	Transplantation of Bilateral Lungs, Allogeneic, Open Approach	73 (1.47%)	11 (0.04%)	0.23%	0.1165

Prediction definition



Covariate setting

All features (**58,405**)

1. Demographics: gender, age group, race, ethnicity

2. Procedure
3. Condition
4. Drug
5. Measurement
6. Condition Era
7. Drug Era

6-domain features (42,199)

8. Procedure count
9. Condition count
10. Drug count
11. Measurement count

12. Condition Group
13. Drug Group

Different subsets of 42,199 features (P-values < 0.0001):

- CHI2 (1,922) 4.5%
- IG (1,917) 4.5%
- SOR (3,266) 7.7 %
- PNF (4,548) 10.7 %

Prediction results

Table 1. Performance of ML algorithms on prolonged opioid use prediction over different feature sets

Model	Covariates	ACC	PPV	SENS.	Prolonged F1-Score	NPV	SPEC.	Not Prolonged F1-Score
Logistic Regression	All 58,405	0.580	0.240	0.792	0.368	0.934	0.541	0.685
Random Forest	All 58,405	0.581	0.242	**0.796**	0.371	**0.935**	0.542	0.686
Adaboost	All 58,405	0.575	0.235	0.774	0.361	0.928	0.538	0.681
Decision Tree	All 58,405	0.561	0.218	0.711	0.334	0.910	0.533	0.672
	Top 4,548 with PNF	0.791	0.324	0.323	0.323	0.876	0.877	0.876
Naïve Bayes	All 58,405	0.794	0.332	0.325	0.328	0.877	0.880	0.878
	Top 4,548 with PNF	**0.845**	**0.498**	0.312	**0.384**	0.882	**0.942**	**0.911**

Table 2. Performance of lasso logistic regression with all covariates over different time-at-risks

TAR (Incidence Rate %)	AUC	PPV	SENS.	Prolonged F1-Score	NPV	SPEC.	Not Prolonged F1-Score
90-365 (29.56%)	0.724	**0.422**	0.728	**0.534**	0.836	**0.581**	**0.686**
180-365 (22.14%)	0.703	0.315	0.727	0.440	0.876	0.551	0.676
90-180 (15.49%)	0.742	0.240	0.792	0.368	0.934	0.541	0.685
90-180 (0.8%) with strict labeling	**0.871**	0.004	**0.933**	0.008	**1.000**	0.491	0.659

Naderalvojoud, B., Hond, A., …, Hernandez-Boussard, T. "Predicting Prolonged Opioid Use Following Surgery Using Machine Learning: Challenges and Outcomes" in *American Medical Informatics Association (AMIA) Annual Symposium*, 2022.

External validation and network study

Objectives

- Evaluate the performance, generalizability, and calibration of the ML models across multisite cohort subgroups: gender, diabetes, depression, obesity

- Evaluate the transportability of ML models based on population differences in the various CDM databases

- Identify common preoperative risk factors predictive of postoperative opioid use over CDM databases

- Incorporate ML models trained on different databases to increase generalizability

External validation and network study

What are we looking for?

- Data partners with CDM databases for running project R package
- We are very open to feedback, changes, and suggestions

What do we anticipate from our data partners?

Using ATLAS to:

- run the project cohort based on the definition setting provided in the project using JSON files,
- run the cohort characterization based on the definition setting provided in the project

Using R to:

- run the project prediction module from project Git repository for:
  - internal validation of the models trained on the local dataset
  - training models and share them for external validation

Thank you, wonderful OHDSI community

OHDSI
Observational Health Data Sciences and Informatics

To join the PORPOISE project, please contact:

Behzad Naderalvojoud, PhD
Postdoctoral Research Fellow
behzadn@stanford.edu

Tina Hernandez-Boussard, PhD MPH, MS, FACMI
Associate Professor of Medicine
boussard@stanford.edu

Catherine M Curtin, MD
Professor of Surgery
ccurtin@stanford.edu