Stanford MEDICINE | Center for Biomedical Informatics Research

Machine Learning for Predicting Patients at Risk of Prolonged Opioid Use Following Surgery

Behzad Naderalvojoud, PhD

Postdoctoral Research Fellow

NIH National Library of Medicine

Research reported in this publication was supported by the National Library Of Medicine of the National Institutes of Health under Award Number R01LM013362. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Background: opioid use

- High-potency analgesic
- Highly addictive
- Serious complications
- Significant morbidity and mortality
- Risk factor for prolonged opioid use and

Is that feasible to identify postoperative patients at risk for prolonged opioid use based on EHRs?

Background: state-of-the-art

- Solutions:
  - Opioid risk assessment tools
  - ML models
- Problems:
  - Non-standardized data from different sources
  - Generalizability and reliability
  - Data bias
  - Evaluation bias

Study objective

- AI to identify prolonged opioid users
- Generalizable, discriminating, calibrated AI

A Network Study on OMOP Databases

PORPOISE

POSTOPERATIVE PROLONGED OPIOID USE

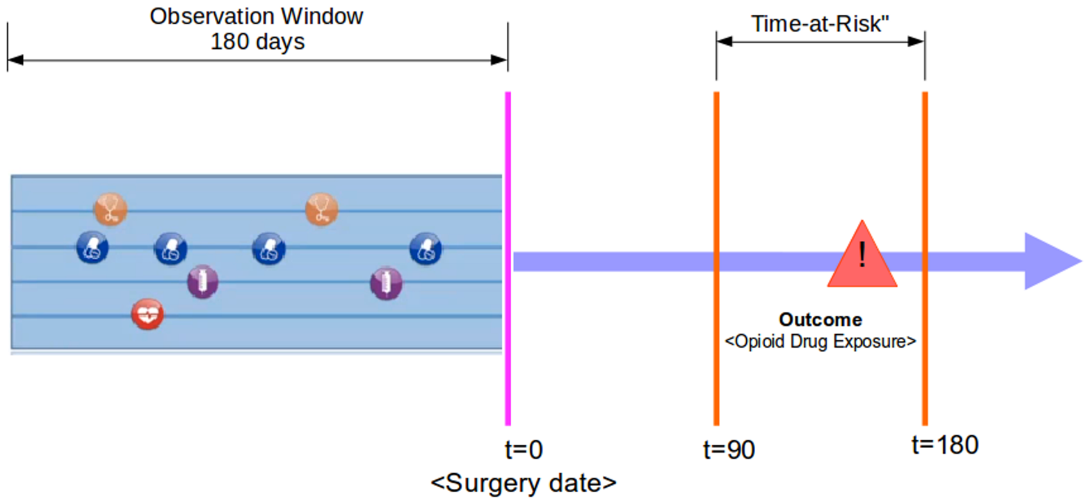
HADES

HEALTH ANALYTICS DATA-TO-EVIDENCE SUITE

- ✓ Transparency
- ✓ Standardized model development and evaluation
- ✓ CDM covariates
- ✓ External validation

Methods

Problem definition:



Model development components:

Data extraction

- Phenotypes
- Population setting
- Covariates
- Lookback period
- …

Data preprocessing

- Observation time
- Loss to follow-up
- Sample size
- Missing data
- …

Model development

- Classifier
- Hyperparameters
- Class imbalance
- Ensemble learning
- …

Model validation

- Validation strategy
- Evaluation measures
- Recalibration
- Model updating
- …

Phenotyping rules

Target Cohort

Patients underwent surgery (first surgery if multiple) (n=121,404)

Surgeries between 2008 and 2019 (n=96,420)

With admission and discharge dates (n=83,179)

At least one opioid prescription (n=82,016)

Opioid prescription within +/- 30 days of surgery (n=81,399)

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

At least 2 encounters before and 2 encounters after surgery (n=44,751)

At least 1 encounter more than 2 years after surgery (n=37,163)

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

Data extraction

- Phenotypes
- Population setting

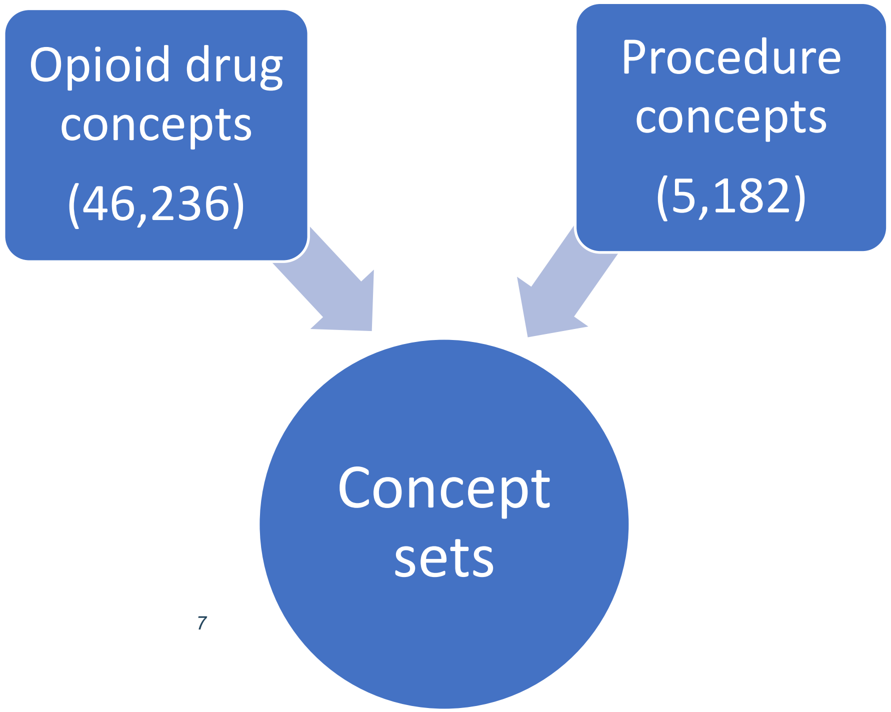
Outcome Cohort

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

Phenotyping rules: 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



Covariate and population settings

Data extraction

- Population setting
- Covariates
- Lookback period

Data preprocessing

- Removing redundant cov.
- Removing infrequent cov.
- Normalizing cov.

All features (58,405)
180 days prior to surgery

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

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

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

12. Condition Group
13. Drug Group

Population setting:

- washoutPeriod = 0,
- firstExposureOnly = FALSE (because it was considered in the cohort selection),
- removeSubjectsWithPriorOutcome = FALSE,
- priorOutcomeLookback = 9999,
- riskWindowStart = 90,
- riskWindowEnd = 180,
- minTimeAtRisk = 90,
- startAnchor = cohort start,
- endAnchor = cohort start,
- requireTimeAtRisk = TRUE,
- includeAllOutcomes = TRUE

Model development and evaluation

Implementation:

- ATLAS, PLP package 4.0.4
- Shiny app:

https://prolonged-opioid-use-prediction.shinyapps.io/shiny-app/

ML models:

- Lasso Logistic Regression
- Random Forest
- AdaBoost
- Decision Tree
- Naive Bayes

Evaluation metrics:

- Model performance:
  ACC, SENS., SPEC.,
  PPV, NPV
- Model discrimination:
  ROC, PRC, AUC, AUPRC
- Model calibration

Hyper parameter:

10-fold cross validation with a grid search strategy

Validation strategy:

- 10-fold cross validation
- Random split:
  80% train, 20% test
- External validation:
  PORPOISR study

PORPOISE

POSTOPERATIVE PROLONGED OPIOID USE

External validation:

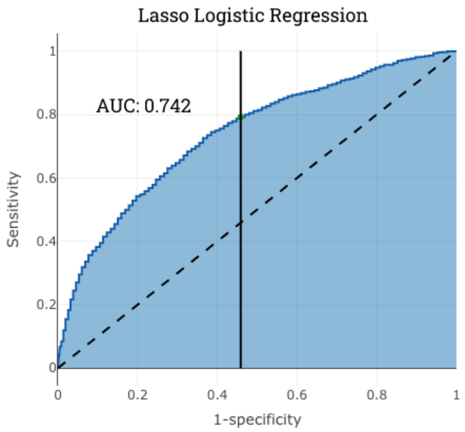
- Codes and study protocol:

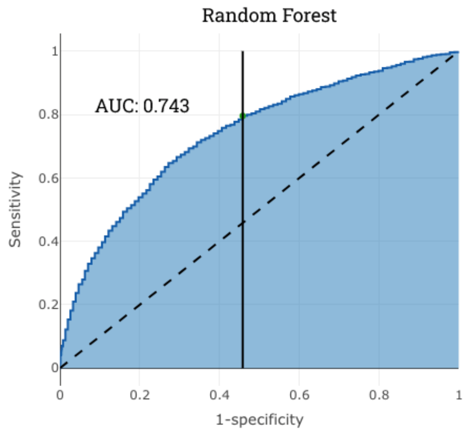
https://github.com/ohdsi-studies/PORPOISE

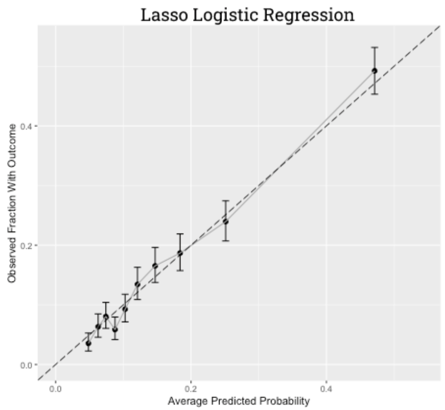
Results

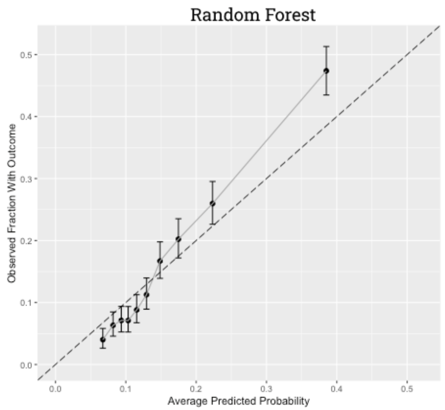
Performance → Discrimination → Calibration

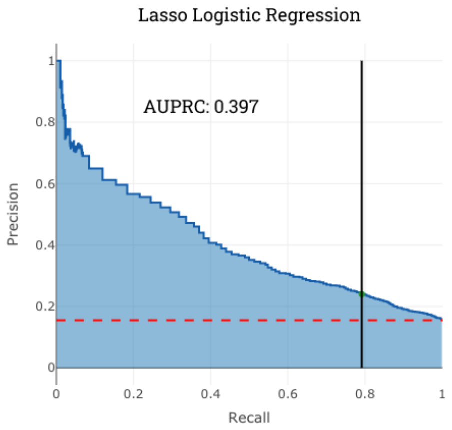
Models	ACC	AUC	AUPRC	SENS.	PPV	SPEC.	NPV
Lasso Logistic Regression	0.580	0.742	0.397	0.792	0.240	0.541	0.934
Random Forest	0.581	**0.743**	**0.399**	**0.796**	0.242	0.542	0.935
AdaBoost	0.575	0.730	0.390	0.774	0.235	0.538	0.928
Decision Tree	0.561	0.676	0.313	0.711	0.218	0.533	0.910
Naïve Bayes	**0.794**	0.680	0.270	0.325	**0.332**	**0.880**	0.877

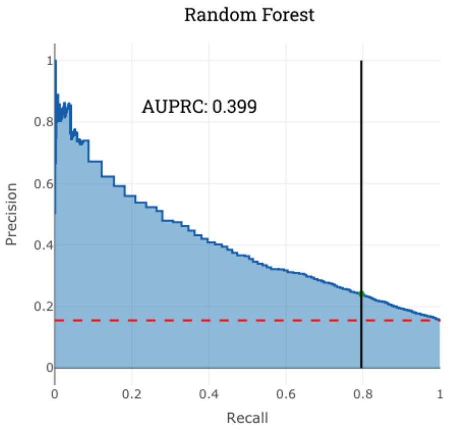








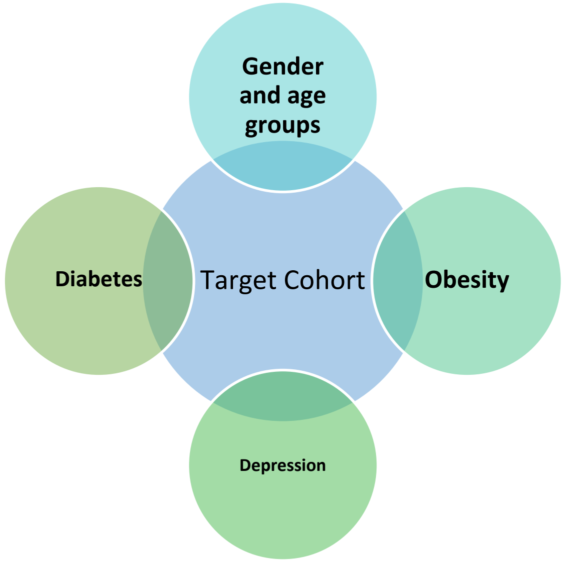




Conclusion and future work

- Conclusion:
  - LR and RF: Highest discrimination and risk calibration
  - NB: higher specificity
  - LR + NB in a single ensemble model: a better balance of sensitivity and specificity

- Future work:
  - External validation across subgroups
  - Evaluate the transportability
  - Ensemble learning
  - Federated learning



Thank you for your attention!

Stanford MEDICINE | Center for Biomedical Informatics Research

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

Tina Hernandez-Boussard, PhD MPH, MS, FACMI
Project PI
boussard@stanford.edu

I would like to thank Tina Seto, Priya Desai, and Dr. Catherine Curtin as collaborators.