Developing and Validating a Machine-learning Algorithm to Predict the Risk of Incident Opioid Use Disorder among OneFlorida+ Patients: A Prognostic Modeling Study

Jabed Al Faysal, MS[1], Weihsuan Lo-Ciganic, PhD[2,3], Walid F. Gellad, MD[2], Yonghui Wu, PhD[1], Christopher A. Harle, PhD[4,6], Khoa Nguyen, PharmD[1], James L. Huang, PhD[1], Gerald Cochran, PhD[5], Debbie L. Wilson, PhD[1], Stephanie A S Staras, PhD[1], Siegfried O.F. Schmidt, MD, PhD[1], Eric I. Rosenberg, MD, MSPH[1], Danielle Nelson, MD[1], Shunhua Yan, Med[1], Gary M. Reisfield, MD[1], William M. Greene, MD[1], Courtney Kuza, PhD[2], Md Mahmudul Hasan, PhD[1]

[1]University of Florida, Gainesville, FL, USA, [2]University of Pittsburgh, Pittsburgh, PA, USA, [3]North Florida/South Georgia Veterans Health System, Gainesville, FL, USA, [4]Indiana University, Indianapolis, IN, USA, [5]University of Utah, Salt Lake City, UT, USA; [6]Regenstrief Institute, Indianapolis, IN, USA

Introduction

- In 2023, nearly 6 million Americans had opioid use disorder (OUD), and 80 thousand people lost their lives for opioid-related overdoses.
- Early identification of individuals at risk of OUD remains challenging due to limitations of traditional screening methods and the absence of personalized risk stratification.
- Machine learning (ML) has shown promise to enhance early identification of OUD and stratifying high-risk patients by capturing complex, non-linear relationships in real-world data (RWD).
- This study aimed to develop and validate ML models to predict the risk of incident OUD among OneFlorida+ patients initiating opioid therapy.
- Our goal is to provide an explainable ML framework leveraging large-scale EHR data to stratify OUD risk at the patient level, improving model transparency and clinical applicability.

Methodology

- **Design:** Retrospective prognostic modeling using 2017–2022 OneFlorida+ electronic health records (EHR).
- **Population:** 182,083 adults (≥18 years) with ≥1 opioid prescription; Index date defined as the first date of opioid prescription.
- **Follow-up:** Sliding-window based (3 months for predictors → followed by 3 months for outcomes) continuous screening.
- **Predictors:** 183 demographic, clinical, medication use, and healthcare utilization updated in sequential 3-month windows.
- **Outcome:** First diagnosis of OUD (identified by ICD-10 codes).
- **Models:** Gradient boosting machine (GBM), Elastic net (EN), Random forest (RF), Least absolute shrinkage and selection operator (LASSO).
- **Evaluation:** AUC-ROC, sensitivity, specificity, PPV, NPV, NNE; explainability via SHAP; following **TRIPOD+AI** guidelines.

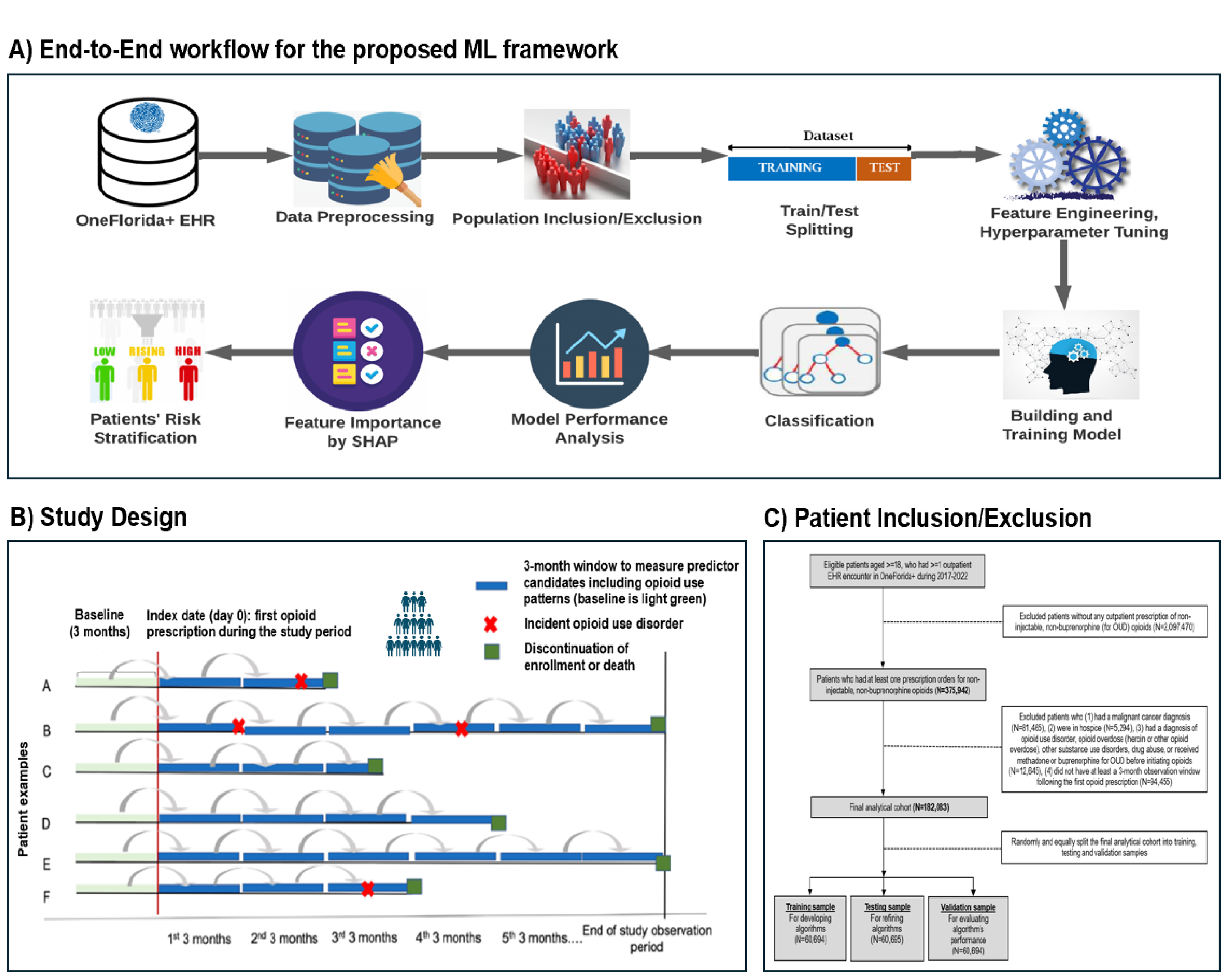

Fig. 1. Overall Study Design and ML Methodology

Results

Patients across train, test, validation sets had similar characteristics (mean age ≈ 55 years; 64% female; 68% White)

Model Performance:

- GBM (AUC = 0.879 (95% CI 0.874–0.884)) outperformed all other ML models. (Fig. 2.A)
- At ~50-60% sensitivity, GBM is most efficient by NNE (~10-15 reviews per true OUD case). (Fig. 2.B)
- GBM controls alerts better, fewer unnecessary alerts per 100 patients compared to other models (Fig. 2.C).

Shapley Additive exPlanations (SHAP) Analysis:

- Most influential risk factors using GBM: Age, No. of outpatient visit, history of back pain, history of other pain, etc. (Fig. 3)

Risk Stratification analysis using GBM (Fig. 4):

- The top decile subgroup (high-risk) had a PPV of 15.06%, NPV of 99.8%, and NNE of 89; captured ~68% of all OUD cases.
- Top 1%: ~1 in 10 develop OUD → highest risk
  Top 2-5%: very high risk → intervene
  Top 6-10%: high risk → monitor/intervene

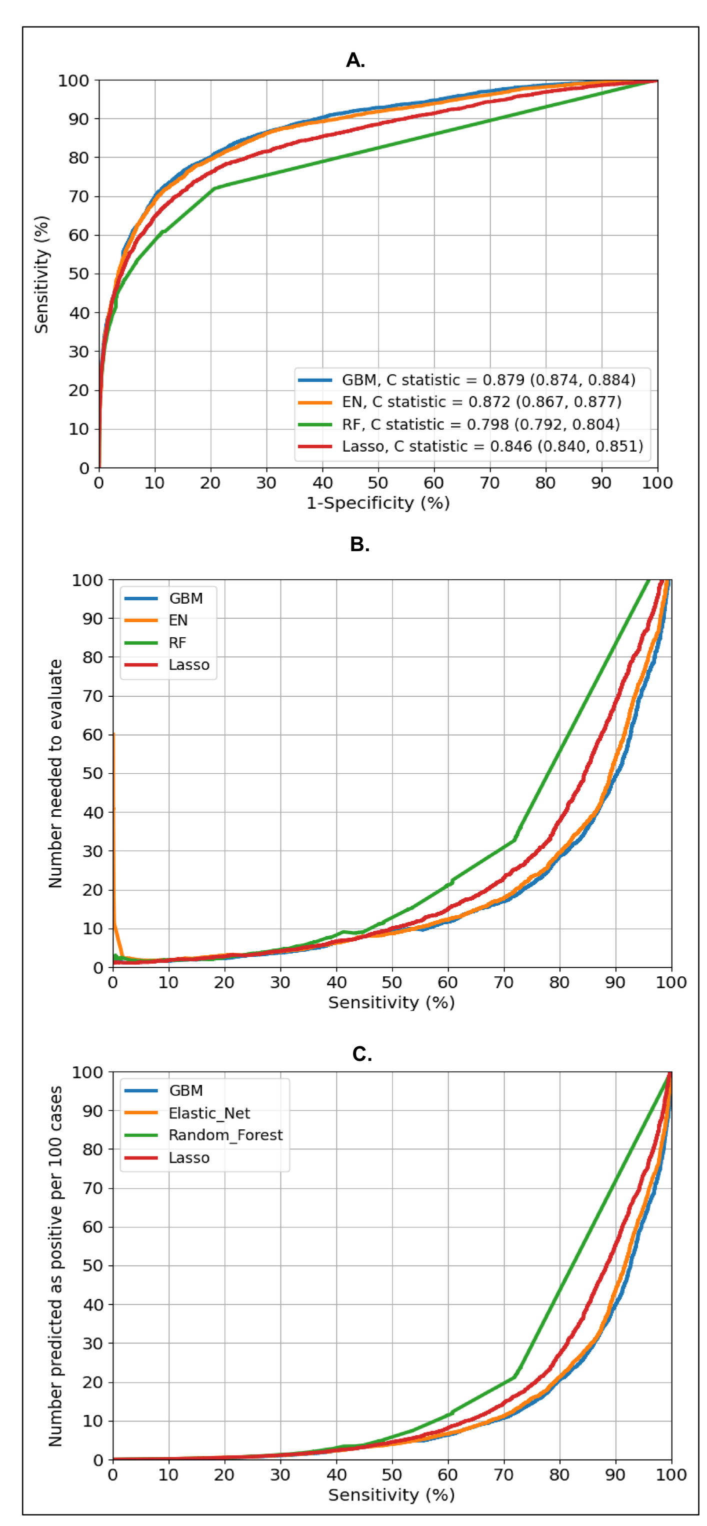

Fig. 2. A) ROC curve; B) Number needed to evaluate (NNE); C) Number predicted positive per 100 patients

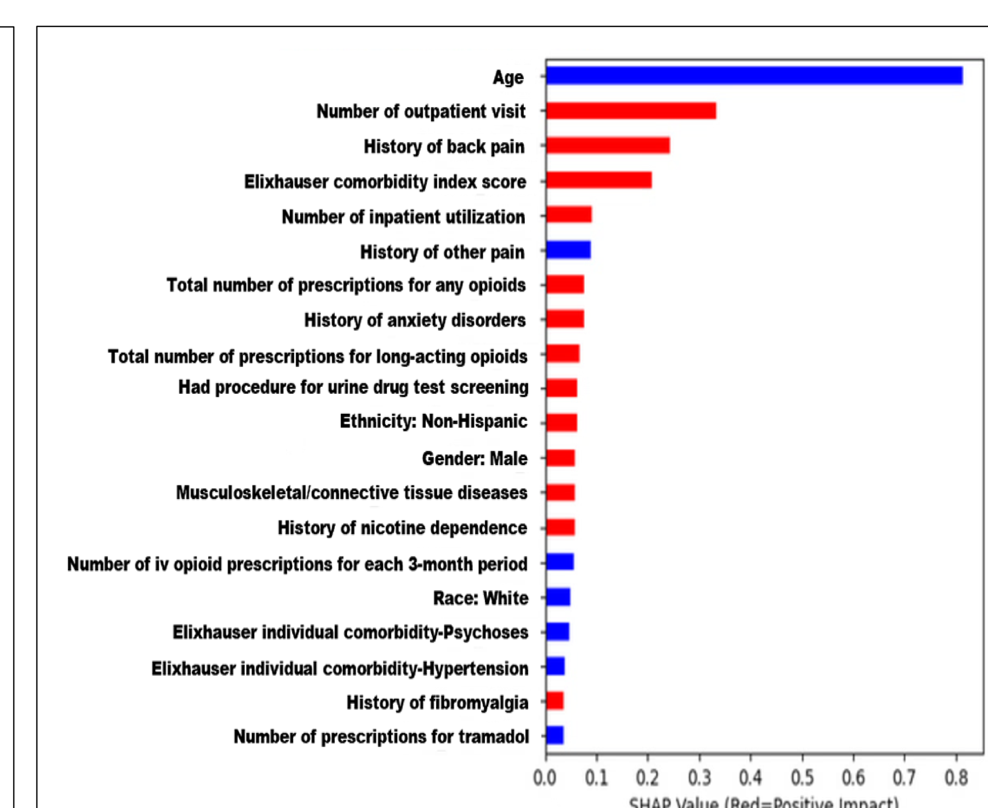

Fig. 3. Variable importance plots using SHAP analysis from GBM model

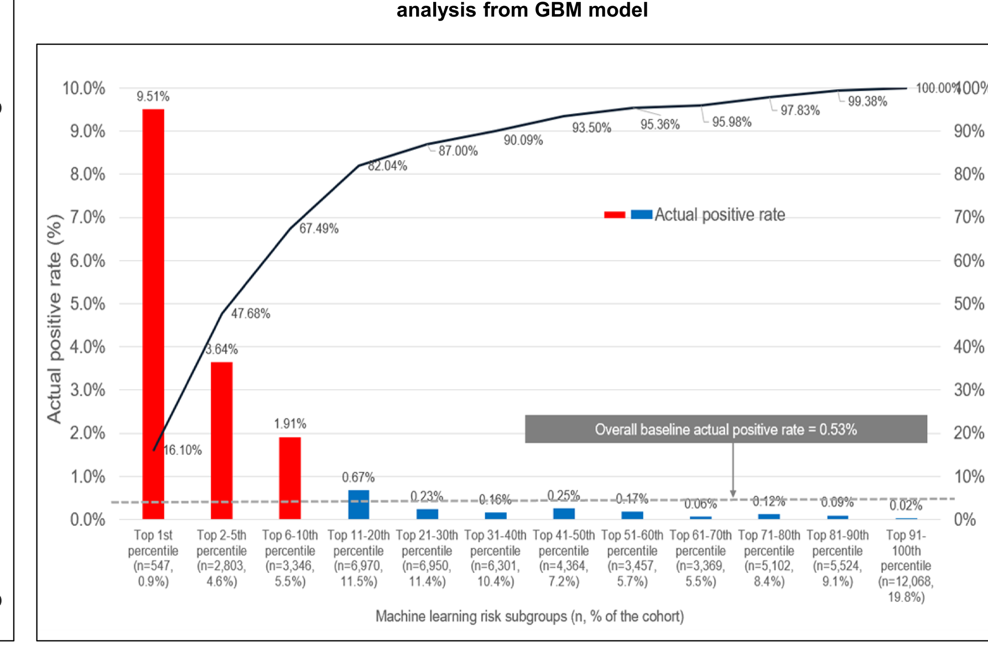

Fig. 4. Patient risk stratification by GBM model in the validation cohort

Discussion

- ML models achieved high performance in predicting short-term OUD risk, outperforming traditional methods.
- Risk stratification identified small subgroups accounting for the majority of OUD cases, enabling targeted intervention.
- High NPV supports safe exclusion of low-risk patients, reduces alert fatigue, and builds clinician confidence in models' low-risk predictions.
- The study was limited by potential unmeasured factors (e.g., SDoH), use of single EHR network, lack of external validation (ongoing work).
- Future work will focus on integrating SDoH factors and evaluating effectiveness in real-world clinical workflow with provider feedback.

Contact: Jabed Al Faysal (faysalj@ufl.edu)