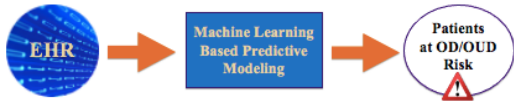


# Opioid Epidemic Study

Xinyu Dong, Advisor: Fusheng Wang  
Data Management and Biomedical Data Analytics Lab

## OD/ODU Prediction with Machine Learning

Critical knowledge is needed for opioid epidemic prevention  
Leveraging large scale electronic health records to identify patients who are at high risk of opioid overdose/opioid use disorder using intelligent machine learning/deep learning methodology  
Detecting most influential features for high risk patients



The US is experiencing opioid epidemic due to the misuse and abuse of opioids

- > 130 opioid overdose death per day
- > 10.3 million people misused opioid prescriptions
- > 47,600 opioid overdose death per year
- > 2.0 million people had opioid use disorder
- > \$504 billion economic burden to combat opioid epidemic



## Prediction Model

> Database includes de-identified EHR data from over 600 participating Cerner client hospitals and clinics in the United States.

> We propose a sequential deep learning model built upon LSTM to predict OUD among patients prescribed with opioid medications in their past health records.

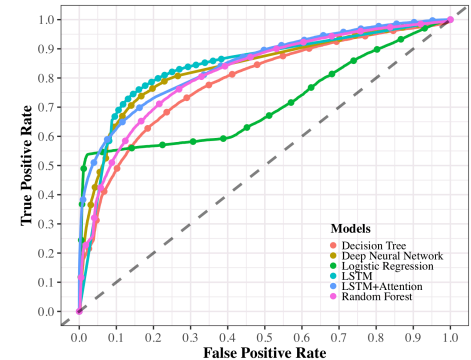
> This sequential model can better represent the progression of diseases and also identify the most important features as risk factors for the diseases.

> Health Facts also includes medication dosage and administration information, vital signs, laboratory test results, surgical case information, other clinical observations, and health systems attributes.

### Quick Facts

- > 5 million patients included in this study
- > 100+ million clinical visits data
- > 1,400+ clinical features used for prediction
- > 6 types of different features (Diagnosis, Procedure, Medication, Laboratory Test, Clinical Event, Demographics)
- > 0.8023 of F1-score, 0.9369 of ROAUC, 0.8184 of precision, 0.7865 of recall scores achieved by our model

## AUC Curves of Different Methods



## Findings

> To evaluate the performance, our model is compared with other baselines including random forest, decision tree, logistic regression, dense neural network.

> The LSTM model achieved a promising result with the best F1 score and can identify important features

> There are 12 features on medications, including the aggregated feature MME, 8 clinical events, 5 diagnoses, and 25 laboratory tests. Among the top 10 features in the list, medications dominate, followed by clinical events and lab tests, e.g. "Pain Scale Score" implies the need for pain management, where opioid is commonly prescribed.