Identifying Individuals at Risk for a Future Opioid Use Disorder Diagnosis with Machine Learning

David M. Simon, PhD[1], Meridith Blevins Peratikos, MS[1,2], Amber R. Watson1, PharmD, Jianhong Shen[1], PhD, Elizabeth Ann Stringer, PhD[1]
[1]axialHealthcare, Nashville, Tennessee; [2]Vanderbilt University Medical Center, Department of Biostatistics, Nashville, Tennessee

BACKGROUND

The growing prevalence of opioid use disorder (OUD) is recognized as a major public health crisis associated with elevated health care costs and poor health outcomes. Previous research has identified numerous risk factors that contribute to whether a patient prescribed opioids will be diagnosed with OUD in the future. Manual review of healthcare records for OUD risk factors is resource intensive due to the diversity of factors and the complexity of interactions found in healthcare data. Machine learning offers great promise for efficiently distilling this complex information into more easily utilized forms.

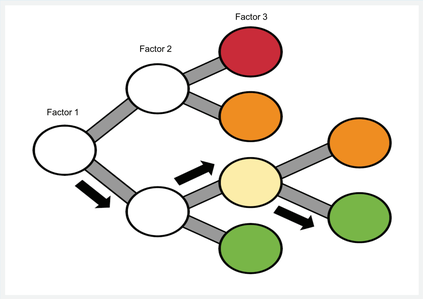
OBJECTIVE

The objective of the study was to generate a predictive machine learning model that identified whether patients would be diagnosed with opioid use disorder within the next 12 months.

METHODS

Patients were selected from commercial and government healthcare claims databases on the basis of two years of continuous insurance eligibility between January 2015 and January 2018.

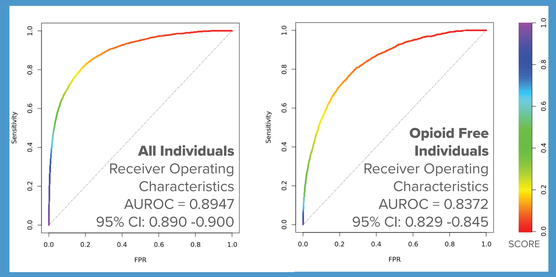
- 23,371 individuals (0.00289%) were diagnosed with new cases of OUD in the outcome year (positive cases).
- We randomly selected a set of 1,051,695 individuals without an OUD outcome as negative cases.
- Model optimized and fit using 80% of available data.
- Evaluated on the remaining 20% of available data.
- xGBoost based machine learning.

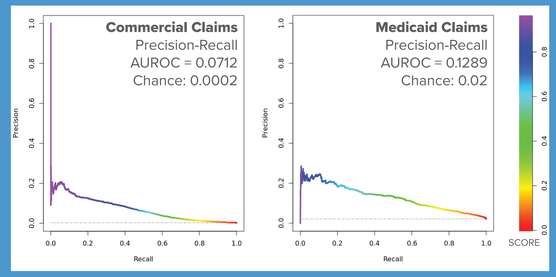


Hundreds of decision trees assess the level of each risk factor to assign patient risk.

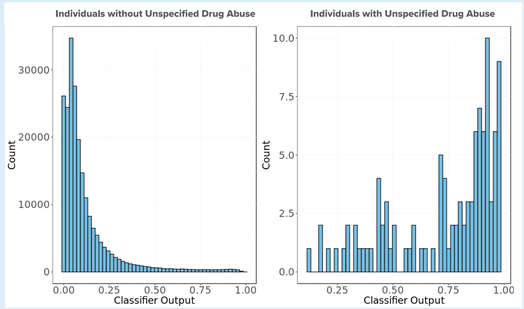
RESULTS

A high level of predictive accuracy was achieved for patients currently on opioids and patients who were not previously prescribed opioids.





Automatic feature selection retained 401 (41%) of the initial 983 features. Rare diagnoses such as unspecified drug abuse (0.087% of training data) were correctly retained by the model. Model precision varied somewhat across data sets.



CONCLUSIONS

The predictive model we developed offers valuable insight for supporting pain management and patient screening in individuals who are currently prescribed opioids. Appropriate intervention with these patients before they are opioid dependent could reduce negative healthcare outcomes. Furthermore, the model achieved substantial sensitivity for patients who are not currently on opioids, supporting the possibility of deployment as a decision support tool when determining whether opioid based treatment is appropriate.

Contact: dsimon@axialhealthcare.com