

# International Journal of Population Data Science

Journal Website: [www.ijpds.org](http://www.ijpds.org)



## Enhanced Depression Phenotyping Using Machine Learning and Electronic Medical Records for Disease Surveillance: A Validation Study in Inpatient Settings

Na Lin<sup>1</sup>, Elliot A. Martin<sup>2</sup>, Seungwon Lee<sup>2</sup>, Allison Grothman<sup>3</sup>, Na Li<sup>1</sup>, Kamala Adhikari<sup>4</sup>, Kiarash Riazi<sup>1</sup>, Cathy A. Eastwood<sup>1</sup>, Danielle A. Southern<sup>5</sup>, and Hude Quan<sup>1</sup>

<sup>1</sup>Centre for Health Informatics, University of Calgary, Calgary, Canada; Department of Community Health Sciences, Cumming School of Medicine, Community Health Sciences, University of Calgary, Calgary, Canada; Libin Cardiovascular Institute of Alberta, Calgary, Canada

<sup>2</sup>Centre for Health Informatics, University of Calgary, Calgary, Canada; Provincial Research Data Services, Health Shared Services Provincial Health Corporation, Calgary, Canada

<sup>3</sup>Faculty of Graduate Studies, University of Calgary, Calgary, Canada

<sup>4</sup>Department of Community Health Sciences, Cumming School of Medicine, Community Health Sciences, University of Calgary, Calgary, Canada; Cancer Prevention and Screening Innovation, Public Health Evidence and Innovation, Primary Care Alberta, Calgary, Canada

<sup>5</sup>Centre for Health Informatics, University of Calgary, Calgary, Canada; Department of Community Health Sciences, Cumming School of Medicine, Community Health Sciences, University of Calgary, Calgary, Canada

### Background

Accurate detection of inpatient depression is essential for surveillance and resource allocation. Traditional International Classification of Diseases (ICD) codes suffer from low sensitivity and delayed availability.

### Objective

To develop and validate machine learning (ML) phenotyping algorithms for depression using unstructured inpatient clinical notes from electronic medical records (EMRs), compared with ICD-10-CA codes and manual chart review.

### Methods

This retrospective cohort study included a derivation cohort (N = 3032 adults, Jan–June 2015) and a temporally independent external validation cohort (N = 10659 adults, 2017–2022) from Calgary acute care hospitals. Both cohorts were deterministically linked to DAD and EMR databases. Manual chart reviews defined the reference standard for depression status. We compared ICD-10-CA codes against three EMR-based algorithms: Keyword Search, Concept Model and Document-Concept Model. The latter two employed a supervised ML classifier (XGBoost) utilizing clinical concepts extracted via the cTAKES natural language processing (NLP) pipeline. Primary measures were prevalence, sensitivity, positive predictive value (PPV), specificity, negative predictive value (NPV), and F1 score.

### Results

In derivation cohort (prevalence: 18.1%), the Concept Model achieved 76.4% sensitivity and 76.7% F1 score, significantly outperforming Keyword Search (F1: 66.7%), Document-Concept Model (F1: 73.7%), and ICD-10-CA codes (sensitivity: 20.0%; F1: 32.6%). In external validation (prevalence: 13.1%), the Concept Model retained robust performance (sensitivity: 92.3%; F1: 80.1%), compared with ICD-10-CA (sensitivity: 6.1%; F1: 11.4%).

### Conclusion

An EMR-based Concept Model using NLP and ML markedly improves depression case detection over ICD coding in inpatient settings. The methods could potentially facilitate psychiatric surveillance and epidemiological research using real-world clinical data.

