## ABSTRACT

## Establishment of a predictive model for antidepressant efficacy based on machine learning

Yiyao Liu<sup>1,2</sup>, Huitong Ni<sup>1</sup>, Teng Zhi<sup>1</sup>, Ziqi Zhao<sup>1</sup>, Xiaoxi Zeng<sup>2</sup>, Ming Hu<sup>1\*</sup>, Zhiang Wu<sup>3\*</sup>

<sup>1</sup>West China School of Pharmacy, Sichuan University, Chengdu 610041, Sichuang Province, China <sup>2</sup>West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu 610041, Sichuang Province, China <sup>3</sup>Yeehong Business School, Shenyang Pharmaceutical University, Beijing 100055, China

**Objective:** Establishing a model can accurately predict the depressive patient's response to different drugs, thereby providing personalized treatment plans to improve treatment outcomes.

**Methods:** Depressive patients who have taken Paroxetine hydrochloride or Venlafaxine hydrochloride or Agomelatine at a hospital in Sichuan Province are the subjects. By analyzing their medical records, medication histories, and basic information, an predictive model is constructed to predict the efficacy of antidepressant medications based on eXtreme Gradient Boosting.

**Results:** For the prediction model of Paroxetine hydrochloride, 52 variables selected by the model were used to construct an efficient predictive model. In the training set, the model achieved an AUC value of 0.6354 and a K-S value of 0.1944, indicating good performance in correctly classifying positive and negative samples and distinguishing different predictive probability thresholds. In the validation set, the AUC was 0.6065, and K-S was 0.1847, confirming the model's effectiveness on new data. Compared to actual clinical data, the efficacy of sertraline hydrochloride was approximately 61.9%. The model's predictions aligned well with real-world data, reinforcing its reliability and practicality. As for venlafaxine hydrochloride, the model achieved an AUC of 0.5745 and K-S of 0.149 on the training set, while the validation set showed an AUC of 0.5298 and K-S of 0.0597. These results suggest that the model's performance in predicting venlafaxine hydrochloride is mediocre. Comparing with clinical data, the efficacy of venlafaxine hydrochloride was approximately 68.9%, indicating a discrepancy between the model's predictions and actual outcomes, possibly due to the uneven distribution of the training set samples. As for agomelatine, due to the insufficient number of agomelatine samples collected at the sample hospital (less than 200), an effective predictive model could not be established.

**Conclusion:** In clinical practice, doctors can make more informed treatment decisions and achieve personalized antidepressant therapy with the assistance of this model.

Key words: predictive model, machine learning, antidepressant efficacy, depressive patients, clinical practice

\*Corresponding Author:

Zhiang Wu, E-mail: wuzhiang@yeehongedu.cn; Ming Hu, E-mail: huming@scu.edu.cn Received: 15 June 2024; Published: 15 July 2024 https://doi.org/10.54844/hd.2024.0012

<sup>2</sup> This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which allows others to copy and redistribute the material in any medium or format noncommercially, as long as the author is credited and the new creations are licensed under the identical terms.