

From Model to Bedside: What Kind of OSA Risk Prediction Tools Do We Need More of? [Letter]

Hongyu Pan¹, Yuchang Fei²

¹Department of Traditional Chinese Medicine, Shangyu People's Hospital of Shaoxing, Shaoxing University, Shaoxing, Zhejiang, 312300, People's Republic of China; ²Department of Integrated Chinese and Western Medicine, The First People's Hospital of Jia Shan, Jia Shan Hospital Affiliated of Jiaxing University, Jiaxing, Zhejiang, 314100, People's Republic of China

Correspondence: Yuchang Fei, Department of Integrated Chinese and Western Medicine, The First People's Hospital of Jia Shan, Jia Shan Hospital Affiliated of Jiaxing University, Jiaxing, Zhejiang, People's Republic of China, Email feiyuchang123456@163.com

Dear editor

We carefully read the recent article “Construction and Validation of a Machine Learning-Based Risk Prediction Model for Sleep Quality in Patients with OSA” by Tong.¹ The authors constructed a model for predicting sleep quality in OSA patients based on the LightGBM algorithm and achieved good performance in the validation set and also identified key predictors using the SHAP method. This study is important in advancing the identification of sleep disorders in OSA patients and the development of clinical decision support tools. However, after reading the full paper in detail, we have a couple of doubts that we would like to take advantage of this format hoping to communicate with the authors.

First, the authors screened and modeled 6 key variables (coffee consumption, weekly exercise, anxiety, depressive symptoms, OSA duration, and ODI) when performing predictive modeling. However, we note that no assessment of covariance (variance inflation factor) was seen in the paper for the variables that were ultimately included in the model.² Indeed, anxiety patients are often accompanied by depressive symptoms in the later stages of life.³ This co-morbid state has strong linear correlations, and simultaneous incorporation into the model without addressing covariance may affect model stability and interfere with the significance of the explanatory variables of the SHAP.

Second, sleep disorders are prevalent in patients with OSAHS, so the prevalence of sleep disorders in the authors' modeling data was as high as 70%, a rate that is consistent with the current state of clinical epidemiology research. In machine learning, however, differences between the two groups may lead to significant category imbalances. In most machine learning algorithms, the model tends to predict the category that accounts for a greater proportion of the population, which in turn reduces the ability to recognize the minority category (good sleep quality). However, the authors do not state whether this is corrected using techniques such as SMOTE oversampling, undersampling, or category weight adjustment.⁴

Finally, this study only uses whether the calibration curve is close to the ideal diagonal as the basis for model assessment in terms of model calibrability assessment. However, it is essentially a subjective visualization tool that cannot quantify the error between the model prediction probability and the actual outcome. In medical prediction modeling research, mainstream quantitative indicators such as the Brier score and the Hosmer-Lemeshow test can more objectively reflect the calibration performance of the model, which can help to improve the credibility and reproducibility of the model in clinical practice.⁵

We highly recognize the authors' application of machine learning methods to the hot clinical issue of predicting sleep disorders in OSA patients. As a reader with dual roles as an OSA patient and a physician, I am deeply impressed by the practical guidance of this study that regular exercise and appropriate noninvasive ventilation therapy play an important role in improving sleep quality. This study encourages us to pay more attention to the lifestyle management and treatment adherence of OSA patients.

Data Sharing Statement

Data sharing is not applicable to this article as no data were created or analysed in this communication.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the author(s) used [ChatGPT 4o] in order to [improve language and readability]. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Author Contributions

Hongyu Pan: Methodology, Formal analysis, Writing - Original Draft; Yuchang Fei: Conceptualization, Methodology, Supervision, Writing - Review & Editing. All authors agreed on the journal to which this communication will be submitted; agreed on the final version accepted for publication; agree to take responsibility and be accountable for the contents of this communication.

Funding

This work was supported by the Chinese Medicine Research Program of Zhejiang Province (no.2023ZF177).

Disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this communication.

References

1. Tong Y, Wen K, Li E, et al. Construction and validation of a machine learning-based risk prediction model for sleep quality in patients with OSA. *Nat Sci Sleep*. 2025;17:1271–1289. doi:10.2147/NSS.S516912
2. Tao X, Ye X. Relationships between vitamin C intake and COPD assessed by machine learning approaches from the NHANES (2017-2023). *Front Nutr*. 2025;12:1563692. doi:10.3389/fnut.2025.1563692
3. Yokomitsu K, Oimatsu R, Toh SHY, Sündermann O. Assessing the efficacy of the INTELLECT cognitive behavioral therapy mobile app for anxiety and depressive symptoms among at-risk Japanese employees: randomized controlled trial. *JMIR Mhealth Uhealth*. 2025;13:e60871. doi:10.2196/60871
4. Rani R, Jaiswal G, Bhushan S, et al. Enhancing liver disease diagnosis with hybrid SMOTE-ENN balanced machine learning models—an empirical analysis of Indian patient liver disease datasets. *Front Med*. 2025;12:1502749. doi:10.3389/fmed.2025.1502749
5. Yuan L, Wang L, Gao J, et al. OncoE25: an AI model for predicting postoperative prognosis in early-onset stage I-III colon and rectal cancer—a population-based study using SEER with dual-center cohort validation. *J Transl Med*. 2025;23(1):695. doi:10.1186/s12967-025-06663-4

Dove Medical Press encourages responsible, free and frank academic debate. The content of the Nature and Science of Sleep 'letters to the editor' section does not necessarily represent the views of Dove Medical Press, its officers, agents, employees, related entities or the Nature and Science of Sleep editors. While all reasonable steps have been taken to confirm the content of each letter, Dove Medical Press accepts no liability in respect of the content of any letter, nor is it responsible for the content and accuracy of any letter to the editor.

Nature and Science of Sleep

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>

Dovepress
Taylor & Francis Group