



# OPEN Comparing machine learning models for osteoporosis prediction in Tibetan middle aged and elderly women

Peng Wang<sup>1,4</sup>, Qiang Yin<sup>2,4</sup>, Kangzhi Ding<sup>1</sup>, Huaichang Zhong<sup>3</sup>, Qundi Jia<sup>1</sup>, Zhasang Xiao<sup>1</sup>✉ & Hai Xiong<sup>1</sup>✉

The aim of this study was to establish the optimal prediction model by comparing the prediction effect of 6 kinds of prediction models containing biochemical indexes on the risk of osteoporosis in middle-aged and elderly women in Tibet. This study adopted a multi-stage cluster random sampling cross-sectional survey method. From January 2022 to January 2024, we obtained biochemical and bone mineral density (BMD) data from high altitudes in Tibet. We built a predictive model of osteoporosis in three steps. First, we performed feature selection to identify factors associated with osteoporosis. Next, the eligible participants were randomly divided into a training set and a test set in a ratio of 8:2. Then, the prediction model of osteoporosis was established based on Random Forest, ANN, XGB, and SVM. Finally, we compared the performance of the prediction models using sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC) to select the best prediction model. Correlation analysis was used to screen indicators with statistical differences from T-score. Finally, Age ( $P < 0.01$ ), LDL-C ( $P < 0.05$ ), UA ( $P < 0.01$ ), AST ( $P < 0.05$ ), CREA ( $P < 0.01$ ), BMI ( $P < 0.01$ ), ALT ( $P < 0.01$ ) were associated with osteoporosis. In train set, the order of AUC from highest to lowest is Random Forest (1.000), XGB (0.887), SVM (0.868), regression (0.801), ANN (0.793) and OSTA (0.739). In test set, the order of AUC from highest to lowest is XGB (0.848), regression (0.801), Random Forest (0.772), SVM (0.755), OSTA (0.739), ANN (0.732). SVM and XGB algorithm models had better screening effect on osteoporosis than OSTA in middle-aged and elderly Tibetan residents in Tibet. Compared with Random Forest, ANN and SVM, the established XGB model had the best prediction ability and can be used to predict the risk of osteoporosis on biochemical indexes. The model needs to be further improved through large sample research.

**Keywords** Osteoporosis, Extreme gradient boosting, Prediction model, Tibet

Osteoporosis is a systemic bone disease characterized by low bone mass, microstructural damage to bone tissue, increased bone fragility, and susceptibility to fractures<sup>1</sup>. In 2001, the National Institutes of Health (NIH) in the United States defined it as a skeletal disorder characterized by decreased bone strength and increased fracture risk<sup>2</sup>. The national epidemiological survey of osteoporosis showed that the prevalence of osteoporosis in people over 50 years old is 19.2%, with 32.1% in females and 6.9% in males; for those over 65 years old, the prevalence is 32.0%, with 51.6% in females and 10.7% in males. According to these epidemiological data, the estimated number of osteoporosis patients in China is about 90 million, with around 70 million being females<sup>3</sup>. In China, with the increase in life expectancy and the continuous growth of the population aged 70 and above, the burden of osteoporosis and fractures is expected to significantly increase. The number of fractures related to osteoporosis is projected to double by 2035<sup>4</sup>.

Currently, dual-energy X-ray absorptiometry (DXA) is the most commonly used method for diagnosing osteoporosis and monitoring changes in bone density over time. However, due to the relatively high cost of DXA, it is not recommended as a routine screening test for the general population. The Osteoporosis Self-Assessment Tool for Asians (OSTA) has been developed and validated as a simple screening tool for osteoporosis<sup>5</sup>. This

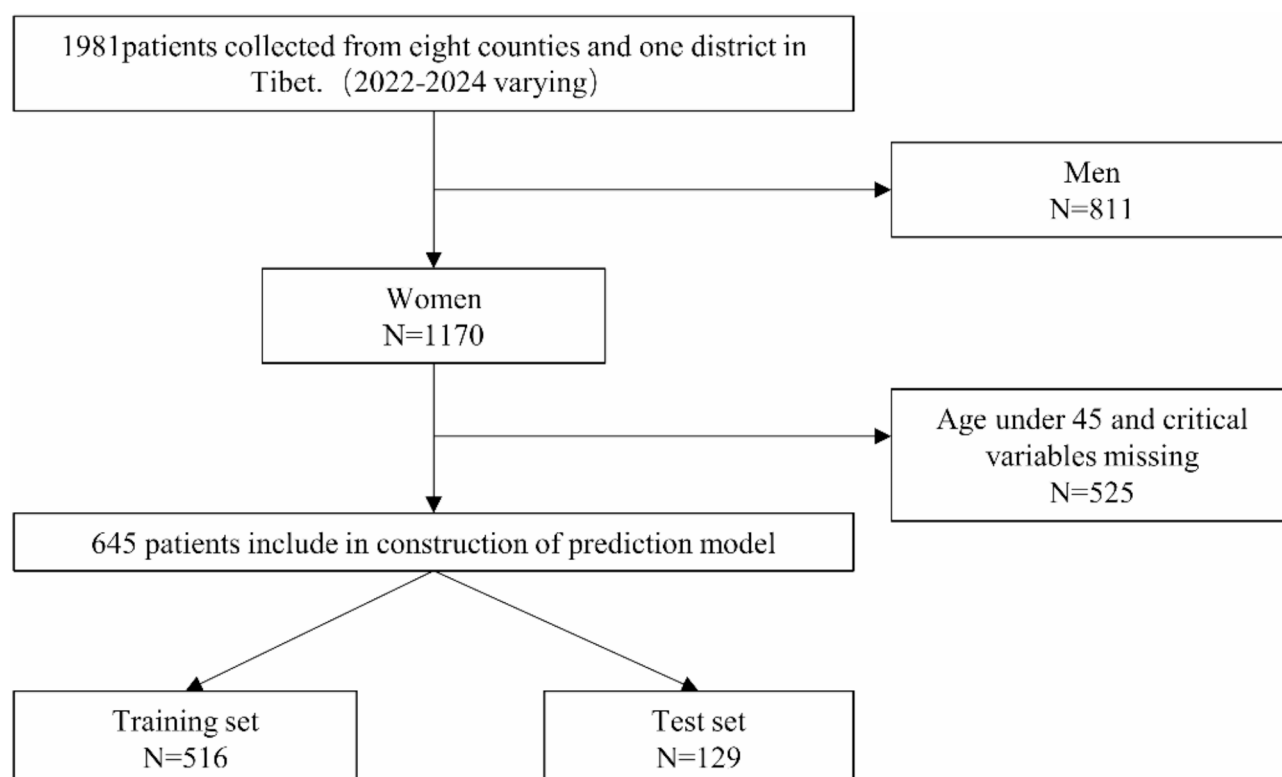
<sup>1</sup>School of Medicine, Tibet University, Lhasa 850000, China. <sup>2</sup>School of Ecology and Environment, Tibet University, Lhasa 850000, China. <sup>3</sup>Hospital Infection Management Department, The First People's Hospital of Shuangliu District, West China Airport Hospital of Sichuan University, Chengdu 610200, China. <sup>4</sup>Peng Wang and Qiang Yin contributed equally to this work. ✉email: 58749403@qq.com; xhxysq@126.com

tool has been widely used to screen osteoporosis in different races and genders<sup>6</sup>. Artificial intelligence (AI) technology is developed based on mathematical modeling. AI software has been applied to different specialties, including epidemiological investigations, drug discovery and diagnostic radiology<sup>7</sup>. XGB is a large-scale machine learning algorithm first formally released in 2016<sup>8</sup>. It is an improvement over Gradient Boosting Decision Tree (GBDT). While a single decision tree is a simple and weak classifier, an ensemble model of decision trees may have better performance, such as Random Forest<sup>9</sup> and GBDT<sup>10</sup>. Compared to Random Forest, XGB utilizes gradient boosting techniques to minimize the loss function<sup>8</sup>, and compared to GBDT, XGB employs a technique called “feature subsampling” to prevent overfitting<sup>8</sup>. XGB algorithm has been widely used in industry but its application in medical research is still limited. Compared to traditional machine learning algorithms, XGB is more up-to-date and complex. This study aims to establish a predictive screening tool for osteoporosis based on XGB algorithm and multicenter large-scale data. SVM<sup>11</sup> and ANN<sup>12</sup> are two machine learning algorithms with different principles from XGB, also commonly used to build classifiers. During the training process of ANN models, there is no need for manually determining weights, which can reduce the subjective factors in the diagnostic process, thereby improving the reliability of diagnosis. This helps make the diagnostic results more effective and objective, contributing to the effective reduction of misdiagnosis and missed diagnosis rates in osteoporosis diagnosis<sup>13</sup>. We also aimed to establish screening models based on these five different types of algorithms for comparison with the OSTA index to evaluate their performance.

In Tibet, due to its vast area and sparse population, BMD measurements are not widely used in some grassroots communities. Patients with low BMD are at higher risk of future fractures and should be actively considered for treatment to reduce this risk. By actively seeking a screening tool for older Tibetan women compared to the traditional OSTA simple screening tool, it provides a new solution for precision medicine strategies.

### Methods data source and participants

The data used in this study were obtained from the Tibetan Chronic and Endemic Disease Prevention on the Qinghai-Tibet Plateau program. The program was carried out in some parts of Tibet from July to October 2022. In this study, prior to the study, residents were invited to partake in health checks conducted at designated medical facilities within the communities. All participants had provided informed consent prior to enrollment in the study. Missing values of continuous features for each patient were replaced with their mean values across all data, and missing values of categorical features for each patient were replaced with their most frequent values across all data. When there was no most frequent value, that patient's data was excluded. After this preprocessing, the total number of women patients considered in the study was 645 (Fig. 1). Inclusion criteria: (1) all Participants were Tibetan; (2) Age range from 45 to 79 years; (3) All Participants signed informed consent. Exclusion criteria: (1) Participants left ulnar radius surgery, disease, fracture; (2) Participants taking drugs that affect bone metabolism, such as glucocorticoid for more than 3 months; (3) Participants suffering from serious



**Fig. 1.** Flow diagram of study population.

heart, liver, kidney and mental diseases. After this preprocessing, the total number of women patients considered in the study was 645 (Fig. 1).

### Definitions and BMD measurement

Normal BMD was defined as a T-score greater than  $-1.0$ . Osteopenia was defined as a T-score less than or equal to  $-1.0$  but greater than  $-2.5$ , while osteoporosis was defined as a T-score less than or equal to  $-2.5$ , following the WHO criteria (1994)<sup>14</sup>. The OSTA index equation:  $0.2 \times [\text{body weight (kg)} - \text{age (years)}]$ <sup>15</sup>. An AUC of 1 indicates a perfectly accurate test for distinguishing individuals, whereas an AUC of 0.5 indicates a useless test. In general, AUC of 0.80–0.90 are considered good, 0.70–0.80 are considered fair, and 0.60–0.70 are considered poor<sup>16</sup>.

All measurements were conducted by trained investigators using equipment that was calibrated in accordance with the manufacturer's instructions. BMD of the left distal forearm was assessed using a dual-energy X-ray absorptiometry densitometer (DXA, AKDX-09 W-I, Shenzhen Xray Electric Co., Shenzhen, China). This assessment was based on a comprehensive database comprising 12,159 radial bone mineral density readings from across China, adhering to the protocol which specifies a precision error of  $\leq 1.0\%$  CV in vivo. The DXA device was transported by vehicle to various study sites during the survey. Quality control procedures were implemented as per the manufacturer's operating manual. Participants' fasting venous blood was collected using a 5-ml vacuum coagulant tube and centrifuged after two hours at room temperature to separate the serum for analysis. Biochemical markers, including Homocysteine (Hcy), total protein (TP), uric acid (UA), creatinine (CREA), total cholesterol (TC), triglycerides (TG), low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C), were quantified by the Tibet Autonomous Region Institute of Tibetan Medicine using a fully automated biochemical analyzer (Hitachi 7600, Hitachi, Tokyo, Japan).

### Feature selection

In the data. There were 14 biochemical indicators. Correlation analysis was used to screen indicators with statistical differences from T-score. Finally, Age ( $P < 0.01$ ), LDL-C ( $P < 0.05$ ), UA ( $P < 0.01$ ), AST ( $P < 0.05$ ), CREA ( $P < 0.01$ ), BMI ( $P < 0.01$ ), ALT ( $P < 0.01$ ) were selected for the development of the prediction models (Fig. 2).

### Statistical analysis

Comparison of the test set and train set data was analyzed using Mann-Whitney U test, using SPSS 25.0 software (IBM SPSS Statistics, <https://www.ibm.com/products/spss-statistics>). Variables that do not conform to the normal distribution were represented by the M (Q1, Q3). Spearman test was used to conduct correlation analysis between biochemical indicators and T-score. Build, train, and validate predictive models using the "XGB" package. Plotting was performed using R Version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria). P value  $< 0.05$  was considered statistically significant.

### Ethical consideration

This study followed the ethical guidelines of the Declaration of Helsinki and was approved by the Ethics Committee of Tibet University. All participants provided informed consent to participate in the study. For minors under 16 years of age, written informed consent was obtained from their parents or legal guardians. Participants signed the informed consent form after receiving information about the content of the project (participants with educational experience signed the form, while those with no educational experience were informed of the content of the project and provided consent by fingerprinting). All participants volunteered to participate in this study.

## Results

### Demographic and clinical features

All eligible participants ( $N=645$ ) were randomly divided into two groups, with a ratio of 8:2, resulting in a train set ( $N_1=516$ ) and a test set ( $N_2=129$ ), there was no statistical difference between the test set and train set considering the 7 characteristics (Table 1). Age (57 (51, 63) and 57 (51, 66) in the train and test set respectively), LDL-C (2.71 (2.15, 3.38) and 2.8 (2.23, 3.27) in the train and test set respectively), UA (296.5 (235, 362.5) and 295 (239, 373) in the train and test set respectively), AST (19 (15, 26) and 20 (15, 25) in the train and test set respectively), BMI (23.6 (21.4, 26.7) and 23.3 (21.1, 26.6) in the train and test set respectively) (Table 1).

### Importance ranking

The importance of the factors included in XGB was as follows: BMI, CREA, LDL-C, ALT, Age, UA, AST (Fig. 3). Among them, the importance ratio of BMI was 0.271, which was the highest among the seven indicators; the importance ratio of age was 0.110; the importance ratio of AST was 0.084, which was the lowest among the seven indicators (Supplement 1).

### SHAP feature dependence plots

Additionally, we employed SHAP dependence plots to further investigate the impact of the top seven factors on the XGB model's prediction of OP risk. The BMI around 20 kg/m<sup>2</sup> had high SHAP values, which had a significant positive impact on the model's prediction results, indicating that the model tended to predict the occurrence of the target event in this range. The color represented the relationship between OP and the feature variables, showing that there was a significant negative correlation between BMI and osteoporosis. It could also be observed that the type of bone mass had some influence on the distribution of SHAP values, especially in regions where SHAP values were high, red dots were more concentrated (Fig. 4a). Age over 70 had high SHAP

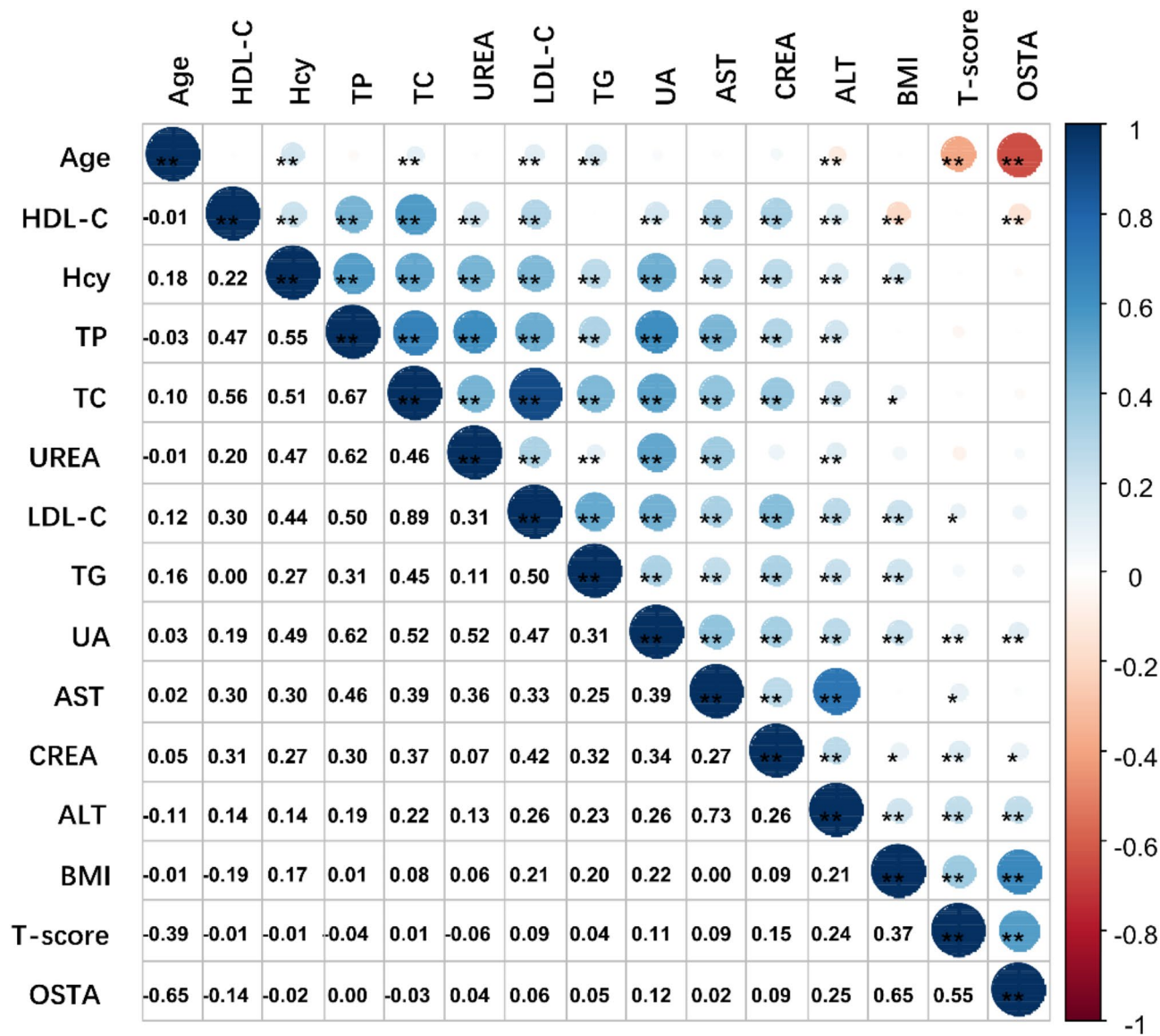
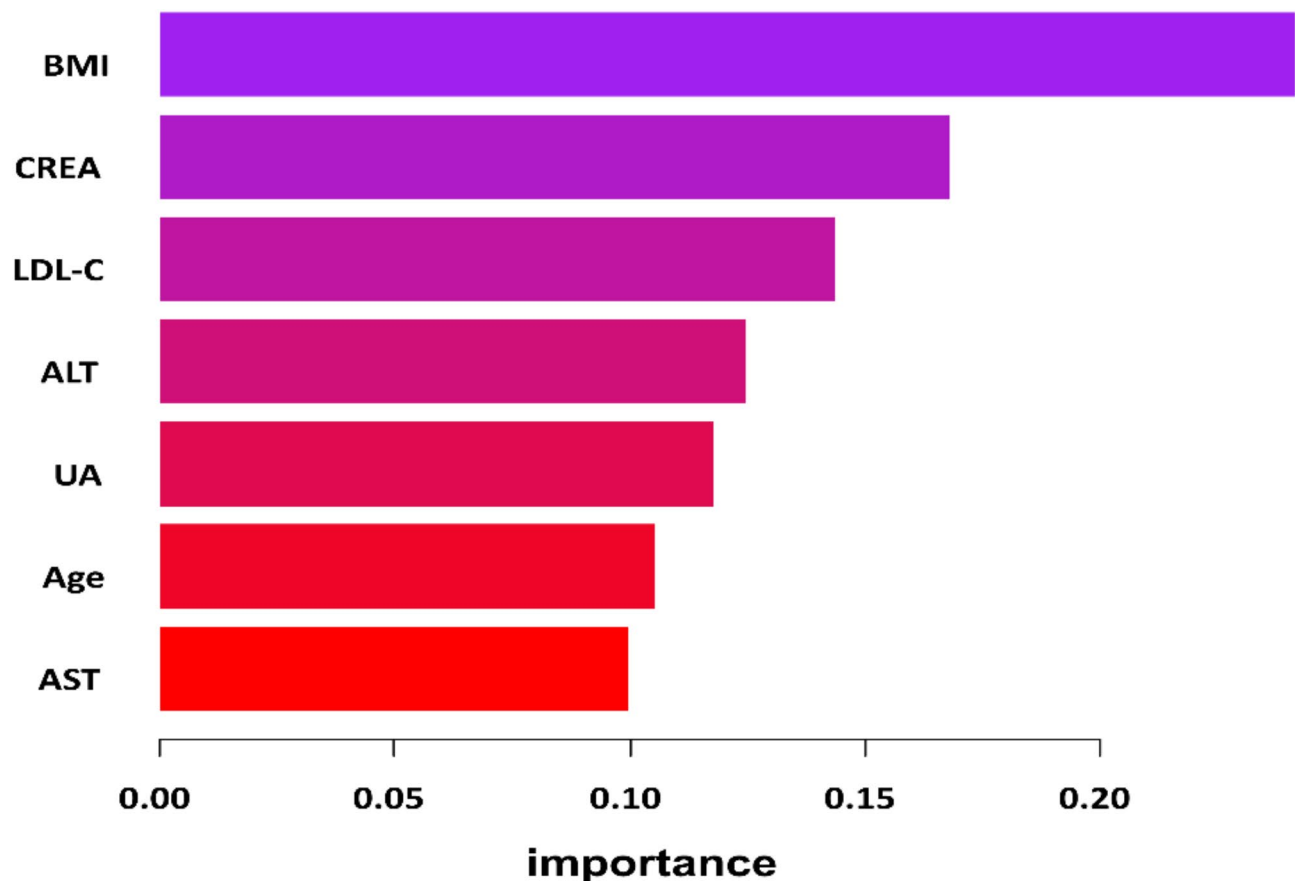


Fig. 2. Correlation between biochemical indicators and T-score, \*P < 0.05; \*\*P < 0.01.

Characteristics	Train set (N1 = 516)	Test set (N2 = 129)	P-value
Age (years)	57 (51, 63)	57 (51, 66)	0.282
LDL-C (mmol/L)	2.71 (2.15, 3.38)	2.8 (2.23, 3.27)	0.672
UA (umol/L)	296.5 (235, 362.5)	295 (239, 373)	0.567
AST (U/L)	19 (15, 26)	20 (15, 25)	0.850
CREA (umol/L)	54.3 (46.4, 63.5)	52.9 ± 13.5	0.158
BMI (kg/m²)	23.6 (21.4, 26.7)	23.3 (21.1, 26.6)	0.583
ALT (U/L)	19 (13, 31)	18 (13.5, 26)	0.184

Table 1. Characteristics of 645 subjects included in the study.



**Fig. 3.** Rank the importance of the included predictors in XGB, all features were listed in order, the importance ratio of all features sums to 1 (or 100%), reflecting their total contribution to the model. Features: feature name; Importance: proportion of importance.

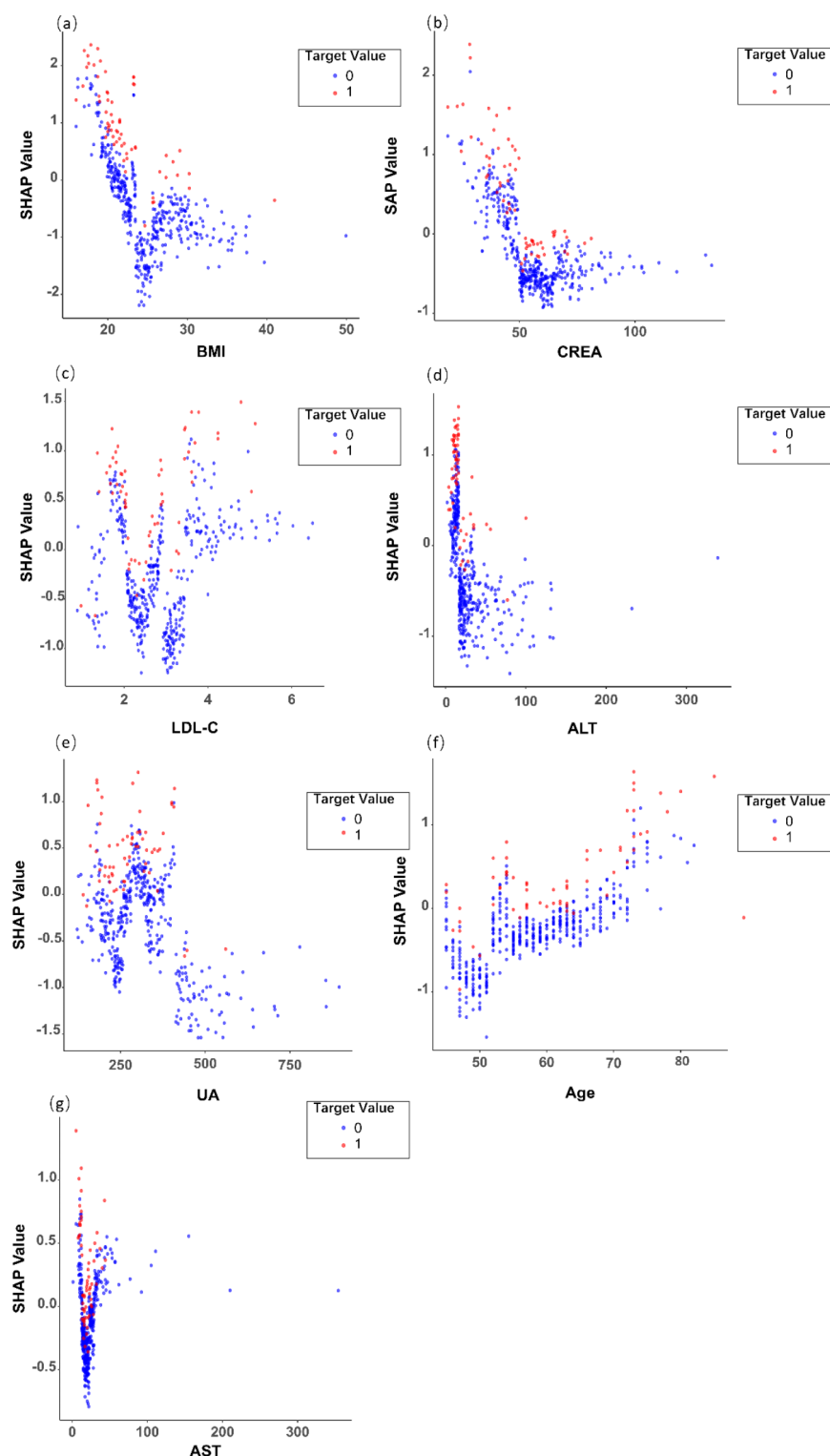
values, which had a significant positive impact on the model's prediction results, indicating that the model tended to predict the occurrence of the target event in this age group. There was a significant positive correlation between age and osteoporosis (Fig. 4f).

### Model evaluation

The order of AUC for the seven indexes included was OSTA (0.739), BMI (0.708), CREA (0.688), ALT (0.682), UA (0.649), Age (0.625), and LDL-C (0.617), AST (0.605) (Fig. 5). In the train set, the order of AUC from highest to lowest was Random Forest (1.000), XGB (0.887), SVM (0.868), regression (0.801), ANN (0.793), and OSTA (0.739) (Fig. 6). In the test set, the order of AUC from highest to lowest was XGB (0.848), regression (0.801), Random Forest (0.772), SVM (0.755), OSTA (0.739), and ANN (0.732) (Fig. 6). The sensitivity (0.889), specificity (0.766), and AUC (0.848 [0.7775, 0.9192]) of XGB were higher than those of OSTA, which had a sensitivity of 0.624, specificity of 0.755, and AUC of 0.739 [0.683, 0.790]. The sensitivity (0.722), specificity (0.811), and AUC (0.755 [0.6376, 0.8729]) of SVM were also higher than those of OSTA (Fig. 6).

### Discussion

In correlation analysis, we could see that Age, LDL-C, UA, AST, CREA, BMI, ALT were correlated with T-score. The relationship between lipid metabolism and bone density had also been confirmed to be significant. Animal experiments had shown that Kindlin-2 expression in adipose tissue increases during aging and in response to a high-fat diet, accompanied by a decrease in bone mass. Kindlin-2 plays a crucial role in maintaining bone homeostasis by modulating the FAS/PPAR $\gamma$ /FABP4/insulin axis in adipocytes<sup>17</sup>. In studies<sup>18</sup> correlating individual lipid parameters with BMD, increases in total cholesterol (TC), triglycerides (TG), and low-density lipoprotein cholesterol (LDL-C) had all been associated with decreased bone density. Elevated AST and ALT may signal liver damage, and studies<sup>19–21</sup> had found that chronic liver disease and impaired liver function increase the risk of osteoporosis. The level of AST/ALT ratio may indirectly reflect osteoporosis<sup>22</sup>, forearm DXA results best reflect the causal effect of LDL-C reduction on BMD<sup>23</sup>. Studies had found that blood uric acid has a protective effect on osteoporosis<sup>24,25</sup>. There is a link between creatinine and osteoporosis at high altitude in Tibet, and this link may be closely related to the local diet<sup>26</sup>. In this study, biochemical indexes associated with osteoporosis were used as characteristic factors to predict the occurrence of osteoporosis.



**Fig. 4.** X-axis represented input variable, Y-axis represented the SHAP value of input variable, namely, the contribution value to the model output. A positive SHAP value indicates that the feature increases the probability of the model's prediction, whereas a negative SHAP value suggests that the feature decreases the probability of the prediction. SHAP dependency plot illustrating the top 7 clinical features in the XGB model. (a) BMI; (b) CREA; (c) LDL-C; (d) ALT; (e) UA; (f) Age; (g) AST.

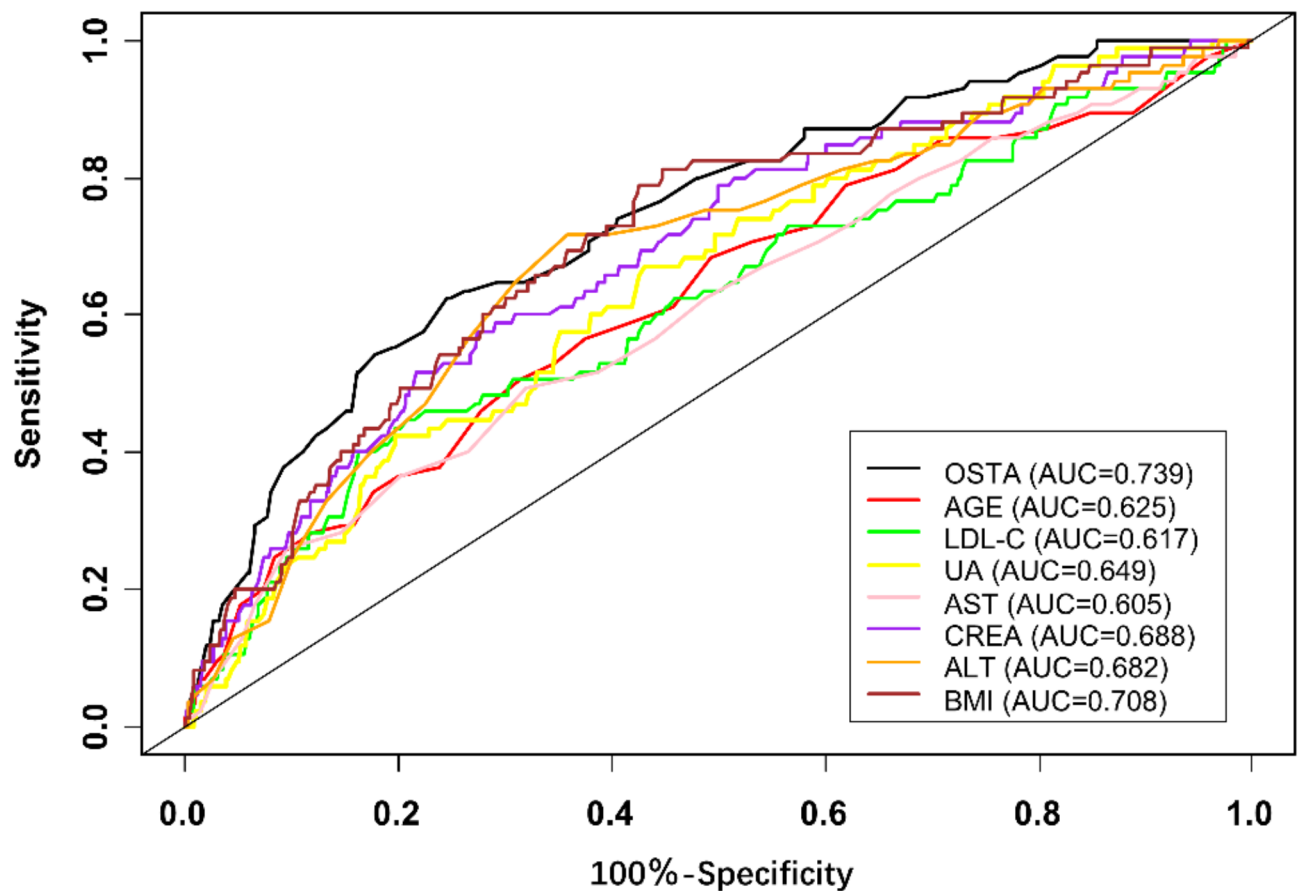


Fig. 5. ROC curve comparison between biochemical indexes and OSTA index.

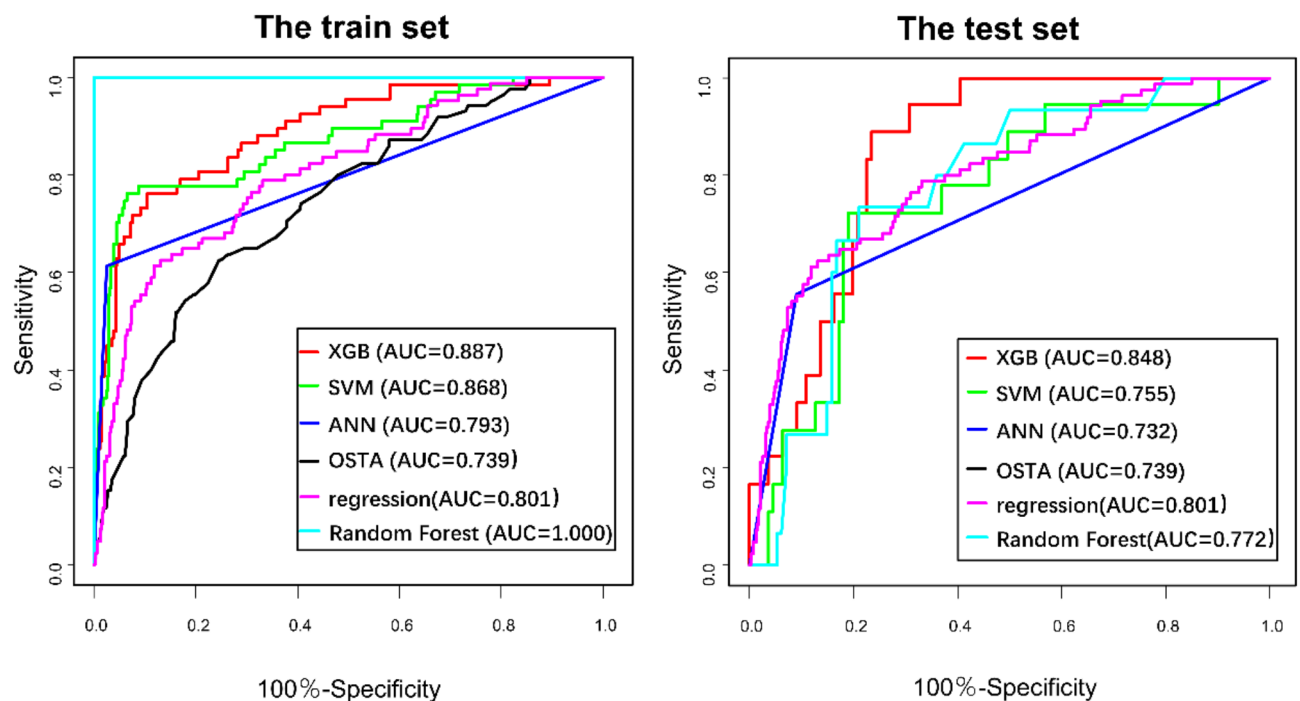


Fig. 6. ROC curves comparison between five models and OSTA index from train and test set.

Model	Data set	Sensitivity	Specificity	AUC (95% CI)
XGB	Train set	0.761	0.895	0.887 (0.8424, 0.9319)
	Test set	0.889	0.766	0.848 (0.7775, 0.9192)
SVM	Train set	0.761	0.935	0.868 (0.8142, 0.9924)
	Test set	0.722	0.811	0.755 (0.6376, 0.8729)
ANN	Train set	0.612	0.976	0.793 (0.6854, 0.8065)
	Test set	0.556	0.910	0.732 (0.5496, 0.7852)
Random Forest	Train set	1.000	1.000	1.000 (1.0000, 1.0000)
	Test set	0.733	0.789	0.772 (0.6580, 0.8852)
regression	ALL	0.612	0.882	0.801 (0.7489, 0.8436)
OSTA	ALL	0.624	0.755	0.739 (0.6830, 0.7960)

**Table 2.** The performance of four prediction models on train set and test set.

This is a novel study to establish a predictive model of osteoporosis based on biochemical indices. In this study, osteoporosis prediction models based on Random Forest, regression, XGB, SVM and ANN were established. In the train and test set, the prediction effect of Random Forest, regression, XGB and SVM algorithm was better than that of OSTA, and XGB algorithm was the best among the five machine learning algorithms (Table 2). In addition to the AUC, the sensitivity and specificity of the proposed XGB model show that it is excellent discrimination. For the train set and the test set, the AUC of XGB model is 0.887 and 0.848, respectively. In the single index forecasting model, the AUC range of the 8 indicators was 0.606 to 0.739, OSTA index is better than other biochemical indicators (Table 2). XGB algorithm has obvious advantages in processing unbalanced data<sup>27</sup>, the results of this paper show that when the data set is in non-equilibrium state, the XGB model can obtain better classification and discrimination effect (AUC = 0.848), and the model performance is better than the SVM and ANN model. SVM is widely used in image processing, text classification and other fields. However, when the input variables are large and the sample set is large, the computational complexity and spatial complexity of SVM will increase sharply, resulting in long training time and large memory resource consumption. SVM stands out for its high accuracy and efficiency in medical applications, which often involve small data sets<sup>28</sup>, SVM, and ANN can use diet and lifestyle data to help predict low BMD<sup>29</sup> this is also may be the reason why the prediction effect of ANN and SVM in this study is poor discrimination.

This is a study based on the population of elderly Tibetan women in the Tibetan region. Due to the influence of natural environment and religious customs, it is challenging to obtain samples in high-altitude areas. This study establishes a screening model for osteoporosis using five algorithms, which is relatively innovative. However, there are limitations to this study. First of all, the cost of blood biochemical indicators is not low, and the information acquisition and screening execution are not as simple and economical as traditional screening tools. Secondly, the performance of the algorithm lacks external verification, and whether it can be used in a wide range is still questionable. Furthermore, the establishment of the algorithm model is based on the general population, and in the population with certain risk factors (suffering from diseases affecting bone metabolism, long-term use of drugs affecting bone metabolism, smoking, alcohol consumption, caffeine intake, etc.), whether XGB is applicable and its screening effect needs to be further verified.

Conclusion

SVM and XGB algorithm models had better screening effect on osteoporosis than OSTA in middle-aged and elderly Tibetan residents in Tibet, Having the XGB algorithm to predict osteoporosis occurrence in adult Tibetan women in Tibet has shown promising predictive effectiveness.

Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Received: 5 September 2024; Accepted: 24 March 2025  
Published online: 31 March 2025

References

1. Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis. *Am. J. Med.* **94** (6), 646–650. [https://doi.org/10.1016/0002-9343\(93\)90218-e](https://doi.org/10.1016/0002-9343(93)90218-e) (1993).  
2. NIH Consensus Development Panel on osteoporosis prevention, diagnosis, and therapy, March 7–29, 2000: highlights of the conference. *S. Med. J.* **94** (6), 569–573 (2001).  
3. Primary osteoporosis diagnosis and treatment guidelines (2022). *Chin. J. Osteopor. Mineral Dis.* **15** (06), 573–611 (2022).  
4. Consensus on the prevention and treatment of osteoporosis in perimenopausal and postmenopausal women. *Chin. J. Clin. Phys.* **48** (08), 903–908 (2020).  
5. Koh, L. K. et al. A simple tool to identify Asian women at increased risk of osteoporosis. *Osteoporos. Int.* **12** (8), 699–705. <https://doi.org/10.1007/s001980170070> (2001).  
6. Chen, C. C. et al. Association of osteoporosis self-assessment tool for Asians (OSTA) score with clinical presentation and expenditure in hospitalized trauma patients with femoral fractures. *Int. J. Environ. Res. Public Health* **13** (10), 995. <https://doi.org/10.3390/ijerph13100995> (2016).

7. Bui, H. M. et al. Predicting the risk of osteoporosis in older Vietnamese women using machine learning approaches. *Sci. Rep.* **12** (1), 20160. <https://doi.org/10.1038/s41598-022-24181-x> (2022).
8. Chen, T. & Guestrin, C. XGB: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).
9. Svetnik, V. et al. Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **43** (6), 1947–1958. <https://doi.org/10.1021/ci034160g> (2003).
10. Natekin, A. & Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurobot.* **7**, 21. <https://doi.org/10.3389/fnbot.2013.00021> (2013).
11. Tang, H., Han, Q. & Yin, Y. Screening of important markers in peripheral blood mononuclear cells to predict female osteoporosis risk using LASSO regression algorithm and SVM method. *Evol. Bioinform.* **18**, 11769343221075014. <https://doi.org/10.1177/11769343221075014> (2022).
12. Meng, J. et al. Artificial neural network optimizes self-examination of osteoporosis risk in women. *J. Int. Med. Res.* **47** (7), 3088–3098. <https://doi.org/10.1177/0300060519850648> (2019).
13. Mengming, Y. et al. Review of the application of artificial intelligence in osteoporosis. *J. Mini-Micro Syst.* **40** (09), 1839–1850 (2019).
14. Xu, H. et al. Osteoporosis and osteopenia among patients with type 2 diabetes aged  $\geq 50$ : role of sex and clinical characteristics. *J. Clin. Densitom.* **23** (1), 29–36. <https://doi.org/10.1016/j.jocd.2019.04.004> (2020).
15. Wu, S. et al. The associations between different sleep patterns and osteoporosis based on the osteoporosis self-assessment tool for Asians. *Arch. Osteoporos.* **15** (1), 164. <https://doi.org/10.1007/s11657-020-00828-y> (2020).
16. Nayak, S. et al. Systematic review and meta-analysis of the performance of clinical risk assessment instruments for screening for osteoporosis or low bone density. *Osteoporos. Int.* **26** (5), 1543–1554. <https://doi.org/10.1007/s00198-015-3025-1> (2015).
17. Tang, W. et al. Targeting Kindlin-2 in adipocytes increases bone mass through inhibiting FAS/PPAR $\gamma$ /FABP4 signaling in mice. *Acta Pharm. Sin. B* **13** (11), 4535–4552. <https://doi.org/10.1016/j.apsb.2023.07.001> (2023).
18. Yang, X. L. et al. Causal link between lipid profile and bone mineral density: A Mendelian randomization study. *Bone* **127**, 37–43. <https://doi.org/10.1016/j.bone.2019.1016/j.bone> (2019).
19. Nakchbandi, I. A. & van der Merwe, S. W. Current Understanding of osteoporosis associated with liver disease. *Nat. Rev. Gastroenterol. Hepatol.* **6** (11), 660–670. <https://doi.org/10.1038/nrgastro.2009.166> (2009).
20. Wibaux, C. et al. Assessing bone status in patients awaiting liver transplantation. *Jt. Bone Spine* **78** (4), 387–391. <https://doi.org/10.1016/j.jbspin.2011.1016/j.jbspin> (2011).
21. Tsai, C. F. et al. Increased incidence of orthopedic fractures in cirrhotic patients: a nationwide population-based study. *J. Hepatol.* **58** (4), 706–714. <https://doi.org/10.1016/j.jhep.2013.07.001> (2013).
22. Yan, H. *Risk Factors Analysis of Osteoporosis in Middle-Aged and Elderly People* (Huazhong University of Science and Technology, 2019).
23. Zheng, J. et al. The effect of plasma lipids and lipid-lowering interventions on bone mineral density: A Mendelian randomization study. *J. Bone Min. Res.* **35** (7), 1224–1235. <https://doi.org/10.1002/jbmr.3989> (2020).
24. Kim, M. H., Lee, H. J. & Osteoporosis Vitamin C intake, and physical activity in Korean adults aged 50 years and over. *J. Phys. Ther. Sci.* **28** (3), 725–730. <https://doi.org/10.1589/jpts.28.725> (2016).
25. Lee, Y. J. et al. The association between oxidative stress and bone mineral density according to menopausal status of Korean women. *Obstet. Gynecol. Sci.* **58** (1), 46–52. <https://doi.org/10.5468/ogs.2015.58.1.46> (2015).
26. Zhong, H. et al. Influencing factors of bone mass abnormalities among postmenopausal women in Tibet, China. *BMC Public Health* **23** (1), 2100. <https://doi.org/10.1186/s12889-023-17015-6> (2023).
27. Fan Ru, X. et al. Establishment of hypertension recognition model based on extreme gradient lifting algorithm. *China Health Stat.* **40** (01), 74–77 (2019).
28. Abdelfattah, S. et al. Lightweight multi-class support vector machine-based medical diagnosis system with privacy preservation. *Sensors (Basel)* **23** (22), 9033. <https://doi.org/10.3390/s23229033> (2023).
29. Yoo, T. K. et al. Osteoporosis risk prediction for bone mineral density assessment of postmenopausal women using machine learning. *Yonsei Med. J.* **54** (6), 1321–1330. <https://doi.org/10.3349/ymj.2013.54.6.1321> (2013).

## Acknowledgements

The authors would like to sincerely thank all the study participants and participating worksites.

## Author contributions

Authors' contributions PW and QY designed the study and wrote the manuscript. HX and ZSX provided analytical ideas and revised the manuscript for important content. All authors contributed to data analysis and the drafting and revisions of the manuscript. All authors read and approved the final version of the manuscript.

## Funding

This research was supported by the Central Financial Support for the Reform and Development of Local Universities (00060585, 00060463, 00060381, 00060695/051) and the “Belt and Road” Science and Technology Innovation of Ali Regional Science and Technology Bureau (18080036). The funding bodies have played no roles in study design, data collection, analysis, and interpretation, and manuscript writing.

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethical approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by Ethics Committee of the Tibet University. Ethics approval number ZDYXLL2024009. All participants provided informed consent to participate in the study. For minors under 16 years of age, written informed consent was obtained from their parents or legal guardians. Participants signed the informed consent form after understanding the content of the project (Participants who were educated took a signature, while those with no educational experience

were informed of the content of the project and attested by fingerprinting). All participants volunteered to participate in this study.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-95707-2>.

**Correspondence** and requests for materials should be addressed to Z.X. or H.X.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025