Article

Discovery Medicine 2025; 37(192): 55–63
https://doi.org/10.24976/Discov.Med.202537192.5

Clinical Application of a Big Data Machine Learning Analysis Model for Osteoporotic Fracture Risk Assessment Built on Multicenter Clinical Data in Qingdao City

Bing Li[1], Yanru Yang[2], Feng Shen[3], Yuelei Wang[3], Ting Wang[3], Xiaxia Chen[3], Chun Lu[4,*]

[1]Qingdao Hiser Hospital Affiliated of Qingdao University (Qingdao Traditional Chinese Medicine Hospital), 266033 Qingdao, Shandong, China
[2]Science and Education Training Department, Zibo Medical Emergency Command Center, 255000 Zibo, Shandong, China
[3]Department of Spinal Surgery, The Affiliated Hospital of Qingdao University, 266000 Qingdao, Shandong, China
[4]Science and Education Department, Zibo Orthopedic Hospital, 255040 Zibo, Shandong, China
*Correspondence: lulu9411@126.com (Chun Lu)
Published: 20 January 2025

Background: Osteoporotic fractures (OPF) pose a public health issue, imposing significant burdens on families and societies worldwide. Currently, there is a lack of comprehensive and validated risk assessment models for OPF. This study aims to develop a model to assess and predict the risk of OPF in Qingdao City, China.

Methods: From January 2021 to January 2023, we recruited 84 osteoporotic patients diagnosed with fractures from Qingdao University Affiliated Hospital, Qingdao Municipal Hospital, Qingdao Hiser Hospital Affiliated of Qingdao University, and Qingdao Central Hospital as the experimental group. In addition, 112 osteoporotic patients without fractures were recruited as the control group. In this study, we employed seven machine learning models, namely Adaboost, random forest (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB), and Gradient Boosting Decision Trees (GBDT), to analyze the risk factors influencing the occurrence of OPF. Next, we plotted receiver operating characteristic (ROC), Precision-Recall (PR), and calibration curves to evaluate the predictive values of the different risk assessment models for OPF.

Results: Among the seven models built based on the training set data, the Adaboost model showed area under the curve (AUC), sensitivity, and specificity values close to 1, indicating the best classification performance. In the test set, the AUC values for the RF, SVM, LR, KNN, NB, AdaBoost, and GBDT models were 0.936, 0.905, 0.88, 0.93, 0.862, 0.939, and 0.859, respectively ($p < 0.001$). All sensitivity and specificity values for these models were higher than 0.8, with sensitivity and specificity values of the Adaboost model closest to 1. Additionally, six models had an area under the Precision-Recall curve (prAUC) values higher than 0.9, except KNN at 0.284 ($p < 0.001$). The calibration curves of the seven models did not significantly deviate from the ideal curve, indicating acceptable discriminative ability and predictive performance of the predictive model. All results showed that trabecular bone score (TBS) was the most important variable affecting the model, followed by osteocalcin (OST) and hunchback.

Conclusions: Given the various clinical data from patients with OPF, we assessed and demonstrated the good predictive performance of our risk predictive models. This model will enable us to take timely intervention measures to reduce the incidence of OPF and improve patient prognosis.

Keywords: osteoporotic fractures; multicenter; machine learning; risk assessment; prediction model

Introduction

Osteoporotic fractures (OPF), also known as fragility fractures, are the most common complications of osteoporosis [1]. They often occur in areas rich in trabecular bone, such as the spine, hip, wrist, and proximal end of the humerus [2]. The clinical manifestations include pain, tenderness, swelling, and functional impairment. The incidence is higher in individuals aged more than 45 years [3], especially in females [4,5]. Approximately one-third of women over 50 years old have experienced osteoporotic fractures [6], severely impacting patients' quality of life and imposing enormous economic burdens on patients, families, and societies [7]. According to surveys, there were 9 million cases of OPF worldwide in 2000 [8]. Because of increasing trends in aging in China, the occurrence of OPF is on the rise. According to data from 2010, the number of patients with OPF in China reached 2.33 million, with hip fractures accounting for 360,000 cases and vertebral fractures accounting for 1.11 million cases. It is estimated that by 2035, the number of OPF will double, reaching 5.99 million cases, and by 2050, it is projected to increase to 5.99

Table 1. Clinical characteristics of the patients.

Characteristic	N, N = 112	Y, N = 84	Test statistic[1]	*p*-value[1]
ESR, median (25%, 75%), (mm/hour)	16.43 (3.57, 34.64)	12 (7.64, 14.35)	–2.528	0.001
OST, median (25%, 75%), (ng/mL)	34.6 (15.09, 58.1)	16.08 (9.39, 18.97)	–4.383	<0.001
Lumbar, median (25%, 75%), (g/cm^3)	–1.34 (–2.05, –0.32)	–1.9 (–2.91, –1.19)	–3.510	<0.001
Fracture_history, n (%)			64	<0.001
N	108 (96%)	39 (46%)	NA	
Y	4 (3.6%)	45 (54%)	NA	
Sex, n (%)			3.8	0.052
F	80 (71%)	70 (83%)	NA	
M	32 (29%)	14 (17%)	NA	
Age, mean (SD), (years)	63 (5)	68 (9)	–4.6	<0.001
BMI, mean (SD)	26.20 (3.19)	23.99 (2.52)	5.4	<0.001
Hunchback, n (%)			94	<0.001
N	89 (79%)	8 (9.5%)	NA	
Y	23 (21%)	76 (90%)	NA	
TBS, mean (SD)	1.53 (0.19)	1.03 (0.14)	22	<0.001

[1]Wilcoxon rank sum test; Welch Two Sample *t*-test; Pearson's Chi-squared test. Abbreviations: ESR, erythrocyte sedimentation rate; OST, osteocalcin; BMI, body mass index; TBS, trabecular bone score; NA, not applicable; SD, standard deviation.

million cases [9]. Therefore, the economic burden associated with OPF should not be overlooked. The total cost in 2010 amounted to 945 million US dollars, and is expected to double by 2035, reaching 25.43 billion US dollars [9]. Early prediction and timely treatment can help reduce the incidence of OPF [10], thereby achieving the goal of lowering the OPF rate.

Currently, the predictive models for OPF mainly include FRAX [11] and QFracture [12] from the United Kingdom, Garvan [13] from Australia, Korean Fracture Risk Score (KFRS) [14] from South Korea, and Fracture Risk Evaluation Model (FREM) [15] from Denmark. However, these models have certain limitations, and there is a lack of risk prediction models for individuals aged more than 60 years with OPF in China [16–18]. Therefore, this study aimed to collect diagnostic and treatment information for OPF patients from Qingdao University Affiliated Hospital, Qingdao Municipal Hospital, Qingdao Hiser Hospital Affiliated of Qingdao University, and Qingdao Central Hospital to develop a risk assessment model. The aim is to provide better guidance for the treatment of patients with OPF and improve preventive and therapeutic strategies for this condition.

Materials and Methods

Study Subjects

This study focused on patients diagnosed with OPF at the Qingdao University Affiliated Hospital, Qingdao Municipal Hospital, Qingdao Hiser Hospital Affiliated of Qingdao University, and Qingdao Central Hospital between January 2021 and January 2023. The diagnostic criteria for OPF followed the guidelines outlined in the Guidelines for Diagnosis and Treatment of OPF [19]. The diagnostic criteria for osteoporosis were determined according to the Chinese Expert Consensus on Diagnostic Criteria for Osteoporosis in Chinese Individuals (3rd Edition, 2014) [20].

The inclusion criteria were as follows: (i) diagnosis of OPF according to the Guidelines for Diagnosis and Treatment of OPF; (ii) complete clinical data; and (iii) age $\geq$55 years.

The exclusion criteria were as follows: (i) patients with major diseases, including diabetes, thyroid disorders, severe gastrointestinal disorders, severe respiratory disorders, chronic kidney diseases, blood disorders, or other diseases affecting bone metabolism; (ii) patients who had used medications such as calcitonin, active vitamin D, bisphosphonates, or hormonal drugs that affect bone metabolism within the past six months; (iii) patients with major organ dysfunction; and (iv) pathological fractures caused by bone tumors, tuberculosis, or infections.

According to the inclusion and exclusion criteria, we recruited 84 patients diagnosed with OPF from four hospitals, constituting the experimental group. We selected 112 osteoporosis patients without fractures who were admitted to the same hospitals during the same period as the control group.

Research Indexes

In clinical research, it is common to use previous studies, statistically significant indices, or indices considered clinically relevant by researchers as a foundation for further investigation. The indices employed in this study included patient admission and discharge records (demographic data such as sex, age, body mass index (BMI), and hunchback situation, erythrocyte sedimentation rate (ESR), osteocalcin

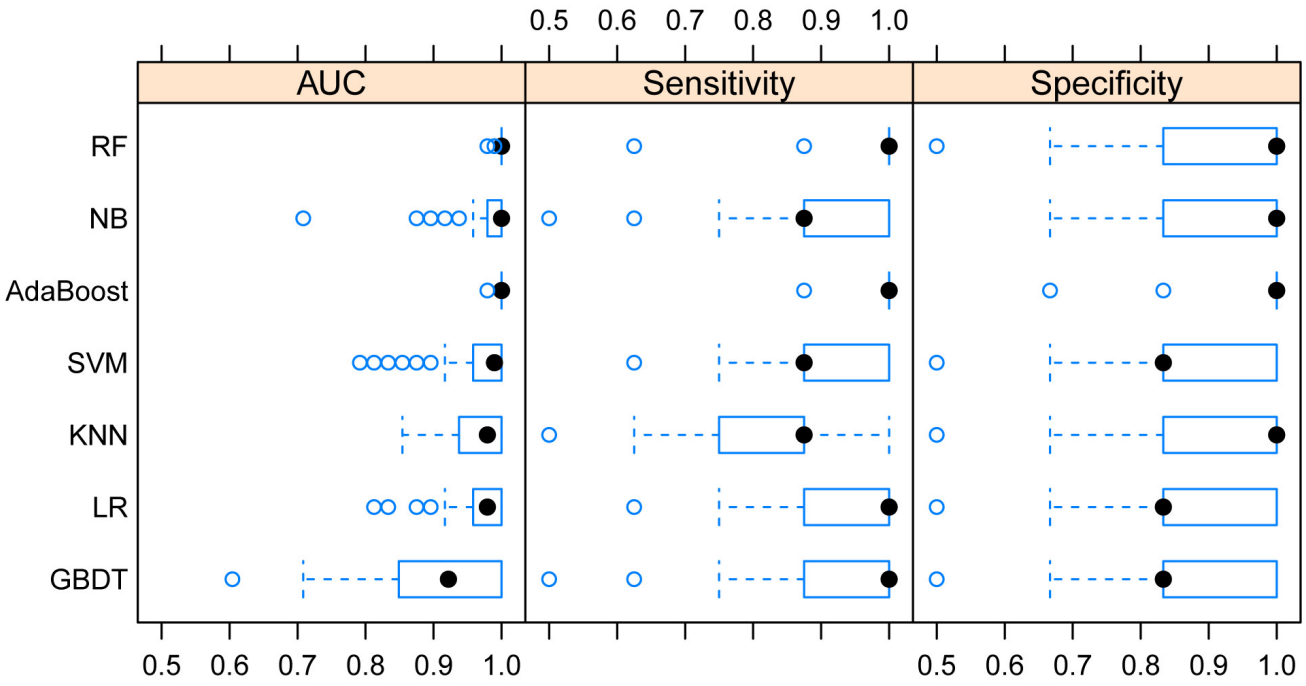

Fig. 1. Box plots of seven risk assessment models for osteoporotic fracture occurrence risk. Abbreviations: AUC, area under the curve; GBDT, Gradient Boosting Decision Trees; KNN, K-Nearest Neighbors; LR, Logistic Regression; NB, Naive Bayes; RF, random forest; SVM, Support Vector Machine.

Table 2. Training datasets for model assessment.

Model	AUC	Accuracy	Precision	Recall	F1	prAUC
Adaboost	0.9996	0.9836	0.9913	0.9812	0.9854	0.8685
GBDT	0.9028	0.8886	0.9070	0.9050	0.9006	0.5124
KNN	0.9684	0.8836	0.9512	0.8462	0.8901	0.2663
LR	0.9671	0.9100	0.9193	0.9300	0.9213	0.8428
NB	0.9806	0.9064	0.9521	0.8850	0.9129	0.8541
RF	0.9984	0.9600	0.9524	0.9850	0.9663	0.8439
SVM	0.9708	0.8893	0.9091	0.9050	0.9025	0.8403

Abbreviations: prAUC, Precision-Recall curve.

(OST), trabecular bone score (TBS), lumbar dual-energy X-ray absorptiometry (DXA) bone density testing, as well as treatment plans and follow-up data.

Risk Assessment Model

Collected data were processed through data cleaning and missing-value imputation using the multiple imputation by chained equations (MICE) method. The data were then normalized, and features with excessively small variances, low label correlations, or high inter-correlations were removed. Next, the data were randomly split into training and testing sets in a 7:3 ratio. In the training set, 10-fold cross-validation and grid search were performed to select the important parameters for seven classification algorithms (Adaptive Boosting (Adaboost), random forest (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Naive Bayes (NB), and Gradient Boosting Decision Trees (GBDT)). The completed training models of these seven algorithms were then used to train and build models on the data, generating predictions from different models. The average accuracy from multiple runs was calculated as the final model score, and the confusion matrix and final model were generated. For each model's training and testing results, receiver operating characteristic (ROC) curves were plotted and the area under the curve (AUC), sensitivity, specificity, the area under the Precision-Recall curve (prAUC), and the calibration curve were computed. The best model was determined by comparing the AUC values, and the contribution of the features to the model was analyzed.

Statistical Methods

The data were subjected to statistical analysis using R 4.2.1 (https://www.r-project.org/, R Foundation for Statistical Computing, Vienna, Austria) and SPSS 23.0 (SPSS Inc., Chicago, IL, USA). Shapiro-Wilk normality test was used

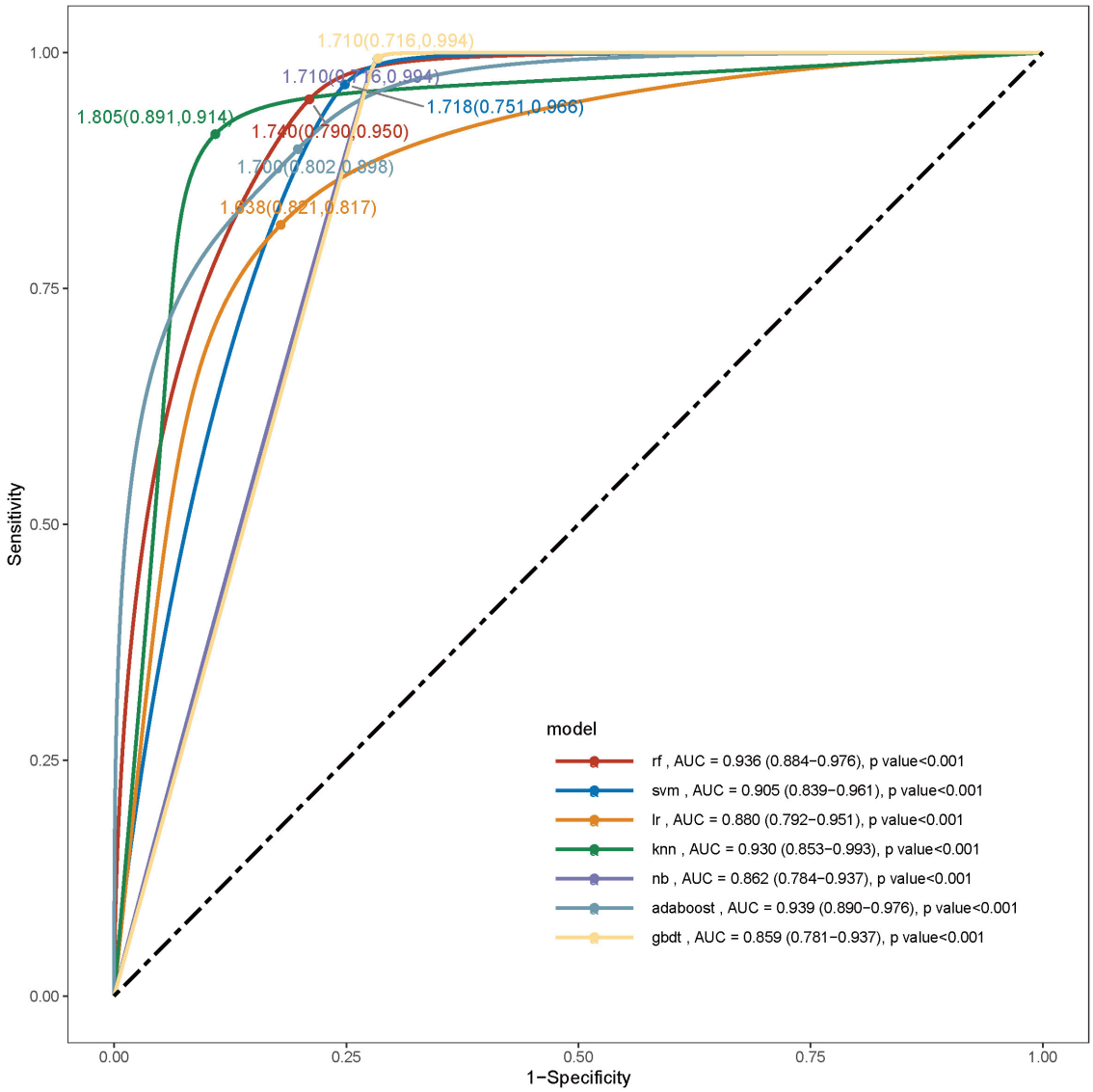

Fig. 2. Receiver operating characteristic (ROC) curves for screening osteoporotic fractures (OPF) risk in seven different models.

to determine whether the data were normally distributed and then homogeneity of variance test was performed. The rank sum test was used for the data that did not meet the normal distribution or homogeneity of variance test. Non-normally distributed continuous variables were presented as median (25%, 75%). Normally distributed continuous variables were shown as mean ± standard deviation ($\bar{x}$ ±SD). Categorical variables were presented as n (%). Single-factor analysis employed an independent samples *t*-test or χ^2 test. A *p*-value of 0.05 or less was considered statistically significant.

Results

Baseline Characteristics of the Patients

After excluding patients with loss to follow-up and missing clinical data (exceeding 20%), 196 participants were included in this study, with their baseline characteristics as follows (Table 1). The variables were ESR, OST, lumbar spine, femur, fracture history, sex, age, BMI, hunchback situation, and TBS. Among the 84 patients in the experimental group, there were 14 males and 70 females. In this group, 45 had a history of fracture, while 39 did not. Additionally, 76 had hunchbacks, while 8 did not. The control group of 112 patients consisted of 32 males and 80 females. In this group, 4 had a history of fracture, while 108 did not. Furthermore, 23 had hunchbacks, while 89 did not.

Establishment of the Risk Assessment Model

This study employed seven models, all of which utilized TBS, OST, hunchback, fracture history, age, BMI, ESR, lumbar, and sex as model variables. It conducted ten repetitions of 10-fold cross-validation, with a grid search for model training on the training dataset. Box plots were

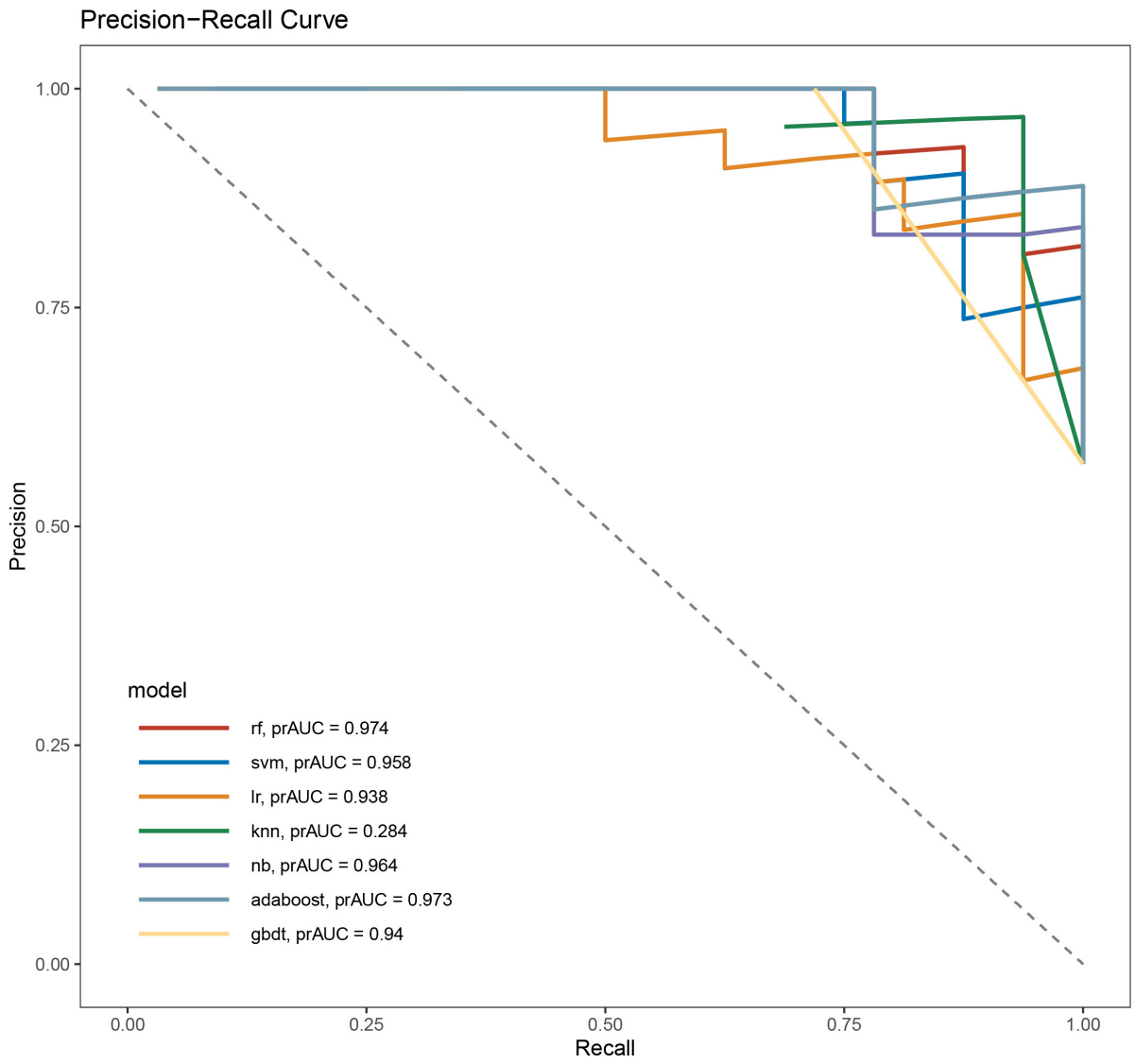

Fig. 3. Precision-Recall (PR) curves for screening OPF risk in seven different models.

generated to illustrate the area under the ROC curve, sensitivity, and specificity of each model (Fig. 1). We observed that the RF, NB, and AdaBoost models exhibited larger AUC values, indicating better classifier performance. The RF, AdaBoost, LR, and GBDT models demonstrated higher sensitivity, suggesting a strong ability to identify positive cases. In contrast, the RF, NB, Adaboost, and KNN models showed higher specificity, indicating a stronger ability to differentiate negative cases. Notably, the AdaBoost model exhibited AUC, sensitivity, and specificity values close to 1, indicating the best classification performance.

Subsequently, the training dataset was used to evaluate the best models, and the AUC, precision, recall, F1 score, and area under the Precision-Recall (PR) curve were calculated for the seven best models (Table 2). We observed that AUC values ranged from 0.9028 to 0.9996, accuracy values ranged from 0.8836 to 0.9836, precision values ranged from 0.9070 to 0.9913, recall values ranged from 0.8462 to 0.9850, F1 scores ranged from 0.8901 to 0.9854, and areas under the PR curve ranged from 0.2663 to 0.8685. These results indicate that the seven models exhibit good discriminative ability.

Evaluation of Feature Differences Between the Clinical Data Training and Validation Sets

Comparing the clinical data between the training and test sets, there were no significant differences in most characteristics such as age, ESR, TBS, and gender distribution. Although OST, BMI and lumbar showed significant differences between the training and test sets, the machine learning model was ultimately able to overcome these differences for accurate predictions, indicating that these models have relatively strong stability (Table 3).

Table 3. Training set and test set feature comparison.

Characteristic	Test	Train	Test statistic[1]	*p*–value[1]
	N = 56	N = 140		
ESR, median (25%, 75%), (mm/hour)	16.07 (10.86, 16.43)	14 (6.21, 22.52)	–0.878	0.4
OST, median (25%, 75%), (ng/mL)	15.97 (6.54, 24.44)	18.72 (11.45, 34.6)	2.014	0.04
Lumbar, median (25%, 75%), (g/cm^3)	–1.83 (–2.79, –1.23)	–1.48 (–2.53, –0.71)	1.968	0.049
Fracture_history, n (%)			0.00	>0.9
N	42 (75%)	105 (75%)	NA	
Y	14 (25%)	35 (25%)	NA	
Sex, n (%)			2.4	0.12
F	47 (84%)	103 (74%)	NA	
M	9 (16%)	37 (26%)	NA	
Age, mean (SD), (years)	66 (8)	65 (8)	1.1	0.3
BMI, mean (SD)	24.58 (2.58)	25.52 (3.27)	–2.1	0.035
Fracture, n (%)			0.00	>0.9
N	32 (57%)	80 (57%)	NA	
Y	24 (43%)	60 (43%)	NA	
Hunchback, n (%)			0.17	0.7
N	29 (52%)	68 (49%)	NA	
Y	27 (48%)	72 (51%)	NA	
TBS, mean (SD)	1.31 (0.30)	1.32 (0.30)	–0.09	>0.9

[1]Wilcoxon rank sum test; Welch Two Sample *t*-test; Pearson's Chi-squared test.

Performance Evaluation on the Risk Assessment Model

ROC Curve

In this test set, the AUC (Fig. 2) of the RF, SVM, LR, KNN, NB, Adaboost, and GBDT models were 0.936, 0.905, 0.88, 0.93, 0.862, 0.939, and 0.859, respectively ($p < 0.001$). Of these, the AdaBoost model achieved the highest AUC value, indicating the best performance.

Precision-Recall (PR) Curve

Using the test dataset, we obtained the prAUC scores for the different models as follows: RF = 0.974, SVM = 0.958, LR = 0.938, KNN = 0.284, NB = 0.964, Adaboost = 0.973, and GBDT = 0.94 ($p < 0.001$) (Fig. 3). Our results indicate that both the RF and AdaBoost models exhibit optimal classification performance.

Calibration Curves

In the calibration curves of the 7 types of models, the RF and AdaBoost models' curves are closest to the ideal curve, with no significant deviation (Fig. 4). This indicates that the RF and Adaboost models' predictive probabilities are more consistent with the actual probabilities, suggesting a higher accuracy in risk assessment modeling.

Importance of the Variables

We selected the AdaBoost model with the best predictive performance and plotted the variable importance graphs (Fig. 5A). According to the ranking of importance, the most influential features were TBS, hunchback, OST, fracture history, ESR, age, BMI, lumbar spine, and sex. Subsequently, we generated a SHapley Additive exPlanations (SHAP) graph to illustrate the variable importance within the validation set (Fig. 5B). In this study, several variables were considered, including TBS, hunchback, OST, fracture history, and ESR. Our findings consistently indicate that TBS has the highest impact on the model, followed by OST and hunchback.

Discussion

OPF represents a significant health burden worldwide, particularly among people aged more than 60 years [2]. Due to the limitations of conventional fracture prediction tools, it is increasingly essential that conducting fracture risk assessments for individuals with OPF and identifying high-risk populations in a timely manner could potentially reduce half of the fracture incidents and alleviate the burden on patients and society [10]. Commonly used risk assessment models for OPF include FRAX from the United Kingdom, QFracture, Garvan from Australia, KFRS from South Korea, and FREM from Denmark. FRAX is widely regarded as the most authoritative model; however, these models have certain limitations [21]. A study has revealed that FRAX is specifically applicable prior to osteoporosis treatment and involves only a limited number of risk factors [18]. The QFracture model has a limited sample size, necessitating further validation of the extrapolated results. However, Garvan lacks large-scale follow-up data to verify its applicability. The KFRS model from South Korea does not include potentially biased factors, such as BMI, fall history, and dietary conditions. The FREM model from

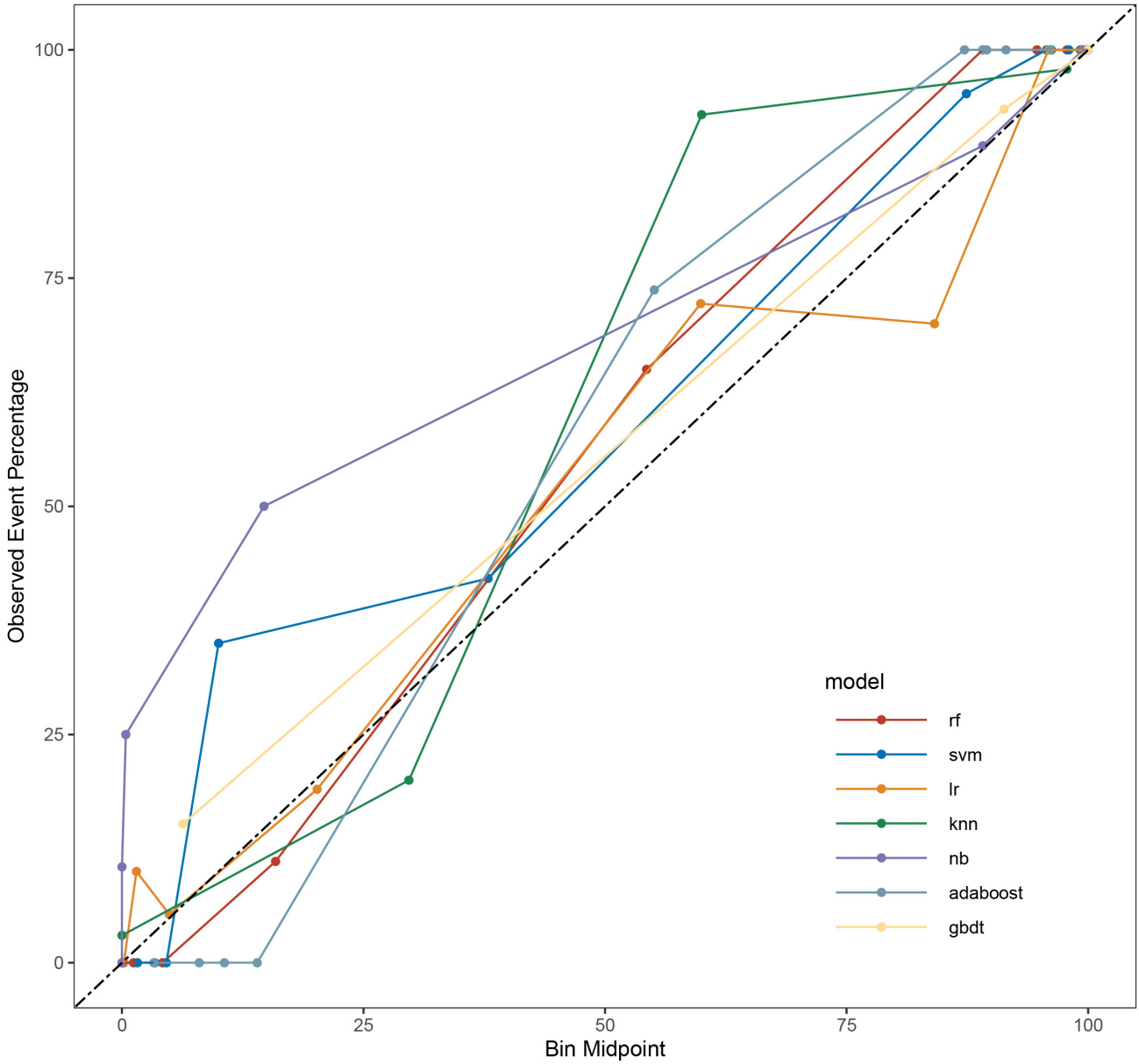

Fig. 4. Calibration curves of the results from seven risk assessment models for screening OPF risk.

Denmark incorporates only a few risk factors, potentially requiring parameter correction during extrapolation. Only several researches on OPF risk assessment in China have been reported. Xia *et al.* [22] established a logistic regression model based on 33 risk factors, with an AUC of 0.817, which can effectively predict the risk of OPF occurrence. Tang *et al.* [23] included 10 risk factors to establish a logistic regression model with an AUC of 0.818, in order to better manage elderly women with osteoporosis and improve screening for high-risk fractures in the future. Although these models demonstrate good discriminative ability, their application to the senior population in China has not been observed. There is still a dearth of risk assessment models based on population between the ages of 45 and 74 data [16]. Therefore, in this study, we aimed to develop a machine learning analysis model to assess and predict the risk of OPF based on various clinical data from patients with OPF who received treatment at four medical centers in Qingdao city.

One of the key advantages of machine learning models is their ability to integrate multiple predictors and identify complex patterns that may not be readily apparent through traditional statistical approaches [24]. In this study, we employed and validated seven machine learning models to develop a risk assessment model in patients with OPF. Our models incorporated a diverse range of features, including demographic variables, ESR, OST, TBS, and DXA bone density testing, allowing for a comprehensive assessment of fracture risk. A comparison of seven machine learning algorithms revealed that the AdaBoost model has the

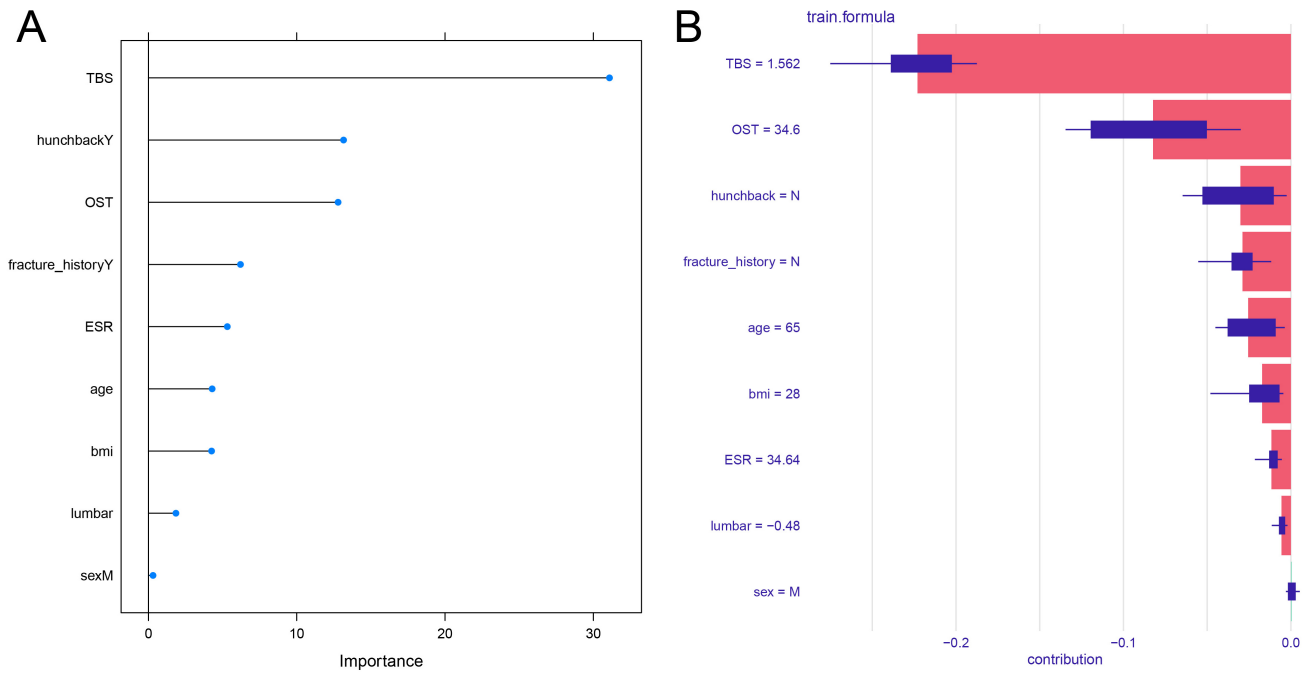

Fig. 5. Variable importance of the Adaboost risk assessment model. (A) Variable importance graph. (B) SHapley Additive exPlanations (SHAP) graph distinguishing between positive and negative.

best performance. To ensure the accuracy of the prediction models, we further carried out the calibration curve analysis. The results showed predicted probability of the models was consistent with the actual probability, indicating that the risk assessment models had high accuracy. Additionally, the AdaBoost algorithm analysis, which has good performance in the predictive model of OPF, showed that TBS is the most influential risk factor variable affecting the performance of the model, followed by OST and hunchback.

Collectively, seven machine learning algorithms are used to establish the risk assessment and prediction model of osteoporosis fractures, with the best performance in the AdaBoost model. All seven models achieved an AUC higher than 0.86, and a prAUC higher than 0.9 (except for the KNN model). TBS is the most important variable affecting the performance of the model, followed by OST and hunchback, thereby providing a new insight into the prediction and intervention of osteoporosis fractures.

However, our research has certain limitations. First, the retrospective and observational nature of our study may introduce inevitable selection bias. Second, our sample size of patients with OPF was limited, potentially affecting the generalizability of the prediction model to other populations. Therefore, a large sample, multi-center study is essential to verify the model's applicability. Finally, certain mechanisms that influence or induce osteoporosis, including estrogen deficiency, were not discussed in this study. Regarding the model's potential in clinical practice, it could potentially serve as a valuable tool for clinicians in assessing the risk of OPF. By providing accurate and personalized risk assessments, the model could facilitate early intervention and preventive strategies, ultimately improving patient outcomes and reducing the burden of OPF.

Conclusions

In conclusion, our study highlights the potential of seven machine learning analysis models in predicting OPF. By harnessing the power of data-driven approaches, we can enhance fracture risk prediction accuracy, improve clinical decision-making, and ultimately, reduce the burden of OPF on individuals and healthcare systems.

Availability of Data and Materials

All the data shown in this study are included in the manuscript.

Author Contributions

BL and CL designed the study; YRY and FS performed the study; YLW, TW and XXC collected and analyzed the data; CL analyzed the data. BL and CL were involved in drafting the manuscript and all authors were involved in revising it critically for important intellectual content. All authors gave final approval of the version to be published. All authors participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

This study was approved by the Medical Ethics Committee of Qingdao Hiser Hospital Affiliated with Qingdao University (2020 Ethics Review (Science) No. 113). Informed consent was obtained from all subjects involved in the study. All experiments in this study were carried out in accordance with the outlines of the Declaration of Helsinki.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Li H, Hu Y, Gan L, Wang Y. Observation of the clinical effectiveness of evidence-based nursing for patients with a vertebral osteoporotic fracture. Journal of Back and Musculoskeletal Rehabilitation. 2023; 36: 517–523.

[2] Clynes MA, Harvey NC, Curtis EM, Fuggle NR, Dennison EM, Cooper C. The epidemiology of osteoporosis. British Medical Bulletin. 2020; 133: 105–117.

[3] Aibar-Almazán A, Voltes-Martínez A, Castellote-Caballero Y, Afanador-Restrepo DF, Carcelén-Fraile MDC, López-Ruiz E. Current Status of the Diagnosis and Management of Osteoporosis. International Journal of Molecular Sciences. 2022; 23: 9465.

[4] Chen P, Li Z, Hu Y. Prevalence of osteoporosis in China: a meta-analysis and systematic review. BMC Public Health. 2016; 16: 1039.

[5] Chadha M, Chaddha R, Divakar H, Kalyan H, Seth S, Shah P. Osteoporosis: epidemiology, pathogenesis, evaluation and treatment. Open Journal of Orthopedics. 2022; 12: 153–182.

[6] International Osteoporosis Foundation. Facts and statistics about osteoporosis and its impact. In ternational Osteoporosis Foundation. Journal of Bone and Mineral Research. 2009; 4: 113–118.

[7] Bolek H, Bolek EC, Tanriover MD. Economic burden of osteoporotic fractures: the tip of the iceberg? The American Journal of Managed Care. 2020; 26: e376.

[8] Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporosis International. 2006; 17: 1726–1733.

[9] Si L, Winzenberg TM, Jiang Q, Chen M, Palmer AJ. Projection of osteoporosis-related fractures and costs in China: 2010-2050. Osteoporosis International. 2015; 26: 1929–1937.

[10] Odén A, McCloskey EV, Johansson H, Kanis JA. Assessing the impact of osteoporosis on the burden of hip fractures. Calcified Tissue International. 2013; 92: 42–49.

[11] Lekamwasam S. The diversity of Fracture Risk Assessment Tool (FRAX)-based intervention thresholds in Asia. Osteoporos Sarcopenia. 2019; 5: 104–108.

[12] Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. BMJ (Clinical Research Ed.). 2009; 339: b4229.

[13] Holloway-Kew KL, Zhang Y, Betson AG, Anderson KB, Hans D, Hyde NK, *et al*. How well do the FRAX (Australia) and Garvan calculators predict incident fractures? Data from the Geelong Osteoporosis Study. Osteoporos Int. 2019; 30: 2129–2139.

[14] Kim HY, Jang EJ, Park B, Kim TY, Shin SA, Ha YC, *et al*. Development of a Korean Fracture Risk Score (KFRS) for Predicting Osteoporotic Fracture Risk: Analysis of Data from the Korean National Health Insurance Service. PLoS ONE. 2016; 11: e0158918.

[15] Rubin KH, Möller S, Holmberg T, Bliddal M, Søndergaard J, Abrahamsen B. A New Fracture Risk Assessment Tool (FREM) Based on Public Health Registries. Journal of Bone and Mineral Research. 2018; 33: 1967–1979.

[16] Li C, Shi Q, Gong YQ, Zhang T, Lu K. Development and validation of a prognostic nomogram for 3-year all-cause mortality risk among elderly patients undergoing surgery for osteoporotic fractures. Frontiers in Medicine. 2024; 11: 1284207.

[17] Tran T, Bliuc D, Pham HM, van Geel T, Adachi JD, Berger C, *et al*. A Risk Assessment Tool for Predicting Fragility Fractures and Mortality in the Elderly. Journal of Bone and Mineral Research: the Official Journal of the American Society for Bone and Mineral Research. 2020; 35: 1923–1934.

[18] Baim S, Blank R. Approaches to Fracture Risk Assessment and Prevention. Current Osteoporosis Reports. 2021; 19: 158–165.

[19] Osteoporosis Group OS, Chinese Medical Association. Guidelines for the diagnosis and treatment of osteoporotic fractures. Chinese Journal of Orthopaedics. 2017; 37: 1–10. (In Chinese)

[20] Zhang Z, Liu Z, Li N, Zhang M, Huang Q, Ma Y, *et al*. Expert consensus on the diagnosis of osteoporosis in Chinese Population. Chinese Journal of Osteoporosis. 2014; 1007–1010. (In Chinese)

[21] Blake GM, Fogelman I. An update on dual-energy x-ray absorptiometry. Seminars in Nuclear Medicine. 2010; 40: 62–73.

[22] Xia P, Jiang Y, Cai F, Peng S, Xu Z. Construction and verification of risk prediction model of osteoporotic fractures in patients with osteoporosis in China. Frontiers in Public Health. 2024; 12: 1380218.

[23] Tang N, Gao L, Song J, Li Y, Song M, Qiu C, *et al*. Risk analysis for subsequent fracture of osteoporotic fractures in Chinese women over age 60: a nationwide cross-sectional study. Scientific Reports. 2024; 14: 13319.

[24] Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: A review. Biotechnology Advances. 2021; 49: 107739.