Clinical Interventions in Aging

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Predicting Hip Osteoporosis with Routine Demographic and Biochemical Data: The Shao HipOsteoRisk Model

Xiangheng Dai[1,*], Weiqi Lu[2,*], Fuzhou Xu[3], Zongping Deng[1], Guorong Xiao[4], Zeping Li[5], Yilv Zhang[3], Ao Liu[3], Weipeng Guo[6], Kunhua Huang[7], Wengang Zhu[8], Junhao Tan[8], Beidi Zhou[9], Chao Lou[10], Qiang Wu[1]

[1]Department of Spinal Surgery, Shaoguan First People's Hospital, Affiliated Hospital of Shaoguan University, Shaoguan, Guangdong, People's Republic of China; [2]Department of Chinese traditional traumatology, Dongguan Shijie Hospital, Dongguan, Guangdong, People's Republic of China; [3]The First School of Clinical Medicine, Guangdong Medical University, Zhanjiang, Guangdong, People's Republic of China; [4]Department of orthopedics, Wengyuan Hospital of Traditional Chinese Medicine, Shaoguan, Guangdong, People's Republic of China; [5]Department of Pain, the Fourth People's Hospital of Nanhai District, Foshan, Guangdong, People's Republic of China; [6]Department of Information, Shaoguan First People's Hospital, Affiliated Hospital of Shaoguan University, Shaoguan, Guangdong, People's Republic of China; [7]Department of Laboratory, Shaoguan First People's Hospital, Affiliated Hospital of Shaoguan University, Shaoguan, Guangdong, People's Republic of China; [8]Department of Osteoarthritis, Yuebei People's Hospital, Shaoguan, Guangdong, People's Republic of China; [9]Department of Prevention and Health Care, Shaoguan First People's Hospital, Affiliated Hospital of Shaoguan University, Shaoguan, Guangdong, People's Republic of China; [10]Department of Trauma Surgery and Orthopedics, AGAPLESION EV. Klinikum Schaumburg, Hannover Medical School, Obernkirchen, Germany

*These authors contributed equally to this work

Correspondence: Chao Lou; Qiang Wu, Email chaolou213@gmail.com; sgwuqiang@sgu.edu.cn

Objective: To develop and externally validate a simple, accessible prediction model for identifying individuals at risk of hip osteoporosis using routine demographic and laboratory data.

Methods: This retrospective study included 7686 adult patients who underwent hip dual-energy X-ray absorptiometry (DXA) at two medical centers in northern Guangdong, China. A total of 4638 patients were used for model development and 3048 for external validation. Predictors were selected using appropriate imputation and regularized regression techniques to ensure stability across datasets. Model performance was evaluated using discrimination, calibration, and clinical utility metrics.

Results: Four routinely available variables—age, sex, body mass index, and the serum albumin-to-alkaline phosphatase ratio—were identified as the key predictors. The final logistic regression model demonstrated strong discrimination, with an area under the curve of 0.9107 in the development cohort and 0.8286 in the external validation cohort. Sensitivity and specificity were both favorable, and calibration showed good agreement between predicted and observed risk across most probability ranges. Decision curve analysis indicated meaningful net clinical benefit across a wide range of threshold probabilities, supporting the model's potential to improve risk stratification in practice.

Conclusion: We developed and validated a practical predictive model for hip osteoporosis based entirely on information commonly obtained during routine clinical care. Because it requires no specialized testing beyond standard laboratory panels, the model offers a low-cost, scalable screening tool—particularly valuable in settings where DXA access is limited. Its strong performance and ease of application suggest that it may help clinicians identify high-risk patients earlier, guide referral for confirmatory DXA scanning, and support more proactive osteoporosis prevention strategies.

Keywords: osteoporosis, risk assessment, predictive model

Introduction

With the rapidly aging population in China, the prevalence of osteoporosis has increased significantly, emerging as a major public health concern.[1] Despite heightened awareness, substantial gaps persist in the diagnosis and management of osteoporosis, with many patients failing to receive timely or appropriate treatment.[2] As a chronic condition with high

https://doi.org/10.2147/CIA.S566021
Received: 8 September 2025
Accepted: 12 January 2026
Published: 20 January 2026

incidence and severe consequences, osteoporosis requires long-term, sequential or combined pharmacological therapies to improve bone mineral density and reduce fracture risk.[3] Hip osteoporosis is particularly concerning, as it often leads to hip fractures associated with high rates of disability and mortality.[4] Globally, approximately nine million osteoporotic fractures occur each year, with hip fractures accounting for more than 30%.[5] Therefore, early identification and intervention in high-risk populations are critical.[6,7] While existing prediction models FRAX primarily focus on assessing the risk of hip fractures,[8] osteoporosis itself is the underlying cause of these events.[9]

The internationally recognized diagnostic standard for osteoporosis is based on dual-energy X-ray absorptiometry (DXA) measurements.[10] In DXA scanning, the anteroposterior lumbar spine assessment typically includes vertebrae L1 to L4, encompassing their posterior elements. However, this region is prone to measurement inaccuracies due to age-related degenerative changes, such as vertebral osteophyte formation, sclerosis of the vertebral bodies and facet joints, and abdominal aortic calcifications.[11] In contrast, DXA scans of the proximal femur—particularly the femoral neck and total hip—are more clinically relevant and reliable for osteoporosis diagnosis. These regions are anatomically stable and less affected by soft tissue interference, making them preferred sites for accurate bone density assessment.[12] Consequently, a predictive model based on hip bone mineral density as the outcome is likely to offer greater diagnostic reliability.

Despite its clinical value, access to DXA remains limited, especially in economically underdeveloped regions. This underscores the urgent need for alternative tools to effectively assess osteoporosis risk. Currently recognized screening tools include the International Osteoporosis Foundation (IOF) One-Minute Osteoporosis Risk Test[13] and the Osteoporosis Self-Assessment Tool for Asians (OSTA).[14] The IOF test is quick, simple, and user-friendly but serves only as a preliminary screening tool and lacks diagnostic capability.[15] OSTA, widely used among Asian women, is limited in scope as it excludes men and relies solely on age and weight, which, while convenient, compromises predictive accuracy.[12] To address these limitations, Cheung et al recently developed and validated the Chinese Osteoporosis Screening Algorithm (COSA).[12] This model, built on the Chinese population in Hong Kong, incorporates fracture history in addition to age and weight. COSA has demonstrated significantly improved predictive performance compared to OSTA and correlates with 10-year hip fracture risk.

Several emerging models have attempted to incorporate biochemical markers or machine-learning techniques to improve osteoporosis prediction. However, many of these approaches rely on specialized assays (eg, bone turnover markers,[16] vitamin D[17] or PTH levels[18]) or advanced imaging inputs[19] that are not routinely available in most clinical settings. Machine-learning–based models often demonstrate good discrimination but require complex data preprocessing and lack interpretability, limiting their practical adoption.[18] In contrast, few models have focused on leveraging simple, universally obtained laboratory tests to create an accessible, clinically transparent prediction tool. This gap highlights the need for a model that balances accuracy, interpretability, and real-world feasibility.

Serum albumin (ALB) and alkaline phosphatase (ALP) are two routinely measured biomarkers with established links to bone physiology. ALB reflects nutritional and inflammatory status, both of which influence bone turnover and fragility.[20] ALP increases during periods of active bone remodeling.[21] The ALB/ALP ratio therefore integrates nutritional status with metabolic activity, offering a physiologically meaningful composite indicator of bone health. Despite its potential relevance, this ratio has rarely been incorporated into osteoporosis prediction models.

Given these limitations, we sought to develop a practical, clinically interpretable model that combines demographic information with simple biochemical markers—including the ALB/ALP ratio—to identify individuals at risk of hip osteoporosis. By using data readily available in routine inpatient and health-check settings, this model aims to provide an accessible screening tool that can support early detection where DXA access is limited.

Methods

Research Design and Data Sources

This study employed a retrospective cohort design using electronic medical record data from two independent tertiary medical centers in northern Guangdong, China. Only hospitalized adult patients were included, as laboratory testing was part of the routine inpatient evaluation and therefore not consistently available for outpatients.

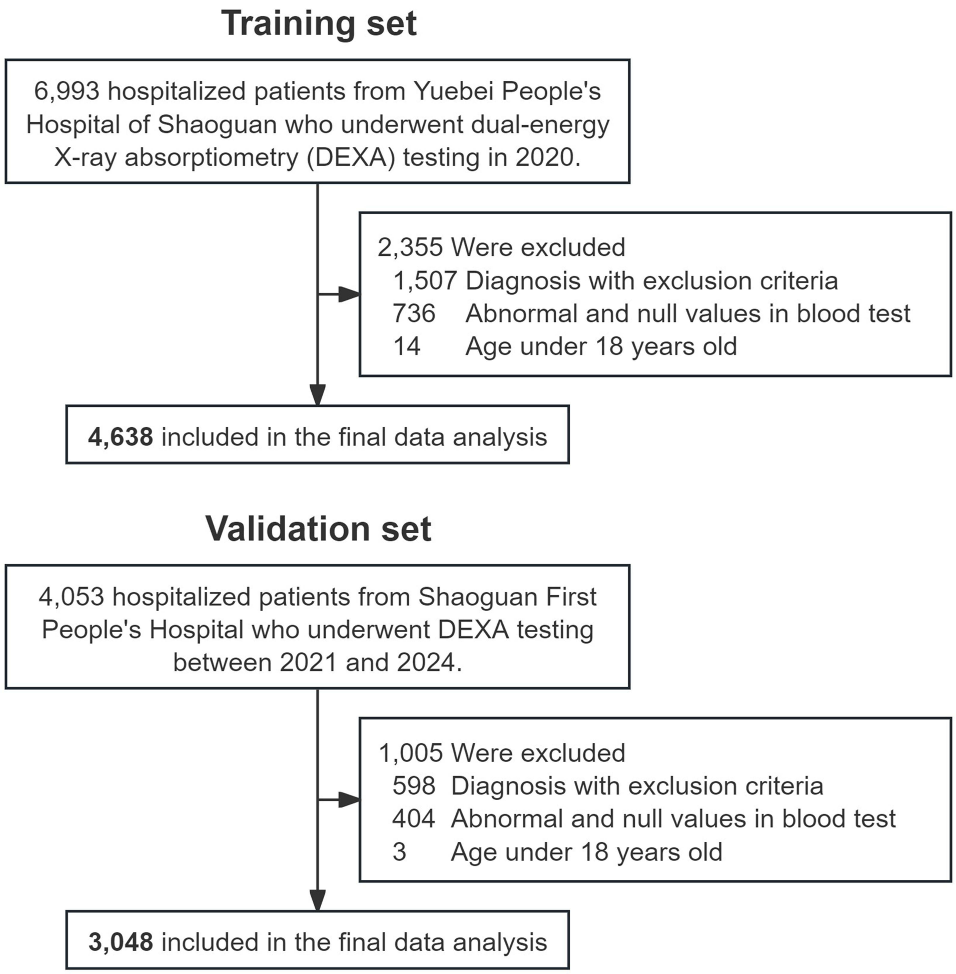

Figure 1 Flowchart of Participant Selection Process.

The training cohort consisted of 6993 inpatients who underwent hip DXA at Yuebei People's Hospital of Shaoguan between January and December 2020. After applying predefined inclusion and exclusion criteria, 4638 patients remained for analysis.

The validation cohort included 4053 inpatients who received DXA testing at Shaoguan First People's Hospital between January 2021 and January 2024, from which 3048 eligible patients were retained. The two hospitals differ in their clinical case mix and patient admission patterns, providing a natural source of population heterogeneity to support external validation.

DXA examinations were performed upon hospital admission. Yuebei People's Hospital used a Hologic DXA system, whereas Shaoguan First People's Hospital used a GE Lunar Prodigy Pro scanner (manufactured in Mexico). Although equipment differed, both institutions follow standardized calibration procedures in accordance with national guidelines.

A detailed flowchart of patient inclusion and exclusion is presented in Figure 1.

Inclusion and Exclusion Criteria

Inclusion Criteria: 1. Age > 18 years; 2. Underwent hip DXA bone mineral density testing; 3. Had liver and kidney function blood tests performed.

Exclusion Criteria: 1. Co-existing conditions that may affect bone metabolism, such as hyperparathyroidism, Cushing's disease, pituitary disorders, hematologic diseases, malignant tumors, autoimmune diseases, or severe mental disorders; 2. Previous or current use of anti-osteoporosis medications (eg, zoledronic acid, calcitriol); 3. Long-term use of medications that may affect bone metabolism, such as corticosteroids, estrogens, anticonvulsants, and anticoagulants; 4. Missing clinical case data; 5. Patients deemed unsuitable for participation in the study by the researchers.

Data Processing and Variable Selection

In the data processing stage, missing values were handled using multiple imputation via the random forest method implemented in the mice package in R (m = 5, method = "rf"), with default settings of 10 iterations and 100 trees per forest. The pre- and post-imputation distributions of variables with missingness were compared and showed no significant differences, confirming that imputation did not distort the underlying data structure. The imputed datasets were analyzed using least absolute shrinkage and selection operator (LASSO) logistic regression implemented in the glmnet package in R. Model tuning was performed with 10-fold cross-validation to determine the optimal regularization parameter λ (both λ_min and the more parsimonious λ_1se). Variables with non-zero coefficients at λ_1se were retained. To ensure stability, the procedure was repeated across five imputed datasets, and consistent predictors were selected for model construction.

Construction of Logistic Regression Model

The logistic regression model was used to evaluate the association between variables and disease occurrence. None of the independent variables were standardized, as their scales were consistent. The model was fitted by maximizing the likelihood function, and variable selection was performed using the AIC criterion. The entirety of the statistical analysis was executed using R version 4.0.3, with a significance criterion of P less than 0.05 applied to the logistic regression analysis.

Model Validation and Evaluation

To evaluate the performance of the logistic regression model, we used ROC curve analysis, DCA curve, and calibration curve. The ROC curve evaluates the sensitivity and specificity of the model, the DCA curve assesses the clinical usefulness of the model under different clinical decision thresholds, and the calibration curve evaluates the model's calibration. All statistical analyses were conducted using the "pROC" and "rms" packages in R. To further assess the stability, internal validity, and potential overfitting of the prediction model, we performed a bootstrap resampling procedure using the rms package in R. A total of 1000 bootstrap resamples were generated from the original training dataset.

Ethical Statement and Data Security

This study has received ethical approval from the Yuebei People's Hospital Medical Ethics Committee (Approval Number: YBSKY-2025-012-001) and the Shaoguan First People's Hospital Ethics Review Committee (Approval Number: Shaoguan First People's Hospital Medical Ethics Review Number: (2024)124). The study was conducted in accordance with the principles of the Declaration of Helsinki. Given that this study utilized medical records obtained from previous clinical care and met the following criteria: 1) the study objectives are of significant importance; 2) the risks to participants are no greater than minimal; 3) waiving informed consent will not adversely affect participants' rights and well-being; and 4) all patient data have been de-identified to ensure patient privacy and safety, we applied for and obtained approval for waiver of informed consent from both Ethics Committees. The study is registered in the National Medical Research Registration and Filing Information System of China under the registration number MR-44-25-010258. As this is a retrospective observational study, it was not applicable for registration on the Clinical trial registration platform.

Statistical Analysis

All statistical analyses were performed using R software. For qualitative variables, the Chi-squared test was used for comparisons. For quantitative variables, Student's *t*-test or the Mann–Whitney *U*-test was applied, depending on the data distribution. The results of the logistic regression model are presented as odds ratios (ORs) with 95% confidence intervals (CIs). All statistical hypothesis tests were conducted using two-sided tests, with a significance level set at $P < 0.05$.

Result

Basic Characteristics of the Data

In this clinical study (Table 1), both training and validation sets show that osteoporosis groups have significantly older ages (training set: osteoporosis group mean age 75 vs non-osteoporosis group 56, P < 0.001; validation set: 73 vs 60, P < 0.001), higher female proportions (training set: 69.20% vs 59.78%, P < 0.001; validation set: 84.74% vs 64.15%, P < 0.001), and lower body mass index (BMI) (training set: 21.93 kg/m^2 vs 23.64 kg/m^2, P < 0.001; validation set: 22.22 kg/m^2 vs 24.02 kg/m^2, P < 0.001) compared to non-osteoporosis groups.

Significant differences (Table 1) in serum biochemical indicators including sodium, calcium, phosphorus, uric acid, total protein, ALB, and ALP are also observed between the two groups in both sets (training set: serum sodium P = 0.003, serum calcium P < 0.001, serum phosphorus P = 0.003, serum uric acid P < 0.001, total serum protein P < 0.001, serum ALB P < 0.001, serum ALP P < 0.001; validation set: serum sodium P < 0.001, serum calcium P = 0.005, serum uric acid P < 0.001, total serum protein P < 0.001, serum ALB P < 0.001, serum ALP P < 0.001).

Moreover, the ratio of ALB/ALP (Table 1) differs significantly between groups (training set: osteoporosis group 0.51 vs non-osteoporosis group 0.62, P < 0.001; validation set: 0.47 vs 0.54, P < 0.001), which may be associated with osteoporosis pathogenesis.

Univariate Analysis Results

In univariate analysis (Table 2), several variables showed significant associations with hip osteoporosis across both training and validation sets.

Table 1 General Characteristics

Hip osteoporosis	Training Set			Validation Set		
	Yes (n=1344)	No (n=3294)	P values	Yes (n=1271)	No (n=1777)	P values
Age (years, SD)	75 (68–82)	56 (49–64)	<0.001	73 (66–80)	60 (54–68)	<0.001
Gender (female, %)	930 (69.20%)	1969 (59.78%)	<0.001	1077 (84.74%)	1140 (64.15%)	<0.001
Body mass index (kg/m^2, SD)	21.93 (19.56–24.24)	23.64 (21.48–25.95)	<0.001	22.22 (20.00–24.67)	24.02 (21.79–26.44)	<0.001
Serum sodium (mmol/L, SD)	143.2 (142–145)	143 (142–145)	0.003	140.8 (138.9–142.2)	140.3 (138.9–141.70)	<0.001
Serum calcium (mmol/L, SD)	2.28 (2.21–2.34)	2.29 (2.23–2.36)	<0.001	2.36 (2.26–2.46)	2.38 (2.3–2.48)	<0.001
Serum phosphorus (mmol/L, SD)	1.07 (0.95–1.18)	1.08 (0.96–1.2)	0.003	1.15 (1.03–1.27)	1.16 (1.05–1.28)	0.051
Serum magnesium (mmol/L, SD)	0.88 (0.83–0.93)	0.88 (0.84–0.93)	0.468	0.88 (0.82–0.95)	0.88 (0.82–0.94)	0.799
Serum uric acid (μmol/L, SD)	303.05 (247.08–373.53)	322 (262–389)	<0.001	292 (237–351.7)	316 (264–379)	<0.001
Serum creatinine (μmol/L, SD)	70 (60–82)	70 (60.08–82)	0.804	64 (54–77)	65 (55–79)	0.238
Total serum protein (g/L, SD)	70.2 (66.6–74.2)	71.1 (67.6–74.9)	<0.001	67.3 (63.5–72.1)	68.7 (64.65–73)	<0.001
Serum albumin (g/L, SD)	40 (37.4–42.7)	41.7 (39.3–44.2)	<0.001	40.1 (37.1–43.1)	42.1 (39.4–44.8)	<0.001
Serum alkaline phosphatase (U/L, SD)	78 (64.05–94)	68 (56–81)	<0.001	85 (70–105)	78 (65–94)	<0.001
ALB/ALP (SD)	0.51 (0.42–0.63)	0.62 (0.50–0.75)	<0.001	0.47 (0.38–0.58)	0.54 (0.45–0.65)	<0.001

Note: Data are presented as median (interquartile range) for continuous variables and n (%) for categorical variables.

Table 2 Univariate Logistic Regression Analysis

	Training Set			Validation Set		
	Univanriate OR (95% CI)	Coefficient	P value	Univanriate OR (95% CI)	Coefficient	P value
Age(years)	1.1762 (1.1649–1.1876)	0.16226	<0.0001	1.1185 (1.1084–1.1286)	0.11197	<0.0001
Gender	0.6615 (0.5779–0.7572)	−0.41321	<0.0001	0.3224 (0.2690–0.3863)	−1.13206	<0.0001
BMI (kg/m^2)	0.843 (0.8257–0.8607)	−0.17075	<0.0001	0.8565 (0.8374–0.8760)	−0.15493	<0.0001
Serum sodium (mmol/L)	1.0226 (0.9965–1.0495)	0.022393	0.0905	1.0278 (1.0008–1.0555)	0.027422	0.0433

(*Continued*)

Table 2 (Continued).

	Training Set			Validation Set		
	Univanriate OR (95% CI)	**Coefficient**	**P value**	**Univanriate OR (95% CI)**	**Coefficient**	**P value**
Serum calcium (mmol/L)	0.2917 (0.1599–0.5321)	−1.23195	0.0001	0.257 (0.1567–0.4217)	−1.35859	<0.0001
Serum phosphorus (mmol/L)	0.5998 (0.4293–0.8378)	−0.51123	0.0027	0.5837 (0.3919–0.8694)	−0.53837	0.0081
Serum magnesium (mmol/L)	0.8479 (0.4070–1.7668)	−0.16495	0.6596	1.1748 (0.7955–1.7348)	0.16106	0.4181
Serum uric acid (μmol/L)	0.9986 (0.9979–0.9992)	−0.0014328	<0.0001	0.9969 (0.9961–0.9978)	−0.0030706	<0.0001
Serum creatinine (μmol/L)	1.0017 (0.9980–1.0055)	0.0017205	0.3684	0.9977 (0.9938–1.0016)	−0.0023351	0.243
Total serum protein (g/L)	0.9749 (0.9640–0.9860)	−0.02538	<0.0001	0.9777 (0.9672–0.9884)	−0.022542	<0.0001
Serum albumin (g/L)	0.8953 (0.8806–0.9103)	−0.11059	<0.0001	0.8948 (0.8791–0.9107)	−0.11118	<0.0001
Serum alkaline phosphatase (U/L)	1.0174 (1.0147–1.0201)	0.017273	<0.0001	1.0034 (1.0014–1.0054)	0.0034152	0.0007
ALB/ALP	0.0382 (0.0257–0.0567)	−3.26566	<0.0001	0.0687 (0.0429–0.1100)	−2.67778	<0.0001

Age was positively associated with osteoporosis (training: OR = 1.1762, 95% CI: 1.1649–1.1876; validation: OR = 1.1185, 95% CI: 1.1084–1.1286), indicating increased risk with older age. Gender was also significantly associated, with males having lower odds of osteoporosis compared to females (training: OR = 0.6615, 95% CI: 0.5779–0.7572; validation: OR = 0.3224, 95% CI: 0.2690–0.3863).

BMI showed a negative association with osteoporosis in both sets (training: OR = 0.843, 95% CI: 0.8257–0.8607; validation: OR = 0.8565, 95% CI: 0.8374–0.8760), suggesting lower BMI increases osteoporosis risk. Serum sodium levels were significantly associated with osteoporosis only in the validation set (OR = 1.0278, 95% CI: 1.0008–1.0555), indicating higher levels may increase risk.

Serum calcium and phosphorus levels both showed significant negative associations with osteoporosis in both sets, meaning lower levels increase risk. Serum uric acid also demonstrated a significant association, with lower levels linked to higher osteoporosis risk (training: OR = 0.9986, 95% CI: 0.9979–0.9992; validation: OR = 0.9969, 95% CI: 0.9961–0.9978).

Total serum protein and its component serum ALB both showed significant associations with osteoporosis. Lower levels of total serum protein (training: OR = 0.9749, 95% CI: 0.9640–0.9850; validation: OR = 0.9777, 95% CI: 0.9672–0.9884) and serum ALB (training: OR = 0.8953, 95% CI: 0.8806–0.9103; validation: OR = 0.8948, 95% CI: 0.8791–0.9107) were associated with increased osteoporosis risk.

Serum ALP showed a slight positive association with osteoporosis in both sets (training: OR = 1.0174, 95% CI: 1.0147–1.0201; validation: OR = 1.0034, 95% CI: 1.0014–1.0054). The ALB/ALP ratio was significantly associated with osteoporosis in both datasets (training: OR = 0.0382, 95% CI: 0.0257–0.0567; validation: OR = 0.0687, 95% CI: 0.0429–0.1100), suggesting a lower ratio increases risk.

Serum magnesium and creatinine did not show significant associations with osteoporosis in either set.

LASSO Regression Analysis for Variable Selection

In this study, some variables had a certain degree of missing data, with the proportion of missing data in the training set ranging from 1.2% to 1.6% and in the validation set ranging from 0.8% to 12.8%. Specifically, in the training set, missing data were observed in Serum phosphorus, Serum magnesium, and Serum uric acid, while no missing data were found in other variables. In the validation set, missing data were observed in Serum calcium, Serum phosphorus, Serum magnesium, Serum uric acid, and Serum creatinine, with other variables remaining complete.

To ensure data integrity and minimize bias, we used the random forest method to perform multiple imputations separately on the training and validation sets, generating five complete datasets for each. The imputed data showed high consistency with the original data distributions, indicating the reliability of the imputation process. Subsequently, LASSO regression was performed on each dataset to identify variables associated with hip osteoporosis.

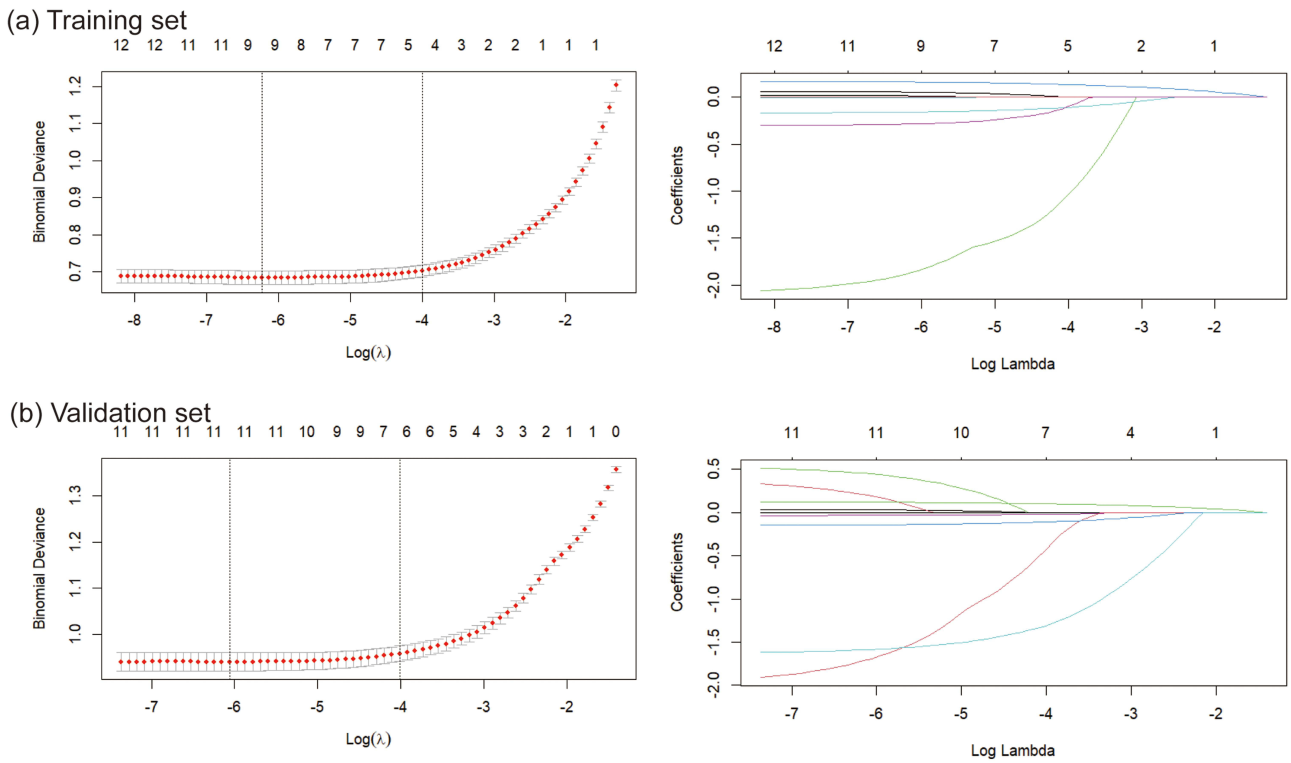

Figure 2 Results of LASSO regression analysis for variable selection. (**a**) Results from one LASSO regression analysis conducted on the training set, showing the selection of four variables: age, sex, BMI, and the ALB/ALP. These variables were consistently selected across all five analyses with non-zero coefficients. (**b**) Results from one LASSO regression analysis conducted on the validation set, showing the selection of six variables. Among them, four variables (age, sex, BMI, and ALB/ALP) were consistently selected across all five analyses, while the remaining two variables varied due to differences in the imputed datasets.

Figure 2a illustrates the results of one LASSO regression analysis on the training set, which identified four variables: age, sex, BMI, and the ALB/ALP. These variables were consistently selected with non-zero coefficients across all five regression analyses.

Figure 2b presents the results of one LASSO regression analysis on the validation set, which identified six variables. Among these, the four variables identified in the training set (age, sex, BMI, and ALB/ALP) were consistently included across all five analyses. However, the remaining two variables varied due to differences in the imputed datasets, likely resulting from the randomness of the imputation process and the varying correlations among the variables. This variability did not substantially impact the overall findings.

It is important to note that variables with missing data (eg, Serum phosphorus, Serum magnesium, Serum uric acid, Serum calcium, and Serum creatinine) were excluded by the LASSO regression analysis. As a result, these variables were not included in subsequent analyses.

Considering the consistency between the training and validation sets, we ultimately selected the four variables identified in the training set (age, sex, BMI, and ALB/ALP) for further analysis. Selecting these consistently identified variables ensures the robustness and generalizability of the model while avoiding potential biases introduced by variability in the imputed datasets. Based on these variables, we constructed a logistic regression model to further evaluate their relationship with hip osteoporosis.

The final logistic regression model (Shao HipOsteoRisk Model) is represented as:

$$f(x) = -6.35634 + 0.16042 * \text{Age} - 0.16679 * \text{BMI} - 0.46871 * \text{Gender} - 2.00693 * \text{ALB/ALP}$$

 https://doi.org/10.2147/CIA.S566021

$$P = \frac{1}{1 + EXP(-f(x))}$$

Among them:

P is the predictive probability of hip osteoporosis;

Age is the age (unit: year);

BMI is body mass index (unit: kg/m^2);

Gender, Gender=1 indicates male gender, Gender=0 indicates female gender;

ALB/ALP is the ratio of serum albumin to serum alkaline phosphatase.

In addition, based on our logistic regression model, we have developed a web-based tool that can generate real-time probabilities of hip osteoporosis, providing a convenient resource for clinical use. The website can be accessed at: https://sgsy-ky.github.io (International users) or http://sgsyy.cn:10011/blog/kyindex (Chinese users).

Model Performance Evaluation-ROC Curve Analysis

Through ROC curve analysis, the area under the curve (AUC) of model in the training set was 0.9107 (Figure 3a), indicating a high identification ability. The AUC in the external validation set was 0.8286 (Figure 3b), indicating consistent performance on an independent dataset and strong generalizability. Internal validation was performed using bootstrap resampling (B=1000). The optimism-corrected C-index was 0.91 (Dxy=0.8208), indicating excellent discrimination. The calibration slope (0.9976) and intercept (−0.0025) were close to the ideal values of 1 and 0, respectively, suggesting minimal overfitting. The mean absolute calibration error was 0.025 and the 90th percentile of absolute error was 0.051, indicating good agreement between predicted and observed risk. The bias-corrected calibration curve almost overlapped with the ideal line (Figure 3c).

To evaluate the added value of the ALB/ALP ratio, we compared the predictive performance of two models—with and without this variable. In the training cohort, the inclusion of ALB/ALP significantly improved the AUC from 0.9066 (95% CI: 0.8984–0.9148) to 0.9107 (95% CI: 0.9027–0.9188) (DeLong's test: Z = 3.97, P < 0.001). This improvement was accompanied by slight increases in sensitivity (from 88.69% to 89.81%), specificity (from 76.62% to 76.75%), and overall accuracy (from 80.12% to 80.53%). In the external validation cohort, the AUC increased from 0.8244 (95% CI: 0.8045–0.8365) to 0.8286 (95% CI: 0.8141–0.8430) with the inclusion of ALB/ALP (DeLong's test: Z = 2.67, P = 0.0076). Notably, this also resulted in improved specificity (from 68.69% to 74.68%) and accuracy (from 73.75% to 75.52%), although sensitivity decreased slightly (from 81.02% to 76.71%). These findings suggest that ALB/ALP contributes to enhanced model identification ability.

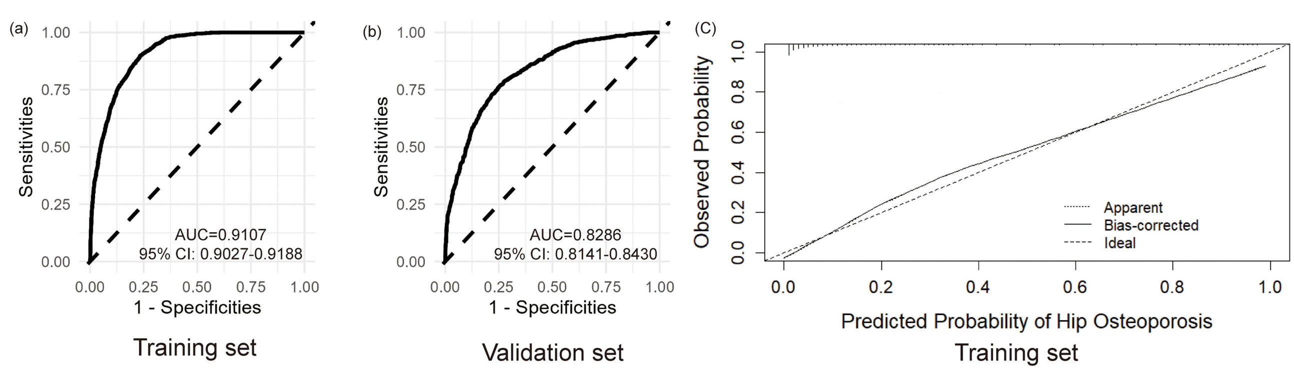

Figure 3 Model performance evaluation using ROC curves and bootstrap calibration. (**a**) ROC curve for the training cohort, with an AUC of 0.9107, indicating excellent discriminative ability. (**b**) ROC curve for the external validation cohort, showing an AUC of 0.8286, demonstrating good generalizability and stable performance in an independent dataset. (**c**) Bootstrap calibration plot based on 1000 resamples. The apparent curve, bootstrap bias-corrected curve, and the ideal reference line (dashed gray line) are shown. The close alignment between the bias-corrected and ideal curves indicates strong calibration and minimal overfitting across the predicted probability range.

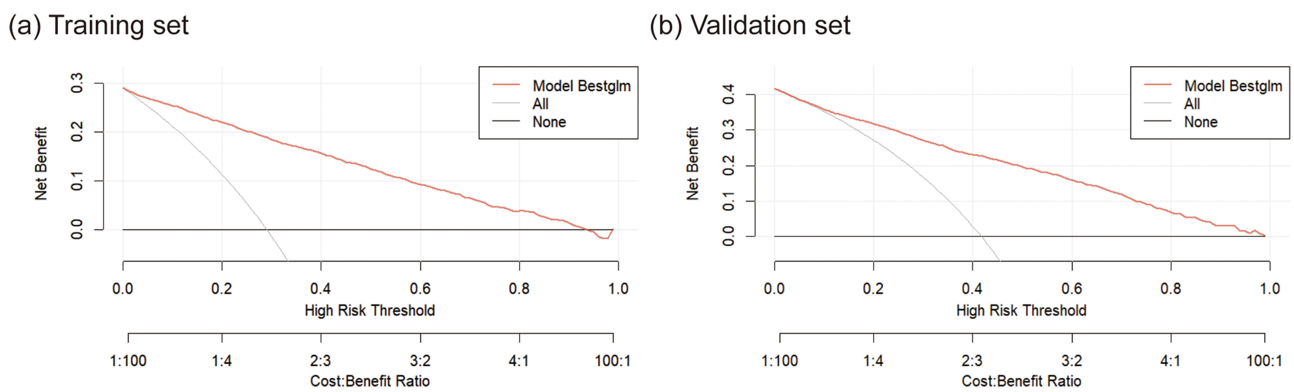

Figure 4 DCA for the logistic regression model. (**a**) DCA for the training set, indicating higher net benefits for threshold probabilities ranging from 0.1 to 0.7. (**b**) DCA for the validation set, showing consistent net benefits for threshold probabilities ranging from 0.15 to 0.6.

Decision Curve Analysis (DCA)

DCA revealed that the logistic regression model provided a higher net benefit across a broad spectrum of predicted probability thresholds, both in the training and validation cohorts. Specifically, in the training set, the model showed notable net benefits when the threshold probability varied between 0.1 to 0.7 (Figure 4a). Similarly, in the validation dataset, a consistent net benefit was evident within the probability threshold range of 0.15 to 0.6 (Figure 4b). These findings suggest that the model has significant potential for clinical application, as it provides meaningful net benefits for decision-making within these threshold ranges.

Calibration Curve Analysis

To evaluate the alignment between predicted probabilities and actual outcomes for hip osteoporosis, calibration curve analysis was performed in both the training and validation datasets.

In the training set (Figure 5a), the predicted probabilities closely aligned with the actual observed outcomes, especially within the mid-risk range (predicted probabilities of 0.2 to 0.6), indicating an excellent calibration performance in this range. Slight deviations were observed at the lower and higher risk extremes, but these did not substantially impact the overall model performance.

In the validation set (Figure 5b), the model demonstrated satisfactory calibration, with the predicted probabilities generally matching the observed outcomes. The alignment was most notable within the risk range of 0.2 to 0.5, where the model displayed good calibration. However, there were minor discrepancies in the lower risk segment (<0.2), suggesting slightly reduced predictive accuracy in this range.

Overall, the calibration curves indicate that the model has a reliable predictive capability, with optimal performance in the mid-risk ranges for both the training and validation datasets, highlighting its robustness and potential clinical value.

Discussion

In this study, we developed and validated a logistic regression model to predict hip osteoporosis using clinical and laboratory parameters. The model was trained on one dataset and validated on an independent cohort. Using LASSO regression, we identified four key predictors: age, sex, body mass index (BMI), and the serum albumin-to-alkaline phosphatase (ALB/ALP) ratio. These variables were subsequently incorporated into the final logistic regression model. The findings are consistent with previous research emphasizing the strong influence of age and sex on osteoporosis risk, with advancing age and female sex recognized as well-established risk factors. BMI was inversely associated with hip osteoporosis, indicating that individuals with lower BMI are at greater risk—a pattern supported by epidemiological studies linking low body weight to reduced bone mineral density.

A key innovation of this study is the incorporation of the ALB/ALP ratio as a predictive variable. Serum ALB is a well-established biomarker of nutritional and inflammatory status, and lower ALB levels have been linked to impaired osteoblast function, reduced bone matrix synthesis, and increased systemic inflammation—factors known to accelerate

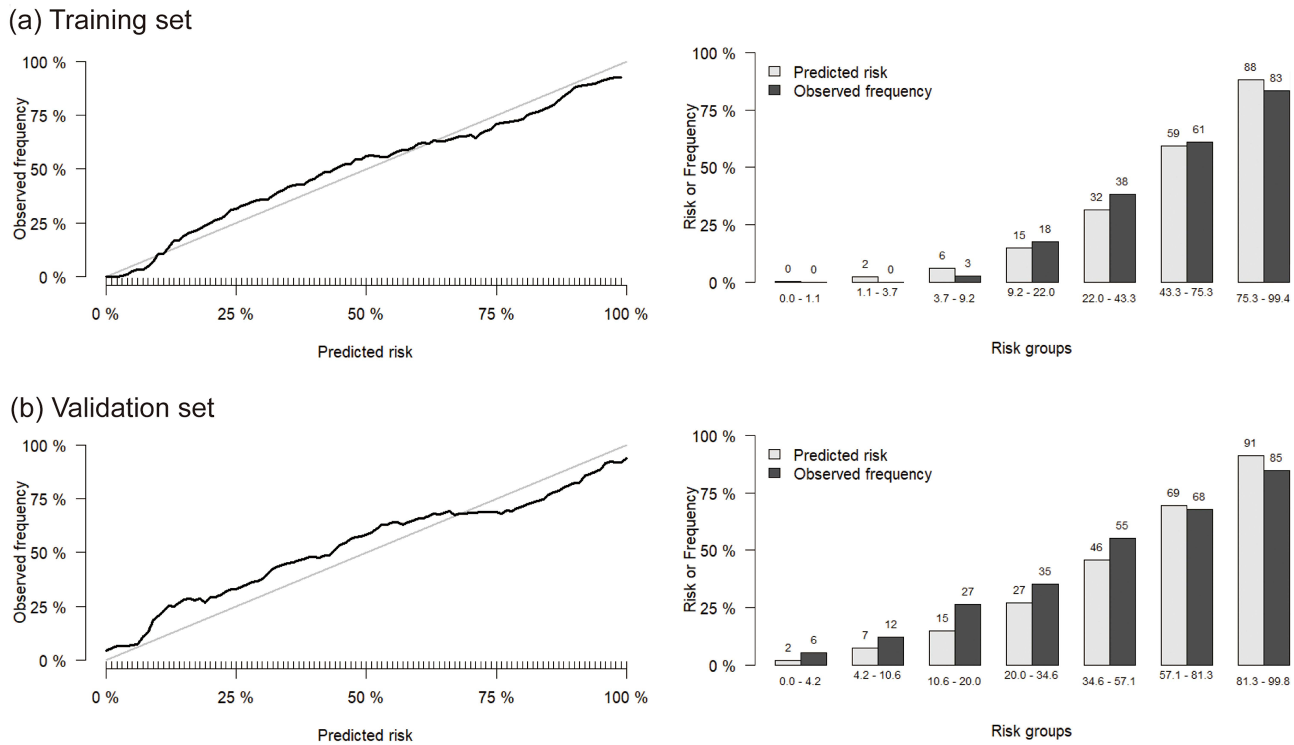

Figure 5 Calibration curves for the logistic regression model. (**a**) Calibration curve for the training set, showing excellent agreement between predicted probabilities and observed outcomes, particularly within the mid-risk range (0.2–0.6). Slight deviations are observed at the lower and higher risk extremes. (**b**) Calibration curve for the validation set, demonstrating satisfactory calibration. The model performs best within the risk range of 0.2–0.5, with minor discrepancies observed in the lower risk segment (<0.2).

bone loss. In contrast, ALP reflects metabolic bone turnover and is often elevated in conditions of increased osteoblastic activity or impaired mineralization. The ALB/ALP ratio therefore serves as an integrated biological indicator that captures both nutritional reserve and the degree of metabolic bone remodeling. A lower ratio may represent a physiologic phenotype characterized by malnutrition and dysregulated bone turnover, both of which contribute to heightened osteoporosis risk.

This ratio also distinguishes our model from widely used osteoporosis screening tools such as OSTA, COSA, and FRAX. OSTA and COSA rely primarily on demographic variables (age, sex, body weight), while FRAX requires multiple clinical risk factors and is designed to estimate fracture probability rather than bone density status.[8] In contrast, our model leverages simple laboratory tests that are universally available, inexpensive, and routinely performed during inpatient evaluations or health examinations. Although the increase in AUC was modest, it remained statistically significant based on DeLong's test, demonstrating the incremental value of the ALB/ALP ratio beyond conventional demographic predictors such as age, sex, and BMI. By grounding the model in both demographic and biochemical indicators, our approach offers a practical, low-burden alternative for early identification of individuals at risk of hip osteoporosis, particularly in settings where access to DXA scans is limited.

ROC curve analysis demonstrated that the model achieved strong discriminatory ability, with high AUC values in both the training and validation cohorts, indicating satisfactory predictive performance. To further assess model stability, we performed internal validation using 1000 bootstrap resamples, which is considered the most rigorous method when temporal validation is not feasible. The bootstrap-corrected estimates showed minimal optimism, with a corrected AUC of approximately 0.9104 and a calibration slope close to 1.0. The near-zero calibration intercept and very small maximum calibration error further confirmed that the model was not overfitted and remained well-calibrated across resampled datasets. These results indicate that the model is stable and reliable, supporting its robustness in different patient samples within similar clinical settings.

DCA further revealed that the model provided substantial net clinical benefit across a range of clinically relevant probability thresholds, underscoring its utility for osteoporosis risk stratification. Moreover, calibration curve analysis showed good concordance between predicted and observed outcomes, particularly within the moderate- to high-risk ranges, supporting the model's reliability and clinical applicability.

Several predictive models for osteoporosis have been proposed in previous studies, incorporating variables such as genetic markers,[22] and advanced imaging techniques.[23] However, many of these models rely on specialized tests that are not routinely available in everyday clinical practice.[24] While population-based samples are commonly used to develop general screening tools, our study deliberately focused on a hospital-based population. This approach is appropriate given the intended application of the model: to provide a practical and accessible tool for identifying the risk of hip osteoporosis among hospitalized patients or individuals undergoing routine health examinations. These individuals are already engaged with the healthcare system, making them ideal candidates for opportunistic screening using routinely collected demographic and laboratory data. The use of hospital-based data also reflects real-world clinical practice, where such a tool would be most effectively implemented. Moreover, hospital datasets often offer higher data completeness and accuracy, benefiting from standardized laboratory testing protocols and detailed clinical documentation—factors that enhance the reliability of model development. Although it is acknowledged that hospital populations may not fully represent the general population, the model was not designed for universal screening. Instead, it serves as a clinical decision-support tool tailored to healthcare settings, where more targeted identification of high-risk individuals is both feasible and beneficial.

To further strengthen the model's reliability and evaluate its generalizability, we performed external validation using data from a second medical center. This step is essential for assessing the model's performance in a different clinical setting, which may vary in terms of patient demographics, diagnostic protocols, and laboratory methodologies. The consistent performance observed across both the development and validation cohorts indicates that the model is robust and transferable across institutional contexts. External validation is a critical component in predictive modeling, as it demonstrates the model's adaptability and supports its potential for broader clinical adoption—particularly in real-world, hospital-based screening and risk stratification programs.

Our online calculator is designed for future in-hospital deployment: a "HIP-OP risk" widget incorporating four factors (age, sex, BMI, and ALB/ALP) embedded within the hospital information system, triggered upon completion of routine blood tests or prior to surgery. It provides risk probabilities via non-interruptive alerts and pre-filled order sets. However, full integration and regulatory review remain future steps, and the current standalone tool serves as a research prototype.

This study has several limitations that warrant consideration. First, the derivation and external-validation cohorts were both recruited from tertiary hospitals in northern Guangdong (latitudes 23–25°N, subtropical climate, almost Han Chinese). Consequently, the β-coefficients for age, BMI and ALP may be over-fitted to this unique environmental and genetic background, and the predicted probability could under-estimate hip-osteoporosis risk in populations at higher latitudes or with different dietary patterns. Second, participants were retrospectively enrolled from in-patient departments where DXA scanning is clinically indicated. Our model may therefore yield false-positive predictions when applied to healthier. Third, certain potentially relevant biochemical markers—such as serum vitamin D and parathyroid hormone levels—were not included due to missing or unavailable data, as these tests were ordered selectively rather than routinely, which may reduce the model's overall comprehensiveness. Future studies with more complete measurements may help further improve the model.

Despite these limitations, the proposed model shows considerable promise as a screening tool for identifying individuals at risk of hip osteoporosis in clinical settings. It may facilitate early detection and enable timely preventive interventions. Future studies should aim to validate the model in more diverse populations, including individuals undergoing routine health examinations and those from various geographic regions, to further evaluate its external applicability. Additionally, expanding the range of included biomarkers and applying advanced machine learning techniques may enhance predictive performance and reveal novel risk factors.

In conclusion, we developed a logistic regression model incorporating age, sex, BMI, and the ALB/ALP ratio to predict hip osteoporosis using readily available clinical information. The model demonstrated strong discriminatory ability, good calibration, and meaningful clinical utility in both the training and validation cohorts, supporting its potential value as a practical tool for hospital-based screening and risk stratification, particularly in settings where DXA access is limited.

To ensure broader applicability, further validation across multiple centers, geographic regions, and multi-ethnic populations will be essential. With adequate external confirmation, this model could be readily integrated into clinical workflows—for example, as an automated alert within hospital information system or as a decision-support tool in primary care—to facilitate early identification of high-risk individuals and improve the timeliness of osteoporosis evaluation and prevention.

Data Sharing Statement

No part of the data reported here has been previously published, presented, or submitted elsewhere. The raw data supporting the findings of this study are not publicly available due to ethical reasons but are available from the corresponding author on reasonable request.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was funded by Shaoguan City Science and Technology Bureau Shaoguan City Social Development Science and Technology Collaborative Innovation System Construction Project (High-level Hospital Construction Research Project) (220602184532348 and 230407098034659).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Cui L, Jackson M, Wessler Z, Gitlin M, Xia W. Estimating the future clinical and economic benefits of improving osteoporosis diagnosis and treatment among women in China: a simulation projection model from 2020 to 2040. *Archiv Osteoporosis*. 2021;16(1):118. doi:10.1007/s11657-021-00958-x
2. Yu F, Xia W. The epidemiology of osteoporosis, associated fragility fractures, and management gap in China. *Archiv Osteoporosis*. 2019;14(1):32. doi:10.1007/s11657-018-0549-y
3. Siu A, Allore H, Brown D, Charles ST, Lohman M. National institutes of health pathways to prevention workshop: research gaps for long-term drug therapies for osteoporotic fracture prevention. *Ann Intern Med*. 2019;171(1):51–57. doi:10.7326/M19-0961
4. Bliuc D, Alarkawi D, Nguyen TV, Eisman JA, Center JR. Risk of subsequent fractures and mortality in elderly women and men with fragility fractures with and without osteoporotic bone density: the Dubbo osteoporosis epidemiology study. *J Bone Miner Res*. 2015;30(4):637–646. doi:10.1002/jbmr.2393
5. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int*. 2006;17(12):1726–1733. doi:10.1007/s00198-006-0172-4
6. Maier GS, Kolbow K, Lazovic D, et al. Risk factors for pelvic insufficiency fractures and outcome after conservative therapy. *Arch Gerontol Geriatr*. 2016;67:80–85. doi:10.1016/j.archger.2016.06.020
7. Lagari VS, Al-Yatama F, Rodriguez G, Berger HR, Levis S. Under-recognition of fractures as osteoporosis indicators. *Geriatrics*. 2019;4(1). doi:10.3390/geriatrics4010009
8. Kanis JA, Hans D, Cooper C, et al. Interpretation and use of FRAX in clinical practice. *Osteoporos Int*. 2011;22(9):2395–2411. doi:10.1007/s00198-011-1713-z
9. Marshall D, Johnell O, Wedel H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *BMJ*. 1996;312(7041):1254–1259. doi:10.1136/bmj.312.7041.1254
10. Camacho PM, Petak SM, Binkley N, et al. American association of clinical endocrinologists/american college of endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis-2020 update. *Endocr Pract*. 2020;26(Suppl 1):1–46.
11. Melton LJ, Atkinson EJ, O'Connor MK, O'Fallon WM, Riggs BL. Bone density and fracture risk in men. *J Bone Miner Res*. 1998;13(12):1915–1923. doi:10.1359/jbmr.1998.13.12.1915

12. Cheung C-L, Li GH, Li H-L, Mak C, Tan KC, Kung AW. Development and validation of the Chinese osteoporosis screening algorithm (COSA) in identification of people with high risk of osteoporosis. *Osteoporos Sarcopenia*. 2023;9(1):8–13. doi:10.1016/j.afos.2023.03.009
13. Foundation IO. International osteoporosis foundation | IOF. 2025 Available from: http://www.iofbonehealth.org/. Accessed November 11, 2025.
14. Koh LK, Sedrine WB, Torralba TP, et al. A simple tool to identify asian women at increased risk of osteoporosis. *Osteoporos Int*. 2001;12 (8):699–705. doi:10.1007/s001980170070
15. Nayak S, Edwards DL, Saleh AA, Greenspan SL. Systematic review and meta-analysis of the performance of clinical risk assessment instruments for screening for osteoporosis or low bone density. *Osteoporos Int*. 2015;26(5):1543–1554. doi:10.1007/s00198-015-3025-1
16. Vasikaran S, Eastell R, Bruyère O, et al. Markers of bone turnover for the prediction of fracture risk and monitoring of osteoporosis treatment: a need for international reference standards. *Osteoporos Int*. 2010;22(2):391–420. doi:10.1007/s00198-010-1501-1
17. Aspray TJ, Bowring C, Fraser W, et al. National osteoporosis society vitamin D guideline summary. *Age Ageing*. 2014;43(5):592–595. doi:10.1093/ageing/afu093
18. Mateo J, Usategui-Martín R, Torres AM, et al. Improving prediction of fragility fractures in postmenopausal women using random forest. *Comput Biol Med*. 2025;196(Pt A):110666. doi:10.1016/j.compbiomed.2025.110666
19. Samelson EJ, Broe KE, Xu H, et al. Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the bone microarchitecture international consortium (BoMIC): a prospective study. *Lancet Diabetes Endocrinol*. 2018;7 (1):34–43. doi:10.1016/S2213-8587(18)30308-5
20. Kunutsor SK, Voutilainen A, Whitehouse MR, et al. Serum albumin and future risk of hip, humeral, and wrist fractures in caucasian men: new findings from a prospective cohort study. *Med Principles Pract*. 2019;28(5):401–409. doi:10.1159/000499738
21. Escobedo-Monge MF, Barrado E, Parodi-Román J, Escobedo-Monge MA, Marcos-Temprano M, Marugán-Miguelsanz JM. Magnesium status and calcium/magnesium ratios in a series of cystic fibrosis patients. *Nutrients*. 2022;14(9):1793. doi:10.3390/nu14091793
22. Sopova J, Krasnova O, Vasilieva G, et al. SNPs in GPCR genes and impaired osteogenic potency in osteoporotic patient lines-based study. *Int J Mol Sci*. 2024;25(24):13594. doi:10.3390/ijms252413594
23. Zhou K, Zhu Y, Luo X, et al. A novel hybrid deep learning framework based on biplanar X-ray radiography images for bone density prediction and classification. *Osteoporos Int*. 2025;36(3):521–530. doi:10.1007/s00198-024-07378-w
24. Amani F, Amanzadeh M, Hamedan M, Amani P. Diagnostic accuracy of deep learning in prediction of osteoporosis: a systematic review and meta-analysis. *BMC Musculoskelet Disord*. 2024;25(1):991. doi:10.1186/s12891-024-08120-7