

Preoperative prediction and perioperative risk updating for adjacent vertebral fracture after percutaneous vertebroplasty: development and internal validation of a two-stage prediction strategy

Linyang Song

Qilu Hospital of Shandong University

Yingnan Li

The Affiliated Hospital of Qingdao University

Xin Liu

Qilu Hospital of Shandong University

Hengfei Qiu

Qilu Hospital of Shandong University

Yanting Wang

Qilu Hospital of Shandong University

Zhiyong Ren

Qilu Hospital of Shandong University

Guangjun Jiao

jiaoguangjun@sdu.edu.cn

Qilu Hospital of Shandong University

Research Article

Keywords: Osteoporotic vertebral compression fracture, Percutaneous vertebroplasty, Adjacent vertebral fracture, Prediction model, Hounsfield unit, Paraspinal muscle, Sagittal imbalance, Bone cement distribution

Posted Date: October 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10976794/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Adjacent vertebral fracture (AVF) is an important complication after percutaneous vertebroplasty (PVP) for osteoporotic vertebral compression fracture (OVCF). Existing prediction models frequently combine variables obtained at different clinical time points, which may limit their applicability for preoperative risk assessment. This study aimed to develop and internally validate a strictly preoperative prediction model and to determine whether perioperative cement-distribution information provides additional prognostic value after PVP.

Methods

We retrospectively included 314 consecutive patients with OVCF who underwent PVP between December 2022 and December 2024 and completed at least 18 months of follow-up. The primary outcome was AVF within 18 months after PVP. Model 1 incorporated six preoperative predictors: age, target vertebral CT Hounsfield units (CT-HU), pre-existing vertebral compression fracture (VCF) count, paraspinal fat infiltration rate (FIR), preoperative sagittal imbalance, and preoperative anterior vertebral compression rate. Model 2 retained these predictors and added cement distribution pattern. Internal validation included 1,000 bootstrap resamples, 100 repetitions of stratified five-fold cross-validation, and fixed stratified 10-fold out-of-fold (OOF) prediction. Discrimination, calibration, Brier score, decision curve analysis, and incremental model performance were evaluated. LASSO and DEXA T-score analyses were performed as sensitivity analyses.

Results

Among 314 patients, 80 (25.5%) developed AVF within 18 months. In Model 1, older age, lower CT-HU, greater pre-existing VCF burden, higher paraspinal FIR, sagittal imbalance, and greater anterior vertebral compression were independently associated with AVF. The apparent AUC was 0.872, the bootstrap-corrected AUC was 0.857, and the 10-fold OOF AUC was 0.854. After addition of cement distribution, contiguous O-shaped and fragmented patterns were associated with higher AVF odds than bilateral H-shaped distribution. Model 2 improved overall model fit (likelihood-ratio $\chi^2=11.99$, $P = 0.0025$), but the increase in OOF discrimination was small (AUC 0.859; $\Delta\text{AUC} = 0.005$, 95% CI – 0.014 to 0.025; $P = 0.622$). Model 1 OOF risk strata showed observed AVF rates of 6.0%, 25.8%, and 63.2% in the low-, intermediate-, and high-risk groups, respectively. Adding DEXA T-score did not materially improve Model 1.

Conclusions

A six-predictor preoperative model integrating bone quality, previous fracture burden, paraspinal muscle degeneration, sagittal alignment, and vertebral morphology showed good internal performance for predicting 18-month AVF after PVP. Cement distribution provided additional perioperative prognostic information and improved overall model fit, but only minimally increased discrimination. These findings support a two-stage framework in which preoperative risk is estimated first and subsequently updated using procedural information. External validation is required before clinical implementation.

1. Introduction

Osteoporosis is a common systemic skeletal disease in older adults, and fragility fractures are associated with pain, functional impairment, reduced quality of life, and increased risks of subsequent fracture and mortality [1]. Osteoporotic vertebral compression fracture (OVCF) is one of the major manifestations of osteoporosis and may be underdiagnosed because some vertebral fractures initially present with few or nonspecific symptoms [2]. For patients with persistent pain or functional limitation despite conservative treatment, vertebral augmentation, including percutaneous vertebroplasty (PVP) and percutaneous kyphoplasty (PKP), is widely used to restore vertebral stability and relieve symptoms [3]. Nevertheless, new vertebral fractures remain an important postoperative event. Adjacent vertebral fracture (AVF), in particular, has received considerable attention because it may lead to recurrent pain, progressive deformity, and further intervention. A previous meta-analysis reported a pooled AVF incidence of approximately 14% after vertebral augmentation [4].

AVF is associated with multiple patient- and imaging-related factors. Previous studies have identified advanced age, poor bone quality, previous vertebral fracture, thoracolumbar location, and several cement-related characteristics as potential risk factors [4, 5]. Biomechanical studies have also shown that vertebral augmentation alters local stiffness and load transfer, which may influence mechanical stress in adjacent vertebrae [6]. Quantitative imaging markers may therefore provide useful information for risk assessment. Computed tomography Hounsfield units (CT-HU) reflect local vertebral bone quality and correlate with bone mineral density and mechanical strength [7]; lower vertebral HU values have also been associated with subsequent fracture after vertebral augmentation [8]. Beyond bone quality, paraspinal muscle degeneration has increasingly been linked to osteoporotic vertebral fracture and refracture, particularly through increased fat infiltration [9, 10]. Spinopelvic alignment may provide additional information, as sagittal imbalance has been associated with AVF after vertebroplasty [11]. Vertebral morphology is also readily available on routine imaging, and the degree of preoperative anterior vertebral compression may complement bone, muscle, and alignment parameters when estimating baseline fracture risk.

In addition to preoperative factors, cement distribution is an important perioperative imaging feature available immediately after vertebral augmentation. Previous studies have described different intravertebral cement distribution patterns, including H-shaped and O-shaped configurations, and have reported associations between cement morphology and postoperative radiographic or clinical outcomes [12]. The continuity, symmetry, and dispersion of cement within the treated vertebra may also be related

to vertebral stability and subsequent fracture [13]. A recent systematic review further showed that cement distribution patterns may influence outcomes after vertebral augmentation, while also highlighting substantial heterogeneity in classification methods across studies [14].

Several prediction models for AVF after vertebral augmentation have been reported. Hijikata et al. developed and internally validated the AVA score using preoperative variables to predict acute AVF [15], whereas Shen et al. developed a model incorporating perioperative imaging characteristics, including bone mineral density, paraspinal muscle fat infiltration, sagittal parameters, and vertebral height restoration [16]. However, a recent systematic review of 13 AVF prediction models found considerable heterogeneity in predictor selection and validation strategies; only two studies performed external validation, and all included studies were judged to have a high risk of bias according to PROBAST [17]. In addition, some existing models combine variables obtained before, during, and after surgery within a single prediction tool, which may obscure the intended prediction time point and limit their use for sequential clinical risk assessment.

Therefore, this study aimed to develop and internally validate a two-stage risk assessment strategy for 18-month AVF after vertebral augmentation. Model 1 was constructed using six preoperative predictors—age, target vertebral CT-HU, number of pre-existing vertebral compression fractures, paraspinal muscle fat infiltration rate, preoperative sagittal imbalance, and preoperative anterior vertebral compression rate—to estimate risk before treatment. Model 2 then incorporated cement distribution pattern to update risk after vertebral augmentation. We further evaluated model performance, the incremental value of perioperative updating, and model robustness using multiple internal validation and sensitivity analyses.

2. Methods

2.1 Study design, participants, and outcome

This single-center retrospective cohort study included consecutive patients with osteoporotic vertebral compression fracture (OVCF) who underwent percutaneous vertebroplasty (PVP) at the Qilu Hospital of Shandong University, between December 2022 and December 2024. All included patients completed at least 18 months of follow-up, and follow-up data were collected through June 2026. The study was reported with reference to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [18]. The study was approved by the institutional ethics committee (No. QLCR20260629), and the requirement for written informed consent was waived because of the retrospective design.

Patients were eligible if they underwent PVP for an OVCF, had a clearly identifiable index vertebra, had sufficient preoperative clinical and imaging data for study assessment, had postoperative imaging allowing assessment of cement-related characteristics, and completed at least 18 months of follow-up. Patients were excluded if follow-up was shorter than 18 months, relevant clinical or imaging data were incomplete, the vertebral fracture was pathological or caused by high-energy trauma, previous spinal

surgery or instrumentation had been performed, or multilevel/non-index fractures precluded the predefined index-level analysis.

A total of 436 patients were screened. Of these, 36 were excluded because of follow-up shorter than 18 months, 28 because of incomplete clinical or imaging data, 22 because of pathological or high-energy vertebral fractures, 17 because of previous spinal surgery or instrumentation, and 19 because of multilevel or non-index fractures unsuitable for the predefined analysis. The final cohort therefore comprised 314 patients. Patient selection and the overall two-stage modeling workflow are shown in Fig. 1.

The primary outcome was a new adjacent vertebral fracture (AVF) occurring within 18 months after PVP. AVF was defined as a new vertebral compression fracture involving the first cranial or caudal vertebra immediately adjacent to the treated level. New fractures were identified by comparing postoperative baseline imaging with follow-up radiographs or computed tomography (CT). Newly developed vertebral height loss, wedge-shaped or biconcave deformity, compression deformity, or new endplate or cortical abnormalities were considered indicative of a new vertebral fracture, with morphologic assessment referring to the semiquantitative principles proposed by Genant et al. [19]. When fracture acuity remained uncertain on radiography or CT, magnetic resonance imaging was used for further confirmation. Bone marrow edema characterized by low signal intensity on T1-weighted images and high signal intensity on T2-weighted or short-tau inversion recovery images was considered supportive of an acute or subacute fracture. Recollapse of the treated vertebra and new vertebral fractures at non-adjacent levels were not counted as primary outcome events.

2.2 PVP procedure and candidate predictors

All patients underwent PVP. Patients were placed in the prone position, and the target vertebral level and pedicle projection were identified using C-arm fluoroscopy. After routine sterile preparation and local infiltration anesthesia, transpedicular access to the fractured vertebra was established under alternating anteroposterior and lateral fluoroscopic guidance. The puncture needle was gradually advanced into the vertebral body while its position and trajectory were monitored fluoroscopically.

Polymethylmethacrylate (PMMA) bone cement was prepared and slowly injected through the working cannula under fluoroscopic monitoring. Cement injection was discontinued when satisfactory intravertebral filling was achieved or when cement extravasation was suspected. After adequate polymerization of the cement, the cannula was withdrawn and the puncture site was dressed. Postoperative radiography and/or CT was used to assess cement distribution and cement leakage.

Demographic, clinical, imaging, and perioperative information was retrieved from electronic medical records, operative records, and the picture archiving and communication system. Sex and the presence of a pre-existing VCF were additionally recorded as descriptive baseline characteristics.

The preoperative candidate predictor set comprised age, body mass index (BMI), dual-energy X-ray absorptiometry (DEXA) T-score, target vertebral CT Hounsfield units (CT-HU), number of pre-existing

vertebral compression fractures (VCFs), paraspinal muscle fat infiltration rate (FIR), thoracolumbar junction location, endplate fracture, preoperative sagittal imbalance, preoperative anti-osteoporosis medication, and preoperative anterior vertebral compression rate. Perioperative or immediate postoperative candidate variables included cement viscosity, cement volume fraction (CVF), cement distribution pattern, cement leakage, and postoperative bisegmental Cobb angle (BCA) correction.

Pre-existing VCF count was defined as the number of vertebral compression fractures already present on preoperative imaging other than the index vertebra. Because the binary presence of a pre-existing VCF and the corresponding fracture count contained overlapping information, VCF count was used as the modeling variable, whereas the binary variable was retained for descriptive analysis.

Definitions, timing, coding, and modeling roles of all candidate predictors are summarized in Supplementary Table S1.

2.3 Imaging measurements and reliability

The principal imaging measurements and representative cement distribution patterns are illustrated in Fig. 2.

Target vertebral CT-HU was measured on preoperative axial CT images. A region of interest (ROI) was placed within the trabecular portion of the target vertebral body and made as large as reasonably possible while avoiding cortical bone, endplates, focal sclerotic areas, bone islands, venous channels, and other structures that could distort attenuation measurements. Similar vertebral HU measurements have been shown to correlate with regional bone mineral density and vertebral mechanical strength [7].

Paraspinal muscle composition was assessed on a predefined axial image. The entire paraspinal muscle boundary was delineated as the total ROI, and threshold-based segmentation was used to identify effective muscle tissue within the ROI. Similar quantitative approaches have previously been used for assessment of paraspinal muscle degeneration and fat infiltration [10].

Muscle proportion was calculated as:

Muscle proportion = threshold-defined muscle area / total paraspinal ROI area

and FIR was calculated as:

$$\text{FIR (\%)} = [1 - \text{threshold-defined muscle area} / \text{total paraspinal ROI area}] \times 100$$

A higher FIR indicated greater paraspinal muscle fat infiltration.

Preoperative global sagittal alignment was evaluated on standing whole-spine lateral radiographs. Sagittal vertical axis (SVA) was defined as the horizontal distance between the C7 plumb line and the posterosuperior corner of the S1 endplate. Preoperative sagittal imbalance was defined as $\text{SVA} \geq 50$ mm, whereas $\text{SVA} < 50$ mm was classified as no sagittal imbalance. The relationship between spinopelvic

sagittal parameters and AVF after PVP has previously been reported [11]. Sagittal imbalance was entered into the models as a binary variable (no = 0; yes = 1).

Pelvic tilt (PT) was additionally measured to characterize pelvic compensation but was not used to define sagittal imbalance and was not included in either primary prediction model.

Preoperative anterior vertebral compression was measured on sagittal imaging. H_U , H_F , and H_L represented the anterior heights of the cranial adjacent vertebra, fractured index vertebra, and caudal adjacent vertebra, respectively. The mean anterior height of the two adjacent vertebrae was used as an internal reference.

The preoperative vertebral compression rate was calculated as:

$$\text{Preoperative vertebral compression rate (\%)} = [1 - H_F / ((H_U + H_L) / 2)] \times 100$$

or equivalently:

$$\text{Preoperative vertebral compression rate (\%)} = [1 - 2H_F / (H_U + H_L)] \times 100$$

A higher value represented greater anterior vertebral height loss.

Cement distribution was assessed on postoperative imaging and classified as bilateral H-shaped, contiguous O-shaped, or fragmented. Bilateral H-shaped distribution was characterized by predominant cement filling on both sides of the vertebral body with varying degrees of connecting cement bridge, whereas contiguous O-shaped distribution consisted of a relatively continuous and concentrated cement mass [12]. A fragmented pattern was operationally defined as multiple separated or discontinuous cement clusters without a distinct continuous bridge. Because cement morphology classifications remain heterogeneous across published studies [14], the same predefined criteria were consistently applied throughout the cohort.

Bilateral H-shaped distribution was used as the reference category in regression analyses, whereas contiguous O-shaped and fragmented distributions were represented by separate indicator variables.

Postoperative BCA correction was calculated as:

$$\text{BCA correction} = \text{preoperative BCA} - \text{postoperative BCA}$$

Endplate fracture and cement leakage were recorded as categorical variables, whereas CVF and BCA correction were retained as continuous variables in candidate analyses.

All major imaging measurements were performed by trained observers using the same predefined measurement protocol. Inter- and intra-observer reliability for continuous variables, including CT-HU, FIR, CVF, preoperative anterior vertebral compression rate, and BCA correction, was assessed using intraclass correlation coefficients (ICCs). Cohen's κ coefficients were used for categorical variables,

including sagittal imbalance, endplate fracture, cement distribution pattern, and cement leakage. For intra-observer assessment, measurements were repeated after an interval without access to the initial measurements.

The inter- and intra-observer reliability results are presented in Supplementary Table S2.

2.4 Model development, internal validation, and sensitivity analyses

Model 1 was developed to estimate AVF risk using information available before PVP. It comprised six preoperative predictors: age, target vertebral CT-HU, pre-existing VCF count, paraspinal FIR, preoperative sagittal imbalance, and preoperative anterior vertebral compression rate. The predictor set was defined on the basis of prior evidence, clinical availability, biological relevance, and avoidance of redundant information. All six predictors were entered simultaneously into a multivariable logistic regression model and were not selected or removed according to univariable P values. Continuous predictors were retained in continuous form rather than being arbitrarily dichotomized.

For clinical interpretability, ORs were expressed per 5-year increase in age, per 10-HU decrease in target vertebral CT-HU, per one additional pre-existing VCF, per 10-percentage-point increase in FIR, and per 10-percentage-point increase in preoperative anterior vertebral compression rate. Preoperative sagittal imbalance was analyzed as yes versus no.

Model 2 was developed as a perioperative risk-updating model. It retained all six Model 1 predictors and additionally incorporated cement distribution pattern. Bilateral H-shaped distribution served as the reference category, with contiguous O-shaped and fragmented distributions entered as dummy variables. Model 1 therefore represented preoperative risk estimation, whereas Model 2 updated risk using cement-distribution information available after PVP.

For both models, the individual predicted probability was calculated as:

$$P = \exp(LP) / [1 + \exp(LP)]$$

where LP represents the corresponding logistic linear predictor. Nomograms were generated from the fitted regression coefficients to facilitate graphical estimation of individual 18-month AVF probability. The complete regression coefficients and model equations are provided in Supplementary Table S3.

All 314 patients were retained for model development, and the cohort was not randomly divided into separate training and test sets. This approach maximized use of the available sample, while model optimism and internal stability were evaluated using resampling and cross-validation [20].

Model discrimination was quantified using the area under the receiver operating characteristic curve (AUC). Overall prediction error was assessed using the Brier score. Calibration was evaluated using the calibration intercept and calibration slope, with ideal values of 0 and 1, respectively.

Internal validation consisted of three complementary procedures. First, 1,000 bootstrap resamples were used to estimate optimism-corrected model performance. Second, repeated stratified five-fold cross-validation was performed with 100 repeats. Third, fixed stratified 10-fold out-of-fold (OOF) predictions were generated. For OOF prediction, each patient's probability was obtained from a model fitted without that patient in the corresponding training folds. The same 10-fold allocation was used for Model 1 and Model 2 to permit paired comparison.

The incremental value of Model 2 over Model 1 was assessed using the Akaike information criterion (AIC), likelihood-ratio testing, and the difference in OOF AUC (Δ AUC). Because predictions from the two models were paired within the same patients, 5,000 paired bootstrap resamples of the OOF predictions were used to estimate the 95% confidence interval and P value for Δ AUC.

Clinical utility was evaluated using decision curve analysis (DCA), comparing the net benefit of Model 1 and Model 2 with treat-all and treat-none strategies across threshold probabilities [21].

Least absolute shrinkage and selection operator (LASSO) logistic regression was performed as a sensitivity analysis rather than as the variable-selection procedure for the primary models. Candidate predictors were standardized before penalized regression, and categorical predictors were represented by indicator variables where appropriate. An L1 penalty was applied, and the penalty parameter λ was selected using 10-fold cross-validation within the glmnet framework [22]. Both the minimum-error value ($\lambda_{\min}$) and the more parsimonious one-standard-error value (λ_{1SE}) were examined.

Two LASSO analyses were conducted. The first included only preoperative candidate predictors, whereas the second included the complete set of preoperative and perioperative candidate variables. These analyses were used to assess the robustness of predictor signals and did not redefine Model 1 or Model 2.

Because DEXA T-score reflects systemic bone density while CT-HU provides a local vertebral bone-quality measure, an additional sensitivity model was fitted by adding DEXA T-score to Model 1. Incremental performance was assessed using apparent AUC, 10-fold OOF AUC, OOF Brier score, AIC, likelihood-ratio testing, and Δ AUC.

For exploratory risk stratification, Model 1 OOF predicted probabilities were categorized as low risk (< 15%), intermediate risk (15% to < 40%), or high risk ($\geq$ 40%). The same thresholds were applied to Model 2 for supplementary comparison. Observed AVF incidence, mean predicted probability, and 95% Wilson confidence intervals were calculated within each group. These thresholds were used for exploratory risk presentation and were not regarded as externally validated clinical decision thresholds.

2.5 Statistical analysis

Continuous variables are presented as mean $\pm$ standard deviation and were compared between the AVF and non-AVF groups using Welch's two-sample t test. Categorical variables are presented as counts and

percentages and were compared using the χ^2 test or Fisher's exact test, as appropriate according to expected cell frequencies. The final modeling dataset contained no missing values.

Multivariable logistic regression was used to model the binary 18-month AVF outcome. Regression effects are reported as odds ratios (ORs), 95% confidence intervals (CIs), and two-sided P values. A two-sided $P < 0.05$ was used as the conventional threshold for statistical significance.

All statistical analyses were performed using R version 4.5 (R Foundation for Statistical Computing, Vienna, Austria). Key packages included pROC for ROC-based analyses, glmnet for LASSO penalized regression, and rms for regression-model presentation. readxl, dplyr, tidyr, ggplot2, patchwork, and scales were used for data handling and graphical output.

3. Results

3.1 Study population and baseline characteristics

Among the 436 patients screened, 122 were excluded according to the predefined eligibility criteria, leaving 314 patients for the final analysis (Fig. 1). During the 18-month follow-up period, 80 patients developed an adjacent vertebral fracture (AVF), corresponding to an event rate of 25.5%, whereas 234 patients remained free of AVF.

Baseline characteristics according to AVF status are summarized in Table 1. Patients who developed AVF were older than those without AVF (81.31 ± 5.39 vs. 74.65 ± 6.01 years, $P < 0.001$). The AVF group also had a lower DEXA T-score (-2.92 ± 0.38 vs. -2.69 ± 0.45 , $P < 0.001$) and lower target vertebral CT-HU values (70.32 ± 15.55 vs. 86.39 ± 15.79 HU, $P < 0.001$). The number of pre-existing VCFs was higher in the AVF group (1.44 ± 1.17 vs. 0.65 ± 0.89 , $P < 0.001$), as was paraspinal FIR ($47.32 \pm 8.61\%$ vs. $37.02 \pm 10.88\%$, $P < 0.001$). Preoperative anterior vertebral compression was also greater in patients who subsequently developed AVF ($40.69 \pm 14.98\%$ vs. $33.29 \pm 14.63\%$, $P < 0.001$).

A pre-existing VCF was present in 76.2% of patients in the AVF group compared with 44.0% in the non-AVF group ($P < 0.001$). Cement distribution pattern also differed significantly between the two groups ($P = 0.017$). Bilateral H-shaped, contiguous O-shaped, and fragmented cement distributions accounted for 27.5%, 47.5%, and 25.0% of the AVF group, respectively, compared with 43.6%, 41.9%, and 14.5% of the non-AVF group.

No significant between-group differences were observed for BMI, sex, thoracolumbar junction location, endplate fracture, preoperative sagittal imbalance, preoperative anti-osteoporosis medication, cement viscosity, CVF, cement leakage, or postoperative BCA correction (all $P > 0.05$).

Table 1
Baseline characteristics according to 18-month adjacent vertebral fracture status.

Variable	No AVF (n = 234)	AVF (n = 80)	P value
Age, years	74.65 ± 6.01	81.31 ± 5.39	< 0.001
BMI, kg/m ²	23.78 ± 3.27	23.63 ± 3.64	0.744
DEXA T-score	-2.69 ± 0.45	-2.92 ± 0.38	< 0.001
Target vertebral CT-HU	86.39 ± 15.79	70.32 ± 15.55	< 0.001
Pre-existing VCF count	0.65 ± 0.89	1.44 ± 1.17	< 0.001
Paraspinal FIR, %	37.02 ± 10.88	47.32 ± 8.61	< 0.001
Preoperative anterior compression, %	33.29 ± 14.63	40.69 ± 14.98	< 0.001
CVF, %	17.91 ± 3.75	18.21 ± 3.36	0.512
Postoperative BCA correction, °	12.40 ± 2.60	12.74 ± 2.57	0.302
Gender			0.140
Female	128 (54.7%)	52 (65.0%)	
Male	106 (45.3%)	28 (35.0%)	
Pre-existing VCF			< 0.001
No	131 (56.0%)	19 (23.8%)	
Yes	103 (44.0%)	61 (76.2%)	
Thoracolumbar junction			1.000
No	57 (24.4%)	19 (23.8%)	
Yes	177 (75.6%)	61 (76.2%)	
Endplate fracture			0.526
No	71 (30.3%)	28 (35.0%)	
Yes	163 (69.7%)	52 (65.0%)	
Preoperative sagittal imbalance			0.106
Table note: Data are presented as mean ± standard deviation or n (%), as appropriate. Continuous variables were compared using Welch's two-sample t test. Categorical variables were compared using the χ^2 test or Fisher's exact test, as appropriate. Cement distribution pattern was compared as an overall three-category variable.			
Abbreviations: AVF, adjacent vertebral fracture; BMI, body mass index; DEXA, dual-energy X-ray absorptiometry; CT-HU, computed tomography Hounsfield unit; VCF, vertebral compression fracture; FIR, fat infiltration rate; CVF, cement volume fraction; BCA, bisegmental Cobb angle.			

Variable	No AVF (n = 234)	AVF (n = 80)	P value
No	123 (52.6%)	33 (41.2%)	
Yes	111 (47.4%)	47 (58.8%)	
Preoperative anti-osteoporosis medication			0.691
No	174 (74.4%)	57 (71.2%)	
Yes	60 (25.6%)	23 (28.7%)	
Cement viscosity			0.401
High	163 (69.7%)	51 (63.7%)	
Low	71 (30.3%)	29 (36.2%)	
Cement distribution pattern			0.017
Bilateral H	102 (43.6%)	22 (27.5%)	
Contiguous O	98 (41.9%)	38 (47.5%)	
Fragmented	34 (14.5%)	20 (25.0%)	
Cement leakage			0.731
No	179 (76.5%)	59 (73.8%)	
Yes	55 (23.5%)	21 (26.2%)	
Table note: Data are presented as mean $\pm$ standard deviation or n (%), as appropriate. Continuous variables were compared using Welch's two-sample t test. Categorical variables were compared using the χ^2 test or Fisher's exact test, as appropriate. Cement distribution pattern was compared as an overall three-category variable.			
Abbreviations: AVF, adjacent vertebral fracture; BMI, body mass index; DEXA, dual-energy X-ray absorptiometry; CT-HU, computed tomography Hounsfield unit; VCF, vertebral compression fracture; FIR, fat infiltration rate; CVF, cement volume fraction; BCA, bisegmental Cobb angle.			

3.2 Reliability of imaging measurements

The major imaging measurements showed good to excellent inter- and intra-observer reliability (Supplementary Table S2).

For continuous variables, inter-observer ICCs ranged from 0.871 to 0.935 and intra-observer ICCs ranged from 0.901 to 0.958. Specifically, the inter- and intra-observer ICCs were 0.935 and 0.958 for target vertebral CT-HU, 0.882 and 0.914 for paraspinal FIR, 0.890 and 0.925 for CVF, 0.885 and 0.918 for preoperative anterior vertebral compression, and 0.871 and 0.901 for postoperative BCA correction.

For categorical variables, inter-observer κ values ranged from 0.832 to 0.891 and intra-observer κ values ranged from 0.876 to 0.932. The corresponding inter- and intra-observer κ values were 0.845 and 0.884 for sagittal imbalance, 0.832 and 0.876 for endplate fracture, 0.856 and 0.895 for cement distribution pattern, and 0.891 and 0.932 for cement leakage. All reliability estimates were statistically significant ($P < 0.001$).

3.3 Development of the preoperative and perioperative models

The multivariable logistic regression results for Model 1 and Model 2 are presented in Table 2.

In the preoperative Model 1, all six predictors were associated with 18-month AVF after multivariable adjustment. Each 5-year increase in age was associated with higher odds of AVF (OR 1.60, 95% CI 1.01–2.54, $P = 0.045$). Each 10-HU decrease in target vertebral CT-HU was associated with an OR of 1.43 (95% CI 1.07–1.92, $P = 0.015$). Each additional pre-existing VCF was associated with an OR of 1.72 (95% CI 1.24–2.41, $P = 0.001$), whereas each 10-percentage-point increase in paraspinal FIR was associated with an OR of 1.65 (95% CI 1.10–2.47, $P = 0.015$).

Preoperative sagittal imbalance was also associated with increased AVF risk (OR 2.20, 95% CI 1.14–4.27, $P = 0.019$). Each 10-percentage-point increase in preoperative anterior vertebral compression was associated with an OR of 1.78 (95% CI 1.39–2.28, $P < 0.001$).

The linear predictor for Model 1 was:

$$LP_1 = -10.909481 + 0.094546 \times \text{Age} - 0.036113 \times \text{CT-HU} + 0.544726 \times \text{VCF count} + 0.049959 \times \text{FIR} + 0.789426 \times \text{Sagittal imbalance} + 0.057727 \times \text{Preoperative anterior compression}$$

The corresponding preoperative nomogram is shown in Fig. 3A.

After cement distribution pattern was added, the effect estimates of the six preoperative predictors remained similar in Model 2. The ORs were 1.63 per 5-year increase in age (95% CI 1.01–2.61, $P = 0.045$), 1.45 per 10-HU decrease in CT-HU (95% CI 1.08–1.96, $P = 0.015$), 1.90 per additional pre-existing VCF (95% CI 1.34–2.69, $P < 0.001$), 1.66 per 10-percentage-point increase in FIR (95% CI 1.09–2.53, $P = 0.018$), 2.37 for sagittal imbalance versus no imbalance (95% CI 1.20–4.68, $P = 0.013$), and 1.83 per 10-percentage-point increase in anterior vertebral compression (95% CI 1.41–2.36, $P < 0.001$).

Using bilateral H-shaped cement distribution as the reference, contiguous O-shaped distribution was associated with higher odds of AVF (OR 2.47, 95% CI 1.13–5.42, $P = 0.024$), while fragmented distribution showed the largest effect estimate (OR 5.27, 95% CI 1.94–14.28, $P = 0.001$).

The Model 2 linear predictor was:

$$LP_2 = -12.064233 + 0.097103 \times \text{Age} - 0.037170 \times \text{CT-HU} + 0.640246 \times \text{VCF count} + 0.050905 \times \text{FIR} + 0.863945 \times \text{Sagittal imbalance} + 0.060183 \times \text{Anterior compression} + 0.903653 \times \text{Contiguous O} +$$

The perioperative updated nomogram is shown in Fig. 3B. Full regression coefficients for both models are provided in Supplementary Table S3.

Observed AVF rates also differed across cement distribution patterns (Fig. S1). AVF occurred in 17.7% (22/124) of patients with bilateral H-shaped distribution, 27.9% (38/136) with contiguous O-shaped distribution, and 37.0% (20/54) with fragmented distribution (overall $P = 0.017$).

Table 2

Multivariable logistic regression analyses for 18-month adjacent vertebral fracture: Model 1 and Model 2.

Predictor	Model 1 OR (95% CI)	Model 1 P	Model 2 OR (95% CI)	Model 2 P
Age, per 5-year increase	1.60 (1.01–2.54)	0.045	1.63 (1.01–2.61)	0.045
Target CT-HU, per 10-HU decrease	1.43 (1.07–1.92)	0.015	1.45 (1.08–1.96)	0.015
Pre-existing VCF count, per 1 increase	1.72 (1.24–2.41)	0.001	1.90 (1.34–2.69)	< 0.001
Paraspinal FIR, per 10% increase	1.65 (1.10–2.47)	0.015	1.66 (1.09–2.53)	0.018
Preoperative sagittal imbalance, Yes vs No	2.20 (1.14–4.27)	0.019	2.37 (1.20–4.68)	0.013
Preoperative anterior compression, per 10% increase	1.78 (1.39–2.28)	< 0.001	1.83 (1.41–2.36)	< 0.001
Cement pattern: Contiguous O vs Bilateral H			2.47 (1.13–5.42)	0.024
Cement pattern: Fragmented vs Bilateral H			5.27 (1.94–14.28)	0.001
Table note: Model 1 included age, target vertebral CT-HU, pre-existing VCF count, paraspinal FIR, preoperative sagittal imbalance, and preoperative anterior vertebral compression rate. Model 2 retained all Model 1 predictors and additionally incorporated cement distribution pattern. Odds ratios are expressed per 5-year increase in age, per 10-HU decrease in CT-HU, per one additional pre-existing VCF, per 10-percentage-point increase in FIR, and per 10-percentage-point increase in preoperative anterior vertebral compression rate. Preoperative sagittal imbalance was coded as yes versus no. Bilateral H-shaped cement distribution served as the reference category.				
Abbreviations: OR, odds ratio; CI, confidence interval; CT-HU, computed tomography Hounsfield unit; VCF, vertebral compression fracture; FIR, fat infiltration rate.				

3.4 Model performance and incremental value of perioperative updating

Model performance is summarized in Table 3 and Fig. 4.

For Model 1, the apparent AUC was 0.872 (95% bootstrap CI 0.830–0.913). After 1,000 bootstrap resamples, the optimism-corrected AUC was 0.857. The mean AUC from 100 repeated five-fold cross-validations was 0.853, and the fixed 10-fold OOF AUC was 0.854 (95% bootstrap CI 0.805–0.896).

Model 2 showed an apparent AUC of 0.882 (95% bootstrap CI 0.839–0.919), a bootstrap optimism-corrected AUC of 0.863, a repeated five-fold cross-validation mean AUC of 0.861, and a 10-fold OOF AUC of 0.859 (95% bootstrap CI 0.809–0.900) (Fig. 4A).

The apparent, bootstrap-corrected, and OOF Brier scores were 0.120, 0.129, and 0.128 for Model 1, and 0.113, 0.124, and 0.124 for Model 2, respectively. OOF calibration intercepts were – 0.015 for Model 1 and – 0.005 for Model 2, while the corresponding calibration slopes were 0.866 and 0.845 (Fig. 4B).

Decision curve analysis showed that both models provided positive net benefit over relevant threshold probabilities compared with the treat-all and treat-none strategies, with substantial overlap between the two model curves (Fig. 4C).

The AIC decreased from 249.8 for Model 1 to 241.8 for Model 2. The likelihood-ratio test comparing the two nested models was significant ($\chi^2 = 11.99$, $df = 2$, $P = 0.0025$), indicating improved overall model fit after inclusion of cement distribution.

The difference in 10-fold OOF AUC between Model 2 and Model 1 was + 0.005. In the paired bootstrap analysis with 5,000 resamples, the estimated Δ AUC was + 0.005 (95% CI – 0.014 to + 0.025; $P = 0.622$).

Table 3

Model performance, internal validation, and incremental comparison of Model 1 and Model 2.

Metric	Model 1	Model 2
Apparent AUC (95% bootstrap CI)	0.872 (0.830–0.913)	0.882 (0.839–0.919)
Bootstrap-corrected AUC (1,000 resamples)	0.857	0.863
Repeated 5-fold CV mean AUC (100 repeats)	0.853	0.861
10-fold OOF AUC (95% bootstrap CI)	0.854 (0.805–0.896)	0.859 (0.809–0.900)
Apparent Brier score	0.120	0.113
Bootstrap-corrected Brier score	0.129	0.124
10-fold OOF Brier score	0.128	0.124
OOF calibration intercept	-0.015	-0.005
OOF calibration slope	0.866	0.845
AIC	249.8	241.8
Likelihood-ratio test: Model 2 vs Model 1	—	$\chi^2=11.99$, df = 2, P = 0.0025
OOF Δ AUC (Model 2 – Model 1)	—	+ 0.005
Paired bootstrap Δ AUC (OOF predictions)	—	+ 0.005 (-0.014 to + 0.025), P = 0.622
Table note: Apparent performance was calculated using the full cohort. Bootstrap optimism correction was based on 1,000 resamples. Repeated stratified five-fold cross-validation was performed with 100 repeats. Fixed stratified 10-fold out-of-fold predictions were used to estimate internally validated discrimination, calibration, and Brier scores. The same fold allocation was used for both models. Model 2 was compared with Model 1 using the Akaike information criterion and likelihood-ratio test. The difference in OOF AUC was additionally evaluated using 5,000 paired bootstrap resamples.		
Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; CV, cross-validation; OOF, out-of-fold; AIC, Akaike information criterion; Δ AUC, difference in AUC.		

3.5 Risk stratification and sensitivity analyses

Using Model 1 OOF predicted probabilities, 149 patients were classified as low risk (< 15%), 89 as intermediate risk (15% to < 40%), and 76 as high risk ($\geq$ 40%) (Fig. 5; Supplementary Table S4).

AVF occurred in 9 of 149 low-risk patients, corresponding to an observed event rate of 6.0% (95% Wilson CI 3.2–11.1%), with a mean predicted probability of 4.5%. In the intermediate-risk group, 23 of 89 patients developed AVF, corresponding to an observed rate of 25.8% (95% CI 17.9–35.8%) and a mean

predicted probability of 26.7%. In the high-risk group, AVF occurred in 48 of 76 patients, corresponding to an observed rate of 63.2% (95% CI 51.9–73.1%) and a mean predicted probability of 66.0%.

When the same thresholds were applied to Model 2, 155 patients were classified as low risk, 81 as intermediate risk, and 78 as high risk. The corresponding observed AVF rates were 7.7% (12/155), 21.0% (17/81), and 65.4% (51/78), respectively.

In the preoperative LASSO sensitivity analysis, five of the six Model 1 predictors were retained at λ_1 SE: age, CT-HU, pre-existing VCF count, paraspinal FIR, and preoperative anterior vertebral compression. Preoperative sagittal imbalance was not retained at this penalty level (Fig. S2; Supplementary Table S5).

When all preoperative and perioperative candidate variables were included, the λ_1 SE model retained age, CT-HU, pre-existing VCF count, paraspinal FIR, preoperative sagittal imbalance, preoperative anterior vertebral compression, and fragmented cement distribution (Fig. S3). Contiguous O-shaped cement distribution and the remaining candidate variables were not retained at λ_1 SE.

In the DEXA sensitivity analysis, adding DEXA T-score to Model 1 increased the apparent AUC from 0.872 to 0.876 and the 10-fold OOF AUC from 0.854 to 0.856. The OOF Brier score changed from 0.128 to 0.126. However, AIC increased from 249.8 to 250.2, and the likelihood-ratio test was not significant ($\chi^2 = 1.62$, $P = 0.204$). The DEXA coefficient was also not statistically significant in the extended model ($\beta = 0.601$, $P = 0.208$), and the increase in OOF AUC was + 0.003 (Supplementary Table S6).

4. Discussion

4.1 Principal findings

In this retrospective cohort of 314 patients undergoing PVP for OVCF, we developed and internally validated a two-stage strategy for predicting 18-month AVF. The preoperative Model 1 incorporated six variables available before treatment: age, target vertebral CT-HU, pre-existing VCF count, paraspinal FIR, preoperative sagittal imbalance, and preoperative anterior vertebral compression rate. The model showed good internal discrimination and acceptable calibration, with a 10-fold OOF AUC of 0.854. After PVP, cement distribution was added to generate Model 2. Both contiguous O-shaped and fragmented distributions were independently associated with AVF, with the strongest association observed for fragmented distribution. However, although Model 2 improved overall model fit, the increase in OOF AUC was small and not statistically significant. These findings support the concept of perioperative risk updating rather than replacement of the preoperative model by a procedurally enriched model.

The 18-month AVF incidence in the present cohort was 25.5%, which was higher than the pooled incidence of approximately 14% reported in a previous meta-analysis [4]. Direct comparison of absolute incidence across studies should nevertheless be made cautiously because of differences in patient age, osteoporosis severity, outcome definitions, follow-up duration, imaging surveillance, and whether asymptomatic radiographic fractures were included. Subsequent vertebral fractures after vertebroplasty

are not uncommon, and previous pooled evidence has also reported substantial post-procedural fracture rates [23]. The event rate observed here should therefore be interpreted within the context of the predefined 18-month adjacent-level outcome.

A notable finding was the strong contribution of pre-existing fracture burden. Rather than modeling previous VCF as a simple yes/no variable, we used the number of pre-existing VCFs, thereby retaining information on cumulative skeletal fragility. This is consistent with the established natural history of osteoporosis, in which a prior vertebral fracture markedly increases the risk of another fracture [24, 25]. The concept of imminent fracture risk further suggests that vulnerability is particularly high following a recent major osteoporotic fracture [26]. Pre-existing VCF count may therefore capture a graded fragility phenotype that is not fully represented by a single bone-density measurement.

Age also remained associated with AVF after adjustment for bone quality, fracture burden, muscle status, sagittal alignment, and vertebral morphology. Rather than representing a single causal pathway, age may integrate several dimensions of vulnerability, including cumulative skeletal deterioration, muscle loss, reduced postural compensation, and lower physiological reserve. Its persistence in both models supports its value as a simple but clinically informative prognostic marker.

4.2 Bone quality, vertebral morphology, and paraspinal muscle status

Lower target vertebral CT-HU was independently associated with AVF in both models. CT-HU can be obtained directly from routine preoperative CT and has established correlations with vertebral BMD and mechanical strength [7]. Opportunistic CT also provides bone-quality information without requiring an additional examination [27]. In patients undergoing vertebral augmentation, lower vertebral attenuation has been associated with subsequent vertebral fracture [8]. More recently, a systematic review and meta-analysis including 29 studies and 7,364 patients showed lower HU values in patients who sustained subsequent fractures after vertebral augmentation and reported clinically meaningful predictive performance for HU [28]. These findings are consistent with the present results and support local vertebral attenuation as a useful component of AVF risk assessment.

DEXA T-score differed significantly between patients with and without AVF at baseline, but its addition to Model 1 produced only minimal changes in OOF discrimination and did not improve overall model fit. This should not be interpreted as evidence against the clinical value of DEXA. DEXA remains central to osteoporosis diagnosis and long-term treatment planning. Rather, in a prediction model that already incorporates local CT-HU, previous fracture burden, muscle quality, sagittal alignment, and vertebral morphology, DEXA appears to provide limited additional prognostic information for this specific outcome. The sensitivity analysis therefore supports retaining the more parsimonious Model 1 while continuing to use DEXA for routine osteoporosis management.

Preoperative anterior vertebral compression rate was another independent predictor. Greater compression may reflect more severe structural failure at the index vertebra and greater alteration of

local vertebral geometry before PVP. Previous studies have not consistently identified fracture morphology as an independent predictor of AVF [29], possibly because of differences in measurement methods, categorical versus continuous treatment of vertebral collapse, and covariate adjustment. The present findings therefore support anterior compression as a prognostic marker within the current multivariable framework, but do not establish vertebral compression itself as a causal determinant of subsequent fracture.

Paraspinal FIR also contributed independently to prediction. Increasing evidence suggests that vertebral fragility is not determined by bone quality alone. A previous meta-analysis reported greater paraspinal muscle fat infiltration in patients with osteoporotic vertebral fracture and refracture [9], while studies in PVP populations have linked paraspinal muscle degeneration with adjacent vertebral refracture [10]. Other investigations have associated fatty degeneration with progressive vertebral collapse [30] and incorporated muscle characteristics into prediction models for subsequent vertebral fracture [31, 32].

Taken together, these findings support a broader bone–muscle fragility phenotype. Greater FIR may reflect reduced muscle quality, impaired dynamic spinal stabilization, lower physical reserve, and altered load sharing. However, FIR should be interpreted as a prognostic imaging biomarker rather than a proven causal factor. Its consistent direction of association in the primary models and its retention in LASSO sensitivity analyses support prognostic relevance, but whether improvement in muscle status reduces AVF remains unknown.

4.3 Sagittal alignment and cement distribution

Preoperative sagittal imbalance provided additional prognostic information after multivariable adjustment. Although the unadjusted baseline comparison was not statistically significant, sagittal imbalance became significant after simultaneous adjustment for age, CT-HU, VCF burden, FIR, and anterior compression. These findings are not necessarily inconsistent because multivariable coefficients reflect the conditional contribution of a predictor after accounting for correlations with other variables.

Baek et al. previously reported an association between spinopelvic balance and adjacent fracture after PVP [11]. More broadly, positive sagittal imbalance has been associated with poorer functional status in adult spinal deformity [33], and SVA is incorporated as an important modifier in the SRS-Schwab classification [34]. In OVCF, vertebral height loss and local kyphosis may shift the trunk anteriorly, while older patients frequently rely on pelvic and lower-extremity compensation to maintain balance. SVA may therefore capture aspects of the global mechanical environment that are not reflected by the index vertebra alone.

The preoperative LASSO analysis did not retain sagittal imbalance at λ_1 SE, whereas it was retained in the all-candidate LASSO analysis. This finding highlights why penalized selection should be interpreted as a sensitivity analysis rather than a definitive test of predictor validity. In addition, dichotomizing SVA at 50 mm improves clinical interpretability but sacrifices continuous information. Future external validation should evaluate continuous SVA and other spinopelvic parameters, including PT and PI–LL.

Cement distribution was the key perioperative variable used to update risk after PVP. Compared with bilateral H-shaped distribution, both contiguous O-shaped and fragmented distributions were associated with greater AVF odds, and observed AVF incidence increased progressively across these three patterns.

Previous studies have suggested that intravertebral cement morphology may affect postoperative outcomes. H-shaped and O-shaped patterns have been associated with different clinical and radiographic results [12], while more diffuse and symmetric cement filling has been linked to more favorable outcomes [13]. A recent systematic review similarly concluded that bilateral, relatively even cement distribution and adequate endplate contact may be associated with better outcomes, while emphasizing substantial heterogeneity in cement classification systems [14].

From a biomechanical perspective, cement distribution may influence how stiffness and load are redistributed within the treated vertebra and adjacent motion segments. Experimental studies have shown that vertebroplasty can alter compression and strain in adjacent discs and vertebral bodies [35]. However, clinical evidence regarding individual cement-related factors remains inconsistent. Intradiscal leakage, for example, has been associated with AVF in some studies and meta-analyses [4], whereas other work found no significant relationship [36]. This inconsistency suggests that a simple leakage/no-leakage variable may not fully characterize the mechanical consequences of cement augmentation.

Importantly, the relatively large ORs associated with cement distribution did not translate into a similarly large improvement in discrimination. Although Model 2 showed better likelihood-based model fit, the OOF AUC increased by only 0.005, and the paired-bootstrap confidence interval for Δ AUC included zero. A predictor may show a strong independent association with the outcome while contributing little additional rank-order discrimination if its information partly overlaps with predictors already present in the base model. Accordingly, Model 2 should be interpreted as a perioperative risk-updating model, not as a substantially superior replacement for Model 1.

4.4 Two-stage prediction and potential clinical use

Several AVF prediction tools have previously been proposed. The AVA score was developed using information available before vertebral augmentation [15], whereas other models have incorporated perioperative imaging variables such as BMD, paraspinal muscle fat infiltration, sagittal parameters, and vertebral height restoration [16]. A recent systematic review identified 13 AVF prediction models, with reported AUCs ranging from 0.72 to 0.988; however, only two studies included external validation, and all were judged to have a high risk of bias according to PROBAST [17].

The central methodological distinction of the present study is the explicit separation of prediction time points. Model 1 addresses the preoperative question of how high a patient's AVF risk is before PVP. Model 2 addresses a subsequent question: whether newly available procedural information changes that risk. This separation avoids incorporating post-procedural information into a nominally preoperative model and provides a clearer intended-use setting for each model.

The validation strategy was also designed to use the available sample efficiently. Rather than performing a single random training-test split, the full cohort was retained for model development, with bootstrap optimism correction, repeated cross-validation, and fixed OOF prediction used for internal validation. Prediction performance was assessed using discrimination, calibration, overall prediction error, and DCA rather than AUC alone [37]. Calibration is particularly relevant because satisfactory discrimination does not guarantee accurate absolute risk estimation [38].

Model 1 showed stable internal discrimination and an OOF calibration intercept close to zero, although the calibration slope below 1 suggests some residual optimism. Model 2 showed similar overall internal performance, but only a small gain in discrimination. These results further support the distinction between statistically significant predictor effects and clinically meaningful improvement in prediction.

The exploratory risk stratification provides an intuitive illustration of potential clinical use. Observed 18-month AVF incidence increased from 6.0% in the low-risk group to 25.8% in the intermediate-risk group and 63.2% in the high-risk group. Such separation may be useful for patient counseling, follow-up planning, and selection of populations for future prospective studies. Nevertheless, the < 15%, 15%–<40%, and $\geq 40\%$ thresholds were derived for exploratory presentation and should not be interpreted as validated treatment thresholds.

From a clinical perspective, Model 1 may complement routine preoperative assessment by integrating several dimensions of vulnerability rather than relying on a single bone-density measurement. Patients with advanced age, low CT-HU, multiple previous VCFs, greater FIR, sagittal imbalance, and marked anterior vertebral compression may represent a particularly fragile phenotype. After PVP, cement distribution may further refine that baseline estimate.

These models should complement, rather than replace, established osteoporosis management. The 2024 ASBMR/BHOF position statement emphasizes individualized assessment based on recent fracture, fracture number and severity, BMD, and other clinical risk factors [39]. The Clinician's Guide similarly highlights the importance of secondary fracture prevention after osteoporotic fracture, including pharmacological treatment, calcium and vitamin D optimization, physical activity, and fall prevention where appropriate [40]. The present model therefore provides a risk-prediction framework rather than a substitute for comprehensive osteoporosis care.

4.5 Strengths and limitations

Several aspects of the present study strengthen its methodological framework. First, the prediction time point was explicitly defined, allowing a clear distinction between preoperative prediction and postoperative risk updating. Second, predictors represented complementary domains of skeletal fragility, including local bone quality, previous fracture burden, muscle degeneration, global sagittal mechanics, and vertebral morphology. Third, continuous predictors were retained in continuous form, and the main models were not determined by univariable significance or LASSO selection alone. Finally, internal

performance was evaluated using multiple complementary approaches, including bootstrap correction, repeated cross-validation, OOF prediction, calibration, Brier score, and DCA.

Several limitations should also be acknowledged. First, this was a retrospective single-center study. Patient selection, treatment practice, imaging protocols, and follow-up patterns may therefore limit generalizability. All patients underwent PVP, and the present model should not be directly extrapolated to PKP populations.

Second, although several internal validation methods were used, independent external validation was not performed. Internal resampling can quantify optimism within the development cohort but cannot fully account for differences in case mix, imaging equipment, measurement protocols, or clinical practice across institutions. This remains particularly important given the high risk of bias reported for many existing AVF prediction studies [17, 41].

Third, the outcome was defined as a binary AVF event within 18 months. This provides a clinically clear prediction horizon but does not incorporate the exact timing of fracture or competing risks. Future studies with sufficiently detailed longitudinal data may benefit from time-to-event modeling.

Fourth, fragmented cement distribution was based on a study-specific operational imaging definition. Although measurement reproducibility was good, cement distribution classifications remain heterogeneous across studies [14]. External application will therefore require either reproducibility of the current definitions or development of more objective quantitative descriptors.

Fifth, SVA was dichotomized at 50 mm. Although this facilitates clinical interpretation, it reduces information compared with continuous modeling and does not capture the full spectrum of spinopelvic compensation. Similarly, FIR is an imaging surrogate for muscle composition and does not directly measure muscle strength or physical function.

Sixth, several potentially relevant factors were not available, including long-term adherence to anti-osteoporosis medication, vitamin D status, bone-turnover markers, fall risk, physical activity, frailty, and some comorbidities. Their inclusion may further refine future models.

Finally, this was a prediction study rather than a causal analysis. Associations between CT-HU, FIR, sagittal imbalance, vertebral compression, cement distribution, and AVF should therefore not be interpreted as evidence that directly modifying any single variable would necessarily reduce fracture risk. Prospective multicenter validation is required before the models or exploratory thresholds are used for clinical decision-making.

5. Conclusions

A preoperative prediction model integrating age, target vertebral CT-HU, pre-existing VCF burden, paraspinal FIR, sagittal imbalance, and anterior vertebral compression showed good internal discrimination and acceptable calibration for predicting 18-month AVF after PVP. Incorporating cement

distribution after PVP improved overall model fit and identified additional perioperative risk information, particularly for fragmented cement distribution, but resulted in only a small and non-significant increase in discrimination. These findings support a two-stage approach that separates preoperative risk estimation from perioperative risk updating rather than combining variables from different prediction time points in a single model. The proposed models and exploratory risk thresholds require prospective multicenter external validation before use in clinical decision-making.

Abbreviations

AIC	
Akaike information criterion	
ASBMR	
American Society for Bone and Mineral Research	
AUC	
Area under the receiver operating characteristic curve	
AVF	
Adjacent vertebral fracture	
BCA	
Bisegmental Cobb angle	
BHOF	
Bone Health and Osteoporosis Foundation	
BMD	
Bone mineral density	
BMI	
Body mass index	
CI	
Confidence interval	
CT	
Computed tomography	
CT-HU	
Computed tomography Hounsfield unit	
CVF	
Cement volume fraction	
DCA	
Decision curve analysis	
DEXA	
Dual-energy X-ray absorptiometry	
FIR	
Fat infiltration rate	
HU	

Hounsfield unit
ICC
Intraclass correlation coefficient
LASSO
Least absolute shrinkage and selection operator
MRI
Magnetic resonance imaging
OOF
Out-of-fold
OR
Odds ratio
OVCF
Osteoporotic vertebral compression fracture
PI–LL
Pelvic incidence–lumbar lordosis mismatch
PKP
Percutaneous kyphoplasty
PMMA
Polymethylmethacrylate
PROBAST
Prediction model Risk Of Bias ASsessment Tool
PT
Pelvic tilt
PVP
Percutaneous vertebroplasty
ROI
Region of interest
STIR
Short-tau inversion recovery
SVA
Sagittal vertical axis
TRIPOD
Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
VCF
Vertebral compression fracture

Declarations

Ethics approval and consent to participate

This retrospective study was approved by the Ethics Committee of Qilu Hospital of Shandong University (Approval No. QLCR20260629). The requirement for written informed consent was waived by the ethics committee because of the retrospective design. All procedures were performed in accordance with the Declaration of Helsinki and applicable institutional ethical standards.

Clinical trial number

not applicable.

Consent for publication

Not applicable. All clinical images included in the manuscript were fully anonymized and contain no identifiable personal information.

Competing interests

The authors declare no competing interests.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available because they contain clinical information that may be potentially identifiable, but are available from the corresponding author on reasonable request and subject to applicable institutional requirements.

Funding

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

L.S. and G.J. conceived and designed the study. L.S. performed the statistical analyses and drafted the manuscript. Y.L., X.L., and H.Q. contributed to clinical data acquisition and imaging assessment. Y.W. and Z.R. contributed to data management, figure preparation, and manuscript organization. G.J. supervised the study and critically revised the manuscript. All authors contributed to the interpretation of the results, reviewed the manuscript, and approved the final version.

References

1. Johnston CB, Dagar M. Osteoporosis in older adults. *Med Clin North Am.* 2020;104(5):873–84. <https://doi.org/10.1016/j.mcna.2020.06.004>.
2. Schousboe JT. Epidemiology of vertebral fractures. *J Clin Densitom.* 2016;19(1):8–22. <https://doi.org/10.1016/j.jocd.2015.08.004>.
3. Prost S, Pesenti S, Fuentes S, et al. Treatment of osteoporotic vertebral fractures. *Orthop Traumatol Surg Res.* 2021;107(1 Suppl):102779. <https://doi.org/10.1016/j.otsr.2020.102779>.
4. Zhang T, Wang Y, Zhang P, et al. What are the risk factors for adjacent vertebral fracture after vertebral augmentation? A meta-analysis of published studies. *Global Spine J.* 2022;12(1):130–41. <https://doi.org/10.1177/2192568220978223>.
5. Ma XL, Xing D, Ma JX, et al. Risk factors for new vertebral compression fractures after percutaneous vertebroplasty: qualitative evidence synthesized from a systematic review. *Spine (Phila Pa 1976).* 2013;38(12):E713–22. <https://doi.org/10.1097/BRS.0b013e31828cf15b>.
6. Berlemann U, Ferguson SJ, Nolte LP, et al. Adjacent vertebral failure after vertebroplasty: a biomechanical investigation. *J Bone Joint Surg Br.* 2002;84(5):748–52. <https://doi.org/10.1302/0301-620X.84B5.11841>.
7. Schreiber JJ, Anderson PA, Rosas HG, et al. Hounsfield units for assessing bone mineral density and strength: a tool for osteoporosis management. *J Bone Joint Surg Am.* 2011;93(11):1057–63. <https://doi.org/10.2106/JBJS.J.00160>.
8. Seuvic FM, Gillet R, Louis M, et al. Association between opportunistic vertebral bone density measurements and new vertebral fractures after percutaneous vertebral cementoplasty: a case-control study. *Eur Radiol.* 2023;33(1):106–15. <https://doi.org/10.1007/s00330-022-08946-6>.
9. Chen Z, Shi T, Li W, et al. Role of paraspinal muscle degeneration in the occurrence and recurrence of osteoporotic vertebral fracture: a meta-analysis. *Front Endocrinol (Lausanne).* 2023;13:1073013. <https://doi.org/10.3389/fendo.2022.1073013>.
10. Chen M, Yang C, Cai Z, et al. Lumbar posterior group muscle degeneration: influencing factors of adjacent vertebral body re-fracture after percutaneous vertebroplasty. *Front Med (Lausanne).* 2023;9:1078403. <https://doi.org/10.3389/fmed.2022.1078403>.
11. Baek SW, Kim C, Chang H. The relationship between the spinopelvic balance and the incidence of adjacent vertebral fractures following percutaneous vertebroplasty. *Osteoporos Int.* 2015;26(5):1507–13. <https://doi.org/10.1007/s00198-014-3021-x>.
12. He S, Zhang Y, Lv N, et al. The effect of bone cement distribution on clinical efficacy after percutaneous kyphoplasty for osteoporotic vertebral compression fractures. *Med (Baltim).* 2019;98(50):e18217. <https://doi.org/10.1097/MD.00000000000018217>.
13. Pan Z, Zhou Q, Yang M, et al. Effects of distribution of bone cement on clinical efficacy and secondary fracture after percutaneous kyphoplasty for osteoporotic vertebral compression fractures. *Front Surg.* 2022;9:1054995. <https://doi.org/10.3389/fsurg.2022.1054995>.
14. Yang K, Zhu X, Sun X, et al. Bone cement distribution patterns in vertebral augmentation for osteoporotic vertebral compression fractures: a systematic review. *J Orthop Surg Res.*

- 2025;20(1):568. <https://doi.org/10.1186/s13018-025-05868-z>.
15. Hijikata Y, Kamitani T, Nakahara M, et al. Development and internal validation of a clinical prediction model for acute adjacent vertebral fracture after vertebral augmentation: the AVA score. *Bone Joint J.* 2022;104–B(1):97–102. <https://doi.org/10.1302/0301-620X.104B1.BJJ-2021-0151.R3>.
 16. Shen S, You X, Ren Y, et al. Adjacent vertebral refracture prediction model based on imaging data after vertebroplasty for osteoporotic vertebral compression fracture. *World Neurosurg.* 2024;190:e548–53. <https://doi.org/10.1016/j.wneu.2024.07.169>.
 17. Sun D, Wen Y, Yu Q, et al. Prediction models for adjacent vertebral fractures after vertebral augmentation: a systematic review and meta-analysis. *Eur Spine J.* 2025;34(5):1631–40. <https://doi.org/10.1007/s00586-025-08785-1>.
 18. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015;350:g7594. <https://doi.org/10.1136/bmj.g7594>.
 19. Genant HK, Wu CY, van Kuijk C, et al. Vertebral fracture assessment using a semiquantitative technique. *J Bone Min Res.* 1993;8(9):1137–48. <https://doi.org/10.1002/jbmr.5650080915>.
 20. Steyerberg EW, Harrell FE Jr, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774–81. [https://doi.org/10.1016/S0895-4356\(01\)00341-9](https://doi.org/10.1016/S0895-4356(01)00341-9).
 21. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak.* 2006;26(6):565–74. <https://doi.org/10.1177/0272989X06295361>.
 22. Friedman JH, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010;33(1):1–22. <https://doi.org/10.18637/jss.v033.i01>.
 23. Ding JK, Zhao B, Zhai YF. Subsequent fractures after vertebroplasty in osteoporotic vertebral fractures: a meta-analysis. *Neurosurg Rev.* 2022;45(3):2349–59. <https://doi.org/10.1007/s10143-022-01755-x>.
 24. Lindsay R, Silverman SL, Cooper C, et al. Risk of new vertebral fracture in the year following a fracture. *JAMA.* 2001;285(3):320–3. <https://doi.org/10.1001/jama.285.3.320>.
 25. Kanis JA, Johnell O, De Laet C, et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone.* 2004;35(2):375–82. <https://doi.org/10.1016/j.bone.2004.03.024>.
 26. Johansson H, Siggeirsdóttir K, Harvey NC, et al. Imminent risk of fracture after fracture. *Osteoporos Int.* 2017;28(3):775–80. <https://doi.org/10.1007/s00198-016-3868-0>.
 27. Pickhardt PJ, Pooler BD, Lauder T, et al. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. *Ann Intern Med.* 2013;158(8):588–95. <https://doi.org/10.7326/0003-4819-158-8-201304160-00003>.
 28. Wang S, Liu D, Kang T, et al. Predictive performance of VBQ scores and Hounsfield units for subsequent fractures after vertebral augmentation: a systematic review and meta-analysis. *Acad Radiol.* 2026;33(5):2141–53. <https://doi.org/10.1016/j.acra.2026.02.005>.

29. Tang B, Liu L, Cui L, et al. Analysis of adjacent vertebral fracture after percutaneous vertebroplasty: do radiological or surgical features matter? *Eur Spine J.* 2024;33:1524–32. <https://doi.org/10.1007/s00586-023-08092-7>.
30. Jeon I, Kim SW, Yu D. Paraspinal muscle fatty degeneration as a predictor of progressive vertebral collapse in osteoporotic vertebral compression fractures. *Spine J.* 2022;22(2):313–20. <https://doi.org/10.1016/j.spinee.2021.07.020>.
31. Xiong Y, Zhang C, Chen X, et al. Prediction of subsequent vertebral fracture after acute osteoporotic fractures from clinical and paraspinal muscle features. *Calcif Tissue Int.* 2024;114(6):614–24. <https://doi.org/10.1007/s00223-024-01209-0>.
32. Zhuo C, Huang M, Wang G, et al. Paraspinal muscle fat infiltration predicts vertebral refracture after kyphoplasty: development and validation of a novel nomogram model. *World Neurosurg.* 2025;201:124325. <https://doi.org/10.1016/j.wneu.2025.124325>.
33. Glassman SD, Bridwell K, Dimar JR, et al. The impact of positive sagittal balance in adult spinal deformity. *Spine (Phila Pa 1976).* 2005;30(18):2024–9. <https://doi.org/10.1097/01.brs.0000179086.30449.96>.
34. Schwab F, Ungar B, Blondel B, et al. Scoliosis Research Society-Schwab adult spinal deformity classification: a validation study. *Spine (Phila Pa 1976).* 2012;37(12):1077–82. <https://doi.org/10.1097/BRS.0b013e31823e15e2>.
35. Nagaraja S, Awada HK, Dreher ML, et al. Vertebroplasty increases compression of adjacent IVDs and vertebrae in osteoporotic spines. *Spine J.* 2013;13(12):1872–80. <https://doi.org/10.1016/j.spinee.2013.06.007>.
36. Lee KA, Hong SJ, Lee S, et al. Analysis of adjacent fracture after percutaneous vertebroplasty: does intradiscal cement leakage really increase the risk of adjacent vertebral fracture? *Skeletal Radiol.* 2011;40(12):1537–42. <https://doi.org/10.1007/s00256-011-1139-x>.
37. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21(1):128–38. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>.
38. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17(1):230. <https://doi.org/10.1186/s12916-019-1466-7>.
39. Cosman F, Lewiecki EM, Eastell R, et al. Goal-directed osteoporosis treatment: ASBMR/BHOF task force position statement 2024. *J Bone Min Res.* 2024;39(10):1393–405. <https://doi.org/10.1093/jbmr/zjae119>.
40. LeBoff MS, Greenspan SL, Insogna KL, et al. The clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int.* 2022;33(10):2049–102. <https://doi.org/10.1007/s00198-021-05900-y>.
41. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med.* 2019;170(1):51–8. <https://doi.org/10.7326/M18-1376>.

Figures

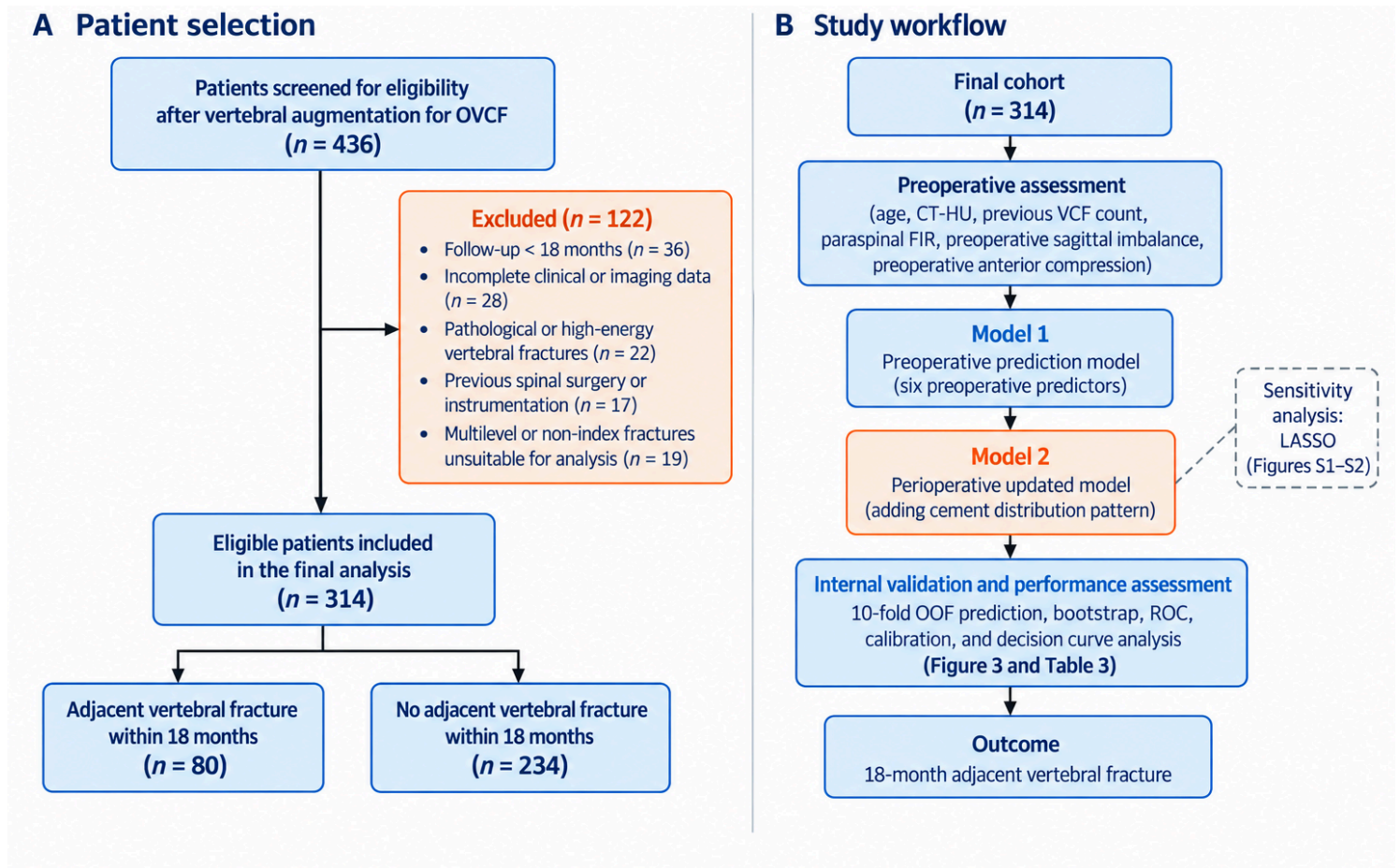


Figure 1

Patient selection and overall two-stage study workflow.

Legend: (A) Patient selection. A total of 436 patients were screened. Among them, 122 were excluded because of follow-up shorter than 18 months (n=36), incomplete clinical or imaging data (n=28), pathological or high-energy vertebral fractures (n=22), previous spinal surgery or instrumentation (n=17), or multilevel/non-index fractures unsuitable for the predefined analysis (n=19). The final cohort comprised 314 patients, including 80 who developed adjacent vertebral fracture (AVF) within 18 months and 234 who did not.

(B) Study workflow. Preoperative clinical and imaging information was used to develop Model 1. Cement distribution available after PVP was subsequently incorporated to construct the perioperative updated Model 2. Both models underwent internal validation and performance assessment.

Abbreviations: AVF, adjacent vertebral fracture; PVP, percutaneous vertebroplasty.

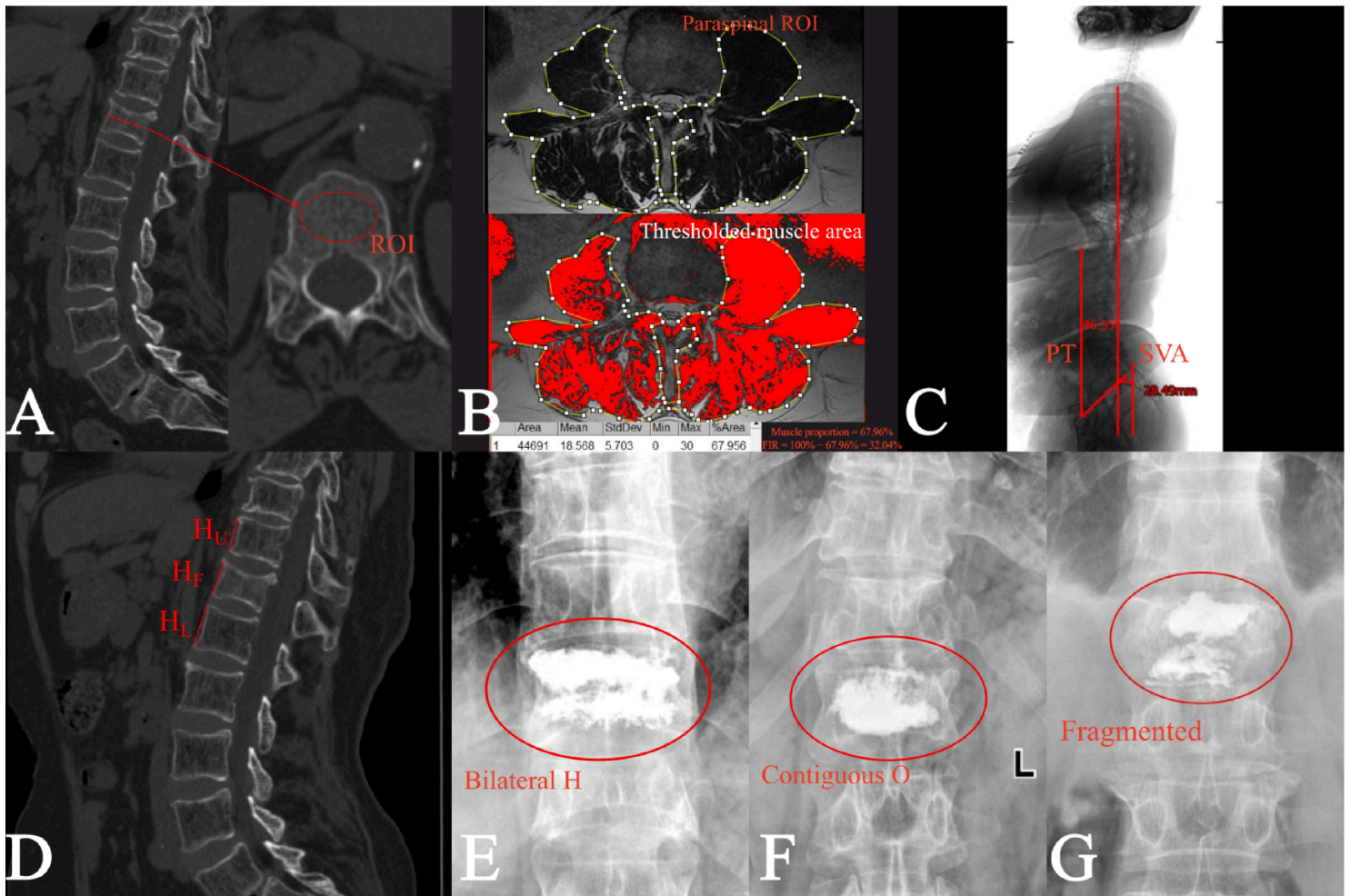


Figure 2

Representative imaging measurements and cement distribution patterns.

Legend: (A) Measurement of target vertebral CT-HU. A cancellous-bone region of interest was placed within the target vertebral body while avoiding cortical bone, endplates, focal sclerosis, bone islands, and venous channels. (B) Measurement of paraspinal fat infiltration rate (FIR). The paraspinal muscle region of interest was delineated and threshold-based segmentation was used to quantify effective muscle area and calculate FIR. (C) Assessment of sagittal alignment on a standing whole-spine lateral radiograph. Sagittal vertical axis (SVA) was measured as the horizontal distance between the C7 plumb line and the posterosuperior corner of S1; SVA ≥ 50 mm was classified as sagittal imbalance. Pelvic tilt was additionally recorded to characterize pelvic compensation. (D) Measurement of preoperative anterior vertebral compression using the anterior heights of the cranial adjacent vertebra (H_U), fractured vertebra (H_F), and caudal adjacent vertebra (H_L). (E–G) Representative bilateral H-shaped, contiguous O-shaped, and fragmented cement distribution patterns, respectively.

Abbreviations: CT-HU, computed tomography Hounsfield unit; FIR, fat infiltration rate; ROI, region of interest; SVA, sagittal vertical axis; PT, pelvic tilt.

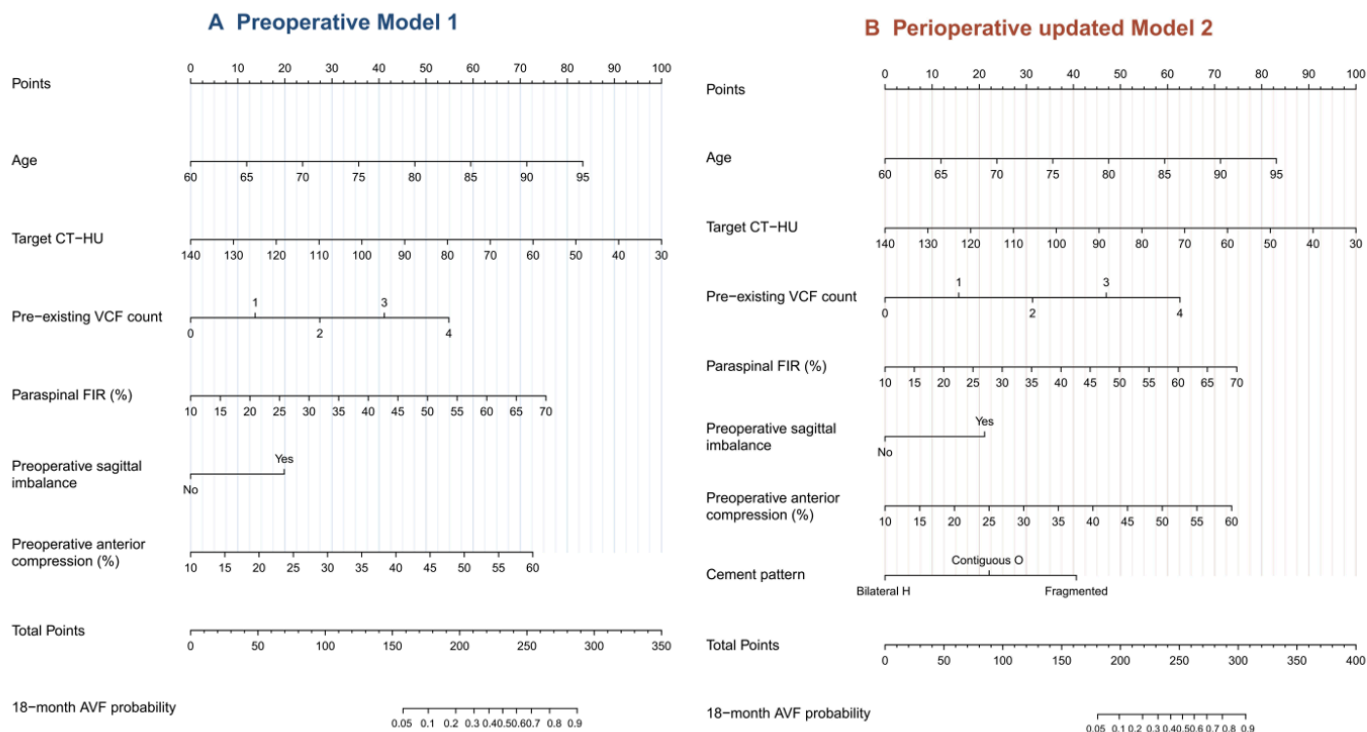


Figure 3

Nomograms for the preoperative Model 1 and perioperative updated Model 2.

Legend: (A) Nomogram for Model 1 incorporating age, target vertebral CT-HU, pre-existing VCF count, paraspinal FIR, preoperative sagittal imbalance, and preoperative anterior vertebral compression rate. (B) Nomogram for Model 2 incorporating all six Model 1 predictors together with cement distribution pattern. Individual predictor values are converted to points, and the total score corresponds to the predicted probability of AVF within 18 months after PVP.

Abbreviations: AVF, adjacent vertebral fracture; CT-HU, computed tomography Hounsfield unit; VCF, vertebral compression fracture; FIR, fat infiltration rate; PVP, percutaneous vertebroplasty.

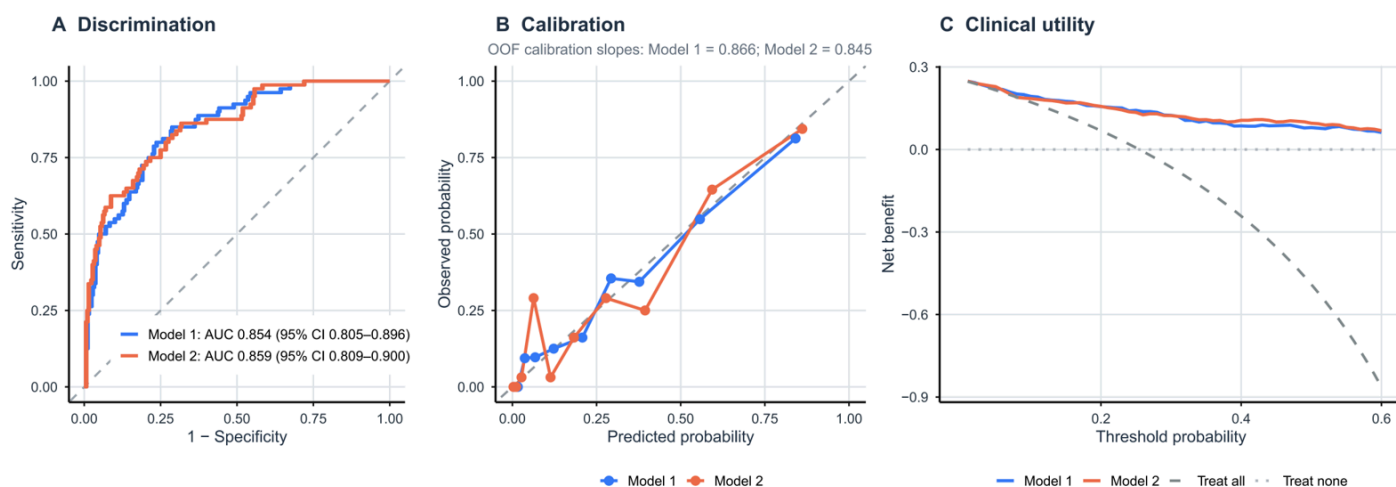


Figure 4

Internal performance and clinical utility of Model 1 and Model 2.

Legend: (A) Receiver operating characteristic curves based on fixed stratified 10-fold out-of-fold predictions. The OOF AUC was 0.854 for Model 1 and 0.859 for Model 2.

(B) Calibration of OOF predicted probabilities. Calibration intercepts were -0.015 and -0.005, and calibration slopes were 0.866 and 0.845 for Model 1 and Model 2, respectively.

(C) Decision curve analysis comparing the net benefit of Model 1 and Model 2 with treat-all and treat-none strategies across threshold probabilities.

Abbreviations: AUC, area under the receiver operating characteristic curve; OOF, out-of-fold; DCA, decision curve analysis.

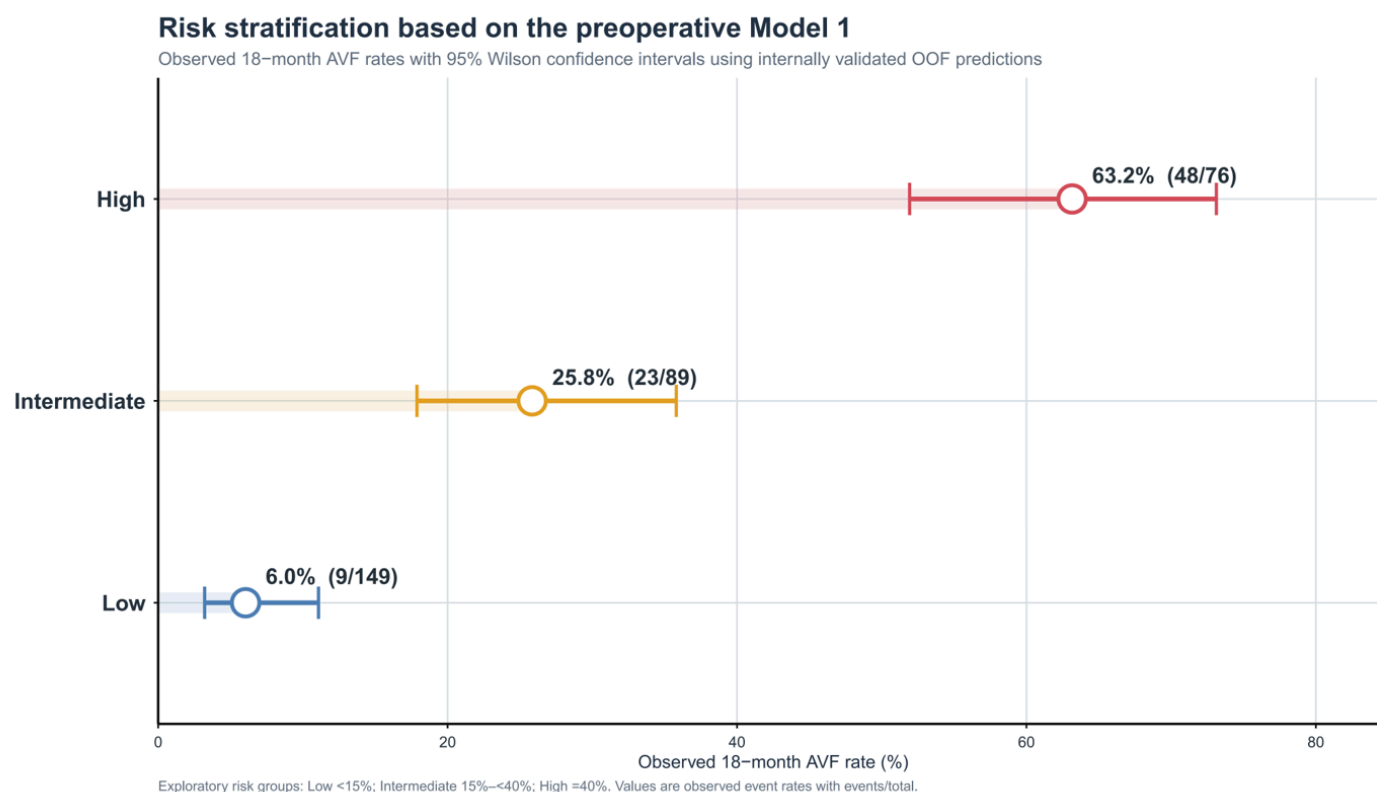


Figure 5

Exploratory risk stratification based on Model 1 out-of-fold predicted probabilities.

Legend: Patients were categorized as low risk (<15%), intermediate risk (15% to <40%), or high risk ($\geq 40\%$) according to Model 1 OOF predicted probabilities. Observed 18-month AVF rates were 6.0% (9/149), 25.8% (23/89), and 63.2% (48/76), respectively. Error bars represent 95% Wilson confidence intervals. The risk thresholds were used for exploratory presentation and should not be interpreted as externally validated clinical decision cutoffs.

Abbreviations: AVF, adjacent vertebral fracture; OOF, out-of-fold.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.png](#)
- [FigureS2.png](#)
- [FigureS3.png](#)
- [SupplementaryTables.docx](#)