Am J Cancer Res 2026;16(6):2189-2214
www.ajcr.us /ISSN:2156-6976/ajcr0173100

Original Article

Time-dependent evolution of osteosarcoma recurrence risk across clinical risk stratification and its prognostic value

Huadong Zhang[1], Xingxing Wang[2], Xiaochen Qiao[1], Chaojian Xu[1], Yi Feng[1]

[1]Department of Orthopedics, Second Hospital of Shanxi Medical University, Taiyuan 030001, Shanxi, China; [2]Special Care Ward, Shanxi Provincial People's Hospital, Taiyuan 030012, Shanxi, China

Received April 6, 2026; Accepted May 19, 2026; Epub June 15, 2026; Published June 30, 2026

Abstract: Background: Patients with osteosarcoma remain at high risk of recurrence after surgery despite multimodal treatment, and previously developed prediction models have rarely considered the time-varying effects of prognostic factors. Methods: We retrospectively analyzed 498 patients with osteosarcoma who received standard treatment at two centres between 2010 and 2019, including a training cohort (n = 341) and an external validation cohort (n = 157). The proportional hazards assumption was assessed using Schoenfeld residuals. For variables with time-dependent effects, piecewise Cox regression models were used. A dynamic prediction model was developed and evaluated using the concordance index, time-dependent receiver operating characteristic curves, calibration curves, and decision curve analysis. Results: During a median follow-up of 52.0 months, 202 patients (40.6%) developed postoperative recurrence, with most recurrences occurring within 12.0 months after surgery. Multivariable analysis showed that axial location, tumor diameter ≥ 10 cm, Huvos grade I-II, positive surgical margins, and incomplete adjuvant chemotherapy were independent risk factors. Time-dependent analyses showed that the effects of these covariates were strongest during 12.0-24.0 months postoperatively and then declined, becoming non-significant beyond 36.0 months in most cases. The one-, three-, and five-year area under the curve values were 0.805, 0.734, and 0.923 in the training cohort and 0.860, 0.775, and 0.890 in the validation cohort, respectively. Risk scores stratified patients into low-, intermediate-, and high-risk groups, with significantly different recurrence-free survival ($P < 0.0001$). Conclusion: This dynamic prediction model showed good discriminatory ability and generalizability and may support risk-adapted surveillance, especially within 12-24 months after surgery.

Keywords: Osteosarcoma, recurrence, prediction model, time-dependent effect, risk stratification, external validation

Introduction

Osteosarcoma is the most prevalent form of bone cancer that develops in children. Despite its rarity in the general population, this cancer's aggressive behavior and tendency to metastasize early make it a major public health threat to young people's lives [1]. The regular treatment at the moment is neoadjuvant chemotherapy, wide surgical resection, and postoperative adjuvant chemotherapy. This course of action has significantly improved patient outcomes. The latest Global Cancer Observatory cancer statistics show that the 5-year survival for non-metastatic osteosarcoma has reached 60%-70% [3]. Nonetheless, despite significant advances in surgical reconstruction techniques and continuous refinements to chemotherapy regimens over three decades, long-term survival benefits appear to have plateaued for osteosarcoma patients and are harder to achieve [4]. For patients with metastasis at diagnosis or with postoperative recurrence, survival is poor, and there is considerable scope for improvement in current management and surveillance.

Postoperative recurrence of the disease and death from it are the commonest causes of treatment failure in osteosarcoma patients and a key pain point in clinical management. About 30%-40% of patients with localized osteosarcoma develop either local recurrence or distant metastasis, most often to the lungs, after receiving standard treatment [5]. When the same

https://doi.org/10.62347/TPOL1062

cancer reappears, outcomes become worse; survival after further disease at 5 years is less than 20-30% [6]. When analyzing the timing of recurrences, it has been shown that the recurrence risk for osteosarcoma is not uniformly distributed over time. In fact, a marked temporal clustering occurs whereby the vast majority of events take place within 2-3 years postoperatively [7]. Failure to apply intensive monitoring during the "time window" of maximum recurrence risk means that the best opportunity for surgical resection of isolated lesions may be missed. On the contrary, over-monitoring during low-risk periods increases patients' radiation dose and costs.

A number of earlier studies have sought biological and clinicopathological markers associated with the prognosis of osteosarcoma to achieve early warning of recurrence risk. According to the National Comprehensive Cancer Network guidelines and European Society for Medical Oncology guidelines, risk factors considered include a large tumor volume, axial skeleton location, advanced Enneking stage, poor Huvos histological response (chemotherapy necrosis rate < 90%), and positive surgical margins [2, 8]. Moreover, some genetic mutations, such as TP53 and RB1, have been proven to be associated with poor prognosis due to improvements in genomics [9]. However, conclusions from existing research are contradictory. The predictive value of some risk factors, such as age and pathological subtype, remains controversial across cohorts. More importantly, current clinical follow-up strategies largely follow a "one-size-fits-all" fixed pattern, such as surveillance every 3 months for the first 2 years postoperatively. Protocols of this sort overlook individual patient risk heterogeneity and cannot address the challenges of precision medicine [2].

Almost all published prognostic prediction models in osteosarcoma are based on traditional Cox proportional hazards (PH) regression from a statistical methodology viewpoint. According to a key assumption of this model, the hazard ratio (HR) of covariates remains constant throughout follow-up [10]. Nevertheless, the biological behavior of tumors evolves dynamically. According to clinical practice and recent studies, certain risk factors have "time-dependent effects". For instance, patients with poor chemotherapy response or positive surgical margins have an extremely high risk of recurrence in the early postoperative period. However, if patients survive beyond 3 years postoperatively, the effect of these factors on late recurrence may significantly diminish or even disappear [11]. When time-dependent features are ignored while Cox models are applied, this not only violates statistical assumptions but also misjudges recurrence risk at a given time point, which decreases model accuracy and clinical usefulness [12]. Presently, there is a dearth of systematic studies on the time-dependent evolution of recurrence risk in osteosarcoma. Such models rarely undergo rigorous independent external validation, which limits their level of evidence and generalizability [13].

In light of this backdrop, this study aims to fill this gap by assessing the time-dependent nature of risk factors for postoperative osteosarcoma recurrence from a large clinical cohort and assessing the PH assumptions of the Cox model. Following this groundwork, we developed and validated a recurrence risk prediction model integrating time-dependent features and evaluated its generalizability in an independent external validation cohort. The ultimate aim of this study is to develop a scientific, rational, and individualized stratified follow-up strategy based on dynamic risk assessment results to achieve early and accurate identification of osteosarcoma recurrence and better allocation of medical resources for improved long-term survival outcomes.

Methods and materials

Sample size calculation

The ultimate goal of this research is to develop a methodologically rational and individualized follow-up strategy based on dynamic risk assessment results to timely and accurately identify osteosarcoma recurrence, thus allowing more effective allocation of medical resources and improved long-term survival. This retrospective cohort study used recurrence-free survival (RFS) as the main endpoint. The sample size was calculated using the Schoenfeld residual method formula for Cox proportional hazards regression models [14]. Thus, the total number of recurrence events was computed using the above formula. The total number of events required was determined through trial and error. The parameters were determined as follows: (1) $Z_{1-\alpha/2}$: the quantile of the standard

normal distribution corresponding to a two-sided significance level of $\alpha = 0.05$, equal to 1.96; (2) $Z_{1-\beta}$: the quantile of the standard normal distribution corresponding to a test power of $1-\beta = 0.80$, equal to 0.84; (3) p_1 and p_2: the sample proportions considered in the two predictor variable groups, which were assumed to be balanced between high- and low-risk populations, i.e., ($p_1 = p_2 = 0.5$); and (4) HR: the expected HR. Taking into account previous prognostic studies of osteosarcoma and clinical experience, a minimum clinically meaningful HR of 1.60 was set [15]. By substituting these parameters, we obtained $E = (1.96 + 0.84)^2/[0.5 \times 0.5 \times (\ln 1.60)^2] \approx 142$ events. Thus, a minimum of 142 recurrence events was required for modelling. According to Riley's method and the guidelines of the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis statement, we aimed to avoid overfitting in multivariable models by ensuring at least 10 outcome events per candidate predictor variable (events per variable ≥ 10) [12]. In view of our research plan to include approximately 15 candidate clinical and pathological predictors, at least 15 × 10 recurrence events were required. Hence, based on both calculation methods, our study required at least 150 recurrence events. According to prior literature, the overall postoperative recurrence rate for osteosarcoma is around 35% to 40% [16]. The recurrence rate was conservatively estimated at 35%, and hence the total required sample size (N) was estimated as follows: N = E/expected recurrence rate = 150/0.35 ≈ 429 patients. In view of anticipated information loss or loss to follow-up of approximately 10% in retrospective data collection, the final target sample size should not be less than 429 ÷ 0.9 ≈ 477 cases. This study eventually included a total of 498 patients, comprising 341 cases in the training cohort and 157 cases in the external validation cohort. A total of 202 recurrence events were noted during follow-up for the entire cohort, with a recurrence rate of 40.6%. The sample size (498 > 477) and recurrence event count (202 > 150) met the requirements for adequate statistical power to develop a Cox regression prediction model with adequate performance.

Study population

Training cohort: We retrospectively collected data from 341 patients with pathologically confirmed primary osteosarcoma at the Department of Orthopedics, Second Hospital of Shanxi Medical University, between January 2010 and December 2018. This cohort was used for model development and internal validation. External validation cohort: We included 157 patients with pathologically confirmed primary osteosarcoma at Shanxi Provincial People's Hospital between January 2012 and June 2019 to assess model generalizability. This study was approved by the Ethics Committee of Second Hospital of Shanxi Medical University. Owing to the retrospective nature of the study and the complete anonymization of all data, the requirement for informed consent was waived. This study was conducted in accordance with the Declaration of Helsinki [17] and relevant ethical guidelines.

Inclusion and exclusion criteria

Inclusion criteria: (1) histopathologically confirmed primary osteosarcoma; (2) age ≥ 18 years and ≤ 70 years; (3) receipt of standard treatment consisting of neoadjuvant chemotherapy, surgical resection, and postoperative adjuvant chemotherapy; (4) surgical approach involving wide excision, including limb-salvage surgery or amputation; (5) complete postoperative follow-up data with a follow-up duration ≥ 24 months or occurrence of recurrence/death during follow-up; and (6) complete clinical and pathological data. Exclusion criteria: (1) secondary osteosarcoma, such as Paget's disease transformation or post-radiation osteosarcoma; (2) low-grade osteosarcoma subtypes, such as parosteal or periosteal osteosarcoma; (3) presence of distant metastasis at initial diagnosis (Stage III); (4) concurrent other primary malignancies; (5) receipt of preoperative radiotherapy; (6) perioperative death within 30 days postoperatively; and (7) loss to follow-up or incomplete follow-up data.

Clinical data collection

We retrospectively collected clinical and pathological data from enrolled patients through electronic medical records. Data included three main categories: demographic characteristics, tumor features, and treatment-related factors. Demographic characteristics included sex (male/female), age at diagnosis (grouped as ≤ 65 years and > 65 years), and body mass index. Body mass index was categorized according to Chinese adult standards as underweight (<

18.5 kg/m^2), normal weight (18.5-23.9 kg/m^2), and overweight/obese (≥ 24.0 kg/m^2). Tumor characteristics included primary site (distal femur, proximal tibia, proximal humerus, other extremity bones, and axial skeleton [spine or pelvis]), maximum tumor diameter (measured by preoperative magnetic resonance imaging [MRI] and grouped as ≤ 10 cm and > 10 cm), tumor volume (calculated using the ellipsoid formula V = π/6 × length × width × height based on MRI-measured dimensions), pathological subtype (conventional type [osteoblastic, chondroblastic, and fibroblastic], telangiectatic, small cell, high-grade surface, and others), Enneking surgical staging (IA, IB, IIA, and IIB), presence of pathological fracture, and presence of skip metastasis, defined as satellite lesions discontinuous with the main tumor mass within the same bone or adjacent joint. Treatment-related factors included neoadjuvant chemotherapy regimen (methotrexate, doxorubicin, and cisplatin protocol; adriamycin and cisplatin protocol; or other regimens), number of neoadjuvant chemotherapy cycles, surgical method (limb-salvage surgery or amputation), surgical margin status (R0 resection, R1 resection, or R2 resection), Huvos tumor necrosis grading (Grade I-IV), further dichotomized as good response (Grade III-IV, necrosis rate ≥ 90%) or poor response (Grade I-II, necrosis rate < 90%), and completion status of postoperative adjuvant chemotherapy (completed or did not complete standard cycles).

Measurement methods

All patients completed standardized imaging and pathological assessments before neoadjuvant chemotherapy and surgery. For imaging assessment, all patients underwent preoperative plain and contrast-enhanced MRI of the primary site using a Siemens Magnetom Skyra 3.0T MRI system (Siemens, Germany). Scanning sequences included T1-weighted imaging, T2-weighted imaging, short tau inversion recovery sequence, and gadolinium contrast-enhanced T1-weighted imaging. Two radiologists with over 5 years of experience in bone tumor imaging independently measured maximum tumor diameter and tumor volume while blinded to patient outcomes. In the event of disagreement in the measurements, discussions were held until consensus was reached. The final values used for the analyses were the average measurements. All patients underwent preoperative high-resolution computed tomography (CT) of the chest and postoperative follow-up high-resolution CT scanning for lung metastasis using a GE Revolution CT scanner (GE Healthcare, USA) with a 1.25-mm slice thickness. Whole-body bone scintigraphy was also performed using the GE Discovery NM/CT 670 SPECT/CT system (GE Healthcare, USA). Approximately three hours after intravenous injection of ^{99m}Tc-methylene diphosphonate, images were acquired to identify bone metastasis and/or skip metastasis. All surgical resection specimens were reviewed independently by two senior pathologists qualified in bone tumor pathology. Histological classification and diagnosis were performed according to the 2020 World Health Organization bone tumor classification. After fixation in 10% neutral formalin, decalcification, and paraffin embedding, specimens were sectioned every 1 cm along the maximum tumor diameter for slide preparation and hematoxylin-eosin staining. The pathologists examined the post-chemotherapy tumor necrosis rate of the specimens and graded the specimens according to the Huvos tumor necrosis grading standard for further statistical analysis.

Follow-up and outcome measures

All patients were followed up postoperatively according to a common protocol: at 3-month intervals during the first 2 years postoperatively, at 6-month intervals during years 3 to 5, and then annually after 5 years. The follow-up content mainly comprised physical examination, imaging evaluation of the primary site (X-ray or MRI), and CT scan of the chest to detect local recurrence and distant metastasis. When indicated, additional imaging studies were performed. The follow-up cut-off date was June 30, 2025. The primary outcome measure was RFS, defined as the time from surgery to the first occurrence of local recurrence or distant metastasis. Tumor recurrence at the primary site or nearby soft tissue was defined as local recurrence. Distant metastasis was defined as metastatic lesions in the lungs, bones, or other distant organs. Recurrence or metastasis events required confirmation by imaging examination, with pathological confirmation through biopsy when necessary. For patients who did not experience recurrence during follow-up, RFS was censored at the date of the last follow-up.

Statistical analysis

Statistical analysis was performed using R software (version 4.5.1, R Foundation for Statistical Computing, Vienna, Austria) and SPSS software (version 27.0, IBM Corporation, USA). R software was mainly used for survival analysis, time-dependent model construction, and result visualization using packages including survival, survminer, rms, timeROC, ggplot2, and forestplot. SPSS software was mainly used for descriptive statistics and basic statistical tests. Continuous variables first underwent normality testing using the Shapiro-Wilk test. Data following a normal distribution were presented as the mean ± standard deviation ($\bar{x}$ ±s), with between-group comparisons using independent samples t-tests. Data not following a normal distribution were presented as the median (interquartile range) [M (IQR)], with between-group comparisons using Mann-Whitney U tests. Categorical variables were presented as frequency and percentage [n (%)], with between-group comparisons using χ^2 tests or Fisher's exact test. RFS was estimated using the Kaplan-Meier method, and survival curves were plotted. Between-group differences were compared using log-rank tests. Cox PH regression models were used for univariable analysis. Variables with $P < 0.10$ were included in multivariable Cox regression analysis using stepwise backward elimination to screen for independent recurrence-related factors. Results were presented as HRs with 95% confidence intervals (95% CIs). We tested the PH assumptions of the Cox model using the Schoenfeld residual method. For variables violating the PH assumptions ($P < 0.05$), we further applied time-dependent Cox regression models or piecewise Cox models. Based on the multivariable analysis results, we constructed nomogram models using the rms package to predict 1-year, 3-year, and 5-year RFS. Model discrimination was assessed using Harrell's concordance index (C-index) and the area under the time-dependent receiver operating characteristic (ROC) curve (AUC). Model calibration was evaluated by plotting calibration curves with internal validation using the bootstrap method with 1000 resamples. Decision curve analysis (DCA) was used to assess the clinical net benefit of the model at different threshold probabilities. We calculated the C-index and AUC in the independent external validation cohort and plotted calibration curves and DCA curves to evaluate model generalizability. Additionally, we performed subgroup analyses stratified by age (≤ 65 years vs. > 65 years), tumor site (extremity bones vs. axial skeleton), and Huvos grade (good response vs. poor response). We handled missing data using multiple imputation methods and compared the consistency of results before and after imputation through sensitivity analyses. Patients were stratified into low-, intermediate-, and high-risk groups based on the tertiles of the risk score distribution in the training cohort. Specifically, the 33rd and 66th percentiles (0.487 and 0.991, respectively) were used as cut-off values, and the same thresholds were applied to the validation cohort. All statistical tests were two-sided, with $P < 0.05$ considered statistically significant.

Results

Patient follow-up outcomes and recurrence overview

This study included 498 patients with osteosarcomas after surgery. The median follow-up time was 52.0 months, with an IQR of 42.0 months. Moreover, the follow-up range was between 2.0-66.0 months. At the last follow-up, recurrence was confirmed in 202 patients, giving an overall recurrence rate of 40.6%. Analysis of the recurrence group showed a median time to recurrence of 12.0 months (range: 2.0-58.0 months), indicating that the recurrence peak was mainly concentrated within the first year postoperatively. Kaplan-Meier survival analysis showed that the median RFS for the entire cohort was not reached because the overall recurrence rate did not exceed 50%. Cumulative RFS rates at 1 year, 3 years, and 5 years postoperatively were 79.1% (standard error, SE: 1.8%), 61.6% (SE: 2.2%), and 58.8% (SE: 2.3%), respectively. Based on the distribution of recurrence events and time-dependent survival estimates, recurrence risk was concentrated during the first 3 years postoperatively, with a decreasing trend thereafter.

Comparison of clinical and pathological characteristics between the recurrence and non-recurrence groups

This study included 498 patients with osteosarcoma: 202 in the recurrence group and 296 in the non-recurrence group. Univariable analysis

Table 1. Comparison of clinicopathological characteristics between the recurrence and non-recurrence groups in patients with osteosarcoma

Characteristic	Total (n = 498)	Recurrence group (n = 202)	Non-recurrence group (n = 296)	χ^2	*P*-value
Age				4.608	0.032
< 65 years	241 (48.39%)	86 (42.57%)	155 (52.36%)		
≥ 65 years	257 (51.61%)	116 (57.43%)	141 (47.64%)		
Sex				1.389	0.239
Male	290 (58.23%)	124 (61.39%)	166 (56.08%)		
Female	208 (41.77%)	78 (38.61%)	130 (43.92%)		
Tumor site				26.636	< 0.001
Extremity/other sites	430 (86.35%)	155 (76.73%)	275 (92.91%)		
Axial skeleton (spine/pelvis)	68 (13.65%)	47 (23.27%)	21 (7.09%)		
Tumor diameter				43.17	< 0.001
< 10 cm	260 (52.21%)	69 (34.16%)	191 (64.53%)		
≥ 10 cm	238 (47.79%)	133 (65.84%)	105 (35.47%)		
Enneking stage				16.064	< 0.001
IIA	229 (45.98%)	71 (35.15%)	158 (53.38%)		
IIB	269 (54.02%)	131 (64.85%)	138 (46.62%)		
Histological subtype				0.004	0.950
Osteoblastic	374 (75.10%)	152 (75.25%)	222 (75.00%)		
Others	124 (24.90%)	50 (24.75%)	74 (25.00%)		
Huvos grade				32.050	< 0.001
I-II (poor response)	274 (55.02%)	142 (70.30%)	132 (44.59%)		
III-IV (good response)	224 (44.98%)	60 (29.70%)	164 (55.41%)		
Surgical method				3.844	0.050
Limb-salvage surgery	421 (84.54%)	163 (80.69%)	258 (87.16%)		
Amputation	77 (15.46%)	39 (19.31%)	38 (12.84%)		
Surgical margin				26.960	< 0.001
Negative (R0)	434 (87.15%)	157 (77.72%)	277 (93.58%)		
Positive (R1/R2)	64 (12.85%)	45 (22.28%)	19 (6.42%)		
Adjuvant chemotherapy completion				16.650	< 0.001
Completed	455 (91.37%)	172 (85.15%)	283 (95.61%)		
Incomplete or refused	43 (8.63%)	30 (14.85%)	13 (4.39%)		
Comorbidity, CCI				1.189	0.275
0	98 (19.68%)	35 (17.33%)	63 (21.28%)		
≥ 1	400 (80.32%)	167 (82.67%)	233 (78.72%)		

Note: Data are presented as n (%). CCI, Charlson Comorbidity Index; R0, no residual tumor; R1, microscopic residual tumor; R2, macroscopic residual tumor; χ^2, chi-square statistic.

revealed significant differences in multiple clinical and pathological characteristics between the two groups. The recurrence group had higher proportions of patients aged ≥ 65 years (P = 0.032), tumors located in the axial skeleton (P < 0.001), tumor diameter ≥ 10 cm (P < 0.001), Enneking stage IIB (P < 0.001), Huvos grade I-II (poor chemotherapy response) (P < 0.001), positive surgical margins (P < 0.001), and failure to complete standard adjuvant chemotherapy cycles (P < 0.001). The two groups showed no significant differences in sex (P = 0.239), pathological subtype (P = 0.950), surgical method (P = 0.050), or Charlson Comorbidity Index (CCI) (P = 0.275) (all P > 0.05) (**Table 1**).

Comparison of baseline characteristics between the recurrence and non-recurrence groups in the training cohort

The training cohort included 341 patients with osteosarcoma: 135 in the recurrence group

Table 2. Comparison of baseline characteristics between the recurrence and non-recurrence groups in the training cohort

Characteristic	Total (n = 341)	Recurrence group (n = 135)	Non-recurrence group (n = 206)	χ^2	*P*-value
Age				0.998	0.318
< 65 years	168 (49.27%)	62 (45.93%)	106 (51.46%)		
≥ 65 years	173 (50.73%)	73 (54.07%)	100 (48.54%)		
Sex				2.169	0.141
Male	193 (56.60%)	83 (61.48%)	110 (53.40%)		
Female	148 (43.40%)	52 (38.52%)	96 (46.60%)		
Tumor site				14.632	< 0.001
Extremity/other sites	297 (87.10%)	106 (78.52%)	191 (92.72%)		
Axial skeleton (spine/pelvis)	44 (12.90%)	29 (21.48%)	15 (7.28%)		
Tumor diameter				23.298	< 0.001
< 10 cm	180 (52.79%)	49 (36.30%)	131 (63.59%)		
≥ 10 cm	161 (47.21%)	86 (63.70%)	75 (36.41%)		
Enneking stage				3.427	0.064
IIA	160 (46.92%)	55 (40.74%)	105 (50.97%)		
IIB	181 (53.08%)	80 (59.26%)	101 (49.03%)		
Histological subtype				0.832	0.362
Osteoblastic	251 (73.61%)	103 (76.30%)	148 (71.84%)		
Others	90 (26.39%)	32 (23.70%)	58 (28.16%)		
Huvos grade				16.566	< 0.001
I-II (poor response)	181 (53.08%)	90 (66.67%)	91 (44.17%)		
III-IV (good response)	160 (46.92%)	45 (33.33%)	115 (55.83%)		
Surgical method				1.848	0.174
Limb-salvage surgery	289 (84.75%)	110 (81.48%)	179 (86.89%)		
Amputation	52 (15.25%)	25 (18.52%)	27 (13.11%)		
Surgical margin				12.215	< 0.001
Negative (R0)	299 (87.68%)	108 (80.00%)	191 (92.72%)		
Positive (R1/R2)	42 (12.32%)	27 (20.00%)	15 (7.28%)		
Adjuvant chemotherapy completion				7.831	0.005
Completed	315 (92.38%)	118 (87.41%)	197 (95.63%)		
Incomplete or refused	26 (7.62%)	17 (12.59%)	9 (4.37%)		
Comorbidity, CCI				2.951	0.086
0	66 (19.35%)	20 (14.81%)	46 (22.33%)		
≥ 1	275 (80.65%)	115 (85.19%)	160 (77.67%)		

Note: Data are presented as n (%). CCI, Charlson Comorbidity Index; R0, no residual tumor; R1, microscopic residual tumor; R2, macroscopic residual tumor; χ^2, chi-square statistic.

and 206 in the non-recurrence group. Baseline characteristic analysis showed significant differences in multiple clinical and pathological features between the two groups. The recurrence group had higher proportions of tumors located in the axial skeleton (P < 0.001), tumor diameter ≥ 10 cm (P < 0.001), Huvos grade I-II (poor chemotherapy response) (P < 0.001), positive surgical margins (P < 0.001), and failure to complete standard adjuvant chemotherapy cycles (P = 0.005). The two groups showed no significant differences in age (P = 0.318), sex (P = 0.141), Enneking stage (P = 0.064), pathological subtype (P = 0.362), surgical method (P = 0.174), or CCI (P = 0.086) (all P > 0.05) (**Table 2**).

Correlation analysis of predictive variables and multicollinearity testing

To assess whether collinearity existed among the 11 candidate predictive variables included

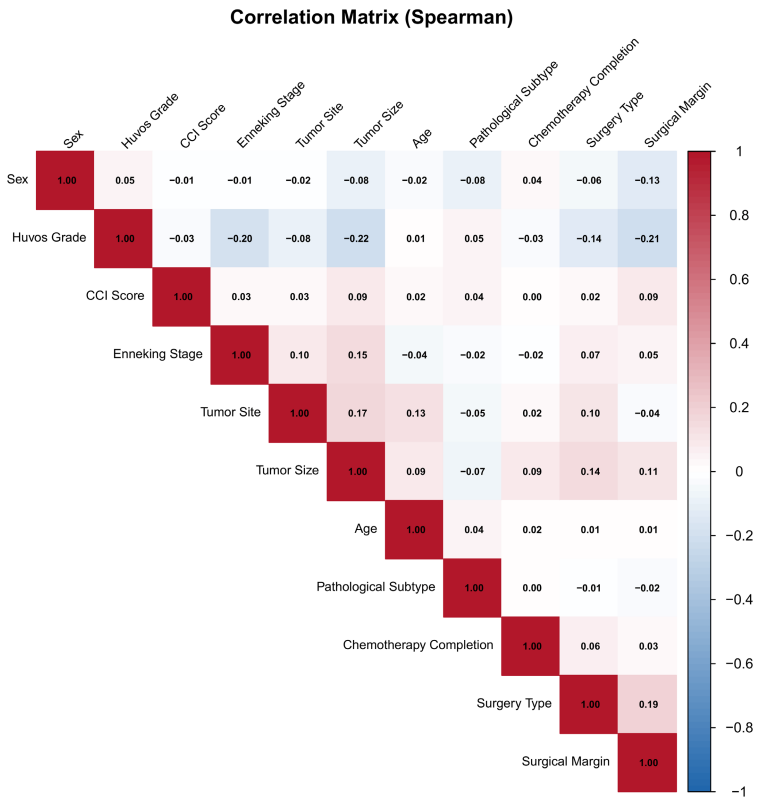

Figure 1. Spearman correlation matrix heatmap of candidate predictor variables in the training cohort. Note: CCI, Charlson Comorbidity Index.

in the model, we performed Spearman correlation analysis and variance inflation factor (VIF) testing. Correlation analysis showed that the absolute correlation coefficients between all variable pairs were less than 0.5. The strongest correlation was observed between Huvos grade and tumor diameter (r = -0.22), indicating no strong correlation among variables (**Figure 1**). VIF testing showed that all variables had VIF values less than 5 (range: 1.013-1.160), which were well below the multicollinearity warning threshold. This indicated no multicollinearity among the included predictive variables, making them suitable for subsequent multivariable regression analysis (**Figure 2**).

Testing the PH assumption of the cox regression model

To verify the applicability of the Cox PH regression model, we tested the PH assumption using the Schoenfeld residual method. In the full multivariable model containing all candidate variables, the global test indicated violation of the PH assumption (global P = 0.0157), with surgical margin (P = 0.027) and adjuvant chemotherapy completion (P = 0.033) showing significant time-dependent effects (**Figure 3**). Accordingly, for these variables that violated the PH assumption, piecewise Cox regression models with time-varying coefficients were applied in subsequent analyses, as prespecified in the statistical analysis plan.

Univariable and multivariable cox regression analysis of risk factors for osteosarcoma recurrence

To explore independent risk factors affecting postoperative osteosarcoma recurrence, univariable and multivariable Cox PH regression analyses were performed on clinical and pathological variables in the training cohort. Univariable Cox regression analysis showed that tumor location in the axial skeleton (P < 0.001), tumor diameter ≥ 10 cm (P < 0.001), Huvos grade I-II (P < 0.001), positive surgical margins (P < 0.001), and incomplete adjuvant chemotherapy (P < 0.001) were significant risk factors for osteosarcoma recurrence. Age (P = 0.348), sex (P = 0.122), Enneking stage (P = 0.056), pathological subtype (P = 0.431), surgical method (P = 0.097), and CCI (P = 0.099) showed no significant association with recurrence risk (P > 0.05). After variables with P < 0.1 in univariable analysis were included in the multivariable Cox regression model, the results showed that tumor location in the axial skeleton (P < 0.001), tumor diameter ≥ 10 cm (P = 0.011), Huvos grade I-II (P = 0.008), positive surgical margins (P < 0.001), and incomplete adjuvant chemotherapy (P < 0.001) were independent risk factors for postoperative osteosarcoma recurrence. Age (P = 0.122), Enneking stage (P = 0.056), and surgical method (P = 0.097) showed no independent predictive value after adjustment for other factors (P > 0.05) (**Figure 4**). Given that some of these five independent risk factors showed evidence of time-dependent effects, we further applied piecewise Cox regression models to estimate period-specific HRs (**Table 3**; **Figure 5**).

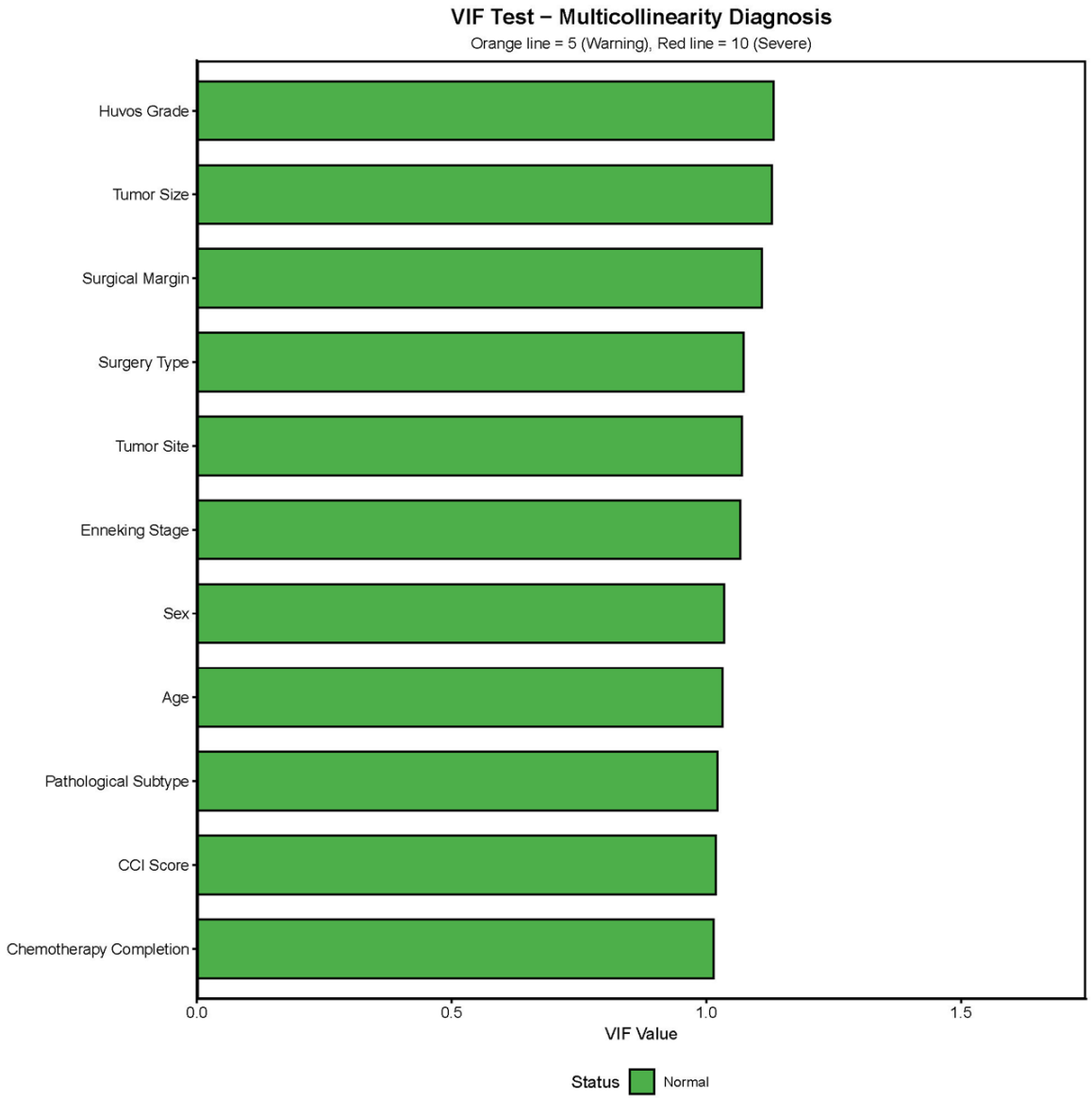

Figure 2. VIF values of candidate predictor variables in the training cohort. Note: VIF, variance inflation factor; CCI, Charlson Comorbidity Index.

Time-dependent effect analysis of independent prognostic factors

Given that some variables showed time-dependent trends in PH assumption testing, we used piecewise Cox regression models to further explore dynamic changes in the HRs of independent prognostic factors over time. Time-dependent analysis revealed that all five independent prognostic factors exhibited significant time-dependent effects. The hazard effect of tumor location in the axial skeleton was strongest in the early postoperative period and gradually attenuated over time. At 0-12 months and 12-24 months, the risk was significantly elevated ($P < 0.001$). At 24-36 months, statistical significance persisted ($P = 0.012$), but after 36 months, the hazard effect was no longer significant ($P > 0.05$). Time-dependency testing showed that this effect changed significantly over time ($P = 0.006$). Tumor diameter ≥ 10 cm was a significant risk factor at 0-12 months ($P < 0.001$), 12-24 months ($P < 0.001$), and 24-36 months ($P = 0.017$), but the effect disappeared after 36 months ($P > 0.05$). Time-dependency testing confirmed effect attenuation over time ($P = 0.008$). Huvos grade I-II (poor chemotherapy response) significantly increased recurrence

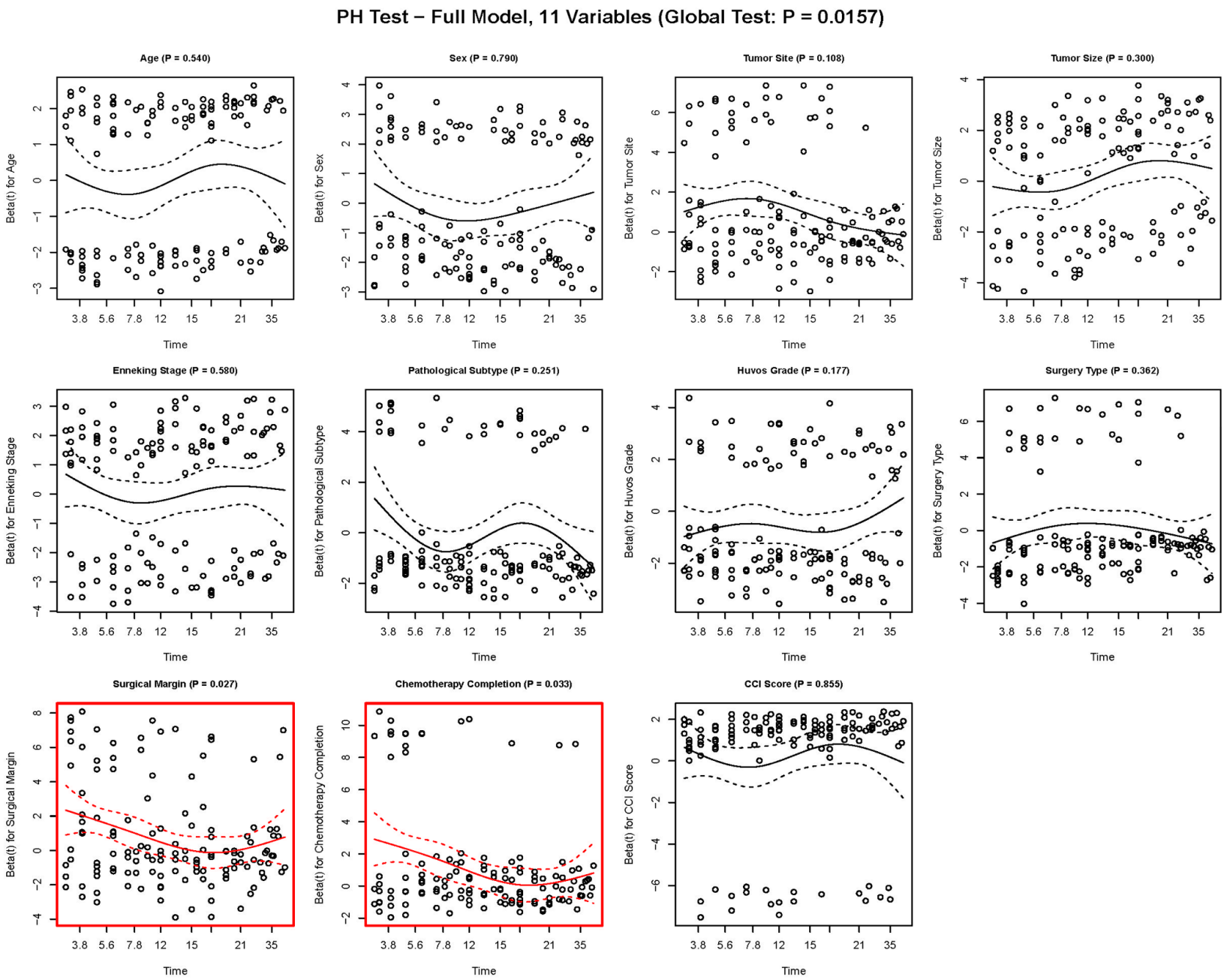
PH Test – Full Model, 11 Variables (Global Test: P = 0.0157)
Age (P = 0.540)
Beta(t) for Age
Sex (P = 0.790)
Beta(t) for Sex
Tumor Site (P = 0.108)
Beta(t) for Tumor Site
Tumor Size (P = 0.300)
Beta(t) for Tumor Size
Enneking Stage (P = 0.580)
Beta(t) for Enneking Stage
Pathological Subtype (P = 0.251)
Beta(t) for Pathological Subtype
Huvos Grade (P = 0.177)
Beta(t) for Huvos Grade
Surgery Type (P = 0.362)
Beta(t) for Surgery Type
Surgical Margin (P = 0.027)
Beta(t) for Surgical Margin
Chemotherapy Completion (P = 0.033)
Beta(t) for Chemotherapy Completion
CCI Score (P = 0.855)
Beta(t) for CCI Score
Time
3.8 5.6 7.8 12 15 21 35

Figure 3. Schoenfeld residual plots for testing the proportional hazards assumption in the full Cox model with 11 variables. Note: PH, proportional hazards; CCI, Charlson Comorbidity Index; Beta(t), time-varying coefficient. Red panels indicate variables that violated the PH assumption.

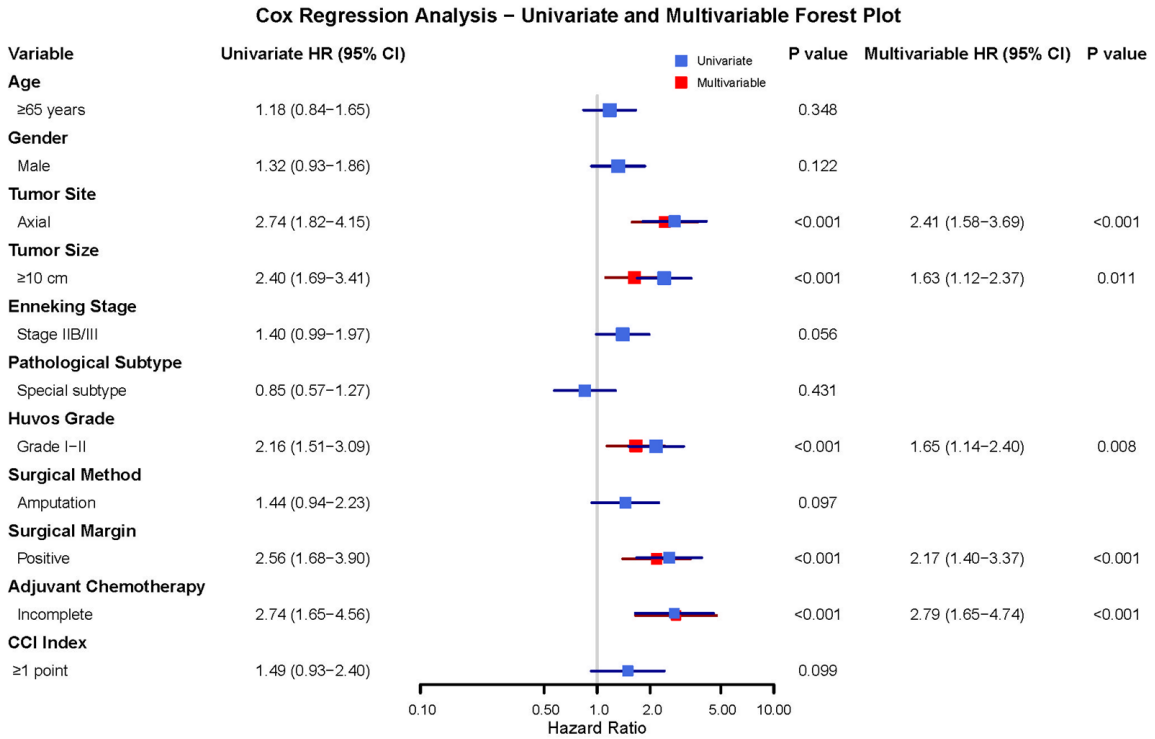

Figure 4. Forest plot of univariable and multivariable Cox regression analyses for osteosarcoma recurrence in the training cohort. Note: HR, hazard ratio; CI, confidence interval; CCI, Charlson Comorbidity Index.

risk at 0-12 months (P = 0.001), 12-24 months (P = 0.001), and 24-36 months (P = 0.036). After 36 months, the effect weakened (P > 0.05), with significant time dependency (P = 0.012). The hazard effect of positive surgical margins was most prominent at 0-12 months (P < 0.001) and 12-24 months (P < 0.001). Significance persisted at 24-36 months (P = 0.029) but was no longer observed after 36 months (P > 0.05), with significant time dependency (P = 0.005). Incomplete adjuvant chemotherapy significantly increased recurrence risk at 0-12 months (P < 0.001), 12-24 months (P < 0.001), and 24-36 months (P = 0.008). After 36 months, the effect disappeared (P > 0.05). Time-dependency testing confirmed an attenuating trend (P = 0.004). These results indicate that the influence of each independent prognostic factor on osteosarcoma recurrence was mainly concentrated in the first 36 months postoperatively, suggesting that the early postoperative period is a critical time for recurrence monitoring (**Table 3**; **Figure 5**).

Comparison of baseline characteristics between the training and external validation cohorts

To assess the reliability of external model validation, we compared the distributions of baseline characteristics between the training cohort (n = 341) and the external validation cohort (n = 157). The results showed no significant differences in any clinical or pathological characteristics between the two groups, including age (P = 0.565), sex (P = 0.276), tumor site (P = 0.472), tumor diameter (P = 0.704), Enneking stage (P = 0.536), pathological subtype (P = 0.256), Huvos grade (P = 0.199), surgical method (P = 0.847), surgical margin (P = 0.599), adjuvant chemotherapy completion (P = 0.237), and CCI (P = 0.789) (all P > 0.05). This indicated good comparability between the training and external validation cohorts, providing a reliable foundation for subsequent external model validation (**Table 4**).

Table 3. Time-dependent effects of independent prognostic factors for osteosarcoma recurrence using segmented Cox regression analysis

Prognostic factor	Time period, months	HR (95% CI)	*P*-value
Huvos grade I-II	0-12	2.65 (1.51-4.65)	0.001
Huvos grade I-II	12-24	2.18 (1.38-3.44)	0.001
Huvos grade I-II	24-36	1.76 (1.04-2.98)	0.036
Huvos grade I-II	36-48	1.35 (0.73-2.49)	0.339
Huvos grade I-II	> 48	1.08 (0.54-2.16)	0.826
Huvos grade I-II	Time-varying test	-	0.012
Incomplete adjuvant chemotherapy	0-12	3.48 (1.98-6.12)	< 0.001
Incomplete adjuvant chemotherapy	12-24	2.95 (1.82-4.78)	< 0.001
Incomplete adjuvant chemotherapy	24-36	2.28 (1.35-3.85)	0.008
Incomplete adjuvant chemotherapy	36-48	1.75 (0.95-3.22)	0.142
Incomplete adjuvant chemotherapy	> 48	1.42 (0.72-2.80)	0.485
Incomplete adjuvant chemotherapy	Time-varying test	-	0.004
Positive surgical margins	0-12	3.92 (1.95-7.88)	< 0.001
Positive surgical margins	12-24	3.24 (1.76-5.97)	< 0.001
Positive surgical margins	24-36	2.15 (1.08-4.28)	0.029
Positive surgical margins	36-48	1.68 (0.72-3.92)	0.231
Positive surgical margins	> 48	1.35 (0.52-3.51)	0.538
Positive surgical margins	Time-varying test	-	0.005
Tumor site, axial skeleton	0-12	3.15 (1.85-5.38)	< 0.001
Tumor site, axial skeleton	12-24	2.68 (1.68-4.28)	< 0.001
Tumor site, axial skeleton	24-36	2.05 (1.25-3.36)	0.012
Tumor site, axial skeleton	36-48	1.58 (0.92-2.72)	0.185
Tumor site, axial skeleton	> 48	1.25 (0.68-2.31)	0.548
Tumor site, axial skeleton	Time-varying test	-	0.006
Tumor diameter ≥ 10 cm	0-12	2.85 (1.62-5.01)	< 0.001
Tumor diameter ≥ 10 cm	12-24	2.42 (1.55-3.78)	< 0.001
Tumor diameter ≥ 10 cm	24-36	1.89 (1.12-3.19)	0.017
Tumor diameter ≥ 10 cm	36-48	1.52 (0.84-2.75)	0.167
Tumor diameter ≥ 10 cm	> 48	1.18 (0.61-2.28)	0.622
Tumor diameter ≥ 10 cm	Time-varying test	-	0.008

Note: HR, hazard ratio; CI, confidence interval. The time-varying test was based on Schoenfeld residuals.

Construction and internal validation of the osteosarcoma recurrence prediction model in the training cohort

Based on five independent prognostic factors screened through multivariable Cox regression analysis, we constructed a postoperative osteosarcoma recurrence risk prediction model. The risk score calculation formula was as follows: Risk score = 0.8810 × tumor site (axial skeleton = 1) + 0.4869 × tumor diameter (≥ 10 cm = 1) + 0.5037 × Huvos grade (I-II = 1) + 0.7747 × surgical margin (positive = 1) + 1.0273 × adjuvant chemotherapy completion (incomplete = 1). We performed internal validation of the model in the training cohort. Time-dependent ROC curve analysis showed that the AUC values for predicting 1-year, 3-year, and 5-year recurrence risk were 0.805 (95% CI: 0.747-0.864), 0.734 (95% CI: 0.681-0.788), and 0.923 (95% CI: 0.887-0.959), respectively, indicating good discriminative ability. DCA showed that within the threshold probability range of 0.1-0.7, the model's net benefit was higher than those of both the "treat all" and "treat none" strategies. At the threshold of 0.3, the net benefit was 0.187, indicating good clinical utility. The calibration curves showed agreement between the predicted and observed risks, with a Brier score of 0.2799.

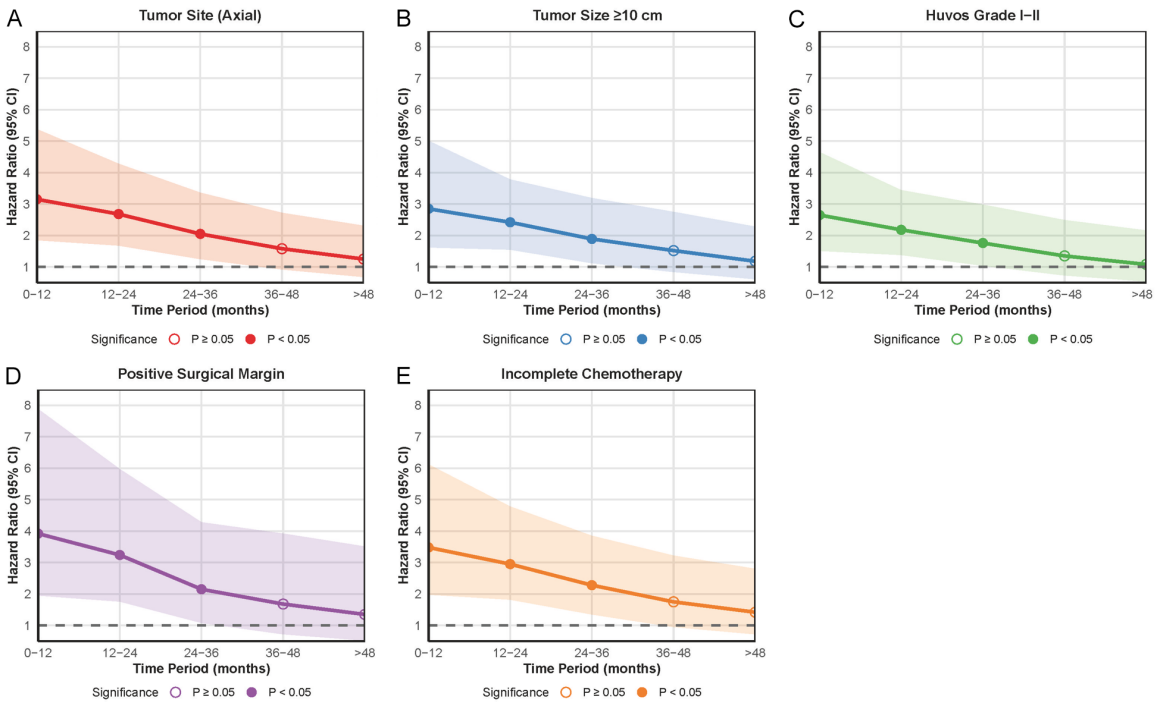

Figure 5. Dynamic changes in HRs of independent prognostic factors across postoperative time periods. A. Dynamic changes in the HR of tumor site, axial skeleton, across postoperative time periods. B. Dynamic changes in the HR of tumor diameter ≥ 10 cm across postoperative time periods. C. Dynamic changes in the HR of Huvos grade I-II across postoperative time periods. D. Dynamic changes in the HR of positive surgical margins across postoperative time periods. E. Dynamic changes in the HR of incomplete adjuvant chemotherapy across postoperative time periods. Note: HR, hazard ratio; CI, confidence interval. The dashed horizontal line indicates HR = 1. Filled circles indicate P < 0.05, and open circles indicate P ≥ 0.05.

Based on the tertiles of the clinical risk score, patients were classified into low-, intermediate-, and high-risk groups. Kaplan-Meier survival analysis showed that RFS differed significantly among the three risk groups (P < 0.0001), confirming the effectiveness of the risk stratification model (**Figure 6**).

Validation of the osteosarcoma recurrence prediction model in the external validation cohort

To assess model generalizability, we independently validated the prediction model in the external validation cohort (n = 157). Time-dependent ROC curve analysis showed that the AUC values for predicting 1-year, 3-year, and 5-year recurrence risk were 0.860 (95% CI: 0.796-0.925), 0.775 (95% CI: 0.703-0.847), and 0.890 (95% CI: 0.791-0.989), respectively. The discriminative ability was comparable to or slightly better than that in the training cohort, indicating good external applicability. DCA showed that within the threshold probability range of 0.1-0.7, the model's net benefit consistently exceeded those of the "treat all" and "treat none" strategies. At a threshold of 0.3, the net benefit was 0.227, further confirming the model's clinical decision-making value. Calibration curves showed agreement between the predicted risk and the observed risk in the external validation cohort, with a Brier score of 0.2600, which was lower than that in the training cohort. This indicated that the model's prediction accuracy remained reliable in external data. After applying the risk stratification criteria established in the training cohort to the external validation cohort, Kaplan-Meier survival analysis similarly showed significant differences in RFS among the low-, intermediate-, and high-risk groups (P < 0.0001), further validating the stability and reproducibility of the risk stratification model across different populations (**Figure 7**).

Development of an individualized follow-up strategy based on risk stratification

Based on the time-dependent characteristics of recurrence risk and risk stratification results,

Table 4. Comparison of baseline characteristics between the training cohort and external validation cohort

Characteristic	Total (n = 498)	Training cohort (n = 341)	External validation cohort (n = 157)	χ^2	*P*-value
Age				0.330	0.565
< 65 years	241 (48.39%)	168 (49.27%)	73 (46.50%)		
≥ 65 years	257 (51.61%)	173 (50.73%)	84 (53.50%)		
Sex				1.188	0.276
Male	290 (58.23%)	193 (56.60%)	97 (61.78%)		
Female	208 (41.77%)	148 (43.40%)	60 (38.22%)		
Tumor site				0.518	0.472
Extremity/other sites	430 (86.35%)	297 (87.10%)	133 (84.71%)		
Axial skeleton (spine/pelvis)	68 (13.65%)	44 (12.90%)	24 (15.29%)		
Tumor diameter				0.144	0.704
< 10 cm	260 (52.21%)	180 (52.79%)	80 (50.96%)		
≥ 10 cm	238 (47.79%)	161 (47.21%)	77 (49.04%)		
Enneking stage				0.382	0.536
IIA	229 (45.98%)	160 (46.92%)	69 (43.95%)		
IIB	269 (54.02%)	181 (53.08%)	88 (56.05%)		
Histological subtype				1.290	0.256
Osteoblastic	374 (75.10%)	251 (73.61%)	123 (78.34%)		
Others	124 (24.90%)	90 (26.39%)	34 (21.66%)		
Huvos grade				1.646	0.199
I-II (poor response)	274 (55.02%)	181 (53.08%)	93 (59.24%)		
III-IV (good response)	224 (44.98%)	160 (46.92%)	64 (40.76%)		
Surgical method				0.037	0.847
Limb-salvage surgery	421 (84.54%)	289 (84.75%)	132 (84.08%)		
Amputation	77 (15.46%)	52 (15.25%)	25 (15.92%)		
Surgical margin				0.276	0.599
Negative (R0)	434 (87.15%)	299 (87.68%)	135 (85.99%)		
Positive (R1/R2)	64 (12.85%)	42 (12.32%)	22 (14.01%)		
Adjuvant chemotherapy completion				1.398	0.237
Completed	455 (91.37%)	315 (92.38%)	140 (89.17%)		
Incomplete or refused	43 (8.63%)	26 (7.62%)	17 (10.83%)		
Comorbidity, CCI				0.072	0.789
0	98 (19.68%)	66 (19.35%)	32 (20.38%)		
≥ 1	400 (80.32%)	275 (80.65%)	125 (79.62%)		

Note: Data are presented as n (%). CCI, Charlson Comorbidity Index; R0, no residual tumor; R1, microscopic residual tumor; R2, macroscopic residual tumor; χ^2, chi-square statistic.

we developed a risk-adaptive individualized follow-up strategy. Since the hazard effects of all independent prognostic factors peaked at 12-24 months postoperatively, this time window was identified as the critical period for recurrence monitoring. For high-risk patients, we suggest close follow-up every 2-3 months during the first 12 months after surgery, with continued relatively intensive surveillance during the 12-24-month recurrence peak. Thereafter, the follow-up interval may be adjusted to every 3-4 months during 24-36 months, every 6 months during 36-60 months, and annually after 60 months. For intermediate-risk patients, follow-up every 3-4 months during the first 24 months may be considered, followed by every 6 months during 24-48 months and annually thereafter. For low-risk patients, follow-up every 4-6 months during the first 24 months is generally sufficient, extending to every 6-12 months during 24-60 months and annually thereafter. In terms of imaging, sur-

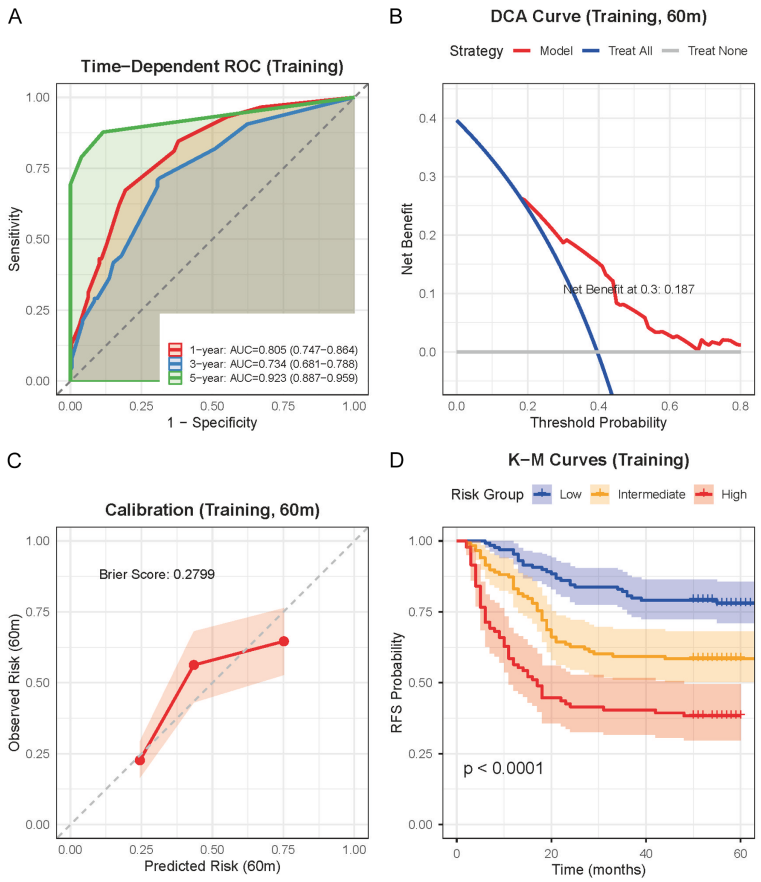

Figure 6. Internal validation of the recurrence prediction model in the training cohort. A. Time-dependent ROC curves showing AUC values for predicting 1-, 3-, and 5-year recurrence. B. DCA comparing the net benefit of the model with "treat all" and "treat none" strategies. C. Calibration curve assessing agreement between predicted and observed recurrence risk at 60 months. D. K-M curves showing differences in RFS among the low-, intermediate-, and high-risk groups. Note: ROC, receiver operating characteristic; AUC, area under the curve; DCA, decision curve analysis; RFS, recurrence-free survival; CI, confidence interval; K-M, Kaplan-Meier.

veillance should primarily rely on chest CT combined with local imaging, such as X-ray or MRI. Positron emission tomography-CT or bone scans should not be considered routine examinations but may be performed selectively in high-risk patients when clinically indicated, particularly during periods of suspected recurrence or ambiguous findings. These follow-up strategies should be interpreted as risk-adaptive suggestions based on recurrence patterns observed in this study rather than fixed clinical recommendations (**Table 5**; **Figure 8**).

Validation of the risk stratification model across different clinical subgroups

To assess the applicability of the risk stratification model across different clinical subgroups, we performed subgroup analyses and interaction testing. Subgroup analysis showed that the risk stratification model demonstrated significant prognostic discrimination ability in most subgroups. In the tumor site subgroup, the risk stratification effect was significant for patients with extremity tumors (n = 430, events = 155) (HR = 2.27, 95% CI: 1.54-3.34, P < 0.001). For patients with tumors in the axial skeleton (n = 68, events = 47), the difference did not reach statistical significance due to the smaller sample size (HR = 1.52, 95% CI: 0.47-4.94, P = 0.482). In the age subgroup, both the < 65 years group (n = 241, events = 86; HR = 2.00, 95% CI: 1.13-3.52, P = 0.017) and the ≥ 65 years group (n = 257, events = 116; HR = 2.58, 95% CI: 1.52-4.37, P < 0.001) showed significant risk stratification effects. In the surgical method subgroup, the risk stratification effect was significant in the limb-salvage surgery group (n = 421, events = 163) (HR = 2.37, 95% CI: 1.57-3.58, P < 0.001). For the amputation group (n = 77, events = 39), the difference did not reach statistical significance due to the limited sample size (HR = 1.77, 95% CI: 0.57-5.49, P = 0.323). In the overall population, the recurrence risk in the intermediate-/high-risk group was significantly higher than that in the low-risk group (HR = 2.31, 95% CI: 1.57-3.40, P < 0.05). The risk stratification model had reliable and consistent prognostic discrimination ability across different clinical subgroups, supporting its robustness and generalizability (**Figure 9**).

Discussion

Key research findings

Osteosarcoma is the most common primary malignant bone tumor. Despite advances in multidisciplinary comprehensive treatment

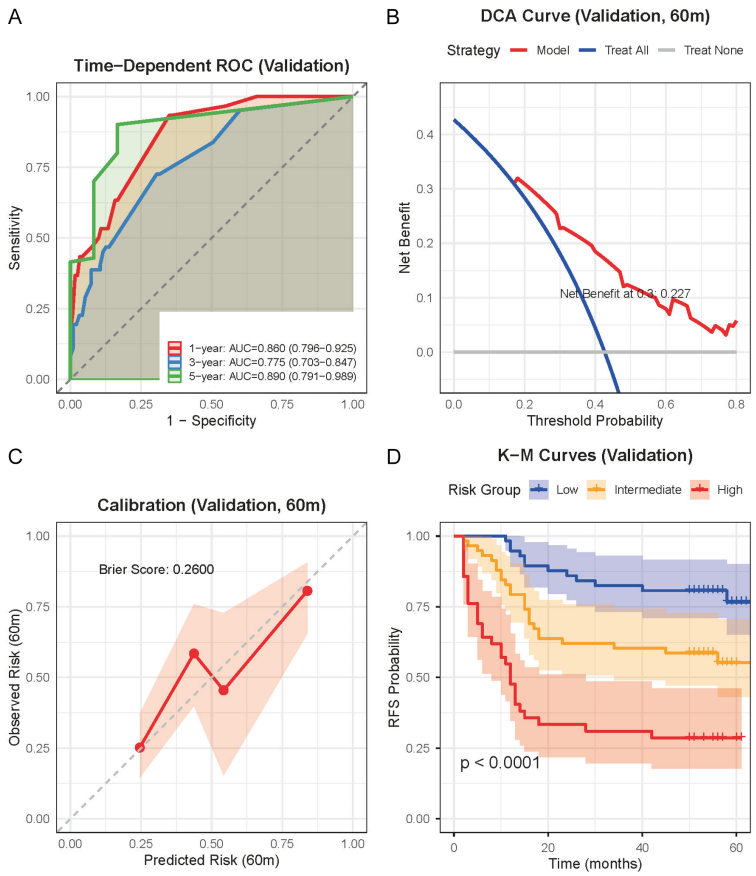

Figure 7. External validation of the recurrence prediction model in the validation cohort. A. Time-dependent ROC curves showing AUC values for predicting 1-, 3-, and 5-year recurrence. B. DCA comparing the net benefit of the model with the "treat all" and "treat none" strategies. C. Calibration curve assessing agreement between predicted and observed recurrence risk at 60 months. D. K-M curves showing differences in RFS among the low-, intermediate-, and high-risk groups. Note: ROC, receiver operating characteristic; AUC, area under the curve; DCA, decision curve analysis; RFS, recurrence-free survival; CI, confidence interval; K-M, Kaplan-Meier.

that have increased 5-year survival at 60%-70%, recurrence remains the main cause of treatment failure and death. In this study, long-term follow-up data from a total of 498 patients with high-grade osteosarcoma were obtained from two centres. This study aimed to systematically explore the risk factors for recurrence and their time-dependent evolution patterns. Recurrence was defined as local recurrence or distant metastasis, mostly in the lungs, confirmed by imaging and/or pathology. Time to recurrence was measured from the date of curative surgery. Due to considerations of statistical power, as the decreasing number of local recurrence events rendered separate modeling for these two events unstable, and clinical operability, as both events trigger intensified treatment and intensive follow-up, local recurrence and distant metastasis were combined as a single composite endpoint. Among all patients who experienced recurrence during the follow-up period, the median time to recurrence was 12 months. Of the 202 recurrence events, 177 (87.6%) occurred within the first 24 months postoperatively. Using multivariable Cox regression and adjusting for confounding variables including age, sex, Enneking stage, and surgical method, we confirmed five independent risk factors: tumor location in the axial skeleton, tumor diameter ≥ 10 cm, Huvos grade I-II response, positive surgical margins, and chemotherapy completion status.

Regarding variable definitions, several points warrant special mention. The 10 cm threshold for tumor diameter was mainly based on previous literature reports and commonly used clinical stratification standards, with ROC analysis in our cohort providing exploratory support. In sensitivity analyses treating tumor diameter as a continuous variable using restricted cubic splines, the results remained directionally consistent.

"Completed standard chemotherapy" was defined as completion of 2 preoperative cycles plus 4 postoperative cycles of the methotrexate, doxorubicin, and cisplatin regimen or an equivalent regimen, with the cumulative dose reaching at least 80% of the planned dose. "Equivalent regimens" and their dose conversions were predefined according to unified treatment protocols at both centers and verified by oncologists. Chemotherapy completion status was determined based on assessment at 6 months postoperatively. For patients still undergoing chemotherapy at 6 months postop-

Table 5. Risk-adaptive follow-up recommendations for patients with osteosarcoma based on recurrence risk stratification

Risk group	Postoperative period, months	Follow-up interval	Monitoring intensity	Recommended imaging
High risk	0-12	Every 2-3 months	Intensive	Chest CT + local MRI ± PET-CT
	12-24 (peak period)	Every 2-3 months	Intensive	Chest CT + local MRI ± PET-CT/bone scan
	24-36	Every 3-4 months	Standard	Chest CT + local MRI
	36-60	Every 6 months	Reduced	Chest CT + X-ray
	> 60	Annually	Minimal	Chest CT + X-ray
Intermediate risk	0-12	Every 3-4 months	Standard	Chest CT + local MRI
	12-24 (peak period)	Every 3-4 months	Standard	Chest CT + local MRI ± bone scan
	24-48	Every 6 months	Reduced	Chest CT + X-ray
	> 48	Annually	Minimal	Chest CT or X-ray
Low risk	0-24	Every 4-6 months	Reduced	Chest CT + X-ray
	24-60	Every 6-12 months	Minimal	Chest X-ray ± CT
	> 60	Annually	Minimal	Chest X-ray

Note: CT, computed tomography; MRI, magnetic resonance imaging; PET-CT, positron emission tomography-computed tomography. The peak recurrence window refers to 12-24 months postoperatively, when recurrence risk is highest.

eratively, those who completed ≥ 4 cycles without dose reduction were classified as "completed"; otherwise, they were classified as "not completed".

To avoid time bias from recurrence within 0-6 months affecting exposure determination, we further performed a 6-month landmark analysis, including only 303 patients from the training cohort without recurrence at 6 months postoperatively, with results directionally consistent with the main analysis. Additionally, sensitivity analysis in extended Cox models treating chemotherapy completion as a time-dependent covariate also supported the main conclusions.

Schoenfeld residual testing indicated that two variables violated the PH assumptions ($P < 0.05$): surgical margin ($P = 0.023$) and adjuvant chemotherapy completion ($P = 0.032$). Accordingly, piecewise Cox regression models were applied to all five independent prognostic factors, with interaction terms constructed as piecewise indicator functions. This yielded period-specific HR estimates for each time period: 0-12 months, 12-24 months, 24-36 months, 36-48 months, and > 48 months.

Analysis of the data showed that the hazard effects of the major risk factors were fairly strong at 12 and 24 months after surgery. The hazard effects started attenuating after 36 months. The pattern of variation differed across factors, and the constructed CIs became wider in later periods; hence, cautious is advisable.

Comparison of risk factors with previous studies

The independent risk factors found in our study were mostly in agreement with previous literature. Rubio-San-Simón et al. [18] represented the Fighting Osteosarcoma Through European Research consortium and conducted a systematic review of 19 studies including 3,245 patients. They confirmed nine prognostic factors affecting post-recurrence survival through a meta-analysis of published studies, including recurrence-free interval (RFI), lesion location, lesion number, lesion size, and resectability. It should be noted that this review was concerned with "prognostic factors after recurrence" rather than "risk factors for recurrence occurrence". Although related, these are not the same. The study revealed that post-recurrence survival was closely related to RFI and complete surgical resection. Indirectly, it confirms that margin status and early monitoring are important in patients. A previous study reported that patients with spinal osteosarcoma have a worse prognosis than those with extremity osteosarcoma [19]. A prediction model built on the Surveillance, Epidemiology, and End Results database found axial skeleton location to be an adverse prognostic indicator. According to Huang et al. [20], tumor site was also an independent prognostic factor in

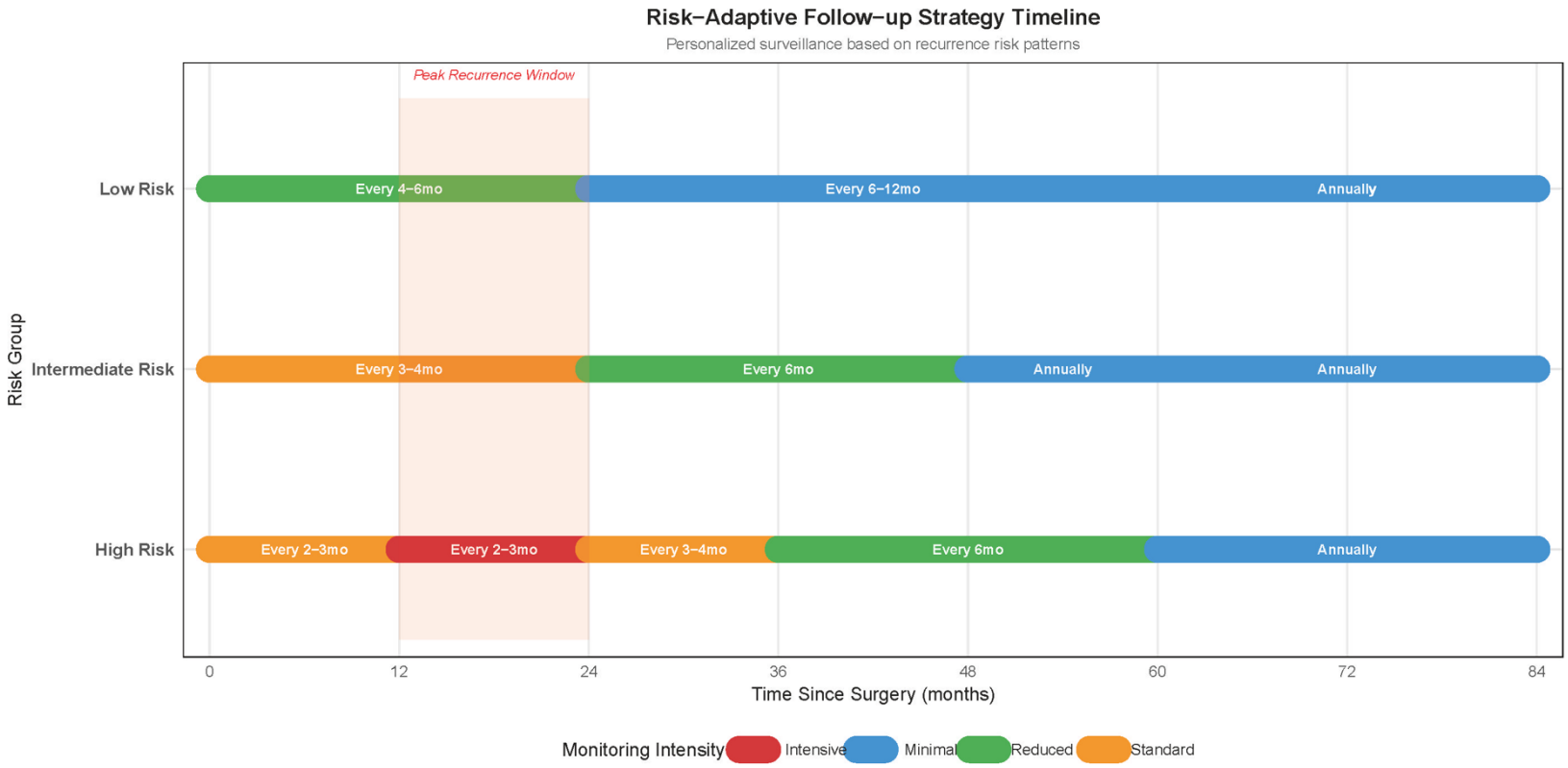

Figure 8. Risk-adaptive follow-up strategy timeline based on recurrence risk patterns. The image shows individualized follow-up strategies for patients in different risk groups, including the high-, intermediate-, and low-risk groups, across different postoperative time periods. Red represents intensive monitoring, orange represents standard monitoring, green represents reduced-frequency monitoring, and blue represents minimal-frequency monitoring. The shaded area marks the peak recurrence window from 12 to 24 months postoperatively. Note: mo, months. The peak recurrence window refers to the period of highest recurrence risk, from 12 to 24 months postoperatively.

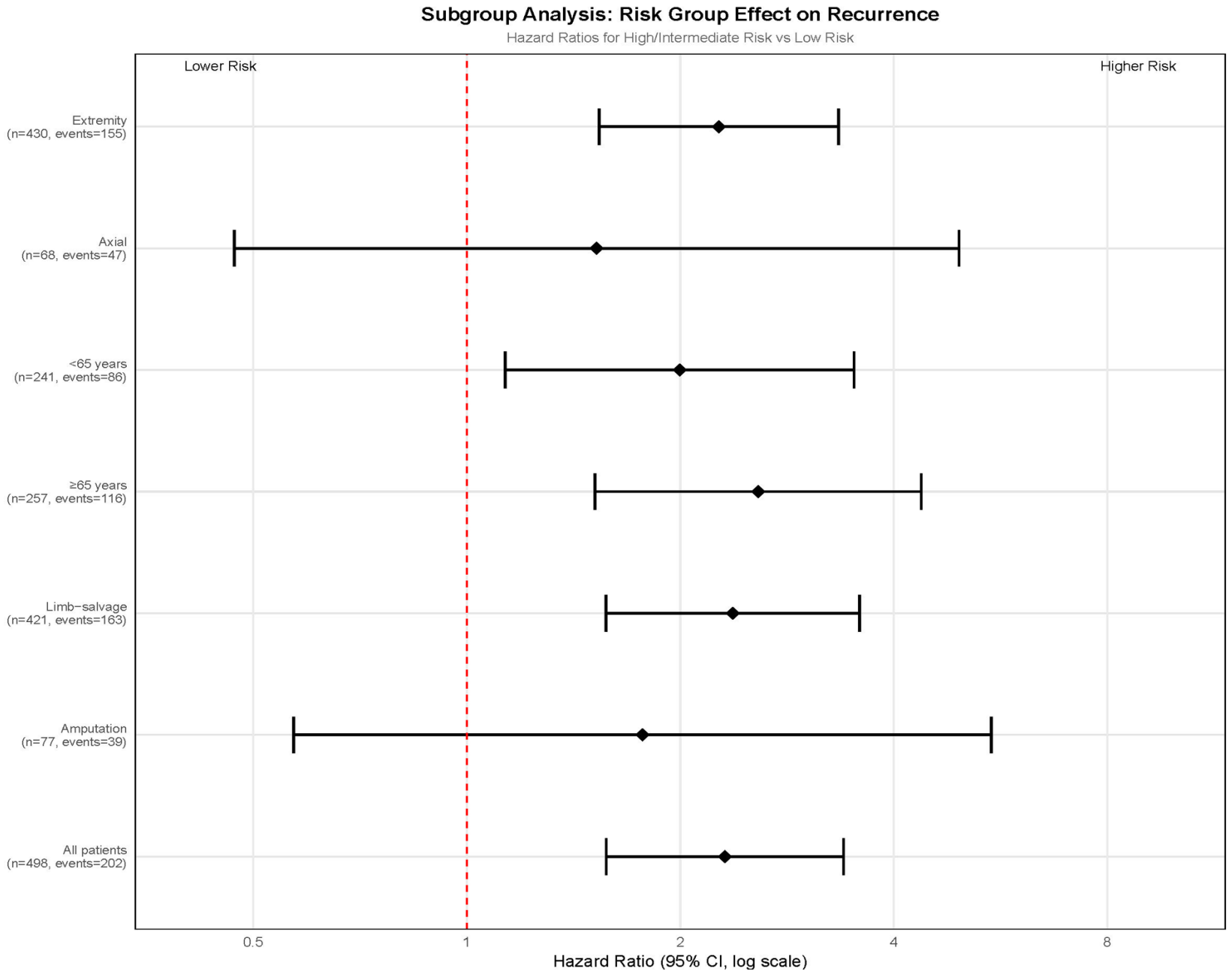
Subgroup Analysis: Risk Group Effect on Recurrence
Hazard Ratios for High/Intermediate Risk vs Low Risk
Lower Risk
Higher Risk
Extremity
(n=430, events=155)
Axial
(n=68, events=47)
<65 years
(n=241, events=86)
≥65 years
(n=257, events=116)
Limb−salvage
(n=421, events=163)
Amputation
(n=77, events=39)
All patients
(n=498, events=202)
0.5
1
2
4
8
Hazard Ratio (95% CI, log scale)

Figure 9. Forest plot of subgroup analysis for the effect of risk stratification on osteosarcoma recurrence. HRs and 95% CIs of the risk stratification model, comparing the intermediate-/high-risk group with the low-risk group across different clinical subgroups. Subgroups included tumor site, age, surgical method, and the overall population. The red dashed line indicates HR = 1, representing the reference line for no effect. Note: HR, hazard ratio; CI, confidence interval; n, sample size; events, number of recurrence events.

extremity osteosarcoma. In our study, patients with osteosarcoma located in the axial skeleton had a significantly higher risk of recurrence than patients with extremity osteosarcoma (HR = 2.34, 95% CI: 1.56-3.51). This may be related to the complex anatomy of the axial skeleton, which makes it difficult to achieve wide surgical margins.

Tumor diameter as a prognostic factor has been widely recognized. Previous literature showed that in patients with osteosarcoma under 60 years old, tumor diameter was an independent factor affecting overall survival [21]. Yu et al. [22] similarly incorporated tumor T stage into prediction models, confirming the close relationship between tumor burden and prognosis.

The prognostic value of chemotherapy response, as reflected by Huvos grading, was reconfirmed in this study. Zhang et al. [23] used dynamic contrast-enhanced MRI radiomics to predict preoperative chemotherapy effects, confirming the importance of chemotherapy response assessment. Previous literature showed that preoperative chemotherapy effects could be predicted through clinical and pathological characteristics [24]. Our findings support the use of chemotherapy response as an important reference indicator for postoperative risk stratification.

Notably, age did not show independent prognostic value in our multivariable analysis, which is inconsistent with some literature reports. Xue et al. [25] found age to be an independent prognostic factor in a large-sample study of 3,566 patients. Fu et al. [26], in their study of elderly patients with osteosarcoma, also emphasized the importance of age. This discrepancy may have several reasons.

First, our cohort comprised patients with primary osteosarcoma from two centers; the median age was 66 years, which may reflect an older patient population than in some prior studies, potentially affecting the generalizability of our findings to younger cohorts. Moreover, age may correlate with treatment intensity, chemotherapy completion, surgical method, and other factors. Its effects may be partially absorbed when predicting risk. Furthermore, differences in age cut-off values across studies may affect the results. Future studies may use restricted cubic splines and similar techniques to investigate the non-linear associations between age and recurrence risk.

Interpretation of time-dependent effects

One important finding of this study was the observation of time-dependent characteristics in the recurrence risk factors of osteosarcoma. Traditional Cox PH models assume that HRs are constant over follow-up. Nonetheless, Schoenfeld residual testing of our data demonstrated that two variables, surgical margin and adjuvant chemotherapy completion, violated the PH assumption ($P < 0.05$; **Figure 3**). Piecewise indicator functions were created as interaction terms. Based on clinical patterns of osteosarcoma recurrence and previous literature, the time periods were set as 0-12 months, 12-24 months, 24-36 months, 36-48 months, and > 48 months. In our sensitivity analyses, we tried an alternative segmentation of 0-6 months, 6-18 months, 18-36 months, 36-48 months, and > 48 months, and the main conclusions were unchanged.

The analysis results showed that the hazard effects of most risk factors were relatively strong at 12-24 months postoperatively, with attenuation trends after 36 months. However, this finding should be interpreted cautiously.

First, time-varying patterns were not completely consistent across different factors. For example, the effect of positive surgical margins attenuated relatively quickly, while the effect of poor chemotherapy response lasted relatively longer. Second, by 36 months, 214 of 341 patients (62.8%) remained in the risk set. The declining number of patients at risk in later periods contributed to wider confidence intervals around HR estimates, and reduced statistical power in later intervals may partly account

for the apparent attenuation of effects beyond 36-48 months.

Third, high-risk patients tended to recur early and exit the cohort, a phenomenon known as "susceptible depletion". This survivor bias may also cause a passive decline in later HRs.

From a biological perspective, this temporal pattern may be related to several factors. The systematic review by Rubio-San-Simón et al. [18] noted that RFI > 24 months was associated with better post-recurrence prognosis, suggesting that recurrence time itself reflects the biological behavior characteristics of tumors. Previous literature showed that the occurrence of lung metastasis in osteosarcoma is related to multiple clinical and pathological factors [27]. As the most common recurrence pattern, the natural course of lung metastasis may be related to the recurrence time distribution we observed.

Additionally, postoperative adjuvant chemotherapy typically lasts 6-12 months. The effect of chemotherapy on micrometastatic lesions and immune function recovery after treatment completion may jointly influence the recurrence time distribution. Previous literature indicated that molecular markers such as miR-34a are related to chemotherapy resistance and recurrence [28], suggesting that molecular mechanisms of chemotherapy sensitivity may affect recurrence temporal patterns. However, these mechanistic explanations currently remain speculative and require further basic and translational research for verification.

Performance and clinical value of the prediction model

Based on these findings, we constructed a risk prediction model integrating time-varying coefficients. The model development cohort consisted of 341 patients from Center A, including 135 recurrence events, and the external validation cohort consisted of 157 patients from Center B, including 67 recurrence events. Linear predictors from the multivariable Cox model combined with baseline hazard functions were used to calculate conditional recurrence probabilities at different time points.

Model discrimination was assessed using time-dependent ROC curves and the corresponding AUC values. Calibration was assessed using Brier scores and calibration plots. The model showed good discrimination in the training cohort, with AUC values for predicting 1-year, 3-year, and 5-year recurrence risk of 0.805 (95% CI: 0.747-0.864), 0.734 (95% CI: 0.681-0.788), and 0.923 (95% CI: 0.887-0.959), respectively (**Figure 6**). The calibration curves showed general agreement between predicted and observed probabilities, with a Brier score of 0.2799, suggesting that the prediction error was within an acceptable range, although this should be interpreted in the context of the event rate of approximately 40%. In the geographic external validation cohort, model performance was comparable, with AUC values of 0.860 (95% CI: 0.796-0.925), 0.775 (95% CI: 0.703-0.847), and 0.890 (95% CI: 0.791-0.989) at 1, 3, and 5 years, respectively, and a Brier score of 0.2600, indicating generalizability (**Figure 7**).

We acknowledge the risk of overfitting. The training cohort included 135 recurrence events. The model comprised five main-effect variables plus two time interaction terms (for surgical margin and adjuvant chemotherapy completion, the two variables that violated the PH assumption), with an effective parameter count of approximately 7. The event-to-parameter ratio (approximately 19:1) met empirical recommendations but overfitting risk cannot be fully excluded. To mitigate this risk, we adopted the following strategies: limiting the degrees of freedom of the time function, using piecewise constants rather than continuous splines, adding interaction terms only for variables that violated the PH assumptions, using bootstrap internal validation, and conducting geographic external validation. The predictive power of the model requires further testing in larger and more diverse populations. In contrast with earlier studies, Feng et al. [29] constructed an osteosarcoma postoperative prognosis nomogram based on the Surveillance, Epidemiology, and End Results database with C-index values of 0.731-0.734. Liu et al. [30] also developed a survival nomogram for osteosarcoma patients using SEER data with external validation. Gao et al. [31] developed a prediction model for survival after surgery and a risk stratification model for osteosarcoma. Yao et al. [32] reported nomograms for patients with high-grade osteosarcoma in China based on time to postoperative metastasis. Our model drew on the experiences documented in these

studies while introducing time-varying coefficients to account for changing risks over time. Notably, the above literature predominantly used overall survival or tumor-specific survival as the endpoint, whereas recurrence was used in our model. Each endpoint has its own clinical relevance and prediction objectives.

Regarding recurrence prediction specifically, Liu et al. [33] used CT radiomics to construct a multicenter model predicting 1-year recurrence risk, with a C-index reaching 0.779. Zhang et al. [34] showed that prognostic models based on RNA markers can predict post-recurrence survival. These studies explored recurrence prediction from radiomics and molecular marker perspectives, while our model, based on routine clinical and pathological variables, is easier to apply in resource-limited clinical settings.

DCA showed that within the threshold probability range of 0.1-0.7, our model had a positive net benefit compared with the "treat all" and "treat none" strategies in both the training cohort (net benefit 0.187 at a threshold of 0.3) and the external validation cohort (net benefit 0.227 at a threshold of 0.3) (**Figures 6**, **7**). In the DCA framework, "intervention" was defined as entering intensive follow-up pathways. The threshold probability corresponded to clinically acceptable recurrence risk levels for triggering intensive follow-up. It should be emphasized that DCA reflects the net benefit of adopting risk-adaptive follow-up strategies rather than the efficacy of specific treatment interventions. These results suggest that the model may assist clinical decision-making regarding individualized surveillance intensity.

Implications for follow-up strategies

Based on the stratification results of the risk prediction model, we offer preliminary recommendations for individualized follow-up strategies. It should be noted that the current European Society for Medical Oncology-European Reference Network for Rare Adult Solid Cancers-European Reference Network on Genetic Tumour Risk Syndromes guidelines and National Comprehensive Cancer Network guidelines already provide framework recommendations for osteosarcoma follow-up. These include regular physical examinations, imaging examinations of the primary site (X-ray/CT/MRI), and chest imaging examinations (X-ray/CT), with allowances for adjustments based on individual risk and clinical needs.

At present, there is insufficient high-quality evidence from prospective cohort studies regarding the optimal monitoring frequency based on quantified risks. According to the systematic review by Rubio-San-Simón et al. [18], RFI and complete surgical resection are important factors affecting post-recurrence prognosis, suggesting that early detection of recurrence and complete resection may help improve prognosis. According to previous literature, risk factors for distant metastasis include age, site, grade, T stage, and surgical method [35]. These factors may serve as references for developing personalized monitoring strategies. According to our findings, we tentatively put forth the following recommendations for clinical reference. Patients predicted by the model to be at high risk, such as those with positive margins, poor chemotherapy response, axial skeleton location, or multiple concurrent risk factors, may be considered for appropriately increased monitoring frequency during the 12-24-month postoperative peak period for recurrence. In low-risk patients, monitoring intervals may be safely extended to minimize unnecessary radiation exposure and medical costs.

The premise of benefit from intensive monitoring strategies is that outcomes can be changed by subsequent intervention. Examples include resectable isolated lung metastases, patients who are candidates for clinical trials, or lesions that can receive local ablation. If subsequent interventions are not effective, early detection may not lead to survival benefit. Surveillance guidelines need to take into consideration the patient's individual risk level, benefit from early diagnosis, risk of radiation exposure, such as whether low-dose CT may be considered, ease of access to medical resources, patient preference, and shared decision-making. It should be reiterated that the these recommendations are derived from retrospective data model predictions without prospective research verification. To apply risk stratification strategies in everyday clinical practice, randomized controlled trials are required to compare the clinical benefits and cost-effectiveness of individualized versus standard follow-up.

Study limitations

This study has several limitations that need to be acknowledged. First, as a retrospective study, it is inherently subject to selection bias

and information bias. Although strict inclusion and exclusion criteria were applied, residual confounding cannot be completely ruled out. In particular, the variable "chemotherapy completion status" may be subject to reverse causation bias. Patients with early recurrence or disease progression may have more difficulty completing standard chemotherapy regimens. To mitigate this bias, we set the determination time point for chemotherapy completion status at 6 months postoperatively and clearly defined operational criteria. To avoid time bias from recurrences occurring within 0-6 months affecting exposure assessment, we performed a 6-month landmark analysis, including only patients without recurrence at 6 months postoperatively; the results were directionally consistent with the main analysis. Sensitivity analysis using extended Cox models with chemotherapy completion treated as a time-dependent covariate also supported the main conclusions. Second, study data came from two centers with limited geographic representation. Treatment protocols, surgical techniques, and follow-up standards may differ between the two centers. Although we adjusted for center effects in the analysis, residual heterogeneity cannot be completely eliminated. Future research could use stratified models or random-effects models to assess coefficient heterogeneity between centers or to test interactions between centers and major risk factors. Additionally, the long follow-up time span (2010-2019) may have seen evolution in chemotherapy protocols and surgical concepts during this period. Third, this study did not incorporate molecular biology markers at the variable level. Studies demonstrate that prognostic indices of the tumor microenvironment predict survival in osteosarcoma patients [36]. According to Tang et al. [37], there are programmed cell death-related genes that have the potential to become prognostic indicators. According to previous literature, molecular subtypes based on protein synthesis predict prognosis and treatment response in osteosarcoma [38]. Integration of these molecular markers may further improve model predictive performance. Fourth, this study combined local recurrence and distant metastasis into a single endpoint, but risk factors and time distributions for the two may differ. Future research could consider separate modeling or the use of competing risk models. Additionally, non-tumor-related deaths as competing risks were not handled separately, which may have some impact on recurrence risk estimation, although this proportion was low in our young patient cohort (< 5%). Fifth, regarding missing data, this study used complete case analysis, excluding patients with missing key variables (~8%). If missingness was not completely random (missing not at random), complete case analysis may lead to underestimation or overestimation of hazard ratios. Future research could consider using multiple imputation or other methods to handle missing data and conduct sensitivity analyses to evaluate the impact of missingness mechanisms. Finally, interpretation of time-dependent effects requires caution. At 36 months, only 32% of patients remained in the risk set. Insufficient statistical power is an important reason for the apparent "weakening" of later effects. In addition, the specification of the time function may affect observed patterns. Sensitivity analyses using alternative segmentation schemes yielded similar results. Hence, in future research, it may be worthwhile to try more flexible modeling methods, such as spline functions, to further verify the results.

Future research directions and conclusions

Future research should be conducted in several directions. It is suggested that multicenter prospective cohort studies be conducted to verify model generalizability. It is also necessary to compare the clinical benefit and cost-effectiveness of individualized follow-up strategies versus standard protocols in randomized controlled trials. In addition, molecular markers can be integrated into prediction models. According to literature [28], microRNAs such as miR-34a are related to chemotherapy resistance and recurrence. Fu et al. [26] found that the fibrinogen-to-albumin ratio can predict prognosis in elderly osteosarcoma patients. According to literature [39], nutritional inflammatory scores, such as the Glasgow Prognostic Score and Controlling Nutritional Status, are linked with osteosarcoma prognosis. The inclusion of these novel markers is expected to improve prediction models. Future studies should investigate the biological mechanisms of time-dependent effects. According to literature [9], the characteristics of the osteosarcoma immune microenvironment impact treatment response and prognosis. The latest ad-

vances in osteosarcoma immunotherapy were summarized by Yu et al. [40]. A thorough understanding of the relationships between immune microenvironment dynamic changes and recurrence time distribution may provide the basis for new treatment strategies. Finally, we propose creating clinical decision tools that are convenient to use by making prediction models available as web calculators or mobile apps for clinical use.

Conclusion

To sum up, we utilized a high-grade osteosarcoma patient cohort from two centers to identify independent risk factors for recurrence. It appears that hazard effects of some factors change over time. The prediction model we established demonstrates certain discrimination and calibration through external validation, which may provide a reference for formulating personalized follow-up plans. The first 24 months postoperatively are an important period for recurrence monitoring. High-risk patients, when effective subsequent interventions are available, may benefit from more intensive monitoring. However, this strategy requires further verification through prospective research.

Disclosure of conflict of interest

None.

Address correspondence to: Yi Feng, Department of Orthopedics, Second Hospital of Shanxi Medical University, No. 382 Wuyi Road, Taiyuan 030001, Shanxi, China. E-mail: fengyi160@126.com

References

[1] Siegel RL, Giaquinto AN and Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024; 74: 12-49.

[2] Biermann JS, Hirbe A, Ahlawat S, Bernthal NM, Binitie O, Boles S, Brigman B, Callan AK, Cipriano C, Cranmer LD, Davis J, Donnelly E, Ferguson M, Graham A, Groundland J, Hess M, Hiniker SM, Hoover-Regan ML, Hornick JL, Jonard B, Kuechle JB, Lindskog D, Mayerson JL, McGarry SV, Morris CD, Olson D, Rose PS, Santana VM, Satcher RL, Schwartz H, Shulman RM, Thorpe SW, Wilky BA, Wustrack RL, Yoon J, Hang LE, Jones F, Sansone N and Lyons M. Bone cancer, version 2.2025, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2025; 23: e250017.

[3] Hosseini H, Heydari S, Hushmandi K, Daneshi S and Raesi R. Bone tumors: a systematic review of prevalence, risk determinants, and survival patterns. BMC Cancer 2025; 25: 321.

[4] Adewuyi E, Chorya H, Muili A, Moradeyo A, Kayode A, Naik A, Odedele T and Opabode M. Chemotherapy, immunotherapy, and targeted therapy for osteosarcoma: recent advancements. Crit Rev Oncol Hematol 2025; 206: 104575.

[5] Brar GS, Schmidt AA, Willams LR, Wakefield MR and Fang Y. Osteosarcoma: current insights and advances. Explor Target Antitumor Ther 2025; 6: 1002324.

[6] Kager L, Zoubek A, Dominkus M, Lang S, Bodmer N, Jundt G, Klingebiel T, Jürgens H, Gadner H and Bielack S; COSS Study Group. Osteosarcoma in very young children: experience of the cooperative osteosarcoma study group. Cancer 2010; 116: 5316-5324.

[7] Strauss SJ, Frezza AM, Abecassis N, Bajpai J, Bauer S, Biagini R, Bielack S, Blay JY, Bolle S, Bonvalot S, Boukovinas I, Bovee JVMG, Boye K, Brennan B, Brodowicz T, Buonadonna A, de Álava E, Dei Tos AP, Garcia Del Muro X, Dufresne A, Eriksson M, Fagioli F, Fedenko A, Ferraresi V, Ferrari A, Gaspar N, Gasperoni S, Gelderblom H, Gouin F, Grignani G, Gronchi A, Haas R, Hassan AB, Hecker-Nolting S, Hindi N, Hohenberger P, Joensuu H, Jones RL, Jungels C, Jutte P, Kager L, Kasper B, Kawai A, Kopeckova K, Krákorová DA, Le Cesne A, Le Grange F, Legius E, Leithner A, López Pousa A, Martin-Broto J, Merimsky O, Messiou C, Miah AB, Mir O, Montemurro M, Morland B, Morosi C, Palmerini E, Pantaleo MA, Piana R, Piperno-Neumann S, Reichardt P, Rutkowski P, Safwat AA, Sangalli C, Sbaraglia M, Scheipl S, Schöffski P, Sleijfer S, Strauss D, Sundby Hall K, Trama A, Unk M, van de Sande MAJ, van der Graaf WTA, van Houdt WJ, Frebourg T, Ladenstein R, Casali PG and Stacchiotti S; ESMO Guidelines Committee, EURACAN, GENTURIS and ERN PaedCan. Electronic address: clinicalguidelines@esmo.org. Bone sarcomas: ESMO-EURACAN-GENTURIS-ERN paedcan clinical practice guideline for diagnosis, treatment and follow-up. Ann Oncol 2021; 32: 1520-1536.

[8] Beird HC, Bielack SS, Flanagan AM, Gill J, Heymann D, Janeway KA, Livingston JA, Roberts RD, Strauss SJ and Gorlick R. Osteosarcoma. Nat Rev Dis Primers 2022; 8: 77.

[9] Han Z, Chen G and Wang D. Emerging immunotherapies in osteosarcoma: from checkpoint blockade to cellular therapies. Front Immunol 2025; 16: 1579822.

[10] Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H and Altman DG; PROGRESS

Group. Prognosis research strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013; 10: e1001381.
[11] Zhang C, Guo X, Xu Y, Han X, Cai J, Wang X and Wang G. Lung metastases at the initial diagnosis of high-grade osteosarcoma: prevalence, risk factors and prognostic factors. A large population-based cohort study. Sao Paulo Med J 2019; 137: 423-429.
[12] Riley RD, Ensor J, Snell KIE, Harrell FE Jr, Martin GP, Reitsma JB, Moons KGM, Collins G and van Smeden M. Calculating the sample size required for developing a clinical prediction model. BMJ 2020; 368: m441.
[13] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L and Logullo P. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024; 385: e078378.
[14] Martin GP, Riley RD, Ensor J and Grant SW. Statistical primer: sample size considerations for developing and validating clinical prediction models. Eur J Cardiothorac Surg 2025; 67: ezaf142.
[15] Tirtei E, Michelsen SW, Haveman LM, Meazza C, Oliveira JF, Rasool A, Palmerini E, Wilson W, Gaspar N, Strauss SJ, Papakonstantinou A and Baecklund F; FOSTER Consortium (Fight OsteoSarcoma Through European Research), work package 4 on trials in newly diagnosed osteosarcoma. Prognostic factors in newly diagnosed high-grade osteosarcoma-a systematic review. Cancer Med 2025; 14: e71044.
[16] Sever N, Şimşek F, Onur İD, Arvas H, Guliyev T, Şakalar T, Çiçek CM, Orman S, Çetin EB, Kayaş K, Akbaş S, Ağyol Y, Güren AK, Erel P, Kocaaslan E, Paçacı B, Tunç MA, Çelebi A, Majidova N, Durnalı A, Şimşek M, Şahbazlar M, Işık S, Arıkan R, Ercelep Ö, Sarı M, Köstek O and Bayoğu İ V. Prognostic factors in high grade osteosarcoma patients who received neoadjuvant therapy and subsequently underwent surgery: data from the turkish oncology group. J Clin Med 2025; 14: 2024.
[17] Franco D'Souza R and Surapaneni KM. New standards, new vision: directing modern medical research as per the 2024 revised declaration of Helsinki. Asian J Psychiatr 2025; 104: 104363.
[18] Rubio-San-Simón A, Wilson W, Sironi G, le Deley MC, Palmerini E, Gaspar N, Piperno-Neumann S, Hecker-Nolting S, Strauss S, van Ewijk R and Meazza C; FOSTER Consortium (Fight OsteoSarcoma Through European Research), work package 3 on recurrent/refractory osteosarcoma trials. Prognostic factors in patients with relapsed high-grade osteosarcoma: a systematic review. Br J Cancer 2025; 133: 1020-1028.
[19] Huang Z, Huang C, Wang Y, Wu Y, Guo C, Li W and Kong Q. Clinical features, risk factors, and prediction nomogram for primary spinal osteosarcoma: a large-cohort retrospective study. Global Spine J 2024; 14: 930-940.
[20] Huang Y, Wang C, Tang D, Chen B and Jiang Z. Development and validation of nomogram-based prognosis tools for patients with extremity osteosarcoma: a SEER population study. J Oncol 2022; 2022: 9053663.
[21] Xu Y, Yan Q, Yang J, Cheng M, Chen J, Yao Y and Liu Y. Constructing a nomogram for predicting pelvic osteosarcoma: a retrospective study based on the SEER database and a Chinese cohort. Br J Hosp Med (Lond) 2024; 85: 1-17.
[22] Yu Y, Wang S, Liu J, Ge J and Guan H. Development and validation of a nomogram to predict long-term cancer-specific survival for patients with osteosarcoma. Sci Rep 2023; 13: 10230.
[23] Zhang L, Gao Q, Dou Y, Cheng T, Xia Y, Li H and Gao S. Evaluation of the neoadjuvant chemotherapy response in osteosarcoma using the MRI DWI-based machine learning radiomics nomogram. Front Oncol 2024; 14: 1345576.
[24] Yang F, Feng Y, Sun P, Traverso A, Dekker A, Zhang B, Huang Z, Wang Z and Yan D. Preoperative prediction of high-grade osteosarcoma response to neoadjuvant therapy based on a plain CT radiomics model: a dual-center study. J Bone Oncol 2024; 47: 100614.
[25] Xue W, Zhang Z, Yu H, Li C, Sun Y, An J, Qi L, Zhang J and Liu Q. Development of nomogram and discussion of radiotherapy effect for osteosarcoma survival. Sci Rep 2023; 13: 223.
[26] Fu Y, Yu Y, Zhou Y, Li T, Xie Y, Wang Y, Ran Q, Chen Y and Fan X. The fibrinogen-albumin ratio as a novel prognostic factor for elderly patients with osteosarcoma. Medicine (Baltimore) 2023; 102: e34926.
[27] Su Z, Huang F, Yin C, Yu Y and Yu C. Clinical model of pulmonary metastasis in patients with osteosarcoma: a new multiple machine learning-based risk prediction. J Orthop Surg (Hong Kong) 2023; 31: 10225536231177102.
[28] Lian H, Zhou Y, Sun Z and Liu K. MicroRNA34a is associated with chemotherapy resistance, metastasis, recurrence, survival, and prognosis in patient with osteosarcoma. Medicine (Baltimore) 2022; 101: e30722.
[29] Feng L, Chen Y, Ye T, Shao Z, Ye C and Chen J. Development and validation of an online prog-

nostic nomogram for osteosarcoma after surgery: a retrospective study based on the SEER database and external validation with single-center data. Transl Cancer Res 2022; 11: 3156-3174.
[30] Liu Z, Xie Y, Zhang C, Yang T and Chen D. Survival nomogram for osteosarcoma patients: SEER data retrospective analysis with external validation. Am J Cancer Res 2023; 13: 900-911.
[31] Gao B, Wang MD, Li Y and Huang F. Risk stratification system and web-based nomogram constructed for predicting the overall survival of primary osteosarcoma patients after surgical resection. Front Public Health 2022; 10: 949500.
[32] Yao Z, Tan Z, Yang J, Yang Y, Wang C, Chen J, Zhu Y, Wang T, Han L, Zhu L and Yang Z. Prognostic nomogram for predicting 5-year overall survival in Chinese patients with high-grade osteosarcoma. Sci Rep 2021; 11: 17728.
[33] Liu J, Lian T, Chen H, Wang X, Quan X, Deng Y, Yao J, Lu M, Ye Q, Feng Q and Zhao Y. Pretreatment prediction of relapse risk in patients with osteosarcoma using radiomics nomogram based on CT: a retrospective multicenter study. Biomed Res Int 2021; 2021: 6674471.
[34] Zhang M, Liu Y and Kong D. Identifying biomolecules and constructing a prognostic risk prediction model for recurrence in osteosarcoma. J Bone Oncol 2020; 26: 100331.
[35] Lu S, Wang Y, Liu G, Wang L, Wu P, Li Y and Cheng C. Construction and validation of nomogram to predict distant metastasis in osteosarcoma: a retrospective study. J Orthop Surg Res 2021; 16: 231.
[36] Wu C, Gong S, Duan Y, Deng C, Kallendrusch S, Berninghausen L, Osterhoff G and Schopow N. A tumor microenvironment-based prognostic index for osteosarcoma. J Biomed Sci 2023; 30: 23.
[37] Tang Z, Zhang Z, Zhao J, Zhang F, Zhang Y, Wen Y, Li M, Sun J, Shi L, Chen W, Li Z, Guo Z and Liu Y. Integrated analysis of multiple programmed cell death-related prognostic genes and functional validation of apoptosis-related genes in osteosarcoma. Int J Biol Macromol 2025; 307: 142113.
[38] Yi C, Liu J, Zhao S, Gong D, Xu B, Li A, Bian E and Tian D. Identification of a pro-protein synthesis osteosarcoma subtype for predicting prognosis and treatment. Sci Rep 2024; 14: 16475.
[39] Yang Y, Zhou C and Ma X. Use of nomogram on nutritional assessment indicators to predict clinical outcomes in patients undergoing surgical resection for high-grade osteosarcoma. Nutr Cancer 2022; 74: 3564-3573.
[40] Yu S and Yao X. Advances on immunotherapy for osteosarcoma. Mol Cancer 2024; 23: 192.