Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A predictive model for adjuvant chemotherapy benefit in locally advanced nasopharyngeal carcinoma integrating post-treatment multi-sequence MRI radiomics and dynamic clinical biomarkers: a retrospective single-center cohort study

Lin Wang
Hainan Medical University
Huali Li
Southern Medical University
Sichao Liang
Nanfang Hospital
Jinmei Li
Hainan Medical University
Yanting Zhang
Nanfang Hospital
Rongrong Shi
Nanfang Hospital
Shenghui Wang
Nanfang Hospital
dejia Liu
Nanfang Hospital
Liaoming Gao
Nanfang Hospital
Zhiting Chen
Tumor Hospital of Guangxi Medical University
Menhan Xie
Hainan Medical University
Xiuling Li
Hainan Modern Women and Children's Hospital
Ting Song
Southern Medical University

Bei Chen

cb18789235270@163.com

Nanfang Hospital

Research Article

Keywords: Nasopharyngeal carcinoma, Radiomics, Adjuvant chemotherapy, Epstein-Barr virus DNA, MRI, Nomogram

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11009295/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

A predictive model for adjuvant chemotherapy benefit in locally advanced nasopharyngeal carcinoma integrating post-treatment multi-sequence MRI radiomics and dynamic clinical biomarkers: a retrospective single-center cohort study

Abstract

Background: The benefit of adjuvant chemotherapy (AC) for locoregionally advanced nasopharyngeal carcinoma (LA-NPC) is uncertain. We aimed to develop a model integrating post-treatment MRI radiomics and Epstein-Barr virus (EBV) DNA to identify patients who benefit from AC.

Methods: We retrospectively enrolled 267 patients with LA-NPC treated at a single institution. A radiomic-clinical signature (RCS) for predicting 2-year progression-free survival (PFS) was constructed using radiomic features extracted from post-concurrent chemoradiotherapy (CCRT) multi-sequence MRI and dynamic EBV DNA data. Model performance was evaluated using Harrell's concordance index (C-index). Patients were stratified into high- and low-risk subgroups based on the RCS, and the survival benefit of AC was assessed using Kaplan-Meier analysis.

Results: The cohort was randomly divided into training (n = 213) and testing (n = 54) sets. The fused RCS model achieved C-indices of 0.899 and 0.719 for predicting 2-year PFS in the training and testing sets, respectively, outperforming models using clinical (0.583) or radiomic (0.695) features alone. Risk stratification by the RCS identified low- and high-risk groups with significantly different survival outcomes (all $p < 0.05$). Subgroup analysis showed that AC significantly improved 2-year PFS only in the high-risk group (94.3% vs. 87.0%, $p = 0.045$), with no benefit observed in low-risk patients.

Conclusions: The RCS model effectively stratifies patients with LA-NPC and may guide personalized AC decisions by identifying high-risk patients likely to benefit.

Trial registration: Not applicable. This was a retrospective observational study and was not registered as a clinical trial.

Keywords

Nasopharyngeal carcinoma; Radiomics; Adjuvant chemotherapy; Epstein-Barr virus DNA; MRI; Nomogram

Background

Nasopharyngeal carcinoma (NPC) exhibits a distinct geographical distribution, with more than 70% of incident cases clustered in East and Southeast Asia. A majority (60-70%) of patients are diagnosed with locoregionally advanced disease (LA-NPC, stage III-IVA), for which concurrent chemoradiotherapy (CCRT) serves as the primary treatment. Despite this, treatment failure occurs in 20-30% of LA-NPC patients, with distant metastases responsible for 60-70% of these failures and representing the leading cause of death [1, 2]. This pattern suggests that systemic micrometastases are often present at diagnosis and may not be adequately controlled by localized CCRT. Consequently, optimizing systemic therapy, specifically by defining the role of adjuvant chemotherapy (AC), is essential to enhance outcomes and mitigate distant metastasis risk in LA-NPC.

Although clinical evidence suggests a benefit for AC, its application is limited. A prospective study in high-risk LA-NPC patients found adjuvant capecitabine reduced treatment failure risk by 45% and severe toxicity by 60% [3]. Similarly, a randomized trial showed improved 5-year failure-free survival with adjuvant capecitabine (78.5% vs. 65.9%, $p = 0.03$) [4]. However, clinical uptake of AC remains constrained, primarily due to suboptimal treatment compliance and significant toxicity. Key factors contributing to low compliance include severe adverse effects associated with traditional platinum-based regimens; historical analyses indicate that only 50-60% of patients complete the intended 3-4 cycles, often due to hematologic toxicities and renal impairment [5, 6].

The identification of high-risk NPC patients for AC in previous studies has been largely based on conventional parameters such as TNM stage and baseline EBV DNA. Despite this, significant prognostic variability is commonly observed within the same stage. A case in point is the notably worse outcome of T1-2N1 patients compared with T2N0 patients, a finding that challenges the precision of the TNM system alone for risk stratification [7]. To address this, the AJCC 8th edition has integrated extranodal extension (ENE) into the N3 category, refining the assessment of aggressiveness [8]. Meanwhile, while pre-treatment EBV DNA is a cornerstone biomarker, reliance on a solitary baseline measurement is suboptimal, as it fails to capture the dynamic

nature of disease risk throughout treatment, potentially leading to misaligned therapy. This underscores the need for dynamic biomarker monitoring, with serial EBV DNA tracking emerging as a key strategy for enabling real-time, personalized risk prediction and treatment adjustment [9].

Beyond conventional clinical factors such as plasma EBV DNA, medical imaging offers a crucial and independent dimension for prognostic stratification and treatment guidance in NPC [10, 11]. Magnetic resonance imaging (MRI), given its excellent soft-tissue contrast and multi-parametric nature, serves as the primary imaging modality [8]. The emerging field of radiomics, which involves the high-throughput extraction and analysis of quantitative image features, is proving valuable for predicting treatment response and risk in NPC. Evidence indicates that integrating MRI-derived radiomic features with clinical data (TNM stage, EBV DNA) enhances the prediction of outcomes in LA-NPC [12-17]. Relevant work includes a multicenter study demonstrating that an MRI radiomics-EBV DNA model effectively stratified prognosis and predicted survival in T3N1M0 patients under varying treatments [18]. Specifically regarding AC, Wu et al. developed a radiomic model based on pre-radiotherapy MRI, which successfully identified a high-risk subgroup with significantly inferior 3-year recurrence-free survival, suggesting a potential role for such models in selecting candidates for adjuvant therapy [19].

However, existing MRI radiomics studies for predicting AC benefit in LA-NPC have important shortcomings. Most prior work has depended on static pre-treatment parameters, including baseline MRI, clinical indicators such as EBV DNA, and TNM staging. Consequently, these models may not adequately reflect the dynamic biological changes that tumors undergo throughout treatment and how these changes influence patient outcomes. Therefore, we designed the present study to develop and internally validate a comprehensive model that fuses multi-sequence MRI features (T1WI, T2WI, and CE-T1WI) obtained after CCRT with serial plasma EBV DNA measurements (tracking levels and trends across the treatment timeline). This integrative approach was intended to enable a more precise identification of high-risk patients who are truly likely to benefit from AC, ultimately advancing toward personalized management in LA-NPC.

Methods

Study population and design

This was a retrospective, single-center prognostic modelling study conducted at Nanfang Hospital, Southern Medical University (Guangzhou, Guangdong, China). The aim was to develop and internally validate a radiomic-clinical signature (RCS) that predicts progression-free survival (PFS) and identifies patients with LA-NPC who are likely to benefit from AC. The study was approved by the Medical Ethics Committee of Nanfang Hospital, Southern Medical University (approval number: NFEC-2024-367) and was conducted in accordance with the Declaration of Helsinki. The requirement for written informed consent was waived by the ethics committee because of the retrospective design and the use of de-identified data.

We included 267 patients with pathologically confirmed, previously untreated LA-NPC (AJCC 8th edition, stage III-IVA) treated between January 2017 and December 2022. Key inclusion criteria comprised: (i) availability of post-CCRT nasopharyngeal MRI scans (1-3 months after treatment, including T1WI, T2WI, and CE-T1WI sequences); (ii) completion of standard systemic therapy (>=1 cycle of induction chemotherapy [IC] and >=1 cycle of concurrent chemotherapy), with patients in the AC group additionally receiving >=2 cycles of traditional AC or >=4 cycles of capecitabine-based AC; and (iii) availability of complete longitudinal plasma EBV DNA data at baseline, post-IC, and post-CCRT. Patients were randomly assigned to a training set (n = 213, 80%) and an independent testing set (n = 54, 20%).

Data collection and EBV DNA kinetics

Comprehensive clinical data were collected, including demographics, TNM stage, treatment details, laboratory parameters (post-CCRT blood counts, liver and renal function), tumor volumes (GTVp, GTVnd), and serial plasma EBV DNA copies. Serial plasma EBV DNA was measured at three predefined timepoints: T0, baseline (pre-treatment); T1, within 2 weeks after induction chemotherapy; and T2, within 4 weeks after concurrent chemoradiotherapy. Based on the dynamic trajectory of EBV DNA across these timepoints, patients were categorized into four distinct kinetic groups according to the following standardized definitions: (1) sustained negative (reference group), EBV DNA undetectable at all three timepoints (T0-, T1-, T2-); (2) rapid clearance, EBV DNA detectable at T0 and undetectable by T1 and at T2 (T0+, T1-, T2-); (3) delayed clearance, EBV DNA detectable at T0 and T1 and undetectable by T2 (T0+, T1+, T2-);

and (4) sustained positive/rebound, EBV DNA detectable at T2 regardless of T0 or T1 status (T2+).

MRI acquisition and tumor segmentation

Post-CCRT axial MRI sequences (CE-T1WI, T1WI, and T2WI) were acquired according to a standardized protocol. Using MIM software (version 3.4; MIM Software Inc., Cleveland, OH, USA), co-registered images were manually segmented. On post-CCRT CE-T1WI, the gross tumor volume of the primary lesion (GTVp) and the gross tumor volume of involved lymph nodes (GTVnd) were delineated slice by slice, with reference to pre-CCRT anatomy and incorporating a 5 mm peritumoral margin. These contours were then propagated to the co-registered T1WI and T2WI sequences, generating three distinct regions of interest (ROIs) per patient. An example is shown in Fig. 1.

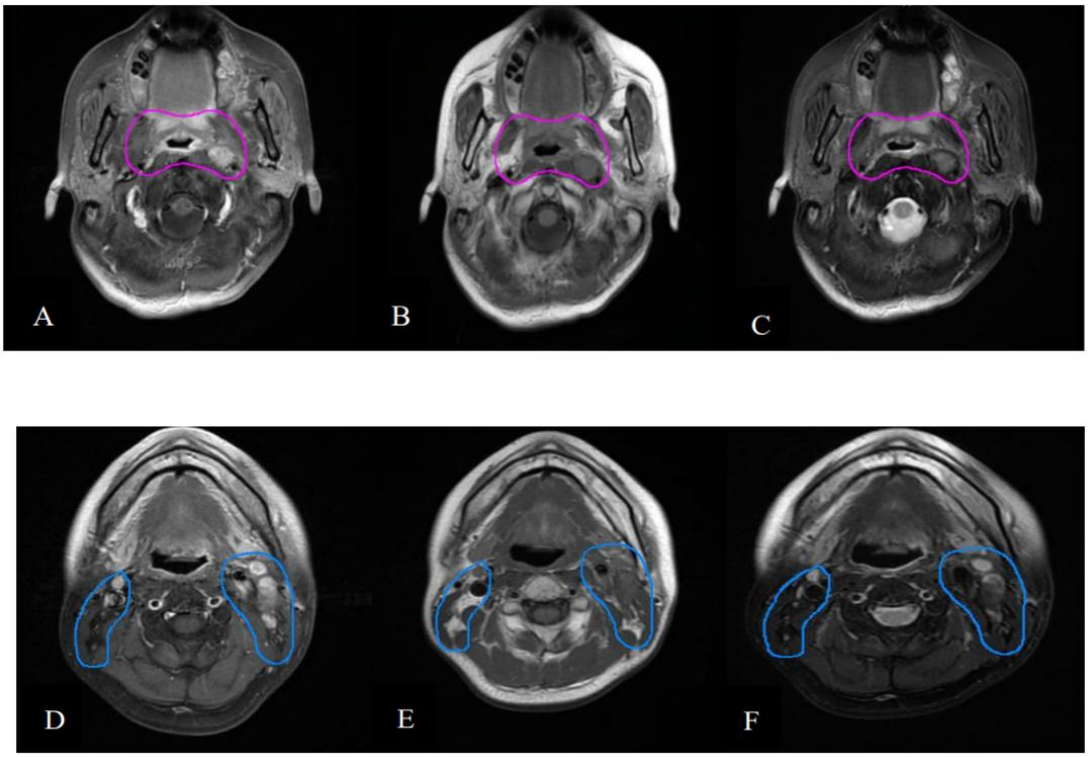


Figure 1. Demonstration of ROI delineation for gross tumor volume. (A-C) GTVp-ROI delineation on CE-T1WI, T1WI, and T2WI sequences. (D-F) GTVnd-ROI delineation on the corresponding sequences. ROI, region of interest; GTVp, gross tumor volume of the primary tumor; GTVnd, gross tumor volume of involved lymph nodes; CE-T1WI, contrast-enhanced T1-weighted imaging; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.

Feature extraction and signature construction

A total of 4,944 radiomic features were extracted from GTVp and GTVnd ROIs on T2WI and CE-T1WI images using the PyRadiomics package. A rigorous, multi-step feature selection pipeline was employed to build a radiomic signature (RS) predictive of PFS: (1) univariable Cox regression ($p < 0.05$); (2) redundancy removal via Spearman correlation (threshold > 0.8); (3) LASSO Cox regression with 5-fold cross-validation; and (4) stepwise AIC selection to mitigate multicollinearity. The final model was constructed using the EasyEnsemble algorithm combined with Cox regression. The radiomic signature risk score (Rscore) for each patient was subsequently calculated as the linear combination of the selected features weighted by their respective LASSO-Cox regression coefficients. The optimal cut-off value of the Rscore for risk stratification was determined by receiver operating characteristic (ROC) curve analysis in the training set, using the Youden index to maximize sensitivity and specificity. The overall workflow of radiomic feature extraction, signature construction, model development, and clinical application is summarized in Fig. 2.

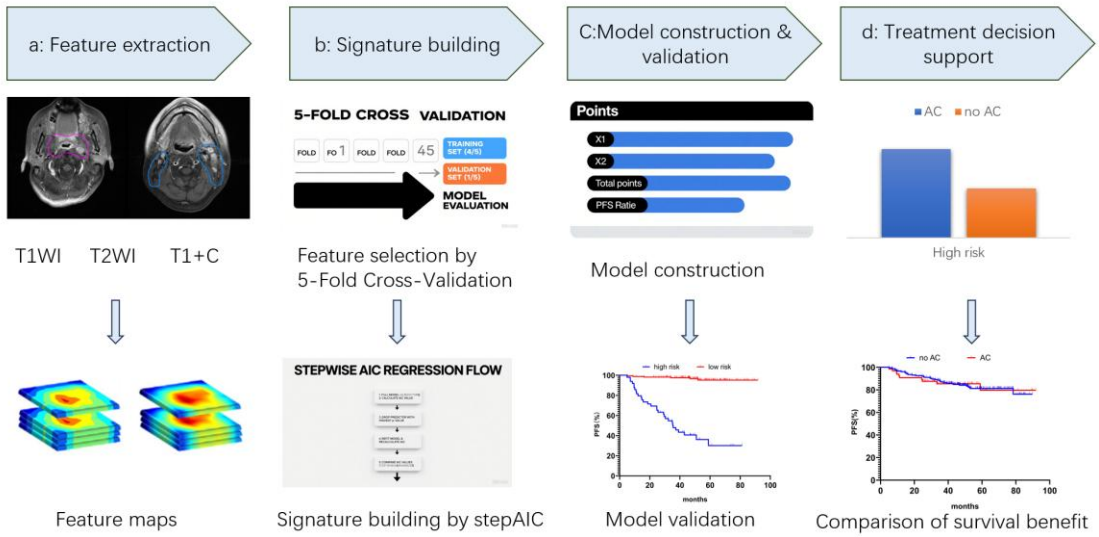


Figure 2. Workflow for establishing the radiomics nomogram. The pipeline comprises post-CCRT MRI acquisition, tumor segmentation, radiomic feature extraction, feature selection, signature construction, model integration, and clinical application. CCRT, concurrent chemoradiotherapy; MRI, magnetic resonance imaging.

Model integration and statistical analysis

The RCS was developed by integrating the RS with clinical features selected via univariable Cox regression ($p < 0.01$) and stabilized through 1000 bootstrap iterations. Model performance was evaluated using Harrell's C-index and visualized with a nomogram. The endpoints assessed were

PFS (primary endpoint), overall survival (OS), locoregional relapse-free survival (LRRFS), and distant metastasis-free survival (DMFS). Statistical analyses were performed using R version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria) and IBM SPSS Statistics version 28.0 (IBM Corp., Armonk, NY, USA). Survival curves were generated using the Kaplan-Meier method and compared with the log-rank test. A two-sided p-value < 0.05 was considered statistically significant. No a priori power calculation was performed: because this was a retrospective analysis of a fixed, consecutively enrolled cohort, the available sample size (n = 267) determined the precision of the estimates.

No generative artificial intelligence or AI-assisted technologies were used in the preparation of this work.

Results

Patient characteristics

The study included 267 patients with LA-NPC treated between 2017 and 2022, comprising 202 patients in the IC+CCRT group and 65 in the IC+CCRT+AC group. Baseline patient characteristics for both the training and testing sets are presented in Table 1, with no statistically significant differences observed between the sets across key demographics, tumor features, or EBV DNA status. Furthermore, a comparison of baseline features between the treatment subgroups (IC+CCRT vs. IC+CCRT+AC), provided in Additional file 1 (Table S1), revealed no significant disparities in factors such as sex, T stage, N stage, or tumor volume.

Table 1. Demographic and key clinical characteristics of the training and testing cohorts.

Characteristics	Training set(N=213)	Testing set(N=54)	*P* value
Age(years)			0.438
>36	163(76.5%)	44(81.5%)	
<36	50(23.5%)	10(18.5%)	
Sex			0.292
Male	146(68.5%)	41(75.9%)	
Female	67(31.5%)	13 (24.1%)	
T stage*			0.846
1/2	50(23.5%)	12(22.2%)	
3	163(76.5%)	42(77.8%)	
N stage			0.89
1/2	124(58.2%)	32(59.3%)	
3	89(41.8%)	22(40.7%)	

TNM stage			0.479
III	95(44.6%)	27(50%)	
IVA	118(55.4%)	27(50%)	
EBV DNA before initial treatment (copies/mL)			0.846
>2000	74(34.7%)	18(33.3%)	
<2000	139(65.3%)	36(66.7%)	
Post-induction EBV DNA (>0 copies/mL)			0.61
Positive	37(17.4%)	11(20.4%)	
Negative	176(82.6%)	43(79.6%)	
Post-chemoradiotherapy EBV DNA (>0 copies/mL)			0.268
Positive	3(1.4%)	2(3.7%)	
Negative	210(98.6%)	52(96.3%)	
Treatment modality			
Patients receiving adjuvant chemotherapy			0.448
No	159(74.6%)	43(79.6%)	
Yes	54(25.4%)	11(20.4%)	
Adjuvant chemotherapy regimen			0.905
Platinum-based doublet regimen	9(16.7%)	2(18.2%)	
Capecitabine regimen	45(83.3%)	9(81.8%)	

* Tumors were staged according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging system.

Construction of the radiomic signature

A total of 20 features were selected to form the final prognostic radiomic signature (Table 2). In terms of feature class, the signature was overwhelmingly composed of wavelet-transformed features (19 of 20), encompassing gray level size zone matrix (GLSZM), gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), and first-order statistics. With respect to MRI sequences, the majority of features (10) originated from CE-T1WI, followed by T1WI (5) and T2WI (5). Spatially, features were equally derived from the primary tumor (GTVp) and the involved lymph nodes (GTVnd), with each region contributing 10 features. A color map illustrating the importance ranking of these features is provided in Fig. 3.

Table 2. Radiomics features selected by LASSO regression and their coefficients.

No.	**Radiomic Feature**	**Coefficient**	**Sequence and Target Volume**
1	T1_GTVnd_original_glszm_SmallAreaHighGrayLevelEmphasis	0.940	T1 GTVnd
2	T1_GTVnd_original_glszm_SizeZoneNonUniformityNormalize	-1.126	T1 GTVnd

	d		
3	T2_GTVnd_wavelet.HHL_glszm_GrayLevelVariance	-0.565	T2 GTVnd
4	T2_GTVnd_wavelet.HLL_glcm_Imc2	0.614	T2 GTVnd
5	T1_GTVnd_wavelet.LLL_glszm_SizeZoneNonUniformityNormalized	0.613	T1 GTVnd
6	T1C_GTVp_wavelet.HLH_glszm_SizeZoneNonUniformity	0.549	T1C GTVp
7	T1C_GTVnd_wavelet.LLL_glszm_SmallAreaLowGrayLevelEmphasis	-0.499	T1C GTVnd
8	T1C_GTVp_wavelet.HHH_firstorder_Mean	0.471	T1C GTVp
9	T1_GTVp_wavelet.HLH_glcm_SumSquares	-0.401	T1 GTVp
10	T1C_GTVnd_wavelet.LHH_glszm_LowGrayLevelZoneEmphasis	-0.400	T1C GTVnd
11	T2_GTVnd_wavelet.LLL_glszm_SmallAreaHighGrayLevelEmphasis	-0.385	T2 GTVnd
12	T1C_GTVp_wavelet.HLL_glszm_GrayLevelVariance	-0.263	T1C GTVp
13	T1C_GTVnd_wavelet.HLH_glrlm_LowGrayLevelRunEmphasis	-0.190	T1C GTVnd
14	T1C_GTVnd_wavelet.LLL_glszm_SizeZoneNonUniformity	-0.179	T1C GTVnd
15	T1_GTVp_wavelet.HLH_glszm_LowGrayLevelZoneEmphasis	-0.146	T1 GTVp
16	T1C_GTVp_wavelet.HHL_glszm_ZoneEntropy	0.328	T1C GTVp
17	T1C_GTVp_wavelet.HHH_glszm_SizeZoneNonUniformity	0.314	T1C GTVp
18	T1C_GTVp_wavelet.	0.313	T1C GTVp

	HLL_glszm_SizeZoneNonUniformityNormalized		
19	T2_GTVp_wavelet.LHL_firstorder_Skewness	0.312	T2 GTVp
20	T2_GTVp_wavelet.HHH_glszm_SmallAreaLowGrayLevelEmphasis	0.296	T2 GTVp

All selected features were derived from wavelet-transformed images (19 of 20) and span GLSZM, GLCM, GLRLM, and First-Order feature classes. Features are equally distributed between the primary tumour (GTVp) and involved lymph nodes (GTVnd), and across T2WI, T1WI, and CE-T1WI (T1C) sequences.

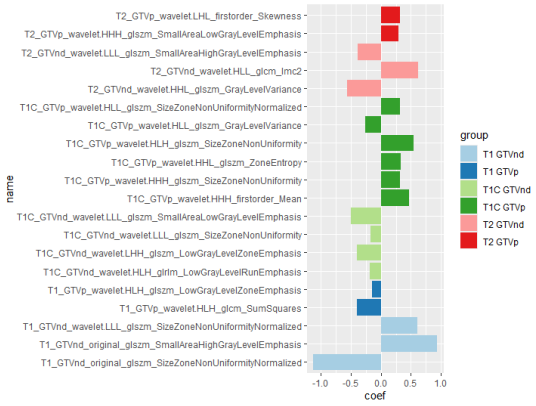


Figure 3. Color map of feature importance rankings derived from LASSO regression coefficients. Positive and negative coefficients indicate risk-increasing and risk-decreasing features, respectively. LASSO, least absolute shrinkage and selection operator.

Construction of the clinical signature

Initial screening via univariable Cox analysis revealed multiple clinical factors with a potential association with PFS. These comprised the RS, overall stage, concurrent chemotherapy cycles, demographics (sex, age), lifestyle history (smoking, alcohol, family history of cancer), laboratory values (neutrophil count, red blood cell count), tumor volumes (GTVp, GTVnd), and EBV DNA parameters (post-CCRT level and dynamic change group). Following this, 1000 bootstrap iterations were applied to refine selection stability. This process yielded a robust set of clinical features, defined by a selection frequency of 500 or more. These stable features were subsequently entered into a multivariable Cox model built with 5-fold cross-validation. The final integrated RCS, visualized as a nomogram in Fig. 4a, included seven components: the RS, age, overall stage, EBV DNA dynamic group, number of concurrent chemotherapy cycles, alcohol history, and neutrophil count, as detailed in Table 3.

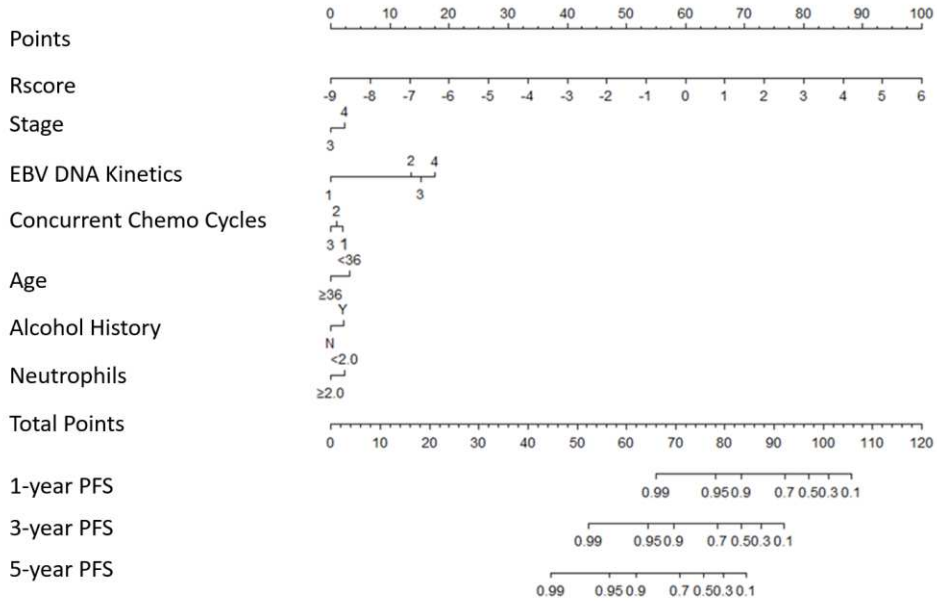


Figure 4a. Prognostic nomogram of the integrated radiomic-clinical signature (RCS). The nomogram estimates 2-year progression-free survival from the radiomic signature, age, overall stage, EBV DNA dynamic group, number of concurrent chemotherapy cycles, alcohol history, and neutrophil count.

Table 3. Univariable Cox analysis with bootstrap feature selection: clinical factors associated with progression-free survival (selection frequency >= 500).

Clinical Factors	HR (95%CI)	*P* value	Selection Frequency
Rscore (≤0.865 vs>0.865)	2.399 (2.13-2.70)	0.0001	1000
Age (≤36 vs>36)	0.58 (0.42-0.79)	0.001	570
TNM stage (III vs IVa)	1.22 (0.91-1.65)	0.007	523
Number of concurrent chemotherapy cycles (1 vs 2-3)	0.68 (0.53-0.86)	0.001	565
Alcohol history (No vs Yes)	1.81 (1.35-2.43)	0.001	584
Neutrophil count (<2 vs ≥2)	2.02 (1.07-3.8)	0.031	528
EBV DNA Kinetics			789
Sustained Negative[1](ref)			
Rapid Clearance[2]	1.13(0.89-1.78)	0.037	
Delayed Clearance[3]	1.43(0.96-1.83)	0.016	
Sustained Positive/Rebound[4]	1.74(1.19-2.85)	0.036	

T0, baseline (pre-treatment); T1, within 2 weeks after induction chemotherapy; T2, within 4 weeks after concurrent chemoradiotherapy. Definitions of EBV DNA kinetic groups:
1: Sustained Negative (ref): EBV DNA undetectable at all three timepoints ($T0^-$, $T1^-$, $T2^-$).
2: Rapid Clearance: Detectable at T0, becomes undetectable by T1 and remains undetectable at T2 ($T0^+$, $T1^-$, $T2^-$).
3: Delayed Clearance: Detectable at T0 and T1, becomes undetectable by T2 ($T0^+$, $T1^+$, $T2^-$).
4: Sustained Positive / Rebound: Detectable at T2, regardless of T0/T1 status ($T2^+$).

Evaluation and validation of models

As summarized in Fig. 4b, the predictive accuracy of the three models was evaluated and compared. In the testing set, the fused RCS model achieved the highest C-index (0.719). This was followed by the radiomic model (RS, C-index: 0.695), whereas the clinical model demonstrated the most limited performance (C-index: 0.583). Decision curve analysis for 1-, 3-, and 5-year PFS (Fig. 4c) further revealed that the combined model maintained a favorable net benefit across clinically meaningful threshold probabilities, particularly in the 5-year prediction, underscoring its enhanced clinical utility for long-term risk stratification.

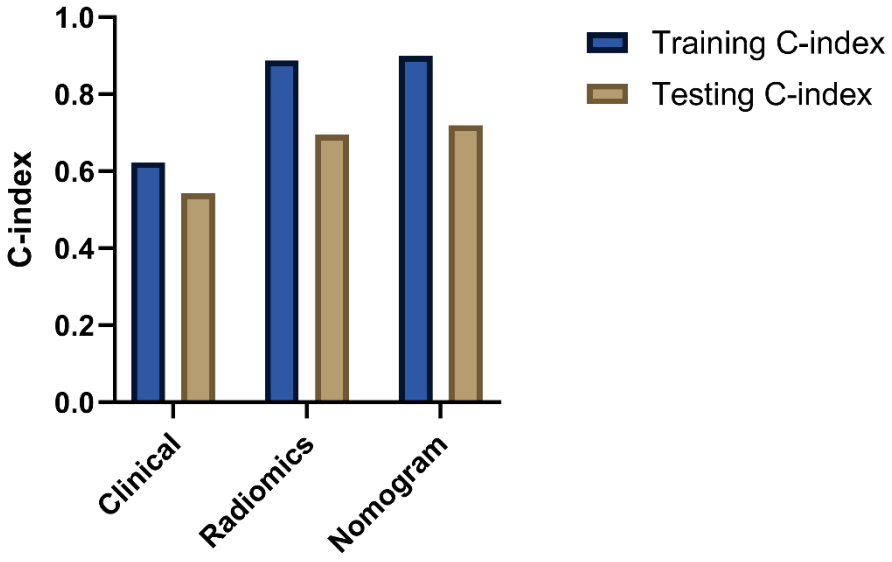


Figure 4b. Assessment and comparison of model predictive performance. C-indices of the clinical, radiomic, and fused radiomic-clinical models in the training and testing sets. C-index, Harrell's concordance index.

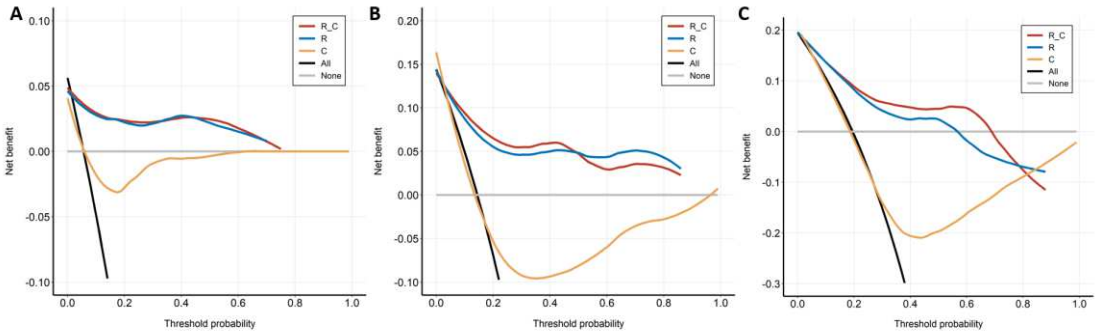
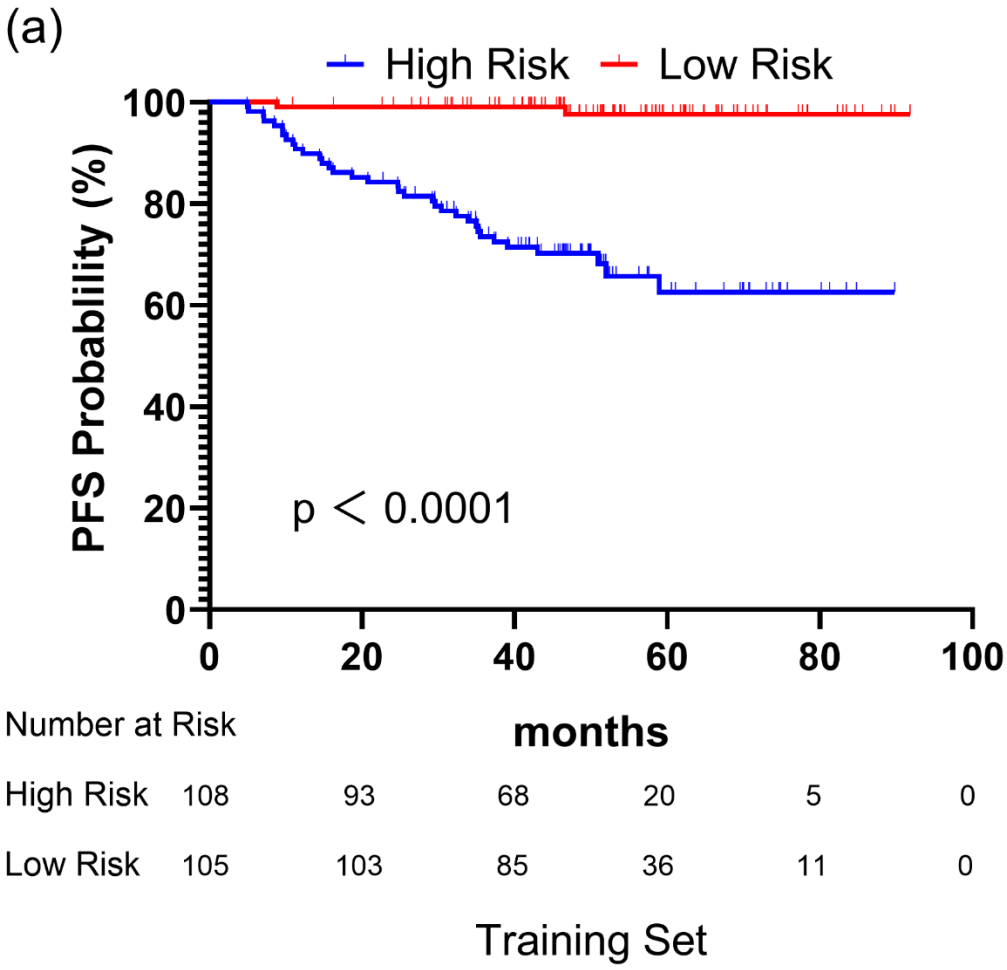


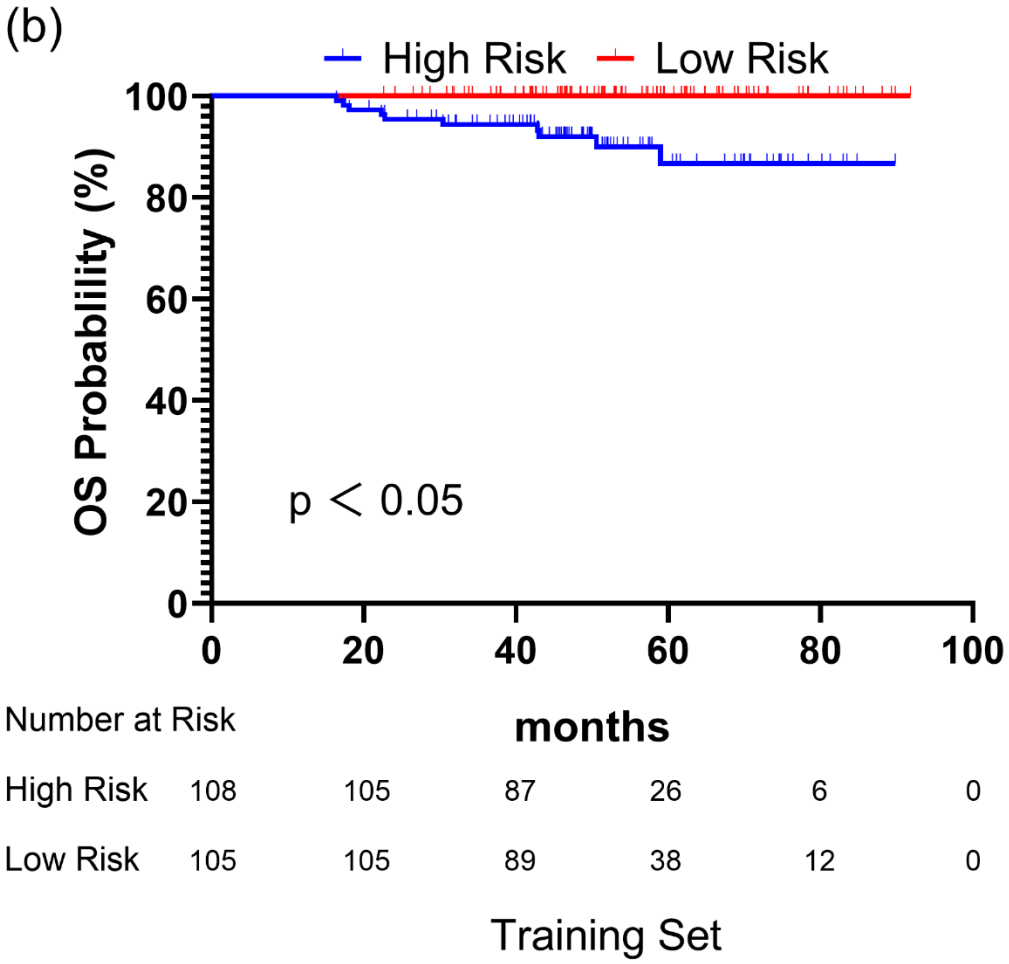
Figure 4c. Decision curves for predicting 1-, 3-, and 5-year progression-free survival (PFS). The fused radiomic-clinical model maintained a favorable net benefit across clinically meaningful threshold probabilities.

Survival outcomes of the RCS model

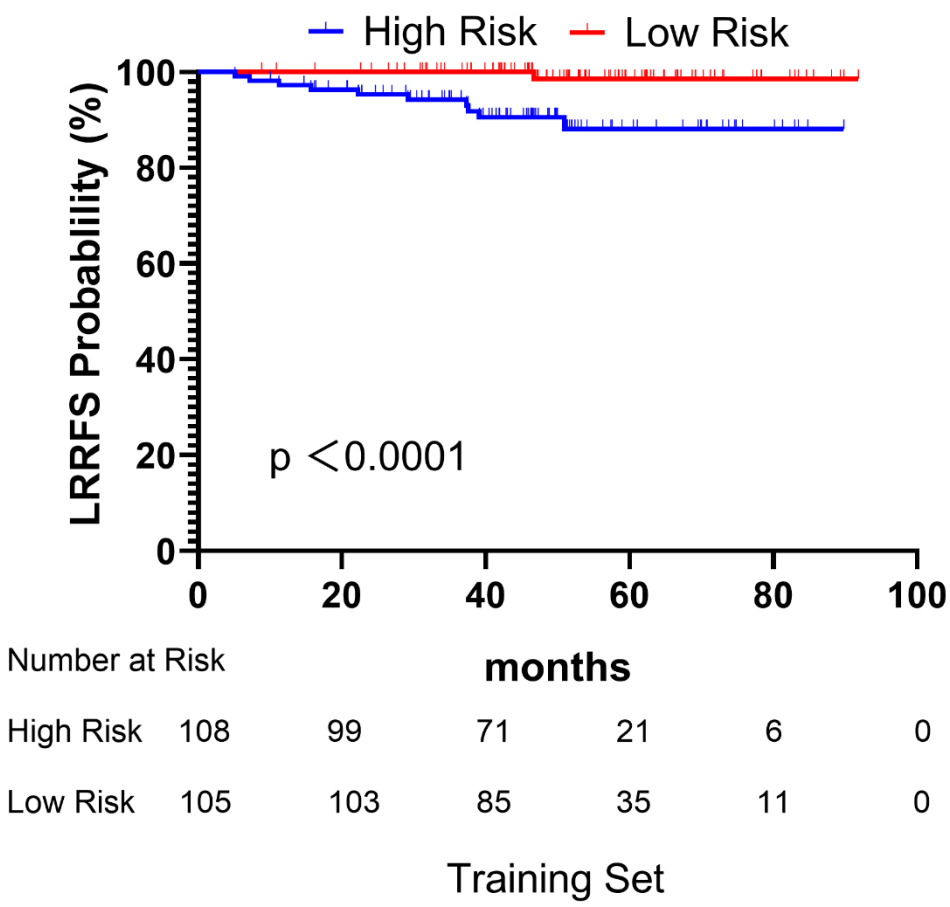
Patients were dichotomized into low- and high-risk groups using the median RCS nomogram score of 60.2 as the cutoff. This stratification revealed pronounced survival disparities in the training set (Fig. 5). The low-risk group exhibited markedly higher 2-year rates for PFS (96.2% vs. 83.3%; $p < 0.0001$), LRRFS (96.2% vs. 86.1%; $p < 0.0001$), OS (99.0% vs. 94.4%; $p = 0.001$), and DMFS (100% vs. 94.7%; $p < 0.0001$). In the independent testing set (Fig. 6), the prognostic value persisted for PFS (97.1% vs. 94.7%, $p = 0.028$) and DMFS (100% vs. 84.6%, $p = 0.025$). Notably, differences in LRRFS and OS were not statistically significant in this cohort.

Kaplan-Meier analyses based on the radiomic signature alone are provided in Additional file 3 (Figure S1).




(b)
High Risk
Low Risk
OS Probablility (%)
100
80
60
40
20
0
p < 0.05
0
20
40
60
80
100
months
Number at Risk
High Risk 108 105 87 26 6 0
Low Risk 105 105 89 38 12 0
Training Set

(c)


High Risk
Low Risk
LRRFS Probablility (%)
100
80
60
40
20
0
p <0.0001
0
20
40
60
80
100
months
Number at Risk
High Risk 108 99 71 21 6 0
Low Risk 105 103 85 35 11 0
Training Set

(d)

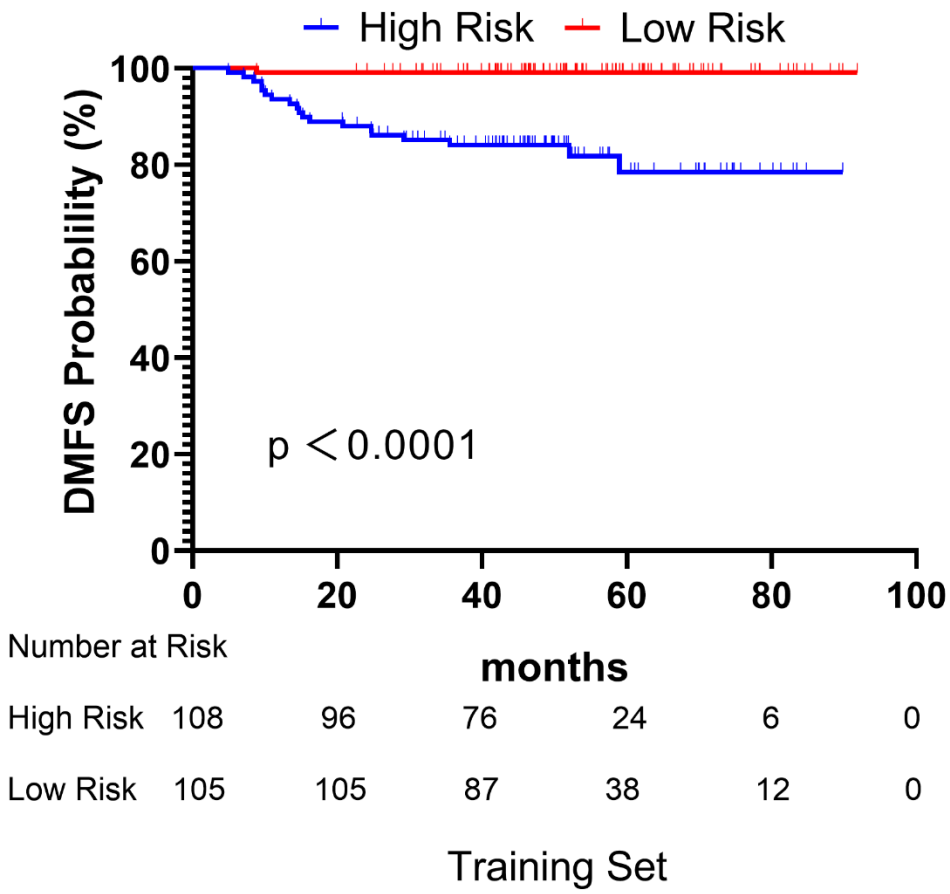
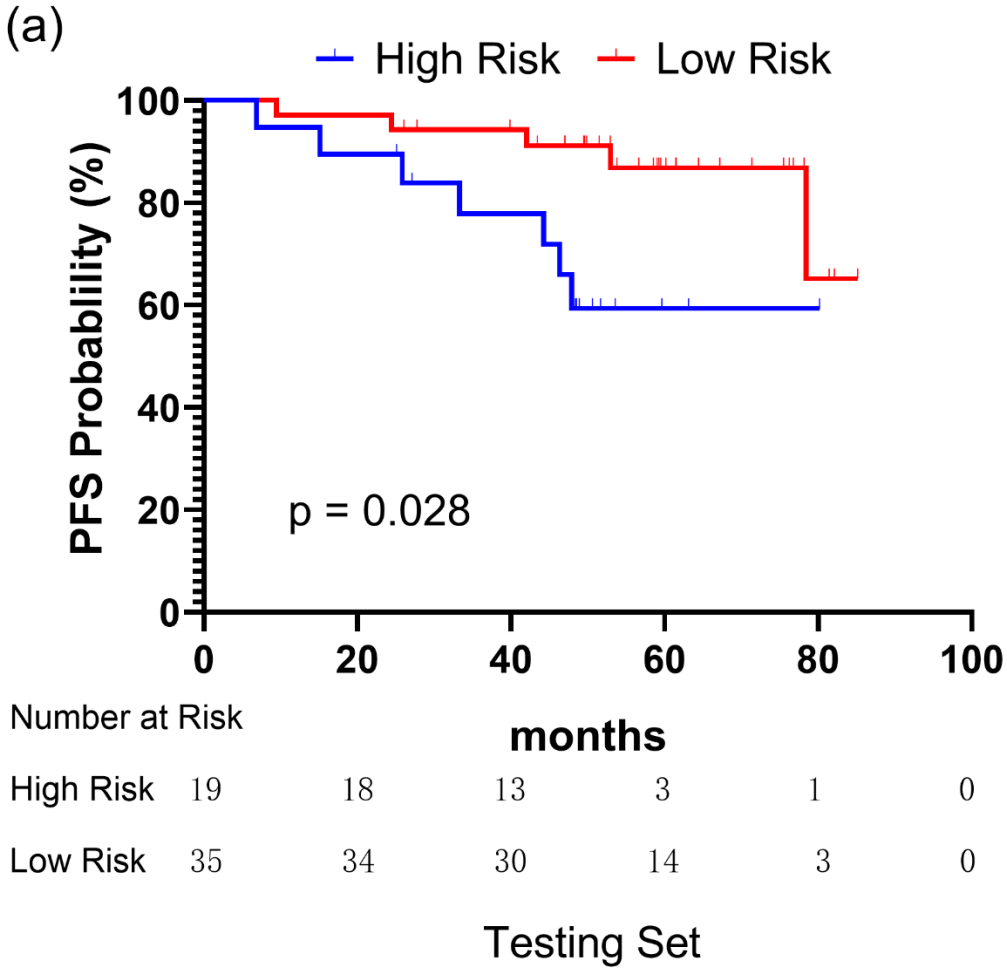
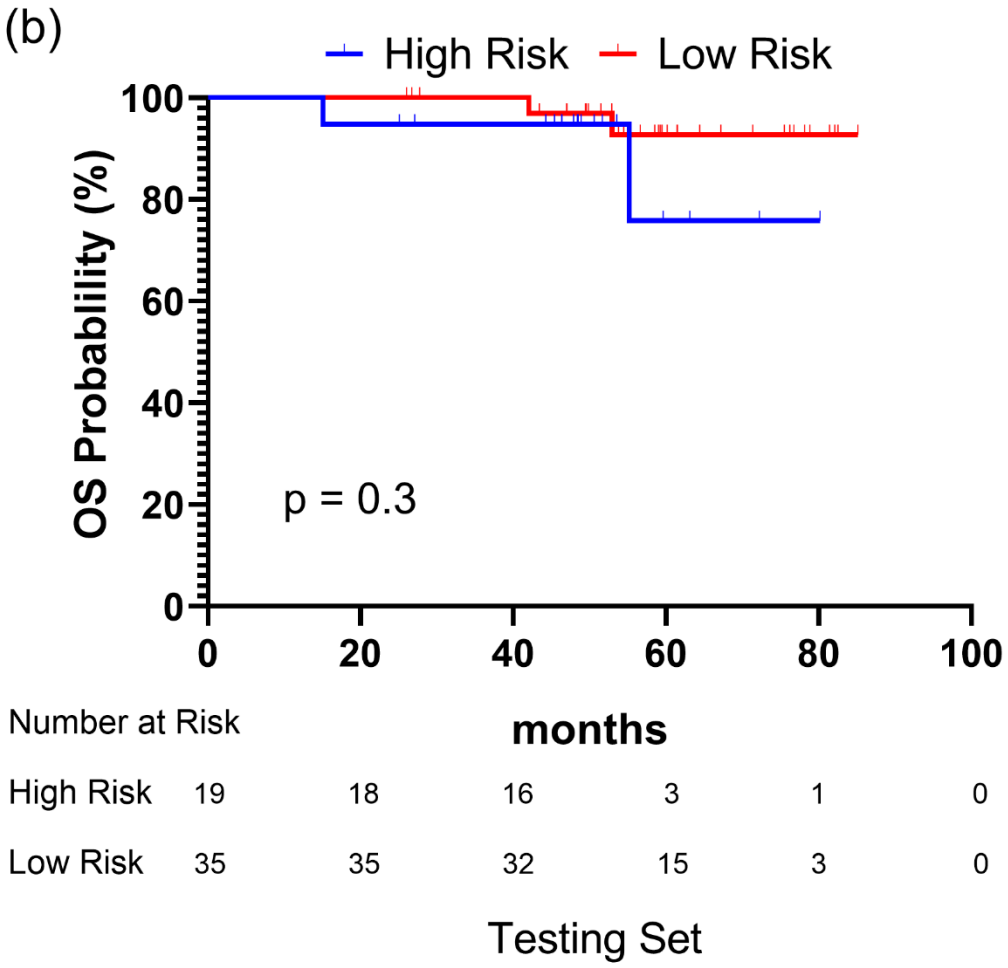
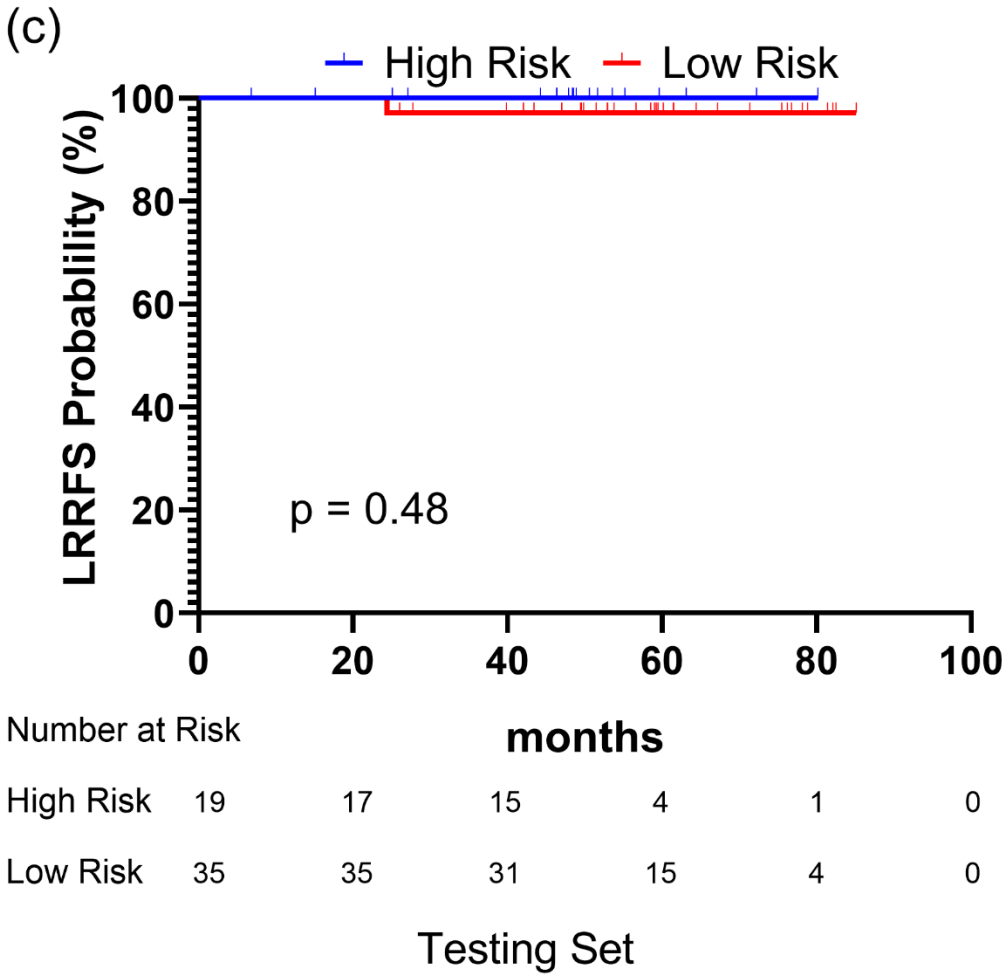


Figure 5. Kaplan-Meier survival curves for patients stratified into low- and high-risk groups in the training set, using the median nomogram score as the cutoff. (a) Progression-free survival (PFS). (b) Overall survival (OS). (c) Locoregional relapse-free survival (LRRFS). (d) Distant metastasis-free survival (DMFS).


(a)
High Risk
Low Risk
PFS Probablility (%)
100
80
60
40
20
0
p = 0.028
0
20
40
60
80
100
months
Number at Risk
High Risk 19 18 13 3 1 0
Low Risk 35 34 30 14 3 0
Testing Set


(b)
High Risk
Low Risk
OS Probablility (%)
100
80
60
40
20
0
p = 0.3
0
20
40
60
80
100
months
Number at Risk
High Risk 19 18 16 3 1 0
Low Risk 35 35 32 15 3 0
Testing Set


(c)
High Risk
Low Risk
LRRFS Probablility (%)
100
80
60
40
20
0
p = 0.48
0
20
40
60
80
100
months
Number at Risk
High Risk 19 17 15 4 1 0
Low Risk 35 35 31 15 4 0
Testing Set

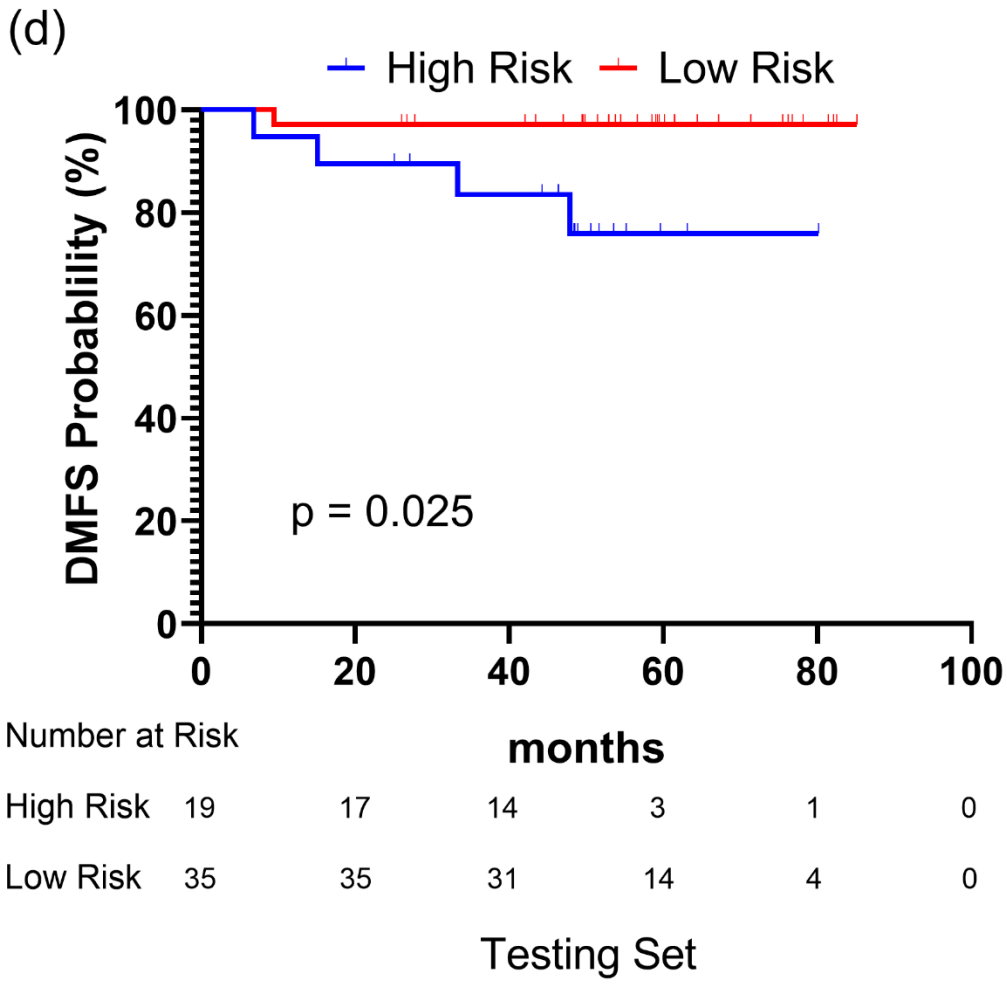
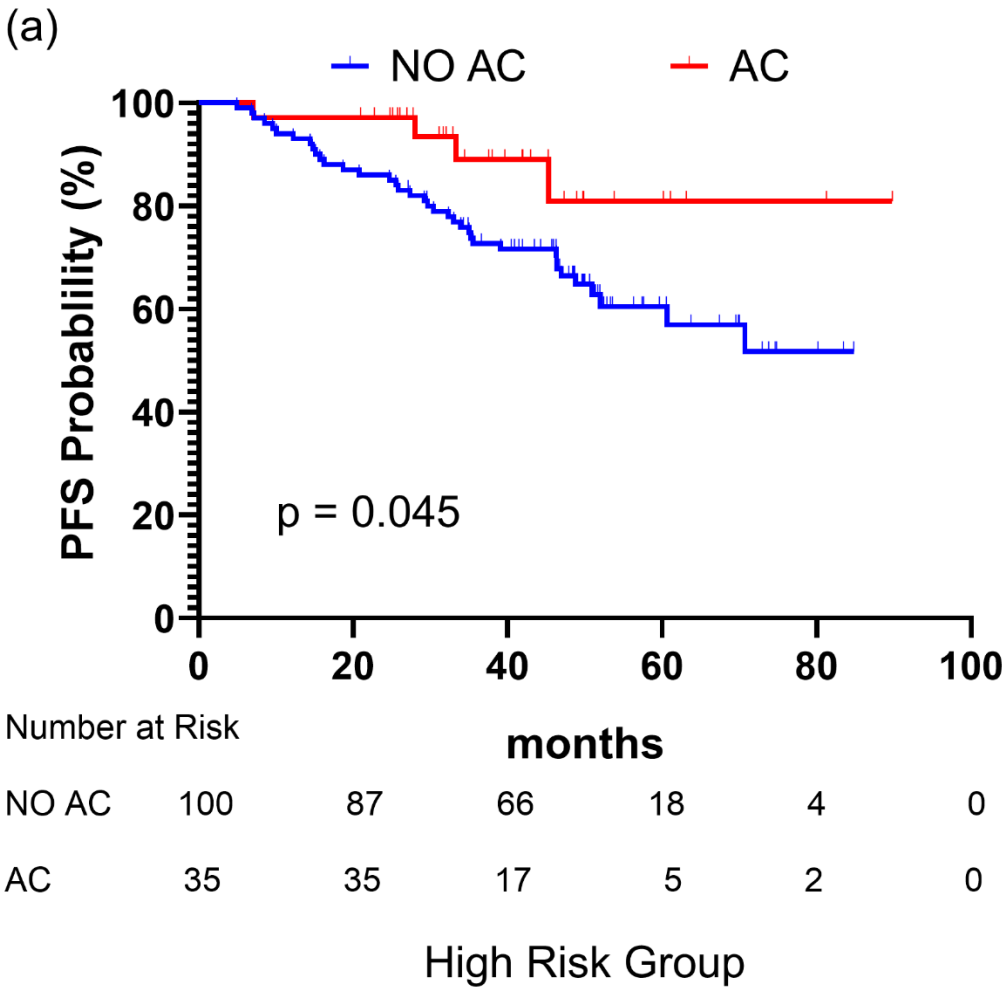
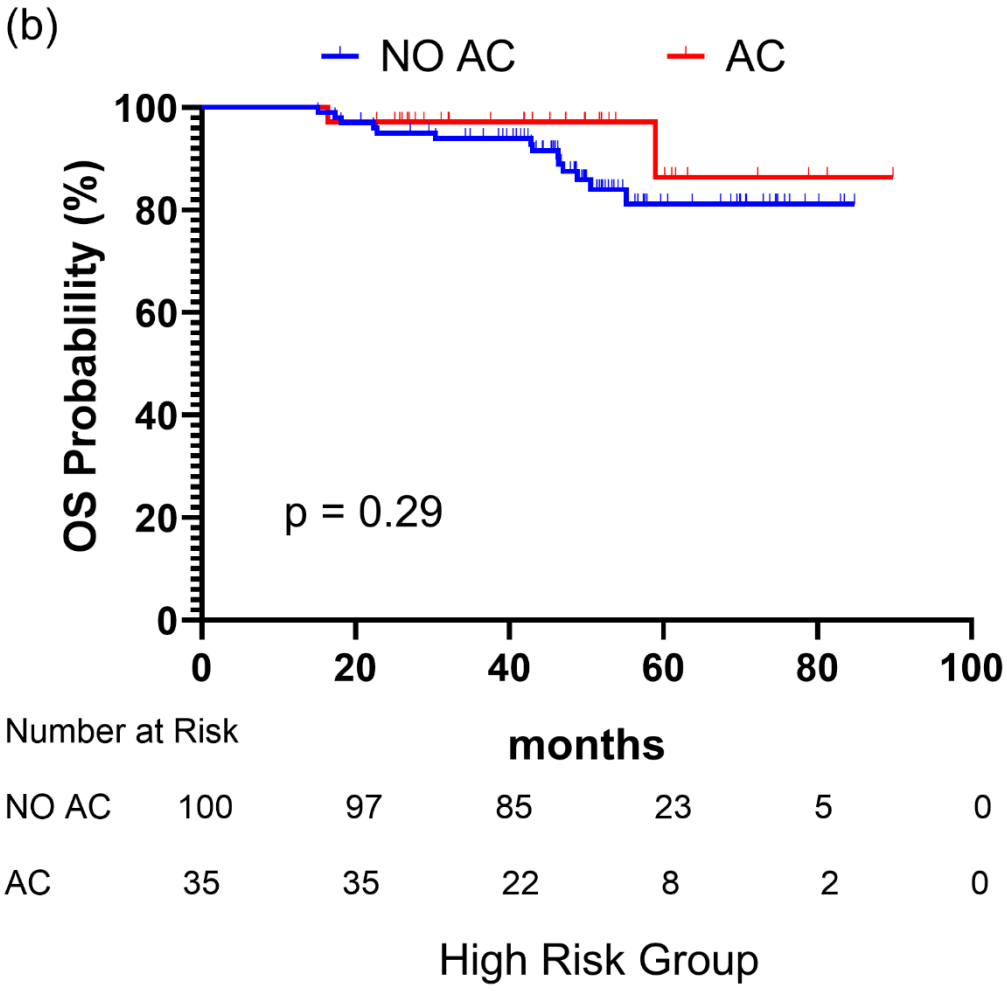


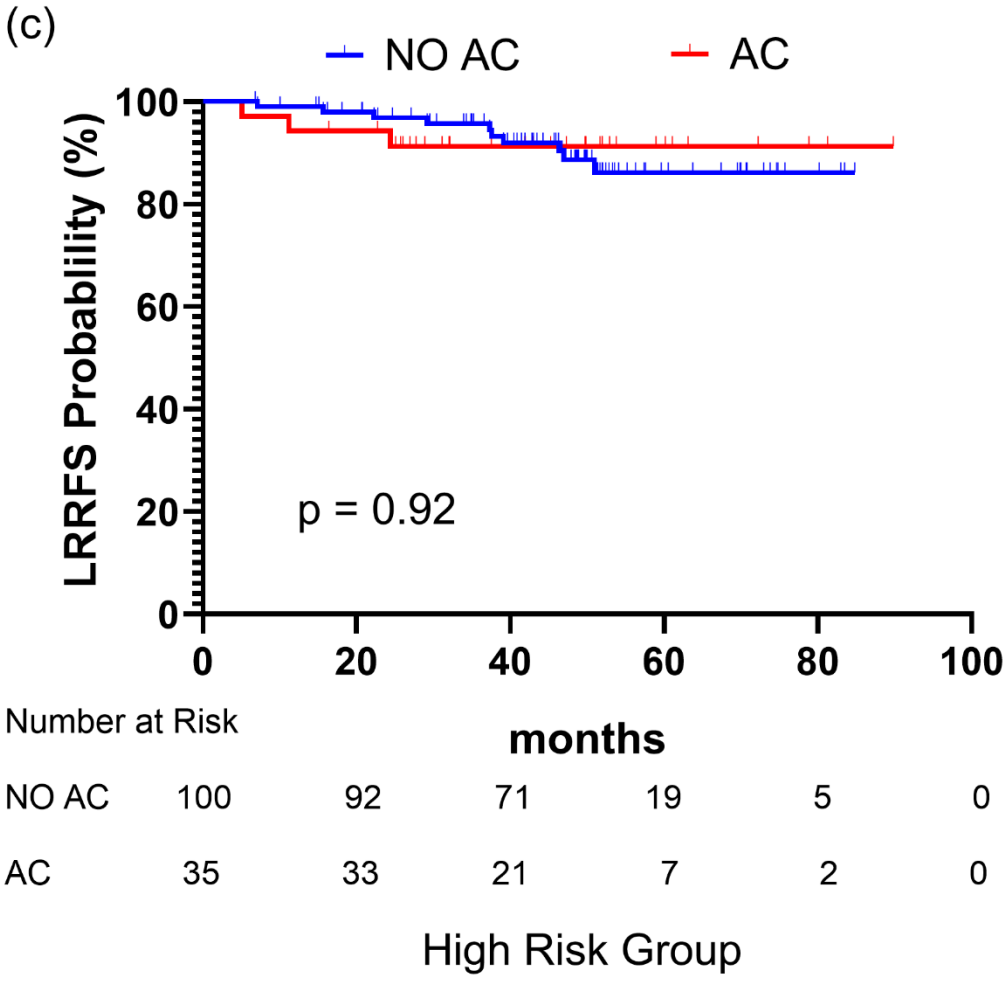
Figure 6. Kaplan-Meier survival curves for the corresponding low- and high-risk groups in the independent testing set. (a) Progression-free survival (PFS). (b) Overall survival (OS). (c) Locoregional relapse-free survival (LRRFS). (d) Distant metastasis-free survival (DMFS).

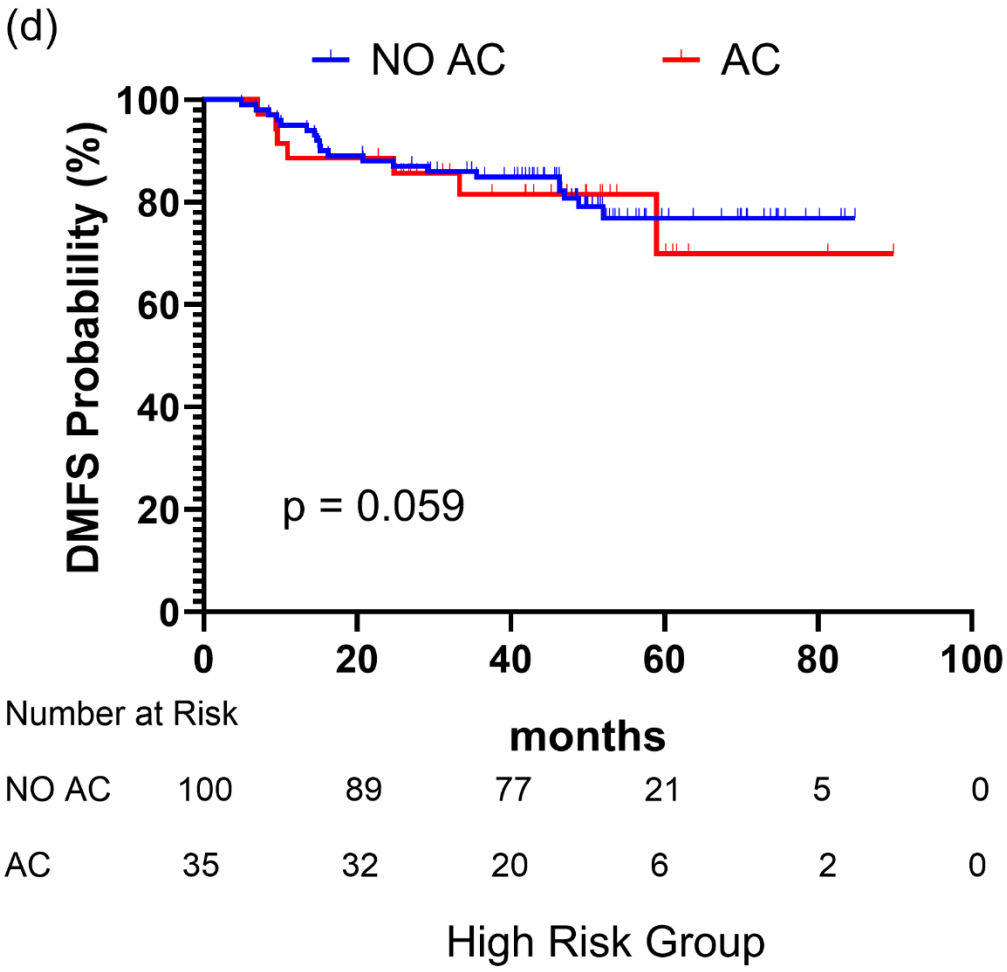
Utility of adjuvant chemotherapy in stratified risk groups

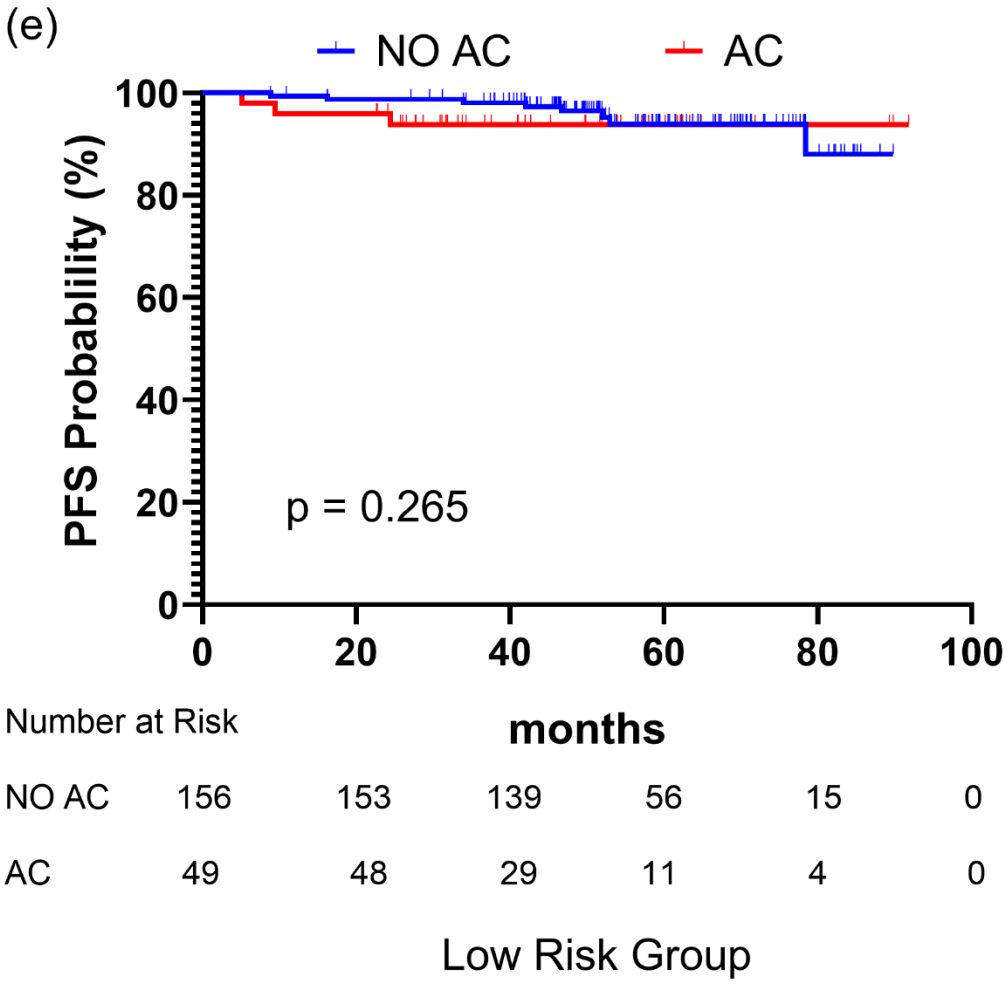
Analysis of AC benefit within the high-risk cohort (Fig. 7) revealed a significant improvement in 2-year PFS (94.3% vs. 87.0%, p = 0.045). Notably, no statistically significant benefit was observed for the other endpoints: LRRFS (88.5% vs. 88.0%, p = 0.92), OS (97.1% vs. 95.0%, p = 0.29), or DMFS (88.0% vs. 88.5%, p = 0.059). In the low-risk cohort, no significant survival differences were found between patients who received AC and those who did not.

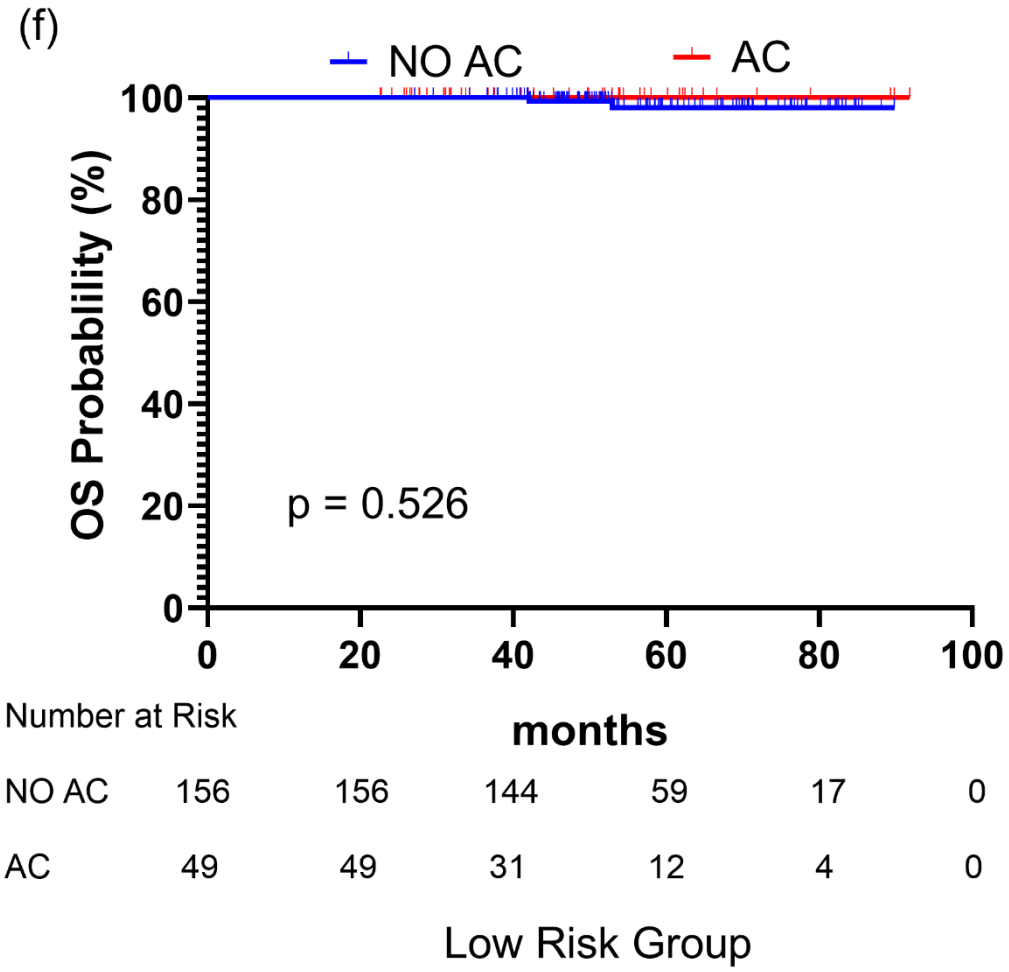

(a)
NO AC
AC
PFS Probablility (%)
100
80
60
40
20
0
p = 0.045
0
20
40
60
80
100
months
Number at Risk
NO AC 100 87 66 18 4 0
AC 35 35 17 5 2 0
High Risk Group


(b)
NO AC
AC
OS Probablility (%)
100
80
60
40
20
0
p = 0.29
0
20
40
60
80
100
months
Number at Risk
NO AC 100 97 85 23 5 0
AC 35 35 22 8 2 0
High Risk Group

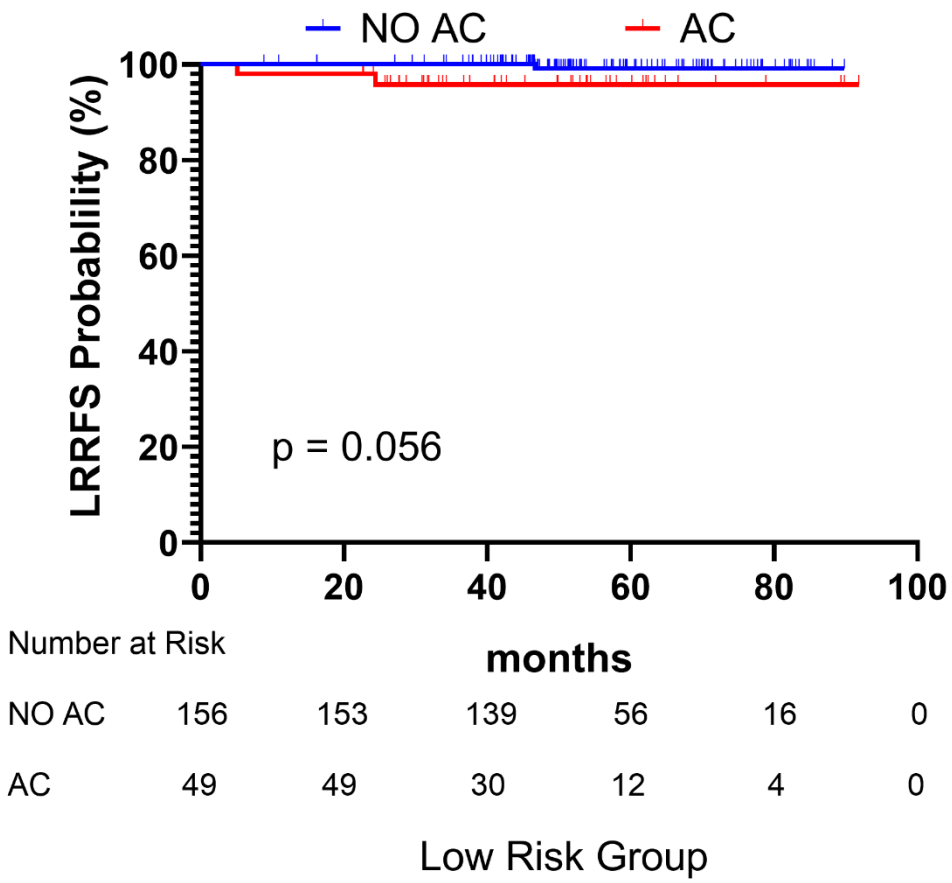

(c)
NO AC
AC
LRRFS Probablility (%)
100
80
60
40
20
0
p = 0.92
0
20
40
60
80
100
months
Number at Risk
NO AC 100 92 71 19 5 0
AC 35 33 21 7 2 0
High Risk Group


(d)
NO AC
AC
DMFS Probablility (%)
100
80
60
40
20
0
p = 0.059
0
20
40
60
80
100
months
Number at Risk
NO AC 100 89 77 21 5 0
AC 35 32 20 6 2 0
High Risk Group


(e)
NO AC
AC
PFS Probablility (%)
100
80
60
40
20
0
p = 0.265
0
20
40
60
80
100
months
Number at Risk
NO AC 156 153 139 56 15 0
AC 49 48 29 11 4 0
Low Risk Group


(f)
NO AC
AC
100
80
60
40
20
0
OS Probablility (%)
p = 0.526
0
20
40
60
80
100
months
Number at Risk
NO AC 156 156 144 59 17 0
AC 49 49 31 12 4 0
Low Risk Group

(g)



Low Risk Group

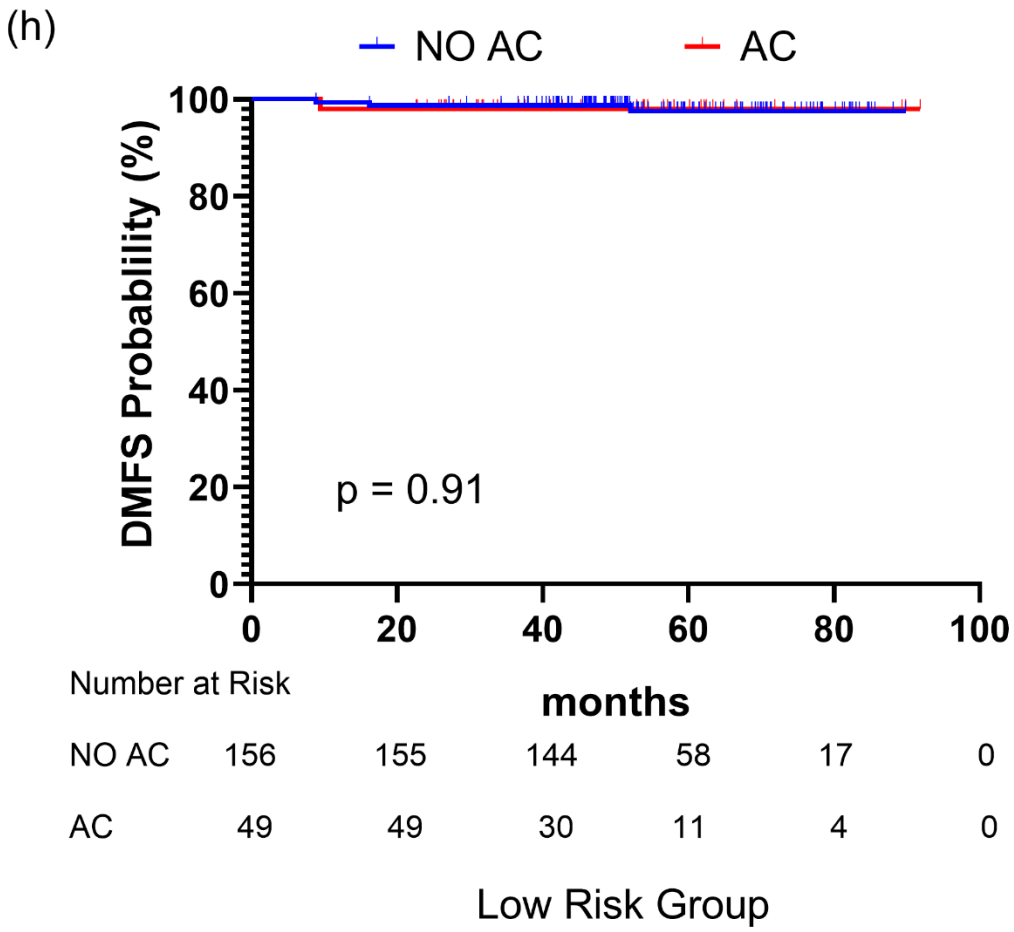


Figure 7. Survival benefit of adjuvant chemotherapy (AC) versus no AC in the high-risk and low-risk cohorts defined by the radiomic-clinical signature nomogram. (a-d) High-risk group: (a) progression-free survival (PFS), (b) overall survival (OS), (c) locoregional relapse-free survival (LRRFS), (d) distant metastasis-free survival (DMFS). (e-h) Low-risk group: (e) PFS, (f) OS, (g) LRRFS, (h) DMFS.

Discussion

By fusing multi-sequence MRI radiomics obtained after CCRT with pivotal dynamic clinical biomarkers, we have developed and internally validated a prognostic nomogram for LA-NPC. The resulting model not only achieves higher predictive accuracy but also incorporates a richer dimensionality of prognostic features, thereby enhancing its potential utility in clinical practice relative to previous methodologies.

The survival benefit of AC for LA-NPC remains uncertain. While pivotal early trials established its efficacy, confirmatory phase III trials in the contemporary intensity-modulated radiotherapy (IMRT) era have yielded inconsistent results, questioning its universal applicability. The clinical

translation of AC is significantly hampered by several factors: suboptimal tolerance leading to high rates of incomplete treatment (approximately 40-76% of patients fail to complete the planned cycles) [5, 6, 20], a performance status often compromised by preceding intensive therapy, and considerable cumulative toxicities that impair quality of life and adherence [21, 22]. This is compounded by the fact that a substantial proportion of patients derive no clear benefit, thereby exposing them to unnecessary toxicity. Consequently, there is a pressing need for precision medicine tools that reliably select the patients most likely to benefit from AC, thereby optimizing its risk-benefit ratio.

To ensure robustness and minimize overfitting, feature selection followed a stringent four-step strategy combined with extensive bootstrap validation. The resulting RCS incorporated seven definitive predictors. More importantly, our RCS model achieved a testing set C-index of 0.719, outperforming the predictive accuracy of a conventional TNM stage-based model (C-index: 0.63) [16], which underscores the added value of integrating radiomic and dynamic clinical data.

Our methodology represents a significant departure from conventional approaches. While models based solely on CE-T1WI achieve a testing C-index of around 0.73, our predictive framework was predicated on the complementary fusion of T1WI, T2WI, and CE-T1WI features. Notably, CE-T1WI contributed roughly 50% of the key features, leveraging its distinct advantage in delineating soft-tissue boundaries to better capture tumor invasion characteristics, thereby enhancing model sensitivity. Moreover, our segmentation protocol explicitly included both the primary gross tumor volume and the involved nodal volumes. This design philosophy is directly informed by the AJCC 8th edition's incorporation of extranodal extension as a critical prognostic determinant, allowing our model to potentially encode imaging signatures associated with this aggressive feature and to offer a more nuanced risk assessment [8].

In the independent testing set, the fused RCS model demonstrated the highest predictive accuracy (C-index: 0.719), validating the advantage of multi-modal fusion over unimodal approaches (radiomics-only: 0.695; clinical-only: 0.583) and over conventional TNM stage-based models. The clinical utility of the RCS was further supported by its effective risk stratification: patients classified as low-risk by the RCS exhibited significantly superior survival outcomes across all evaluated endpoints in the training set.

This study addresses a central controversy in the management of LA-NPC by dissecting AC benefit within prognostically distinct subgroups. In doing so, we demonstrate that low-risk patients derive no significant survival benefit from AC, supporting the rationale for its omission to avoid undue toxicity. Critically, however, AC significantly improved PFS in the high-risk subgroup. The principal advancement of our model is its capacity for precise risk stratification, which underpins personalized AC decision-making. This represents a shift away from reliance on conventional clinicopathological factors alone. Moreover, by integrating multi-sequence MRI radiomics, our approach overcomes the limitations of single-sequence analyses, yielding a more robust and reproducible prognostic tool.

Several limitations should be acknowledged. First, the single-center retrospective design may introduce selection bias. Second, the sample size is limited, especially in the high-risk subgroup of the testing set (n = 19). Third, manual tumor contouring carries inherent inter-observer variability. Fourth, a statistically significant difference in the number of concurrent chemotherapy cycles was observed between the training and testing sets ($p = 0.033$; Additional file 1: Table S1), which may represent an additional source of bias and could affect the generalizability of the model. To address these limitations, future efforts are warranted: (1) multicenter, large-scale prospective validation to establish generalizability; (2) development of artificial intelligence (AI)-driven auto-segmentation to improve reproducibility and efficiency; (3) incorporation of dynamic biomarkers (e.g., serial radiomics and circulating tumor DNA [ctDNA]) to enable adaptive risk assessment; and (4) exploration of personalized adjuvant therapies, such as immunotherapy for high-risk patients. Promisingly, the model maintained robust performance in the limited test set, supporting its potential for clinical translation.

Conclusions

In this retrospective single-center study, we developed and internally validated a radiomic-clinical signature that fuses post-CCRT multi-sequence MRI radiomics with dynamic plasma EBV DNA kinetics and routinely available clinical variables. The RCS effectively stratifies patients with LA-NPC into distinct prognostic groups and identifies a high-risk subgroup that derives a significant PFS benefit from AC, whereas low-risk patients appear to gain no survival advantage and could therefore be spared unnecessary toxicity. These findings provide a practical,

imaging-based framework for personalizing adjuvant treatment decisions in LA-NPC. Prospective multicenter validation is warranted before routine clinical implementation.

List of abbreviations

AC: adjuvant chemotherapy.

AI: artificial intelligence.

AIC: Akaike information criterion.

AJCC: American Joint Committee on Cancer.

CCRT: concurrent chemoradiotherapy.

CE-T1WI: contrast-enhanced T1-weighted imaging.

CI: confidence interval.

C-index: Harrell's concordance index.

ctDNA: circulating tumor DNA.

DMFS: distant metastasis-free survival.

EBV: Epstein-Barr virus.

ENE: extranodal extension.

GLCM: gray level co-occurrence matrix.

GLRLM: gray level run length matrix.

GLSZM: gray level size zone matrix.

GTVnd: gross tumor volume of involved lymph nodes.

GTVp: gross tumor volume of the primary tumor.

HR: hazard ratio.

IC: induction chemotherapy.

IMRT: intensity-modulated radiotherapy.

LA-NPC: locoregionally advanced nasopharyngeal carcinoma.

LASSO: least absolute shrinkage and selection operator.

LRRFS: locoregional relapse-free survival.

MRI: magnetic resonance imaging.

NPC: nasopharyngeal carcinoma.

OS: overall survival.

PFS: progression-free survival.

RCS: radiomic-clinical signature.

ROI: region of interest.

ROC: receiver operating characteristic.

RS: radiomic signature.

Rscore: radiomic signature risk score.

T1WI: T1-weighted imaging.

T2WI: T2-weighted imaging.

Declarations

Ethics approval and consent to participate

This retrospective study was approved by the Medical Ethics Committee of Nanfang Hospital, Southern Medical University (approval number: NFEC-2024-367) and was conducted in accordance with the Declaration of Helsinki. The requirement for written informed consent was waived by the ethics committee because of the retrospective design and the use of de-identified patient data.

Consent for publication

Not applicable. This manuscript does not contain data from any individual person.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available because of patient privacy and institutional restrictions, but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (grant numbers 81972895 and 81101821), the Natural Science Foundation of Guangdong Province (grant number 10451001002004732), and the Health Science and Technology Innovation Joint Program of Hainan Province (grant number WSJK2025QN061). The funding bodies had no role

in the design of the study; in the collection, analysis, or interpretation of data; or in the writing of the manuscript.

Authors' contributions

LW and HL contributed equally to this work. LW: conceptualization, data curation, methodology, software, formal analysis, investigation, and writing - original draft. HL: conceptualization, data curation, formal analysis, investigation, methodology, and writing - review and editing. SL: software and visualization. JL: resources and data curation. YZ: data curation and writing - review and editing. RS: formal analysis, software, and validation. SW: resources and project administration. DL: investigation and writing - review and editing. LG: software and validation. ZC: software and validation. MX: software and validation. XL: resources and validation. TS: investigation and writing - review and editing. BC: conceptualization, funding acquisition, project administration, supervision, and writing - review and editing. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Authors' information (optional)

Not applicable.

References

1. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. Lancet. 2019;394(10192):64-80. doi:10.1016/S0140-6736(19)30956-0
2. Bossi P, Chan AT, Licitra L, Trama A, Orlandi E, Hui EP, et al. Nasopharyngeal carcinoma: ESMO-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2021;32(3):452-65. doi:10.1016/j.annonc.2020.12.007
3. Chen YP, Liu X, Zhou Q, Yang KY, Jin F, Zhu XD, et al. Metronomic capecitabine as adjuvant therapy in locoregionally advanced nasopharyngeal carcinoma: a multicentre, open-label, parallel-group, randomised, controlled, phase 3 trial. Lancet. 2021;398(10297):303-13. doi:10.1016/S0140-6736(21)00023-5
4. Miao J, Wang L, Tan SH, Li J, Yi J, Ong EHW, et al. Adjuvant capecitabine following concurrent chemoradiotherapy in locoregionally advanced nasopharyngeal carcinoma. JAMA Oncol. 2022;8(12):1776-85. doi:10.1001/jamaoncol.2022.4651
5. Chen L, Hu C, Chen X, Hu G, Cheng Z, Sun Y, et al. Adjuvant chemotherapy in patients with locoregionally advanced nasopharyngeal carcinoma: long-term results of a phase 3

multicentre randomised controlled trial. Eur J Cancer. 2017;75:150-8. doi:10.1016/j.ejca.2017.01.002

6. Chen L, Hu CS, Chen XZ, Hu GQ, Cheng ZB, Sun Y, et al. Concurrent chemoradiotherapy plus adjuvant chemotherapy versus concurrent chemoradiotherapy alone in patients with locoregionally advanced nasopharyngeal carcinoma: a phase 3 multicentre randomised controlled trial. Lancet Oncol. 2012;13(2):163-71. doi:10.1016/S1470-2045(11)70320-5
7. Chen YP, Wang ZX, Chen L, Liu X, Tang LL, Mao YP, et al. A Bayesian network meta-analysis comparing concurrent chemoradiotherapy followed by adjuvant chemotherapy, concurrent chemoradiotherapy alone and radiotherapy alone in patients with locoregionally advanced nasopharyngeal carcinoma. Ann Oncol. 2015;26(10):2050-61. doi:10.1093/annonc/mdv328
8. Pan J, Mai H, Ng WT, Hu C, Li J, Chen X, et al. Ninth version of the AJCC and UICC nasopharyngeal cancer TNM staging classification. JAMA Oncol. 2024;10(11):1627-40. doi:10.1001/jamaoncol.2024.1947
9. Lv J, Xu L, Li Z, Lin L, Wu C, Quan T, et al. Longitudinal on-treatment circulating tumor DNA as a biomarker for real-time dynamic risk monitoring in cancer patients: the EP-SEASON study. Cancer Cell. 2024;42(8):1401-14. doi:10.1016/j.ccell.2024.06.001
10. Liu Y, Yan W, Qi X, Zhang Y, Wang K, Qu Y, et al. Significance of longitudinal Epstein-Barr virus DNA combined with multipoint tumor response for dynamic risk stratification and treatment adaptation in nasopharyngeal carcinoma. Cancer Lett. 2024;605:217276. doi:10.1016/j.canlet.2024.217276
11. Liu Y, Yan W, Chen Y, Miao J, Zhang H, Wang J, et al. Treatment response-adapted risk index model for survival prediction and adjuvant chemotherapy selection in nonmetastatic nasopharyngeal carcinoma. npj Digit Med. 2025;8:564. doi:10.1038/s41746-025-01918-2
12. Jiang Y, Liang Z, Chen K, Li Y, Yang J, Qu S, et al. A dynamic nomogram combining tumor stage and magnetic resonance imaging features to predict the response to induction chemotherapy in locally advanced nasopharyngeal carcinoma. Eur Radiol. 2023;33(3):2171-84. doi:10.1007/s00330-022-09086-1
13. Zhao L, Gong J, Xi Y, Xu M, Li C, Kang X, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. Eur Radiol. 2020;30(1):537-46. doi:10.1007/s00330-019-06427-z
14. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2017;23(15):4259-69. doi:10.1158/1078-0432.CCR-17-0105
15. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749-62. doi:10.1038/nrclinonc.2017.141
16. Liu Q, Ma L, Ma H, Yang L, Xu Z. Establishment of a prognostic nomogram for patients with locoregionally advanced nasopharyngeal carcinoma incorporating clinical characteristics and dynamic changes in hematological and inflammatory markers. Front Oncol. 2022;12:1032213. doi:10.3389/fonc.2022.1032213

17. Yang K, Tian J, Zhang B, Li M, Xie W, Zou Y, et al. A multidimensional nomogram combining overall stage, dose volume histogram parameters and radiomics to predict progression-free survival in patients with locoregionally advanced nasopharyngeal carcinoma. Oral Oncol. 2019;98:85-91. doi:10.1016/j.oraloncology.2019.09.025
18. Zhong L, Dong D, Fang X, Zhang F, Zhang N, Zhang L, et al. A deep learning-based radiomic nomogram for prognosis and treatment decision in advanced nasopharyngeal carcinoma: a multicentre study. EBioMedicine. 2021;70:103522. doi:10.1016/j.ebiom.2021.103522
19. Wu Q, Chang Y, Yang C, Liu H, Chen F, Dong H, et al. Adjuvant chemotherapy or no adjuvant chemotherapy? A prediction model for the risk stratification of recurrence or metastasis of nasopharyngeal carcinoma combining MRI radiomics with clinical factors. PLoS One. 2023;18(5):e0287031. doi:10.1371/journal.pone.0287031
20. Al-Sarraf M, LeBlanc M, Giri PG, Fu KK, Cooper JS, Vuong T, et al. Chemoradiotherapy versus radiotherapy in patients with advanced nasopharyngeal cancer: phase III randomized Intergroup study 0099. J Clin Oncol. 1998;16(4):1310-7. doi:10.1200/JCO.1998.16.4.1310
21. Yip PL, Lee AWM, Chua MLK. Adjuvant chemotherapy in nasopharyngeal carcinoma. Lancet Oncol. 2023;24(8):713-5. doi:10.1016/S1470-2045(23)00290-7
22. Wang B. Adjuvant chemotherapy in locoregionally advanced nasopharyngeal carcinoma: finding the suitable population. Oral Oncol. 2022;124:105544. doi:10.1016/j.oraloncology.2021.105544

Additional files

Additional file 1: Table S1. Treatment regimens, and hematologic, hepatic, renal, and body composition parameters in the training and testing sets. (DOCX)

Additional file 2: Table S2. Prognostic radiomics features selected by the LASSO-Cox model. (DOCX)

Additional file 3: Figure S1. Kaplan-Meier survival curves for patients stratified into low- and high-risk groups in the training and testing sets, using the median Rscore as the cutoff. (a) Progression-free survival (PFS), (b) locoregional relapse-free survival (LRRFS), (c) overall survival (OS), and (d) distant metastasis-free survival (DMFS) in the training set; (e) PFS, (f) LRRFS, (g) OS, and (h) DMFS in the testing set. (DOCX)

Supplementary Files

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

- Additionalfile1TableS1.docx
- Additionalfile2TableS2.docx
- Additionalfile3FigureS1.docx