BMC Cancer

RESEARCH

Open Access

Early prediction of progression-free survival of patients with locally advanced nasopharyngeal carcinoma using multi-parametric MRI radiomics

Lian Jian[1†], Cai Sheng[2†], Huaping Liu[1†], Handong Li[1], Pingsheng Hu[1], Zhaodong Ai[1], Xiaoping Yu[1*] and Huai Liu[1*]

Abstract

Purpose Prognostic prediction plays a pivotal role in guiding personalized treatment for patients with locoregionally advanced nasopharyngeal carcinoma (LANPC). However, few studies have investigated the incremental value of functional MRI to the conventional MRI-based radiomic models. Here, we aimed to develop a radiomic model including functional MRI to predict the prognosis of LANPC patients.

Methods One hundred and twenty-six patients (training dataset, $n=88$; validation dataset, $n=38$) with LANPC were retrospectively included. Radiomic features were extracted from T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), contrast-enhanced T1WI (cT1WI), and diffusion-weighted imaging (DWI). Pearson correlation analysis and recursive feature elimination or Relief were used for identifying features associated with progression-free survival (PFS). Five machine learning algorithms with cross-validation were compared to develop the optimal single-layer and fusion radiomic models. Clinical and combined models were developed via multivariate Cox regression model.

Results The clinical model based on TNM stage achieved a C-index of 0.544 in the validation dataset. The fusion radiomic model, incorporating DWI-, T1WI-, and cT1WI-derived imaging features, yielded the highest C-index of 0.788, outperforming DWI-based (C-index = 0.739), T1WI-based (C-index = 0.734), cT1WI-based (C-index = 0.722), and T1WI plus cT1WI-based models (C-index = 0.747) in predicting PFS. The fusion radiomic model yielded the C-index of 0.786 and 0.690 in predicting distant metastasis-free survival and overall survival, respectively. However, the addition of TNM stage to the fusion radiomic model could not improve the predictive power.

Conclusion The fusion radiomic model demonstrates favorable performance in predicting survival outcomes in LANPC patients, surpassing TNM staging alone. Integration of DWI-derived features into conventional MRI radiomic models could enhance predictive accuracy.

†Lian Jian, Cai Sheng and Huaping Liu contributed equally to this work.

*Correspondence:
Xiaoping Yu
yuxiaoping@hnca.org.cn

Full list of author information is available at the end of the article

Keywords Nasopharyngeal carcinoma, Magnetic resonance imaging, Radiomics, Machine learning, Progression-free survival

Introduction

Nasopharyngeal carcinoma (NPC) exhibits a unique geographical distribution, with the highest incidence rates reported in east and southeast Asia [1]. Approximately 70-80% of NPC patients are classified as having locoregionally advanced NPC (LANPC) [1]. Moreover, the 5-year survival rates for LANPC patients who undergo chemoradiotherapy remain a persistent challenge due to locoregional recurrence and distant metastasis, typically ranging from 70 to 80% [2]. Thus, accurately predicting the prognosis of LANPC patients holds significant clinical importance, enabling the optimization of treatment strategies, enhancement of patient care, and ultimately leading to improvements in survival and quality of life.

In clinical practice, the tumor-node-metastasis (TNM) classifcation system serves as a crucial tool for clinical decision-making and prognostic evaluation in LANPC. While patients classified under the same category are generally treated with similar therapeutic strategies, their clinical outcomes can vary significantly [3]. This disparity indicates that the TNM system predominantly focuses on the tumor's relationship with surrounding tissues and organs, neglecting intratumor heterogeneity. As a result, its capacity to accurately predict risk stratification markers is limited. In contrast, magnetic resonance imaging (MRI), renowned for its superior soft-tissue resolution, provides more precise visualization of microscopic lesions and the exact extent of abnormalities. Presently, MRI is extensively utilized for tumor staging, image guidance, and follow-up through visual interpretation. However, the use of MRI predominantly revolves around qualitative structural information.

Radiomics, as a computational methodology, involves the extraction of quantitative features from medical images to reveal underlying tumor characteristics [4]. These features function as biomarkers that signify tumor biology and histological properties. Through the correlation of these radiomic signatures with clinical outcomes, predictive models are formulated to anticipate treatment response and survival. Many studies have utilized MRI radiomic models to predict the prognosis of LANPC patients [5, 6, 7, 8, 9, 10, 11, 12, 13]. Nevertheless, most of these studies are based on conventional MRI sequences and have not explored the incremental value of diffusion-weighted imaging (DWI) to the predictive models.

Machine learning (ML) algorithms are transforming the prognosis prediction landscape for LANPC [14], which identify intricate patterns and features within medical images, surpassing human capabilities. By integrating multimodal imaging data, ML models enable precise and robust prognostication. Combining radiomics with ML presents a synergistic approach to LANPC prognosis prediction. By merging the wealth of quantitative information from radiomics with the advanced pattern recognition capabilities of ML, comprehensive and personalized prognostic models are developed. These models provide clinicians with invaluable decision support, guiding tailored treatment strategies for LANPC patients.

In this current study, we aimed to develop and validate a machine learning-based radiomic model including DWI-derived features to predict the progression-free survival (PFS) of LANPC patients. This tool has the potential to enhance treatment decision-making, improve patient counseling, and ultimately lead to better clinical outcomes. In addition, we explored the added value of DWI compared to conventional MRI in a radiomic modeling approach.

Materials and methods

Patient cohort

This retrospective cohort study was approved by our Ethics Committee and the requirement to obtain informed written consent was waived. The study was performed in accordance with the Declaration of Helsinki. Between August 2015 and November2018, the medical records of 232 NPC patients were collected; the inclusion criteria were as follows: [1] patients with pathologically confirmed newly diagnosed NPC; [2] patients aged ≥ 18 years; [3] Eastern Cooperative Oncology Group performance status of 0–1; [4] patients were classified as III-IVa according to the 8th edition of AJCC system; [5] patients who received concurrent chemoradiotherapy (CCRT) after 2 or 3 cycles of induction chemotherapy (IC); and [6] MRI examinations were performed before IC regimens. The exclusion criteria were as follows: [1] patient has concomitant severe diseases of organs such as the heart, lungs, and brain; [2] patient has concomitant primary tumors in other organs; [3] previous anti-cancer therapy before MRI scans; [4] incomplete MR images and low quality images; and [5] lost to follow up. The patient enrollment process is illustrated in Fig. 1. Finally, a total of 126 patients were included for analysis.

All patients receive 2–3 cycles of platinum-based doublet or triplet chemotherapy regimens (such as TP, GP, PF, TPF) for induction chemotherapy. The dose of platinum-based agents (e.g., cisplatin/oxaliplatin/ nedaplatin) is 60–80 mg/m^2, administered intravenously every 3 weeks as one cycle. Doublet regimens consist of: platinum agents combined with paclitaxel liposome

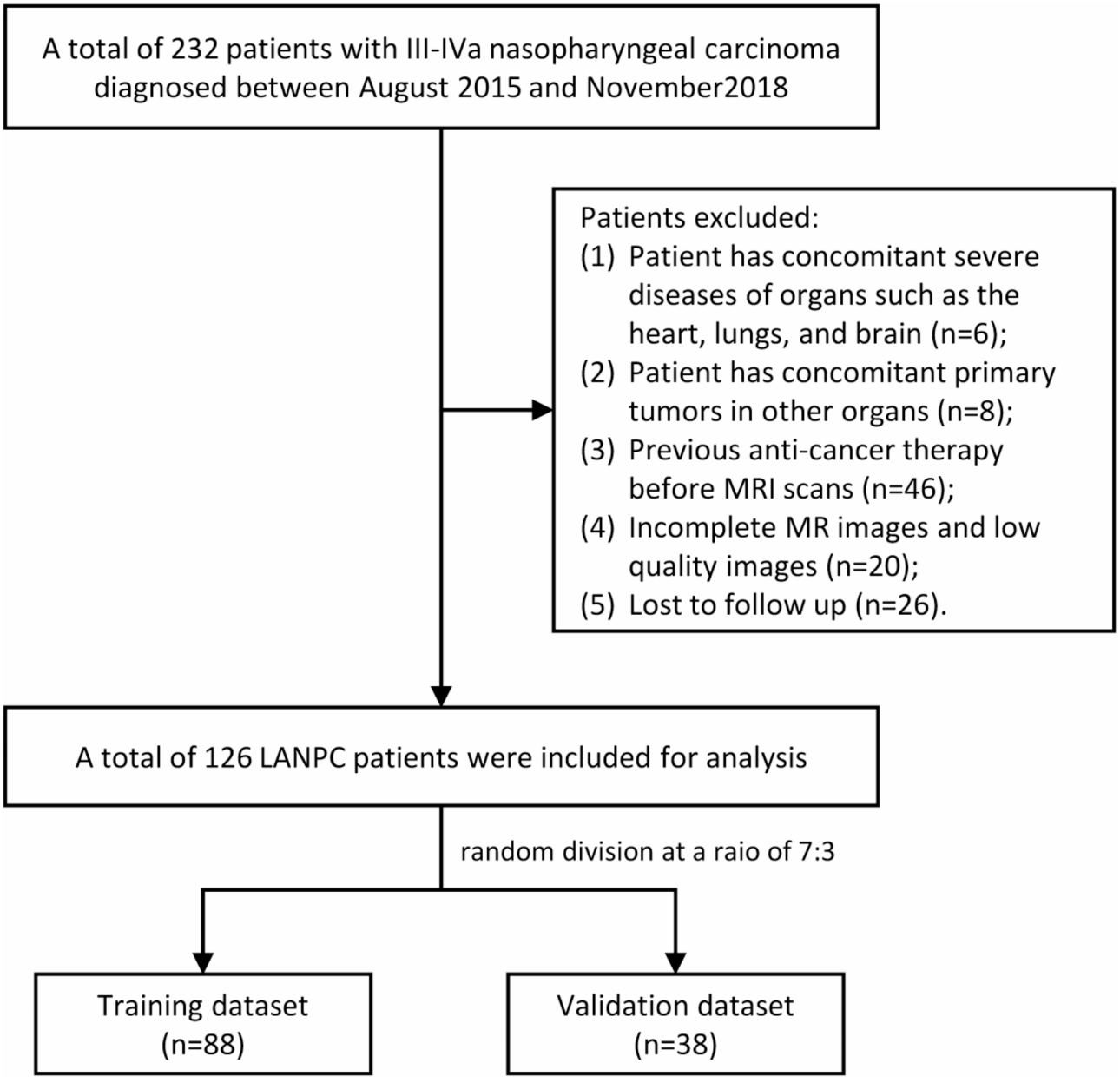

Fig. 1 The flowchart of patient inclusion

(135–175 mg/m² intravenous infusion); platinum agents combined with docetaxel (60–75 mg/m² intravenous infusion); platinum agents combined with fluorouracil (300–500 mg/m²continuous intravenous infusion for 3–5 days). Triplet regimens consist of: platinum agents combined with paclitaxel liposome (135–175 mg/m² intravenous infusion) + fluorouracil (300–500 mg/m² continuous intravenous infusion for 3–5 days); platinum agents combined with docetaxel (60–75 mg/m² intravenous infusion) + fluorouracil (300–500 mg/m² continuous intravenous infusion for 3–5 days). After IC therapy, all patients underwent definitive intensity-modulated radiation therapy (IMRT) using a 6MV-X linear accelerator. IMRT was delivered with the patient immobilized in a thermoplastic mask for head, neck, and shoulder positioning. CT simulation was performed with a slice thickness of 3 mm, initiated after the completion of the final session of IC. The radiation doses were as follows: PGTVnx and PGTVrpn received 68–72 Gy/ 30–33 fractions, PGTVnd received 66–70 Gy/30–33 fractions, PCTV1 received 60–62 Gy/30–33 fractions, and PCTV2 received 50–56 Gy/30–33 fractions. Specifically, the gross tumor volume (GTV) was outlined based on the tumor regression area post-IC, with particular attention to including the pre-treatment skull base invasion area within the GTV for radical dosing. Simultaneously, ensuring

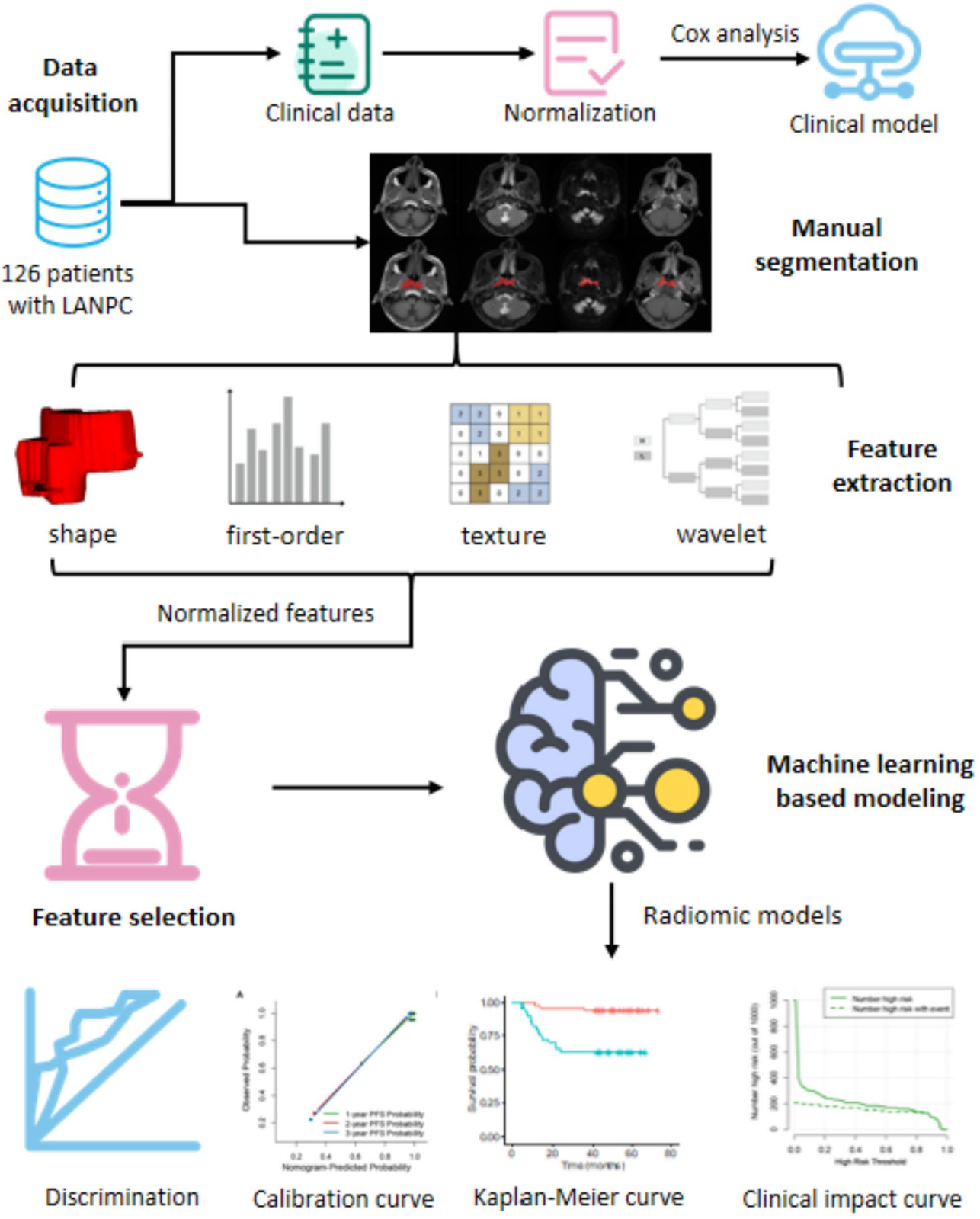

Fig. 2 Radiomic workflow for predicting prognosis in LANPC. The key steps of the radiomics approach encompass clinical and MRI data acquisition, manual segmentation of MRI images, feature extraction and selection, model development, and comprehensive evaluation of the model in terms of discrimination, calibration, and clinical utility

Table 1 Patient characteristics

Characteristics	Training dataset (*n* = 88)	Validation dataset (*n* = 38)	*P*-value
Sex			0.134
Male	69 (78.4)	25 (65.8)	
Female	19 (21.6)	13 (34.2)	
Age (years)	43.3 ± 9.1	46.0 ± 9.2	0.128
Overall stage			0.639
III	48 (54.5)	19 (50.0)	
IV	40 (45.5)	19 (50.0)	
Histology, WHO Type			0.275
I	2 (2.3)	0	
II	28 (31.8)	17 (44.7)	
III	58 (65.9)	21 (55.3)	
EBV-DNA (copies/mL)			0.585
≤400	46 (52.3)	18 (47.4)	
>400	26 (29.5)	10 (26.3)	
Unknown	16 (18.2)	10 (26.3)	
IC regimens			0.853
TP	19 (21.6)	8 (21.1)	
DP	19 (21.6)	10 (26.3)	
DPF	45 (51.1)	19 (50.0)	
TPF	5 (5.7)	1 (2.6)	
IC cycles			0.880
2	52 (59.1)	23 (60.5)	
3	36 (40.9)	15 (39.5)	
Treatment response			0.287
Responder	75 (85.2)	35 (92.1)	
Non-responder	13 (14.8)	3 (7.9)	
Median PFS (months)	49.1 (41.4, 58.7)	48.8 (41.9, 54.7)	0.776

Abbreviations: EBV DNA, Epstein-Barr virus DNA; IC, Induction chemotherapy; TP, paclitaxel liposome + cisplatin/nedaplatin; DP, docetaxel + Cisplatin / Nedaplatin; DPF, docetaxel + cisplatin/nedaplatin + 5-fluorouracil; TPF, paclitaxel liposome + cisplatin/nedaplatin + 5-fluorouracil; PFS, progression-free survival

coverage of the entire tumor area prior to IC within the low-risk clinical target volume (CTV). In this study, the radiotherapy plan follows the 2018 international guidelines for delineating clinical target volumes in NPC [15]. During IMRT, patients received 1–3 cycles of platinum-based chemotherapy concurrently. The specific drug and dosage were as follows: cisplatin at a dose of 80–100 mg/m^2, administered intravenously every 21 days as one cycle of chemotherapy.

After completion of initial treatment, each patient was followed up for every 3 months in the first 2 years, then every 6 months in the third to fifth year, and annually thereafter. The primary endpoint of this study is PFS, which is defined as the time from the initiation of treatment to the date of disease progression or death from any cause. The secondary endpoints include distant metastasis-free survival (DMFS) and overall survival (OS). DMFS is defined as the period from the start of treatment until the occurrence of recurrence or metastasis, while OS is defined as the duration of time from the initiation of treatment until death from any cause.

Clinical characteristics were collected from medical records, including age, sex, TNM stage, WHO Type, Epstein-Barr virus DNA (EBV-DNA), chemoradiotherapy regimens, and treatment response to IC. TNM staging was performed following the 8th edition of the American Joint Committee on Cancer (AJCC)/Union for International Cancer Control (UICC) staging system [16]. The WHO types were classified as follows: keratinizing squamous cell carcinoma (type I), nonkeratinizing differentiated carcinoma (type II), and nonkeratinizing undifferentiated carcinoma (type III) [17]. The clinical treatment response to IC was defned as stable disease (SD), progressive disease (PD), partial response (PR), or complete response (CR) based on the Response Evaluation Criteria in Solid Tumors 1.1 (RECIST) criteria [18]. Patients were categorized into responders (CR/PR) and non-responders (SD/PD).

MRI image acquisitions

The MRI examinations were performed on a 1.5-T MRI scanner (Optima® MR360, GE Healthcare). Standard head coil was used for scanning. The parameters for each sequence were listed as follows: T1WI, repetition time (TR) / echo time (TE) = 570/8 ms, slice thickness = 5 mm, field of view (FOV) = 220 mm, matrix = 320 × 192; axial T2WI, TR/TE = 6289/8 ms, slice thickness = 5 mm, FOV = 220 mm, matrix = 320 × 192; cT1WI, TR/TE = 365/3 ms; slice thickness = 5 mm, interslice gap = 1 mm, FOV = 220 mm, matrix = 290 × 192; DWI (a single-shot diffusion-weighted spin-echo echo-planar sequence), b values = 0, 800 s/mm^2, TR/TE = 10,000/92 ms, slice thickness = 5 mm, FOV = 240 mm, matrix = 160 × 160. For the post-contrast acquisition, Gadodiamide (Omniscan™, GE Healthcare) was administrated intravenously at a rate of 3.5 ml/s and a dosage of 0.1 mmol/kg body weight, immediately followed by a rapid 20 ml normal saline solution flush.

Radiomic analysis pipeline

The radiomic workflow for predicting prognosis in LANPC is shown in Fig. 2.

Tumor segmentation

We used open-source ITK-SNAP software (version 3.6.0, www.itksnap.org) for the manual segmentation of MRI images (Fig. 1). Tumor segmentation was performed on the tumor slice-by-slice by a radiologist (L.J., with 6-year experience in NPC diagnosis) and subsequently reviewed by a board-certified radiologist (X.P.Y., > 10-year experience).

Table 2 The selected radiomic features for radiomic models

Models	Features
T1WI-based model	
	original_shape_Flatness
	original_ngtdm_Strength
	log-sigma-6-0-mm-3D_glszm_SmallAreaHighGrayLevelEmphasis
	log-sigma-6-0-mm-3D_ngtdm_Coarseness
	wavelet-LHL_firstorder_Entropy
	wavelet-LHL_glcm_SumSquares
	wavelet-LHH_firstorder_Energy
	wavelet-LHH_glszm_SizeZoneNonUniformityNormalized
	wavelet-HLL_glszm_HighGrayLevelZoneEmphasis
	wavelet-HLH_glszm_HighGrayLevelZoneEmphasis
	wavelet-HHH_firstorder_Energy
	wavelet-HHH_firstorder_Minimum
T2WI-based model	
	log-sigma-4-0-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis
	wavelet-HHL_glszm_GrayLevelNonUniformityNormalized
cT1WI-based model	
	log-sigma-6-0-mm-3D_glszm_SizeZoneNonUniformity
	original_glszm_SmallAreaHighGrayLevelEmphasis
	wavelet-LHL_glszm_SizeZoneNonUniformity
	wavelet-LLH_glszm_SizeZoneNonUniformityNormalized
	wavelet-LLH_glszm_SizeZoneNonUniformity
	original_glszm_SmallAreaEmphasis
	wavelet-HHH_glcm_MaximumProbability
	wavelet-HHH_glcm_Imc2
	wavelet-HLL_glcm_MaximumProbability
	wavelet-HHL_glszm_SizeZoneNonUniformity
	log-sigma-6-0-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis
DWI-based model	
	original_shape_Flatness
	original_ngtdm_Strength
	log-sigma-6-0-mm-3D_glszm_SmallAreaHighGrayLevelEmphasis
	log-sigma-6-0-mm-3D_ngtdm_Coarseness
	wavelet-LHL_firstorder_Entropy
	wavelet-LHL_glcm_SumSquares
	wavelet-LHH_firstorder_Energy
	wavelet-LHH_glszm_SizeZoneNonUniformityNormalized
	wavelet-HLL_glszm_HighGrayLevelZoneEmphasis
	wavelet-HLH_glszm_HighGrayLevelZoneEmphasis
	wavelet-HHH_firstorder_Energy
	wavelet-HHH_firstorder_Minimum

Image preprocessing

Utilizing the N4 bias correction algorithm in 3D Slicer software to mitigate MRI field inhomogeneity artifacts, thereby reducing the effects of radiofrequency field non-uniformity and inherent MR equipment variations. Employing a selected grayscale normalization algorithm to standardize MRI grayscale values to the range of [0, 255], mitigating grayscale discrepancies across MRI sequences due to variations in patients, acquisition times, and parameter settings, ensuring accurate and reliable texture analysis. Subsequently, resampling the regions of interest to a uniform size (1*1*1) prior to feature extraction.

Feature extraction

Using the open-source package Pyradiomics v3.0, radiomic features were extracted from the original MRI, log-filtered, and wavelet-transformed images of each sequence. This software complies with the standards established by the Image Biomarker Standardization Initiative [19]. The features include the following types: [1] shape-based features; [2] first-order statistical features; [3] texture features: grey-level co-occurrence matrix (GLCM), grey-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), gray-level dependence matrix (GLDM), and neighborhood grey-tone difference matrix (NGTDM); [4] wavelet features. A total of 962 image features were extracted from each MR sequence.

Feature selection and model construction

Univariate and multivariate Cox proportional hazards analyses with stepwise regression method were performed to identify independent indicators of PFS. The variables with P values < 0.05 in both univariate and multivariate analyses were included in the clinical models. The clinical model was constructed using Cox regression analysis based on these selected independent variables.

Radiomic feature selection and model construction were based on the training dataset. Firstly, Pearson correlation analysis (PCC) was performed on the extracted radiomic features from each MR sequence, and the feature set with low redundancy was preliminarily obtained (correlation coefficient threshold was set at 0.99) [20]. After preliminary screening, the radiomic feature values were standardized by Z-score algorithm. Recursive feature elimination (RFE) or Relief was used for the final feature selection. Five ML algorithms were compared to develop optimal single-layer and fusion radiomic models, including Support Vector Machine (SVM), Random forest (RF), Naive Bayes (NB), Gaussian process (GP) or Linear discriminant analysis (LDA). The feature selection and ML algorithm were optimized by cross-validation, and the optimal radiomic model was evaluated in the validation dataset. The performance evaluation metrics of the predictive model were C-index, calibration curve, and clinical impact curve. The calibration curve assesses the accuracy of a predictive model by plotting the observed event rates against the predicted probabilities. Ideally,

Table 3 The performance of clinical and radiomic models for PFS prediction

Models	C-index (95%CI)			
	Training dataset	*P*-value	Validation dataset	*P*-value
Clinical model	0.722 (0.698–0.746)	< 0.001*	0.544 (0.487–0.601)	< 0.001*
DWI-based radiomic model	0.786 (0.759–0.813)	< 0.001*	0.739 (0.712–0.766)	0.013*
T1WI-based radiomic model	0.875 (0.861–0.889)	< 0.001*	0.734 (0.709–0.759)	0.004*
cT1WI-based radiomic model	0.924 (0.910–0.938)	0.752	0.722 (0.697–0.747)	< 0.001*
T2WI-based radiomic model	0.611 (0.587–0.635)	< 0.001*	0.598 (0.541–0.655)	< 0.001*
T1WI- plus cT1WI-based radiomic model	0.927 (0.900-0.954)	0.692	0.747 (0.720–0.774)	0.038
Fusion radiomic model	0.921 (0.907–0.935)	Ref.	0.788 (0.763–0.813)	Ref.
Combined model	0.945 (0.931–0.959)	0.012*	0.788 (0.763–0.813)	1

this curve should closely follow the 45-degree diagonal line, indicating that the observed event rates align with the predicted probabilities. The calibration curve assesses the accuracy of a predictive model by plotting the observed event rates against the predicted probabilities. Ideally, this curve should closely follow the 45-degree diagonal line, indicating that the observed event rates align with the predicted probabilities. The threshold was set by X-tile 3.6.1 (Yale University, USA) [21] to divide the high- and low-risk groups. The Log-rank test was used to evaluate the significance of the survival differences between the two groups.

The combined models integrating significant clinical variables and the optimal radiomic model were developed via multivariate Cox proportional-hazards model.

Statistical analysis

Comparison of the clinicopathologic features and DNA methylation data between the training and validation cohorts used a t-test, Chi-squared test, or Mann–Whitney U-test, where appropriate. For model construction, the clinical models were built using univariate and multivariate Cox regression analyses. The radiomic model was constructed using ML algorithms, and the rad-scores were yielded by performing Cox regression analysis on the predicted probabilities generated by the model. The combined model was established by incorporating significant clinical factors, and fusion rad-scores. All predictive models were developed on the training cohort and verified in a held-out validation cohort. Python package (v 3.7.6) was employed for all statistical analyses. A two-tailed *P*value < 0.05 was considered to indicate statistically signifcant differences.

Results

Patient characteristics

The mean age of the 126 patients were 40.1 years ± 9.2, 94 (74.6%) were males. The 3-year PFS, DMFS, and OS rates of the entire cohort was 80.2%, 85.7%, 86.5%, respectively. The median follow-up time was 83 months. The demographic and clinical characteristics of patients are presented in Table 1. No statistically significant differences between the training and validation datasets were observed in sex, age, overall stage, WHO type, EBV-DNA, IC regimens, IC cycles, treatment response, and PFS distribution.

Performance of the predictive models

After Cox regression analyses of clinical variables, only TNM stage was the independent variable significantly associated with PFS (hazard ratio [HR] = 7.05, 95% confidence interval [CI]: 2.07–24.05, $P = 0.002$). The HRs and 95%CI for sex, age, WHO type, EBV-DNA, IC regimens, IC cycles, and IC response were 0.440 (95%CI: 0.101–1.912, $P = 0.273$), 1.026 (95%CI: 0.975–1.080, $P = 0.328$), 1.130 (95%CI: 0.461–2.771, $P = 0.789$), 0.911 (95%CI: 0.497–1.670, $P = 0.763$), 1.302 (95%CI: 0.758–2.237, $P = 0.339$), 2.449 (95%CI: 0.949–6.320, $p = 0.064$), 0.694 (95%CI: 0.160–3.019, $P = 0.626$). The TNM stage achieved a C-index of 0.722 (95%CI: 0.679–0.765) in the training dataset and 0.544 (95%CI: 0.487–0.601) in the validation dataset.

Table 2 shows the useful radiomic features utilized in the single-layer radiomic models. In the DWI-based radiomic model, the optimal model was developed utilizing LDA classifier, integrating 12 features chosen via the PCC and RFE methods, comprising eight wavelet, two first-order, and two texture-based features. This model attained a C-index of 0.786 (95%CI: (0.759–0.813) in the training dataset and 0.739 (95%CI: 0.712–0.766) in the validation dataset (Table 3). For the T1WI-based radiomic model, the most effective model was constructed using LDA classifier, incorporating 12 features selected through the PCC and RFE methods, including eight wavelet and four texture features. This model obtained a C-index of 0.875 (95%CI: 0.861–0.889) in the training dataset and 0.734 (95%CI: 0.709–0.759) in the validation dataset (Table 3). In the cT1WI-based radiomic model, the optimal model was built using GP classifier based on seven wavelet and four texture features selected by PCC and Relief. This model achieved a C-index of 0.924 (95%CI: 0.910–0.938) in the training dataset and 0.722 (95%CI: 0.697–0.747) in the validation dataset (Table 3). In the T2WI-based radiomic model, the

optimal model was developed using NB classifier based on one texture and one wavelet feature selected by PCC and Relief. This model yielded a C-index of 0.611 (95%CI: 0.587–0.635) in the training dataset and 0.598 (95%CI: 0.541–0.655) in the validation dataset (Table 3). Among the single-layer radiomic models, the DWI-based model achieved the highest C-index, followed by the T1WI-, cT1WI-, and T2WI-based models.

The combination of T1WI and cT1WI yielded a C-index of 0.927 (95%CI: 0.900-0.954) in the training dataset and 0.747 (95%CI: 0.720–0.774) in the validation dataset. While the fusion radiomic model based on DWI, T1WI, and cT1WI achieved higher C-index than single-layer radiomic models, with a C-index of 0.921 (95%CI: 0.907–0.935) in the training dataset and 0.788 (95%CI: 0.763–0.813) in the validation dataset (Table 3). The results suggest that the DWI provides complementary information beyond conventional MRI sequences. The T2WI-based radiomic model was not included in the fusion radiomic model. However, the addition of TNM stage failed to achieve a higher C-index (0.788, 95%CI: 0.763–0.813 in the validation dataset).

The calibration curves illustrate an excellent agreement between the PFS probabilities predicted by the fusion radiomic model and the actual observed PFS, as shown in Fig. 3A and B. The time-dependent ROC curves for the fusion radiomic model, presented in Fig. 3C and D, indicate that the model's predictive AUC improves over time (up to 60 months). The clinical impact curve reveals fluctuations in the number of high-risk individuals across different thresholds (Fig. 3E and F). Furthermore, the Kaplan-Meier curves demonstrate that the fusion radiomic model effectively stratifies patients into high-risk and low-risk subgroups with distinct PFS outcomes, surpassing the performance of other models (Fig. 4). In contrast, both the clinical model and the T2WI-based radiomic model failed to distinguish between high-risk and low-risk groups (Fig. 4).

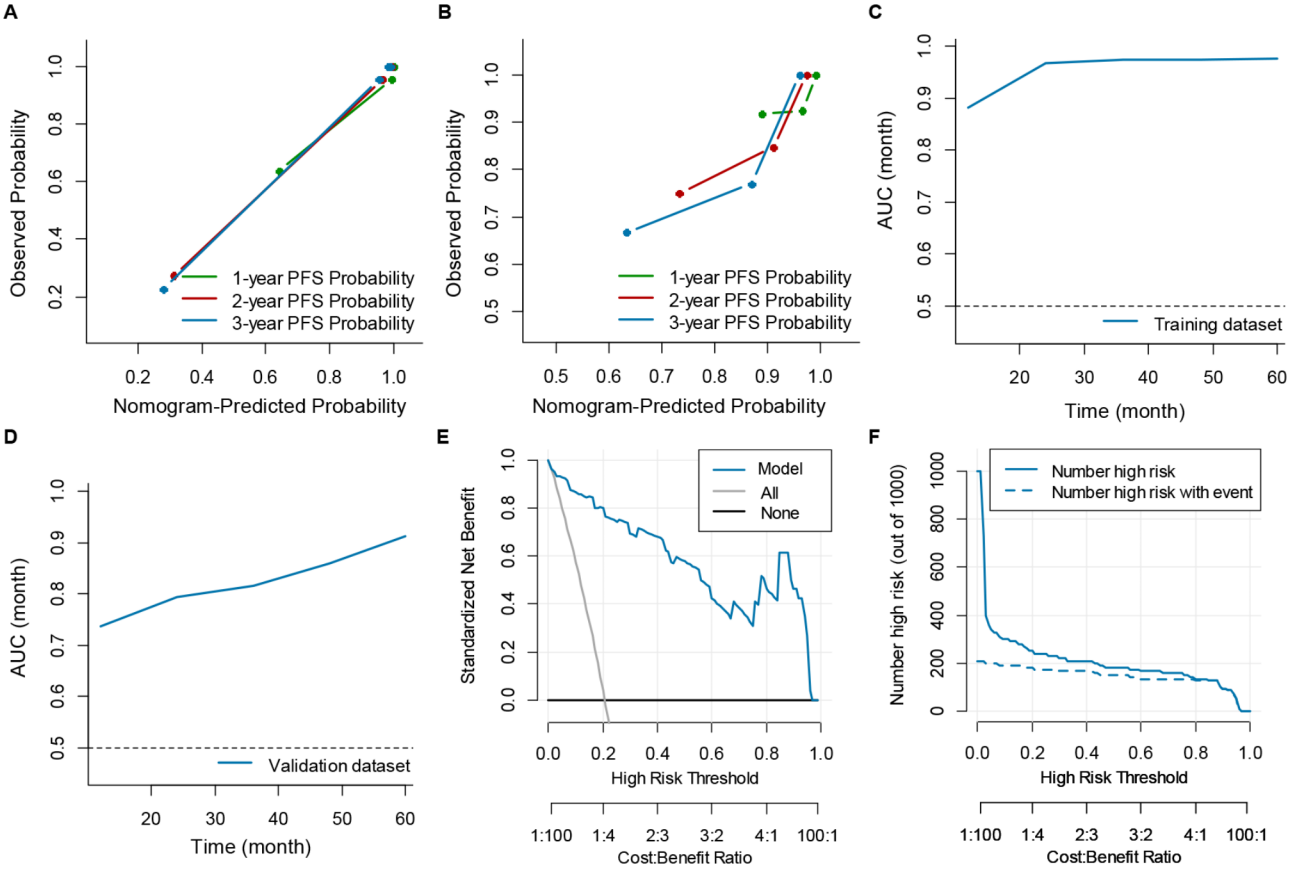

Fig. 3 Calibration curves and clinical impact curve. (**A**-**B**) Calibration curves in the training and validation datasets, respectively. The curves assess the agreement between predicted probabilities and observed outcomes, with closer alignment to the diagonal indicating better calibration. (**C**-**D**) Time-dependent ROC curves in the training and validation datasets, respectively. These curves evaluate the model's discrimination ability at specific time points, reflecting sensitivity and specificity over time. (**E**) Decision curve analysis quantifying the net benefit of the model across threshold probabilities. The curve compares the model's predictions against default strategies (treat all or treat none), with higher net benefit indicating greater clinical utility. (**F**) Clinical impact curve illustrating the relationship between predicted risk and actual clinical outcomes. The curve demonstrates how risk thresholds influence patient management decisions

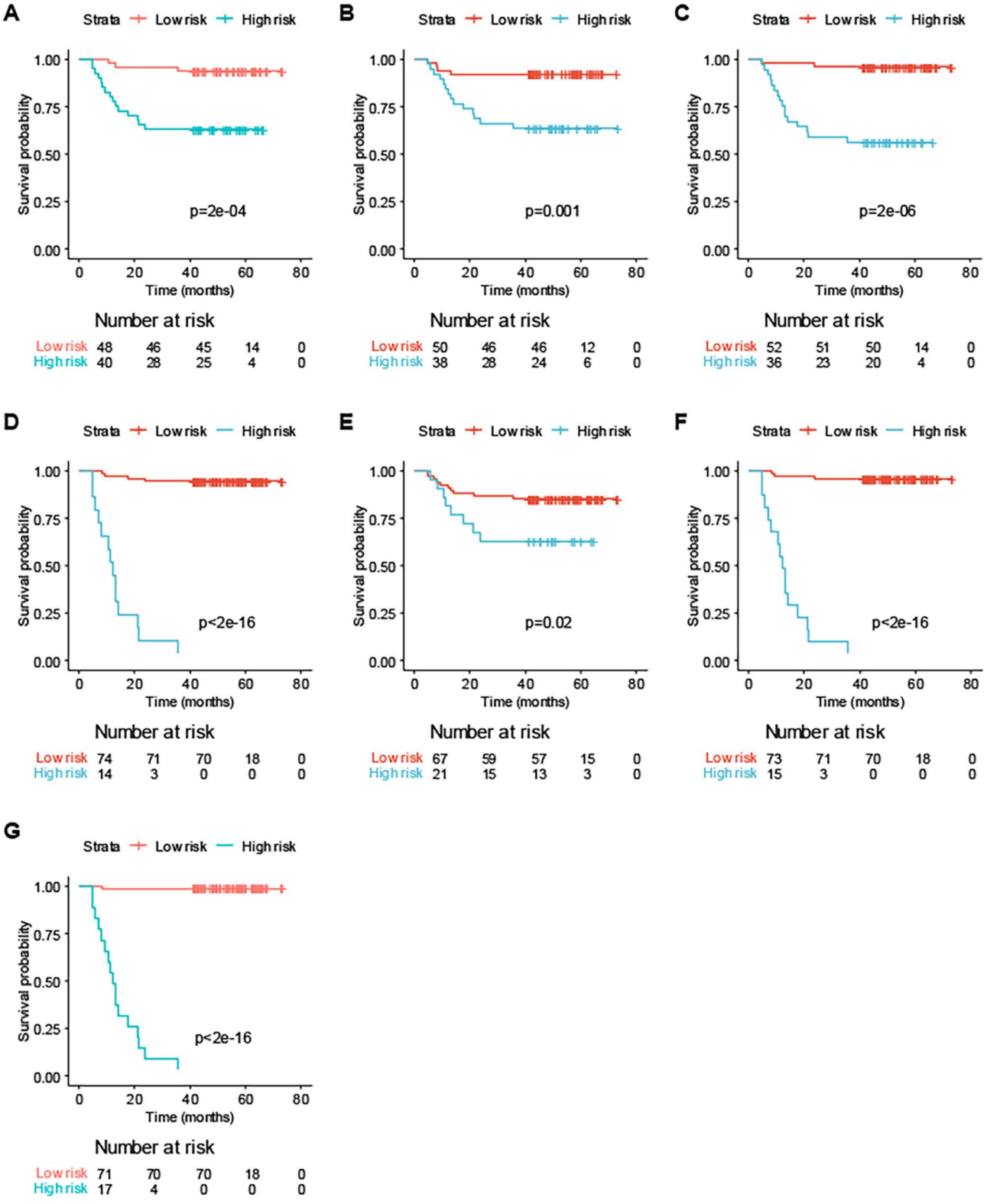

Fig. 4 Kaplan-Meier survival curves of high- and low-risk groups stratified by various models. (**A**-**G**) Training dataset: Kaplan-Meier curves for the clinical model, DWI-based radiomic model, T1WI-based radiomic model, cT1WI-based radiomic model, T2WI-based radiomic model, T1WI- plus cT1WI-based radiomic model, and fusion radiomic model, respectively. (**H**-**N**) Validation dataset: Kaplan-Meier curves for the clinical model, DWI-based radiomic model, T1WI-based radiomic model, cT1WI-based radiomic model, T2WI-based radiomic model, T1WI- plus cT1WI-based radiomic model, and fusion radiomic model, respectively

Table 4 presents the predictive performance of the fusion radiomic model, achieving a C-index of 0.888 (95% CI: 0.863–0.913) in the training dataset and 0.786 (95% CI: 0.757–0.815) in the validation dataset for predicting DMFS. For OS predictions, the model recorded a C-index of 0.776 (95% CI: 0.739–0.813) in the training dataset and 0.690 (95% CI: 0.666–0.714) in the validation dataset. Figure 5 illustrates the Kaplan-Meier survival curves for high- and low-risk groups, as stratified by the fusion radiomic model for both DMFS and OS predictions.

Discussion

Prognostic prediction is crucial for guiding personalized treatment strategies for patients with LANPC. However, few studies have explored the additional benefit of incorporating functional MRI data into conventional MRI-based radiomic models. In this study, we aimed to develop a radiomic model that integrates DWI information to predict prognosis in LANPC patients. The results demonstrate that DWI-based radiomic model achieved the highest accuracy among the single-layer radiomic models. The fusion of DWI, T1WI, and cT1WI-derived features showed the best predictive performance. Adding the TNM stage to the fusion radiomic model failed to yield any incremental value.

Several studies [5, 6, 7, 8, 9, 10, 11, 12, 13] have investigated the potential of MRI-based radiomic models in prognostic prediction of LANPC patients, aiming to improve patient stratification and guide personalized treatment strategies. These studies have demonstrated the feasibility and effectiveness of utilizing MRI radiomics to extract quantitative imaging features that capture tumor heterogeneity and other relevant biological characteristics. The accuracy of radiomic models were superior to clinical variables in predicting survivals. However, only several studies [22, 23, 24] have explored the added benefit of integrating functional MRI into conventional MRI radiomic models. Unlike conventional MRI, DWI provides valuable information about tissue cellularity and microstructural changes within tumors. By incorporating DWI features, MRI-based radiomic models can better capture tumor heterogeneity and reflect underlying biological characteristics, thereby improving the accuracy of prognostic prediction.

Our results demonstrate the feasibility and effectiveness of utilizing MRI radiomics for prognostic prediction in LANPC. By extracting a wide array of quantitative imaging features from MRI images, we could capture nuanced tumor characteristics and spatial heterogeneity, which are known to correlate with disease progression and patient outcomes [25, 26]. This highlights the utility of MRI-based radiomics as a non-invasive tool for prognostic stratification in LANPC, complementing traditional clinical factors, such as TNM stage and EBV-DNA level. Furthermore, the integration of multi-parametric MRI data, including conventional and functional MRI, holds promise for further improving prognostic models in LANPC patients. By leveraging complementary information from different MRI sequences, multi-parametric radiomics approaches can provide a more comprehensive assessment of tumor biology and microenvironment, leading to enhanced prognostic stratification and personalized treatment planning [27, 28]. The incorporation of ML algorithms enabled robust prognostic modeling based on MRI radiomic features. These models exhibited favorable predictive performance, showcasing the ability of machine learning techniques to harness the rich information contained within MRI images and extract meaningful prognostic insights. Our findings underscore the potential of ML-driven approaches in enhancing prognostic accuracy and guiding clinical decision-making in LANPC management. For example, by utilizing a cutoff value of 0.6 derived from the output of the optimal fusion radiomic model, the model effectively identifies high-risk patients. This capability enables individualized risk assessments, guiding treatment decisions to ensure that high-risk patients receive more intensive therapy, while low-risk patients may benefit from de-escalation strategies, thereby optimizing both therapeutic efficacy and safety.

Our study has some limitations. First, its retrospective design inherently comes with limitations, including reliance on existing medical records and the potential for incomplete or missing data, which could impact the accuracy and reliability of our findings. Second, the relatively small sample size may restrict the generalizability of the results, a challenge also encountered in prior studies. However, we conducted a held-out validation to mitigate bias and enhance result robustness. In the future, we plan to expand our cohort by including additional patients from multiple institutions. This will not only enhance the statistical power of our analysis but also improve the representativeness of our findings across diverse populations. Third, there is a need for standardized imaging protocols and feature extraction methodologies to ensure reproducibility and generalizability across different institutions. Fourth, although tumor segmentation was performed by experienced radiologists, the manual process may introduce subjectivity. The adoption of automated segmentation techniques could enhance reproducibility. Given the scale of our current dataset, we opted

Table 4 The performance of the fusion radiomic model for DMFS and OS prediction

Models	C-index (95%CI)	
	Training dataset	Validation dataset
DMFS	0.888 (0.863–0.913)	0.786 (0.757–0.815)
OS	0.776 (0.739–0.813)	0.690 (0.666–0.714)

for manual segmentation with expert review to ensure high-quality annotations. Nevertheless, we acknowledge the potential advantages of deep learning-based or semi-automated segmentation methods, which could be investigated in future studies to further improve reproducibility. Fifth, the radiomic models did not include

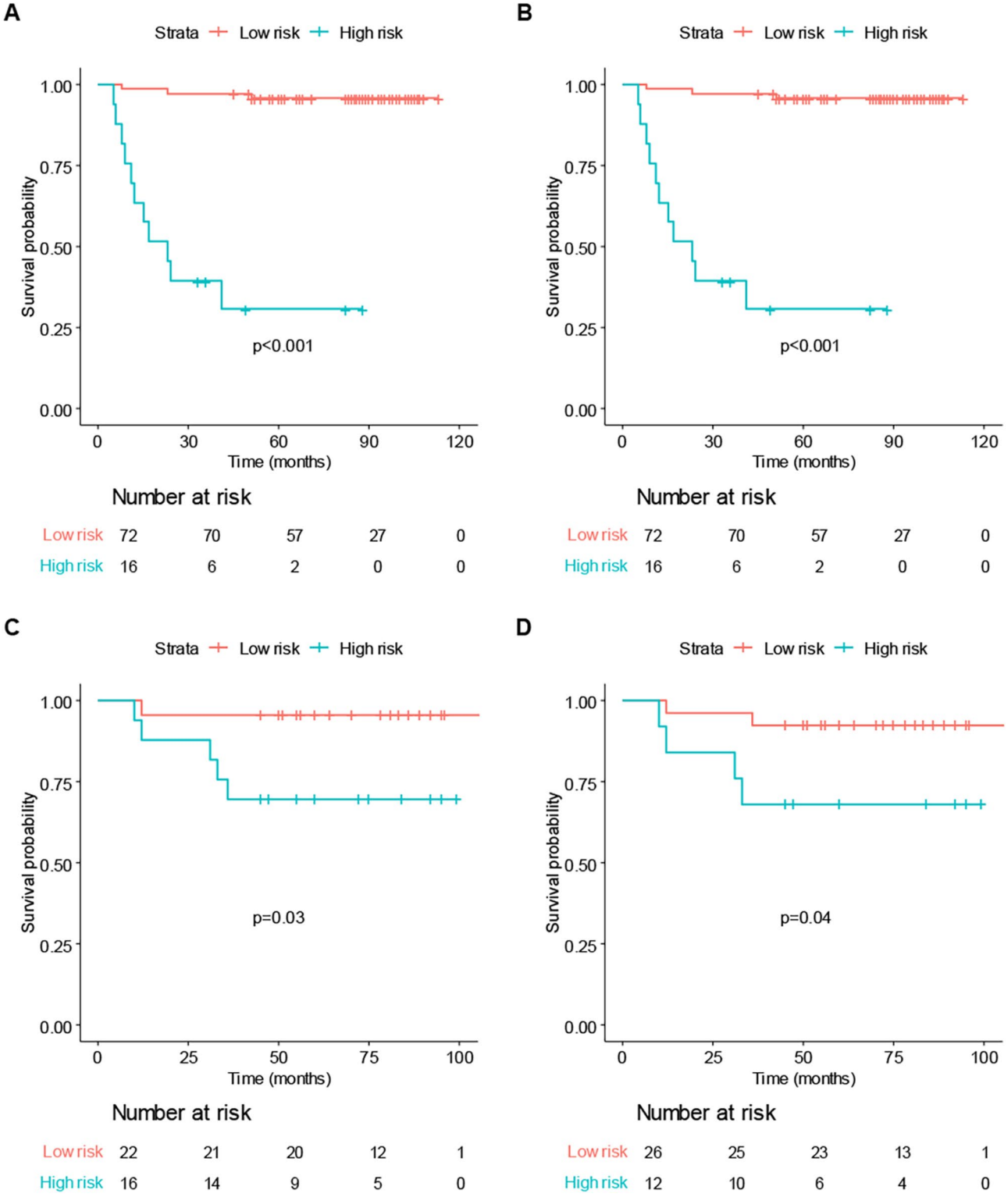

Fig. 5 Kaplan-Meier survival curves of high- and low-risk groups stratified by the fusion radiomic model for DMFS and OS prediction. (**A**-**B**) training dataset: DMFS, OS; (**C**-**D**) validation dataset: DMFS, OS

the post-treatment MRI images to enhance the predictive performance in the evaluation of PFS. Sixth, the biological interpretation of radiomic features remains challenging and warrants further investigation. Exploring associations with digital pathology features, performing radiology-pathology co-registration, and assessing biological pathways or genomic correlations could offer valuable insights into the underlying biology and mechanisms driving the observed radiomic patterns [29, 30]. Finally, external validation of the acquired data is essential to confirm the reliability and reproducibility of the findings. Future prospective studies should strive to validate the results using independent datasets to ensure the generalizability of the MRI-based radiomic models. Addressing these challenges will be pivotal in fully leveraging the potential of MRI radiomics to enhance prognostic prediction and patient management in LANPC.

In conclusion, our study demonstrates that the multiparametric MRI radiomic model has the potential to assess the prognostic risk in LANPC patients. The radiomic model, which includes T1WI, cT1WI, and DWI-derived features, offers a noninvasive approach to predict survival outcomes. It can also help to distinguish high-risk patients from low-risk patients. This model has the potential to contribute to the field of precision medicine for LANPC, fulfilling the ultimate goal of personalized treatment for patients.

Abbreviations

LANPC	Locoregionally advanced nasopharyngeal carcinoma
PFS	Progression-free survival
T1WI	T1-weighted imaging
T2WI	T2-weighted imaging
cT1WI	Contrast-enhanced T1WI
DWI	Diffusion-weighted imaging
TNM	Tumor-node-metastasis
MRI	Magnetic resonance imaging
ML	Machine learning
CCRT	Concurrent chemoradiotherapy
IC	Induction chemotherapy
IMRT	Intensity-modulated radiation therapy
GTV	Gross tumor volume
CTV	Clinical target volume
EBV-DNA	Epstein-barr virus DNA
AJCC	American joint committee on cancer
SD	Stable disease
PD	Progressive disease
PR	Partial response
CR	Complete response
RECIST	Response Evaluation Criteria in Solid Tumors
TR	Repetition time
TE	Echo time
FOV	Field of view
PCC	Pearson correlation analysis
RFE	Recursive feature elimination
SVM	Support vector machine
RF	Random forest
NB	Naive bayes
GP	Gaussian process
LDA	Linear discriminant analysis

Acknowledgements
Not applicable.

Author contributions
LJ, ZDA, and HL: manuscript preparation and literature research. LJ, CS, and HPL: data collection. LJ, HDL, and PSH: data analysis and interpretation. LJ and HPL: study conception and design, manuscript review and guarantor of integrity of the entire study. All authors have read and approved the final manuscript.

Funding
This work was supported by Grants from the National Natural Science Foundation of China (Grant No. 81502658), the Key Research and Development Program of Hunan Province (No. 2021SK51117), the Natural Science Foundation of Hunan Province (No.2023JJ60327), Scientific Research Project of Hunan Provincial Health Commission (W20243197), Hunan Provincial Health High-Level Talent Scientific Research Project(No. 20240304101).

Data availability
All data generated or analysed are included in this article.

Declarations

Ethics approval and consent to participate
The study was reviewed and approved by the Institutional Ethics Committee of Hunan Cancer Hospital ethics committee. Written Informed consent was waived by the institutional review Board.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Hunan Cancer Hospital, Central South University, No.283, TongzipoRoad, Changsha 410013, China
[2]the Affiliated Changsha Central Hospital, University of South China, No.283, TongzipoRoad, Changsha 410004, China

Received: 12 May 2024 / Accepted: 10 March 2025
Published online: 21 March 2025

References

1. Chen YP, Chan ATC, Le QT, et al. Nasopharyng Carcinoma Lancet. 2019;394(10192):64–80.
2. Liu Y, Sun S, Zhang Y, et al. Predictive function of tumor burden-incorporated machine-learning algorithms for overall survival and their value in guiding management decisions in patients with locally advanced nasopharyngeal carcinoma. J Natl Cancer Cent. 2023;3(4):295–305.
3. Chen YP, Ismaila N, Chua MLK, et al. Chemotherapy in combination with radiotherapy for Definitive-Intent treatment of stage II-IVA nasopharyngeal carcinoma: CSCO and ASCO guideline. J Clin Oncol. 2021;39(7):840–59.
4. Bera K, Braman N, Gupta A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. Nat Rev Clin Oncol. 2022;19(2):132–46.
5. Qiang M, Li C, Sun Y, et al. A prognostic predictive system based on deep learning for locoregionally advanced nasopharyngeal carcinoma. J Natl Cancer Inst. 2021;113(5):606–15.
6. Zhang B, Tian J, Dong D, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res. 2017;23(15):4259–69.
7. Zhang L, Dong D, Li H, et al. Development and validation of a magnetic resonance imaging-based model for the prediction of distant metastasis before initial treatment of nasopharyngeal carcinoma: A retrospective cohort study. EbioMedicine. 2019;40:327–35.
8. Zhang B, Luo C, Zhang X, et al. Integrative scoring system for survival prediction in patients with locally advanced nasopharyngeal carcinoma: A retrospective multicenter study. JCO Clin Cancer Inf. 2023;7:e2200015.
9. Zhang L, Wu X, Liu J, et al. MRI-Based Deep-Learning model for distant Metastasis-Free survival in locoregionally advanced nasopharyngeal carcinoma. J Magn Reson Imaging. 2021;53(1):167–78.

10. Zhang L, Zhou H, Gu D, et al. Radiomic nomogram: pretreatment evaluation of local recurrence in nasopharyngeal carcinoma based on MR imaging. J Cancer. 2019;10(18):4217–25.
11. Zhang B, Ouyang F, Gu D, et al. Advanced nasopharyngeal carcinoma: pre-treatment prediction of progression based on multi-parametric MRI radiomics. Oncotarget. 2017;8(42):72457–65.
12. Shen H, Yin J, Niu R, et al. MRI-based radiomics to compare the survival benefit of induction chemotherapy plus concurrent chemoradiotherapy versus concurrent chemoradiotherapy plus adjuvant chemotherapy in locoregionally advanced nasopharyngeal carcinoma: A multicenter study. Radiother Oncol. 2022;171:107–13.
13. Yang K, Tian J, Zhang B, 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.
14. Zhang B, He X, Ouyang F, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. Cancer Lett. 2017;403:21–7.
15. Lee AW, Ng WT, Pan JJ, et al. International guideline for the delineation of the clinical target volumes (CTV) for nasopharyngeal carcinoma. Radiother Oncol. 2018;126(1):25–36.
16. Amin MAJCC. Cancer staging manual. 8th ed. Berlin/Heidelberg, Germany: Springer. 2018.
17. Ou SI, Zell JA, Ziogas A, et al. Epidemiology of nasopharyngeal carcinoma in the united States: improved survival of Chinese patients within the keratinizing squamous cell carcinoma histology. Ann Oncol. 2007;18(1):29–35.
18. Schwartz LH, Seymour L, Litière S, eet al, et al. RECIST 1.1 - Standardisation and disease-specific adaptations: perspectives from the RECIST working group. Eur J Cancer. 2016;62:138–45.
19. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for High-Throughput image-based phenotyping. Radiology. 2020;295:328–38.
20. Pearson's Correlation Coefficient. In. In: Kirch W, editor. Encyclopedia of public health. Dordrecht: Springer. 2008. https://doi.org/10.1007/978-1-4020-5614-7_256
21. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004;10:7252–9.
22. Hu Q, Wang G, Song X, et al. Machine learning based on MRI DWI radiomics features for prognostic prediction in nasopharyngeal carcinoma. Cancers (Basel). 2022;14(13):3201.
23. Guo Y, Dai G, Xiong X, et al. Intravoxel incoherent motion radiomics nomogram for predicting tumor treatment responses in nasopharyngeal carcinoma. Transl Oncol. 2023;31:101648.
24. Li WZ, Wu G, Li TS, et al. Dynamic contrast-enhanced magnetic resonance imaging-based radiomics for the prediction of progression-free survival in advanced nasopharyngeal carcinoma. Front Oncol. 2022;12:955866.
25. Moskowitz CS, Welch ML, Jacobs MA, et al. Radiomic analysis: study design, statistical analysis, and other bias mitigation strategies. Radiology. 2022;304(2):265–73.
26. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures. They Are Data Radiol. 2016;278(2):563–77.
27. Liu T, Dong D, Zhao X, et al. Radiomic signatures reveal multiscale intratumor heterogeneity associated with tissue tolerance and survival in re-irradiated nasopharyngeal carcinoma: a multicenter study. BMC Med. 2023;21(1):464.
28. Jing B, Deng Y, Zhang T, et al. Deep learning for risk prediction in patients with nasopharyngeal carcinoma using multi-parametric MRIs. Comput Methods Programs Biomed. 2020;197:105684.
29. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. Radiology. 2021;298(3):505–16.
30. Sun Q, Chen Y, Liang C, et al. Biologic pathways underlying prognostic radiomics phenotypes from paired MRI and RNA sequencing in glioblastoma. Radiology. 2021;301(3):654–63.

Publisher's note