

Radiomic analysis based on machine learning of multi-sequences MR to assess early treatment response in locally advanced nasopharyngeal carcinoma

Science Progress

2025, Vol. 108(2) 1–27

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00368504251338930

journals.sagepub.com/home/sci



Lei Qiu^{1,2,*}, Yinjiao Fei^{1,*}, Yuchen Zhu^{1,2,*}, Jinling Yuan^{1,2},
Kexin Shi^{1,2} , Mengxing Wu^{1,2}, Gefei Jiang^{1,2}, Xingjian Sun^{1,2},
Jinyan Luo¹, Yurong Li³, Weilin Xu¹, Yuandong Cao¹
and Shu Zhou¹

Abstract

Objective: The prediction of early response in locally advanced nasopharyngeal carcinoma (LA-NPC) after concurrent chemoradiotherapy (CCRT) is important for determining the need for timely consolidation therapy. We developed a radiomic analysis of

¹Department of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China

²The First School of Clinical Medicine, Nanjing Medical University, Nanjing, Jiangsu, China

³Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

*Co-first authors. These authors contributed equally to this work.

Corresponding authors:

Shu Zhou, Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, Jiangsu, China.
Email: zhoushu164086035@126.com

Yuandong Cao, Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, Jiangsu, China.
Email: yuandongcao@163.com

Weilin Xu, Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, Jiangsu, China.
Email: weilinxu2018@163.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

multi-sequences MR based on machine learning (ML) to assess early response in LA-NPC after CCRT.

Methods: This study retrospectively enrolled 104 LA-NPC patients, randomly divided into training (70%) and test (30%) cohorts. Radiomic features were extracted from five MR sequences (T1, T1C, T2, DWI, and ADC). Feature selection was performed using Pearson's correlation coefficient and LASSO regression to reduce redundancy. ML algorithms were compared to develop models, with suboptimal sequences excluded from the multi-sequence MR fusion model. A combined model integrating the fusion and clinical model was developed using logistic regression, and its diagnostic effectiveness was evaluated using receiver operating characteristic (ROC) analysis.

Results: In the mono-sequence MR analysis, T1 demonstrated the lowest discriminative capacity (AUC = 0.505), followed by T2 (AUC = 0.738). Consequently, we developed a fusion model incorporating ADC, DWI, and T1C features while excluding T1 and T2. In the test cohort, the combined model outperformed both the clinical (AUC = 0.852) and fusion (AUC = 0.886) models, achieving superior effectiveness (AUC = 0.900). Shapley Additive Explanations (SHAP) analysis identified `lbp_GrayLevelVariance_ADC` as the most influential predictive feature.

Conclusions: A combined model, which merges clinical and multi-sequences MR radiomics model, showed good performance for predicting early response of LA-NPC after CCRT.

Keywords

Multi-sequence, radiomics, locally advanced nasopharyngeal carcinoma, early response, machine learning, SHAP

Introduction

Nasopharyngeal carcinoma (NPC) is a malignant epithelial neoplasm that arises in the nasopharynx and exhibits a significant prevalence in East and Southeast Asia.¹ The diagnosis of NPC is often made at a locally advanced stage, with estimates indicating that this is the case in 75 to 90% of instances due to its deep localization and vague clinical presentation.^{2,3}

Concurrent chemoradiation (CCRT) is the backbone treatment in locally advanced nasopharyngeal carcinoma (LA-NPC).⁴ Following CCRT, 58%-89% of patients diagnosed with LA-NPC attain a clinical complete response (CR) within a few months.⁵⁻⁷ Early tumor regression has been identified as an independent prognostic indicator for overall survival (OS) and progression-free survival (PFS) in individuals diagnosed with NPC.^{8,9} Besides, research conducted by Huang et al. revealed that early intervention in NPC patients with radiographic residual disease resulted in improved short-term efficacy and long-term survival outcomes.¹⁰ Therefore, it is essential to predict individuals with LA-NPC who still have residual lesions after undergoing systematic treatment as early as possible. This enables the implementation of timely consolidation or salvage

therapy. A prior investigation conducted in 2023 established a magnetic resonance imaging (MRI)-based clinical radiomics nomogram aimed at forecasting the early response to tumor treatment.¹¹ However, in order to enhance the predictive performance, this study intends to augment the sample size and integrate novel machine learning techniques that leverage multi-sequence MR data.

Radiomics refers to the process of extracting high-dimensional data from medical images, making them amenable for analysis and mining purposes. It has facilitated the evolution of medical imaging from a tool primarily used for diagnostic purposes to one that serves as a clinical decision support system within the framework of personalized medicine.¹² Although radiomics studies have shown encouraging outcomes in forecasting treatment response in various tumor types, it is worth noting that the majority of conventional radiomics approaches predominantly concentrate on mono-sequence (e.g. MRI contrast-enhanced T1-weighted imaging in Bramen et al.; CT imaging in Jiang et al.)^{13–15} Multi-sequences MR imaging, in contrast to traditional radiomics, integrates two or more sequences into a unified system. This integration is fundamentally motivated by the complementary nature of MR sequences — while mono-sequence MR may excel in specific tasks (e.g. T1-weighted for anatomical structure, T2-weighted for edema detection), it inherently lacks the capacity to capture multifaceted biological information required for complex clinical diagnoses. By synergizing sequences, multi-sequence MRI overcomes individual limitations and provides a holistic characterization of tissues through high spatial resolution, improved soft tissue contrast, and the capacity to deliver molecular-level biological information with high sensitivity.¹⁶

This systemic superiority has been consistently validated in recent studies, where multi-sequence models demonstrated significantly enhanced performance compared to mono-sequence approaches in predicting rectal cancer response to chemoradiotherapy, predicting disease-free survival in early-stage squamous cervical cancer, improving pre-operative assessment of bladder cancer, and enabling precise evaluation of rectal cancer preoperative immunescore.^{17–22}

We searched for relevant multi-sequence MR studies in the field of NPC. In the studies conducted by Cai and Shi, they integrated all radiomic features from T1-weighted images (T1-WI), contrast-enhanced T1-weighted images (T1-C), and T2-weighted images (T2-WI) sequences to predict the survival outcome or early treatment response of NPC patients, but they ignored the role of ADC and DWI.^{23,24} Zhang's study, compared the radiomic features obtained from individual T1-C, T2, and combined T1-C and T2 images, concluding that the combined images had better predictive efficacy.²⁵ However, the study did not consider other imaging sequences and lacked an in-depth discussion on the selection and integration of the T1-C and T2 sequences. Lu et al.'s research utilized PET/CT alongside integrated T1, T1-C, and T2 radiomic models to forecast the survival outcomes of patients diagnosed with NPC.²⁶ Similarly, the study did not provide a discussion on the rationale behind the selection of these specific sequences. Previous researches have shown that radiomic models utilizing Apparent Diffusion Coefficient (ADC) and Diffusion Weighted Imaging (DWI) MRI sequences possess the capability to accurately forecast the individual prognoses of different malignancies, including malignant bone marrow lesions and prostate cancer.^{27,28} Nonetheless, the effectiveness

of radiomic models utilizing DWI and ADC sequences in assessing the early treatment response of NPC remains uncertain. Unfortunately, while prior studies predominantly focused on structural sequences (T1/T2/T1C), these studies above did not thoroughly examine the significance of DWI and ADC sequences in predicting outcomes. Additionally, they lacked in-depth discussion regarding the rationale behind selecting specific combinations of T1, T2, and T1-C sequences. Further study is warranted to explore the potential contributions of DWI and ADC sequences in outcome prediction and to provide a comprehensive understanding of the selection criteria for specific imaging sequence combinations.

This study unveils a combined model that fuses clinical data with multi-sequence MR radiomics, powered by machine learning, to forecast the early response of patients with LA-NPC after undergoing CCRT.

Materials and methods

Study patients

We declare that this study was conducted in strict compliance with the ethical principles outlined in the Declaration of Helsinki (revised in 2024).

From January 2020 to November 2023, we conducted a retrospective study of consecutive patients with pathologically confirmed NPC, involving a dataset of 219 patients. The stage of the disease was classified using the tumor node metastasis (TNM) staging, as delineated in the eighth edition of the American Joint Committee on Cancer (AJCC) guidelines.

Individuals who satisfied the specified criteria were incorporated into this study²⁹:

Inclusion criteria: (a) Nasopharyngeal squamous cell carcinoma was confirmed through pathological examination; (b) Classified as stage III to IVA based on pre-CCRT MRI; (c) MRI images of the nasopharyngeal neck region were obtained both prior to and following treatment, encompassing T1-WI, T1-C, T2-WI, DWI, and ADC; (d) All patients underwent radical CCRT; (e) Complete clinical data.

Exclusion criteria included: (a) Inadequate MRI quality resulting from issues such as motion artifacts, blurring, or discontinuities in the images; (b) Background of past treatments for NPC; (c) History of prior malignancy; (d) The presence of concurrent immune system disorders or a background of prolonged hormone medication use; (e) Disease progression after induction chemotherapy (IC) or receiving treatment elsewhere after IC.

In total, 104 patients diagnosed with LA-NPC were taken into this study. We have de-identified all patient details. The patient selection process is illustrated in Figure 1.

Treatment protocols

The IC regimen employed in this study consisted of a taxane and platinum combination (TP), with intravenous administration of docetaxel at a dosage of 75 mg per square meter or paclitaxel at a dosage of 135–175 mg per square meter, in conjunction with nedaplatin or cisplatin at a dosage of 80 mg per square meter on day 1, given once every 3 weeks for a

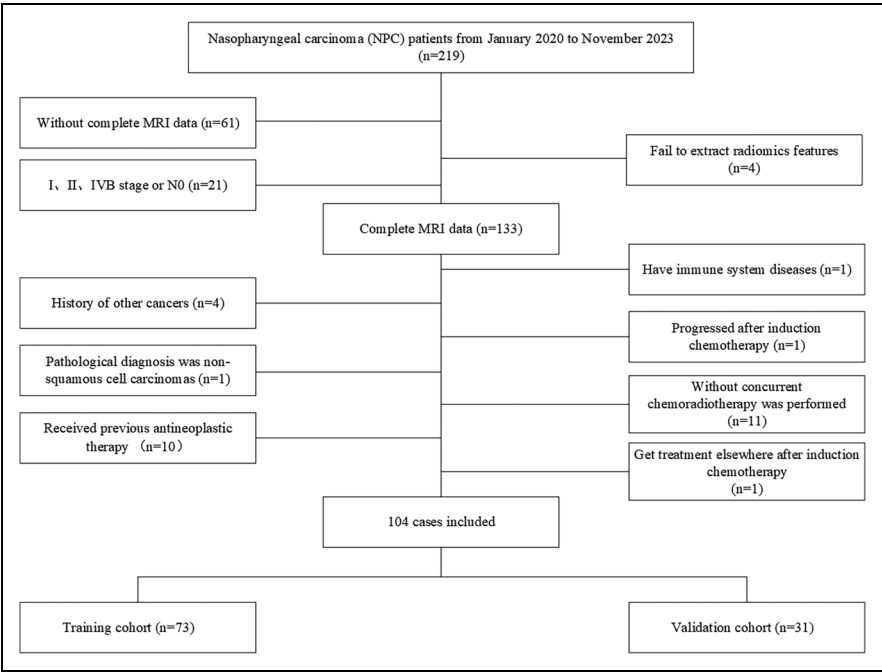


Figure 1. Flowchart for the inclusion of patients.

total of 2–3 cycles. Following the conclusion of IC, it was advised that CCRT be conducted within a timeframe of three to four weeks. Radiation therapy (RT) in this study was administered using intensity-modulated radiotherapy (IMRT) mode, utilizing 6 MV photon irradiation. The recommended doses for the planning target volumes (PTVs) derived from the gross tumor volume of the nasopharyngeal carcinoma (GTVnx), gross tumor volume of cervix node (GTVnd), clinical target volume 1 (CTV1), and clinical target volume 2 (CTV2) were 66–70 Gy, 64–70 Gy, 60–62 Gy, and 54–56 Gy, respectively, in 30–33 fractions. Concurrent with RT, chemotherapy was administered intravenously with either cisplatin or nedaplatin at a dosage of 80 mg per square meter on the first day, repeated every 3 weeks. The number of chemotherapy cycles was tailored according to specific physical condition of each patient. It is advised that an MRI examination be conducted within one to three months following the completion of CCRT.

Clinical characteristics acquisition

Clinical characteristics prior to treatment were gathered from the health information system (HIS). These characteristics included age, sex, weight, height, smoking status, drinking habits, family history, EBV DNA level, white blood cell count (WBC), platelet count

(PLT), neutrophil count, lymphocyte count, monocyte count, neutrophil to lymphocyte ratio (NLR), platelet to lymphocyte ratio (PLR), lymphocyte to monocyte ratio (LMR), serum albumin (Alb), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), and d-dimer levels. Furthermore, the volume of the nasopharyngeal tumor, maximal axial diameter of positive lymph node (N-axial-length), maximum coronal length of positive lymph node (N-cor-length), and the total volume of lymph nodes (N-total-volume) were obtained by tracing the region of interest (ROI) using ITK-SNAP, an open-source software (www.itk-snap.org).

Tumor response criteria

Tumor response was conducted by comparing MRI images obtained prior to treatment with those obtained one to three months after the completion of CCRT. This evaluation adhered to the Response Evaluation Criteria in Solid Tumors 1.1 (RECIST 1.1) guidelines and was performed by two radiologists respectively.³⁰ Any discrepancies identified by the two radiologists were figured out through thorough discussion and consensus.

The Cohen's Kappa coefficient(κ) is a metric used to evaluate inter-rater agreement in categorical tasks, particularly by accounting for agreement due to chance. In medical imaging, it quantifies consistency among clinicians in tasks such as tumor segmentation or diagnostic classification, ensuring data reliability.³¹ The metric ranges from 0 to 1, with higher values indicating stronger consistency. Cohen's Kappa was used for inter-user agreement of observers.

$$\text{Formula} \quad \kappa = \frac{Po - Pe}{1 - Pe}$$

The responses were categorized into four groups: CR, partial response (PR), stable disease (SD), or progressive disease (PD). Compared with individuals who achieved CR after CCRT, the prognosis of patients with residual tumor is worse.^{9,32} Subsequently, the patients were categorized into the CR group, and non-CR group encompassing patients with PR, SD, and PD.

MRI image acquisition

Image acquisition was performed using GE 750 W 3.0 T MR or Philips Ingenia 3.0 T. Axial T1WI, T2WI, ADC, T1C and DWI pictures for each patient were obtained through PACS (Carestream, Ontario, Canada) in DICOM format. The manual segmentations of ROI were carried out utilizing ITK-SNAP by one radiation oncologists specializing in NPC radiotherapy with 3 years of experience. The accuracy of the segmentations was further validated by a senior radiation oncologist, who resolved any discrepancies through discussion. Cohen's Kappa was used for inter-user agreement of observers. In our study, we focused on segmenting the primary nasopharyngeal tumor region, while metastatic lymph nodes were not included.

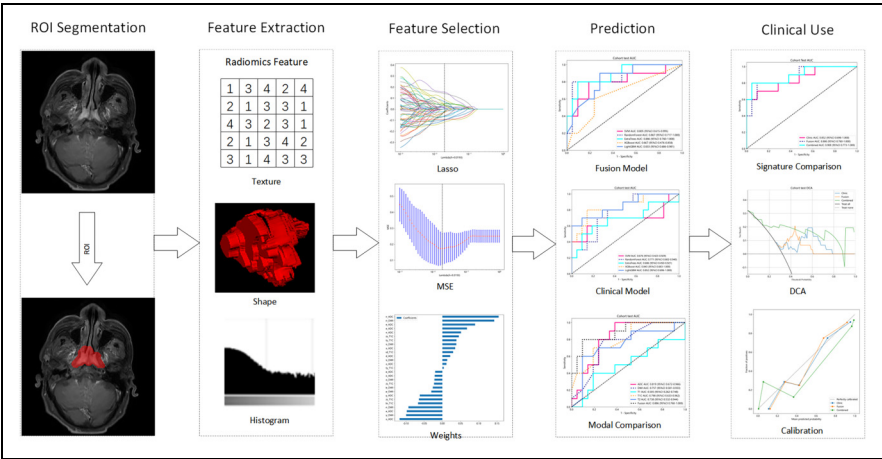


Figure 2. The flow chart of the process radiomics ROI, region of interest; Lasso, least absolute shrinkage and selection operator; T1-C, contrast-enhanced T1-weighted images; ADC, apparent diffusion coefficient; DWI, diffusion weighted imaging; MSE, mean square error; ROC, receiver operating characteristic; DCA, decision curve analysis; AUC, area under the curve.

Due to the potential variation in machine parameters, all our data underwent sampling at a 1 mm ×1 mm ×1 mm resolution, thus achieving a consistent spacing across all datasets. Figure 2 presents a visual representation of the radiomics process flowchart.

Radiomics model

For increasing the reliability and reproducibility of the results, we added radiomic quality score (RQS) report as a relevant evaluation method.³³ (Supplement Table 1)

Feature extraction. In our analysis, we meticulously extracted radiomic features that corresponded to mono-sequences (T1, T1C, T2, ADC, and DWI). Our process involved categorizing the handcrafted features into three primary types: (I) Geometry, encompassing the three-dimensional shape characteristics of tumor; (II) Intensity, involving the first-order statistical distribution of voxel intensities in the tumor; and (III) Texture, detailing the patterns or higher-order spatial distributions of intensities. To enhance feature representation, we also applied image transformations such as Laplacian of Gaussian (LoG) and Wavelet to the original data. Various techniques were utilized for extracting texture features, such as the gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-tone difference matrix (NGTDM) were employed. The extraction of all features was conducted utilizing the pyradiomics tool (version 3.0.1), adhering to the standards established by the Imaging Biomarker Standardization Initiative (IBSI).

Feature selection. In our methodology, we initially standardized all extracted features using the Z-score normalization technique, ensuring a uniform distribution for subsequent analysis. This was followed by a t-test statistical evaluation of these normalized features, wherein we selectively retained those exhibiting a p-value less than 0.05, indicative of statistical significance.

To further refine our feature set and reduce redundancy, we employed Pearson's correlation coefficient to assess collinearity among the features. In instances where a pair of features demonstrated high linear correlation (similarity coefficient greater than 0.9), we strategically retained only one of the two. This step was crucial in minimizing redundancy while preserving the integrity and diversity of the feature set.

The model is instrumental in feature selection as it applies penalties to the absolute values of the regression coefficients. To optimize this model, we determined the best regularization parameter, λ , through a rigorous 10-fold cross-validation process. This approach involved evaluating various λ values to identify the one that minimized the cross-validation error, thereby ensuring the most effective feature selection. The final model included only those features with non-zero coefficients, representing the most relevant and informative predictors for our study. This comprehensive approach to feature selection and model optimization underscores the robustness and reliability of our findings.

Model building

Radiomics model. In our study, we employed the LASSO technique for feature selection and developed a risk model utilizing a range of machine learning algorithms, specifically SVM, ExtraTrees, XGBoost, Random Forest, and LightGBM. Comparative performance assessments were carried out among each mono-sequence. Feature-level fusion, also referred to as early fusion, entails consolidating all features from various modalities into a single feature vector.³⁴ This process integrates multi-modal features into a unified feature vector prior to model training, creating a comprehensive representation that preserves inter-modal relationships. It allows the model to simultaneously learn from all available imaging data while maintaining the inherent correlations between different sequences. In this study, we utilized the early fusion method to perform feature fusion. Feature fusion was performed on multi-sequences MR features, and the fused features were used to assess the effectiveness of multi-sequences versus mono-sequence approaches. Model robustness was ensured through 5-fold cross-validation in the training dataset, with hyperparameters optimized via Grid-Search. To evaluate the effectiveness of our multi-instance learning approach, we compared two ensemble methods: maximum and average values.

Clinical model. Clinical characteristics were quantitatively represented and subjected to analysis through machine learning algorithms, including SVM, ExtraTrees, XGBoost, Random Forest, and LightGBM, similar to the Radiomics Model.

Combined model. We incorporated radiomics into the clinical model to improve its clinical applicability. The diagnostic performance of the model was evaluated using the area under receiver operating characteristic curves (AUC-ROC) as an evaluative metric. We first selected the best-performing machine learning algorithm based on AUC and used its predicted probabilities as the Rad-Signature. The Rad-Signature was then combined with selected clinical features to form a combined model. Finally, we employed Logistic Regression (LR) to integrate the Rad-Signature and clinical features, leveraging their complementary strengths to build a robust and clinically applicable model. Decision curve analysis (DCA) was employed to evaluate the clinical utility of our predictive models, while calibration curves and the Hosmer-Lemeshow test were utilized to assess the calibration of the models.

SHAP visualization

The Shapley Additive Explanation (SHAP) method quantifies each feature's contribution to the model, ensuring a fair and consistent assessment of feature importance. In this study, SHAP analysis was employed to interpret the model's feature contribution.

Statistical analysis

The dataset was split randomly, with 70% allocated for training and the remaining 30% for testing. To evaluate patient clinical characteristics, we performed statistical analyses using the independent sample t-test, Mann-Whitney U test, and chi-squared test for categorical variables. As detailed in Supplement Table 2, the analysis showed no significant statistical differences between the training and testing groups because almost all features had p-values greater than 0.05. This confirms an unbiased distribution of data across the two datasets. Significant clinical features were discerned utilizing both univariate and stepwise multiple regression analyses. The analyses were conducted utilizing Python version 3.7.12 in conjunction with the statsmodels library version 0.13.2. Additionally, the construction of machine learning models was conducted utilizing the scikit-learn version 1.0.2 interface. Scikit-learn provided a comprehensive and user-friendly framework for implementing various machine learning algorithms and conducting model training, evaluation, and prediction tasks.

Results

Patients

Patients were categorized into the CR group (n = 63, 60.6%), and non-CR group (n = 41, 39.4%). The dataset was divided into training and test sets with a ratio of 7:3. In the training set, there were 73 patients, consisting of 42 CR cases (57.5%) and 31 non-CR cases (42.5%). The test set comprised 31 patients, with 21 CR cases (67.8%) and 10 non-CR cases (32.3%). A previous study reported a 65.57% CR rate in LA-NPC patients treated with the TP regimen, which closely aligns with our current findings.³⁵

Cohen's kappa value was calculated to analyze the observers' agreement for tumor response evaluation ($k=0.92>0.8$) and handcraft tumor segmentation consistency ($k=0.85>0.8$), indicating almost perfect agreement. This strong inter-observer agreement further validates the reliability and reproducibility of our segmentation methodology.

Feature statistics

In this study, we conducted an extensive extraction of radiomic features, yielding a detailed collection of 1834 manually crafted features derived from multiple sequences. The aforementioned features were subsequently categorized into three distinct groups: shape, comprising 14 features; first-order, encompassing 360 features; and a diverse array of texture features. After evaluating the performance of different sequences, we noted that T1 and T2 were less effective. Consequently, we focused on the remaining three sequences, aggregating a total of 5502 features. The extraction was performed using a custom-developed program in Pyradiomics, further details of which are available at <http://pyradiomics.readthedocs.io>. Supplement Figure 1 visually displays the distribution of these handcrafted features across the various categories.

Evaluation of rad model

Feature selection. In the final analysis stage, we used a LASSO logistic regression model to select features with nonzero coefficients, ensuring inclusion of only the most relevant features. The range of λ values explored in this research extended from 0.001 to 1. Ultimately, the final λ value selected during the 10-fold cross-validation process was $\lambda=0.0193$. Figure 3 presents the coefficients and mean standard error (MSE) obtained from 10-fold validation, offering valuable insights into the model's performance and the significance of the features.

Metrics in combined modal. During the training phase, we applied 5-fold cross-validation and Grid Search to enhance the optimization of the model's hyperparameters, selecting the best parameters based on test dataset performance. The model was then trained on the entire training dataset for improved robustness and accuracy, with specific hyperparameters listed in Supplement Table 3.

In our fusion model comparison, the ExtraTrees classifier outperformed others, showing the highest AUC. (Table 1) It attained an AUC of 0.972 in the training dataset and 0.886 in the test dataset, indicating superior accuracy and generalization capabilities compared to SVM, RandomForest, XGBoost, and LightGBM. (Figure 4) The high AUC values, particularly in the test dataset, underscore the clinical potential of the ExtraTrees-based fusion model, thereby supporting personalized clinical decision-making. Additionally, the robust performance across both training and test datasets suggests that the model is reliable and applicable in real-world clinical settings, where generalizability is critical. The model effectively discriminates high-risk residual disease patients, enabling timely treatment intensification and improving prognosis. Specific metrics for each modality are shown in the Supplement Figure 2.

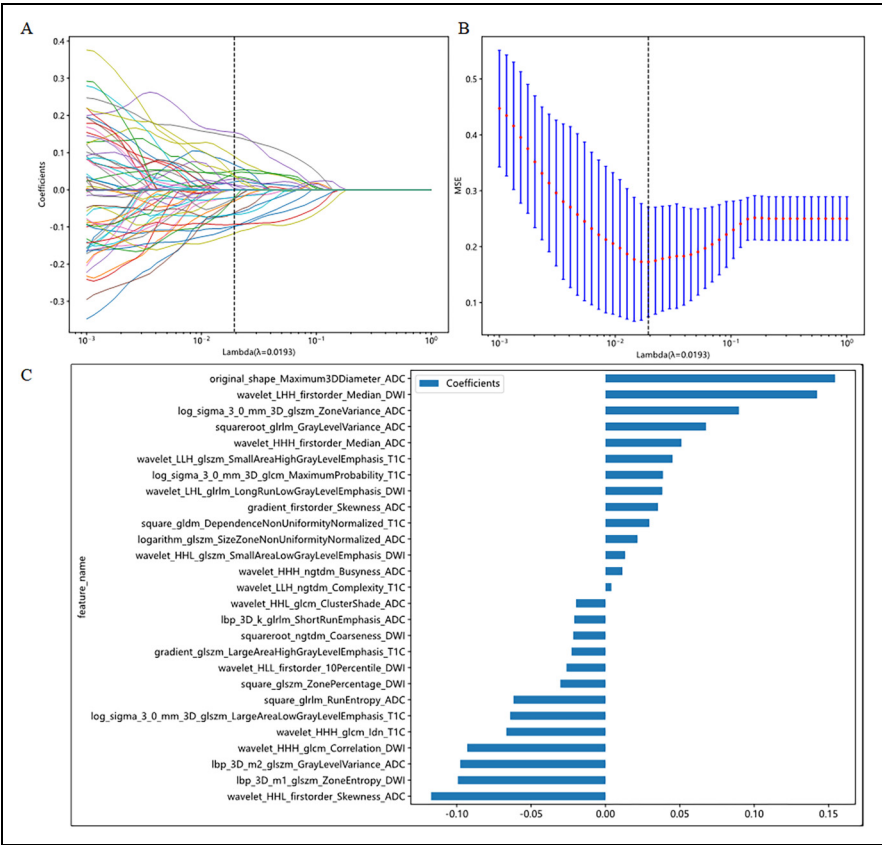


Figure 3. Radiomic feature selection based on LASSO algorithm. Ten-fold cross-validated coefficients (A) and MSE (B) and the histogram of the Rad score based on the selected features (C).

Comparison of radiomics different sequences. In the comparison of different imaging sequences, ADC was proven to be the most effective, exhibiting the highest AUC values. In the training dataset, ADC reached an AUC of 0.951, significantly outperforming T1 (AUC: 0.952), T1C (AUC: 0.767), T2 (AUC: 0.863), and DWI (AUC: 0.868). Notably, in the test dataset, T1's performance drastically declined (AUC: 0.505), while ADC maintained the highest AUC of 0.819 among individual sequences. (Table 2)

Based on these results, T1 and T2 features were excluded in the fusion model. The fusion model, integrating features from the selected sequences, demonstrated improved performance, achieving an AUC of 0.972 in the training dataset and 0.886 in the test dataset, demonstrating its superiority over mono-sequence approaches and validating the decision to omit T1 and T2 features. (Figure 5)

The Delong's test demonstrated that the fusion model achieved statistically significant performance improvement over the T1 model in the test cohorts ($p < 0.05$). These results confirm the clinical value of multimodal integration in our combined model, underscoring

Table 1. Comparative performance metrics of machine learning models for fusion model.

	Train			Test						
	SVM	Random Forest	Extra Trees	XGBoost	Light GBM	SVM	Random Forest	Extra Trees	XGBoost	Light GBM
Accuracy	0.957	0.929	0.929	0.857	0.857	0.774	0.871	0.839	0.677	0.742
Sensitivity	0.929	0.929	0.893	0.679	0.857	0.700	0.700	0.700	0.500	0.800
Specificity	0.976	0.929	0.952	0.976	0.857	0.810	0.952	0.905	0.762	0.714
PPV	0.963	0.897	0.926	0.95	0.8	0.636	0.875	0.778	0.500	0.571
NPV	0.953	0.951	0.93	0.82	0.9	0.850	0.870	0.864	0.762	0.882
AUC	0.995	0.986	0.972	0.879	0.927	0.805	0.867	0.886	0.667	0.833
95% CI	0.986-1.000	0.967-1.000	0.940-1.000	0.791-0.966	0.868-0.986	0.615-0.995	0.717-1.000	0.760-1.000	0.476-0.858	0.686-0.981

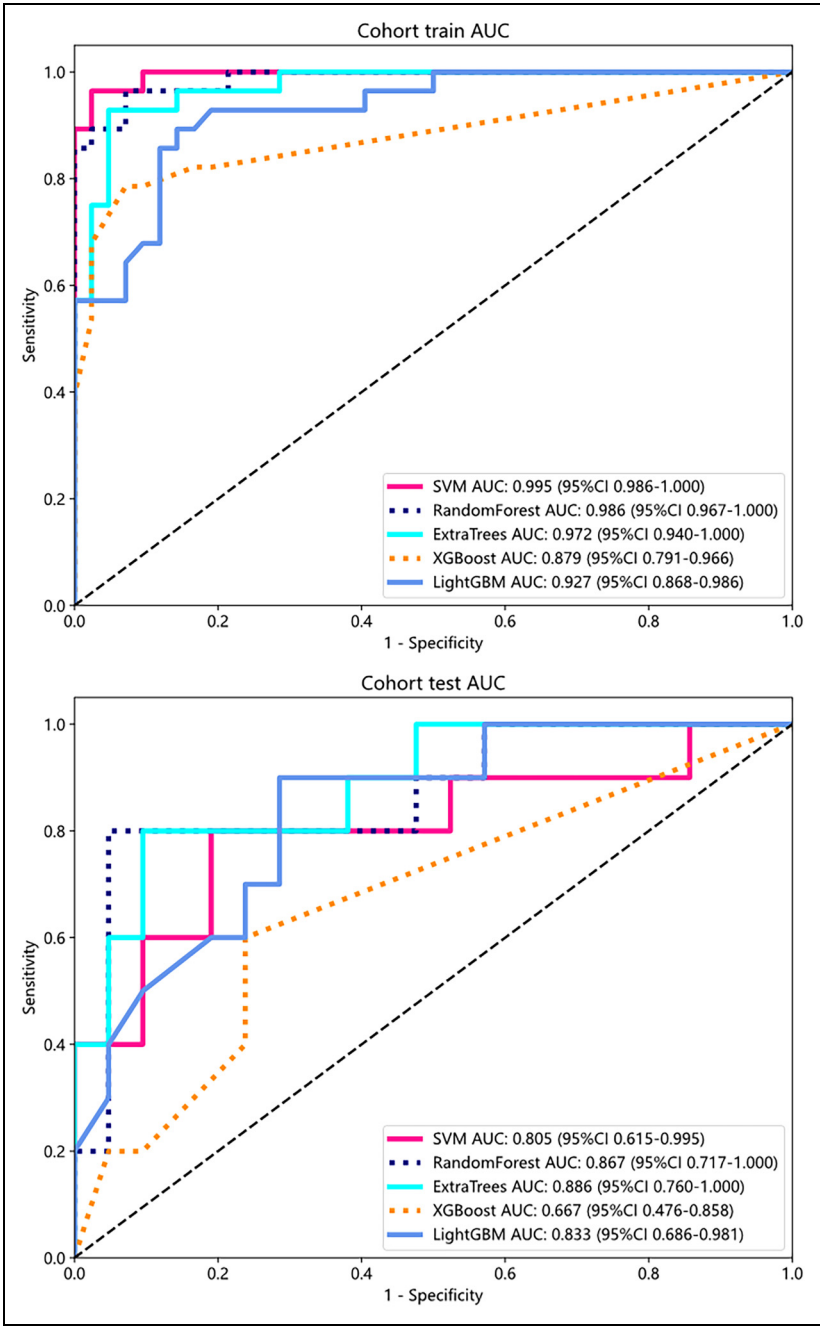


Figure 4. ROC curves for training and test sets of fusion model.

Table 2. Task specific prediction performance of different models.

	Train					Test						
	ADC	DWI	T1	T1C	T2	Fusion	ADC	DWI	T1	T1C	T2	Fusion
Accuracy	0.871	0.757	0.886	0.657	0.800	0.929	0.710	0.710	0.645	0.742	0.742	0.839
Sensitivity	0.786	0.667	0.833	0.476	0.810	0.929	0.571	0.667	0.762	0.762	0.810	0.857
Specificity	1.000	0.893	0.964	0.929	0.786	0.929	1.000	0.800	0.400	0.700	0.600	0.800
PPV	1.000	0.903	0.972	0.909	0.850	0.951	1.000	0.875	0.727	0.842	0.810	0.900
NPV	0.757	0.641	0.794	0.542	0.733	0.897	0.526	0.533	0.444	0.583	0.600	0.727
AUC	0.951	0.868	0.952	0.767	0.863	0.972	0.819	0.757	0.505	0.798	0.738	0.886
95% CI	0.9074- 0.9940	0.784- 0.9512	0.901- 1.0000	0.654- 0.8805	0.769- 0.9568	0.9401- 1.0000	0.672- 0.9661	0.580- 0.9334	0.261- 0.7477	0.633- 0.9618	0.532- 0.9441	0.7597- 1.0000

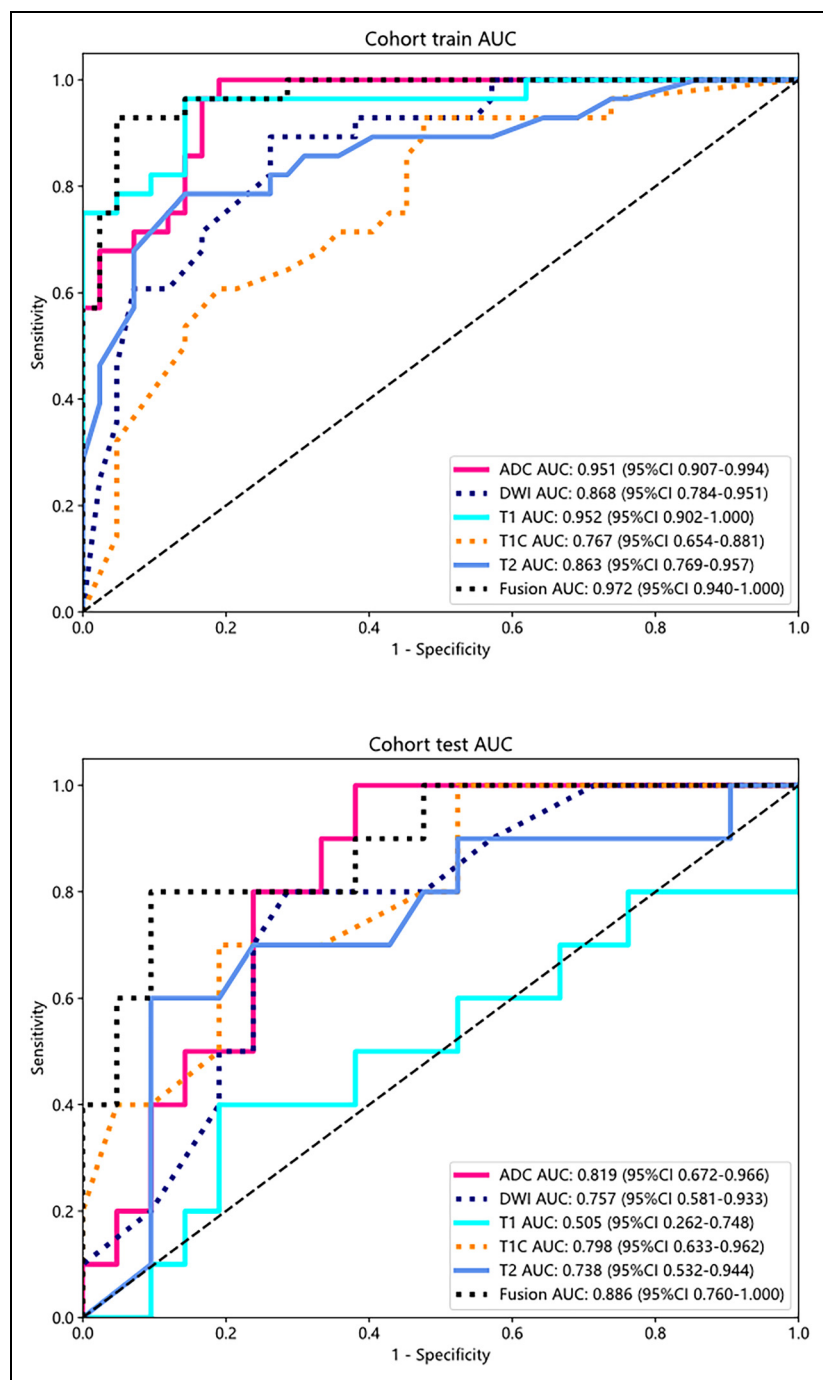


Figure 5. ROC curve of different sequences in train and test cohort. The fusion model, integrating features from the selected sequences(T1C, ADC and DWI), showed enhanced performance with an AUC of 0.972 in the training cohort and 0.886 in the test cohort.

its superior diagnostic performance compared to single-modality approaches. (Supplement Figure 3)

Model comparison

The Combined model, which merges clinical and multi-sequences MR radiomics model, exhibited higher accuracy, sensitivity, specificity, PPV, and NPV in both the training (0.957, 0.939, 0.976, 0.963, 0.953) and test (0.871, 0.7, 0.923, 0.875, 0.87) datasets, outperforming the Clinic and Fusion models. This demonstrates the model's reliability in accurately classifying cases, maintaining robust performance in identifying both positive and negative cases across datasets.

In the training cohort, the clinical, fusion, and combined models demonstrated false positive rates (FPR) of 21.4%, 4.8%, and 2.4%, and false negative rates (FNR) of 17.9%, 10.7%, and 7.1%, respectively. Corresponding values in the test cohort were 9.5%, 9.5%, and 4.8% for FPR, and 40%, 30%, and 30% for FNR. The combined model consistently reduced both FPR and FNR across training and test cohorts, thereby optimizing the trade-off between overtreatment and missed diagnoses - a critical balance for clinical deployment.

Moreover, the Combined model, demonstrated superior performance in terms of AUC. In the training dataset, while the Clinic model achieved an AUC of 0.892 and the Fusion model reached 0.972, the Combined model excelled with an AUC of 0.990, indicating a substantial improvement. Similarly, in the test dataset, the Combined model outperformed both the Clinic (AUC: 0.852) and Fusion models (AUC: 0.886), achieving a remarkable AUC of 0.900. (Supplement Table 4)

These metrics collectively underscore the robustness of the Combined model, highlighting its superior performance not only in terms of AUC but also in other critical clinical performance indicators. This comprehensive evaluation further validates the decision to integrate clinical and multi-sequence MR radiomics features, as it enhances the model's overall diagnostic accuracy and reliability. (Figure 6)

Calibration. The Hosmer-Lemeshow (HL) test evaluates the fit between predicted probabilities and actual outcomes, where lower values indicate better calibration, contrary to initial assumptions. In our research, the Combined model demonstrated superior calibration, as indicated by the HL test values of 0.883 and 0.605 for the training and test datasets, respectively. The observed low values indicate a significant alignment between the predicted results and the actual outcomes, thereby underscoring the model's reliability. (Supplement Figure 4)

Model explanation and visualization

SHAP provided a quantitative explanation for the fusion model. The beeswarm plot (Figure 7A) reveals feature contributions by displaying the distribution of SHAP values across samples, where features are vertically ranked by global importance. Red and blue colors indicate positive or negative effects of features on prediction outcomes, with point

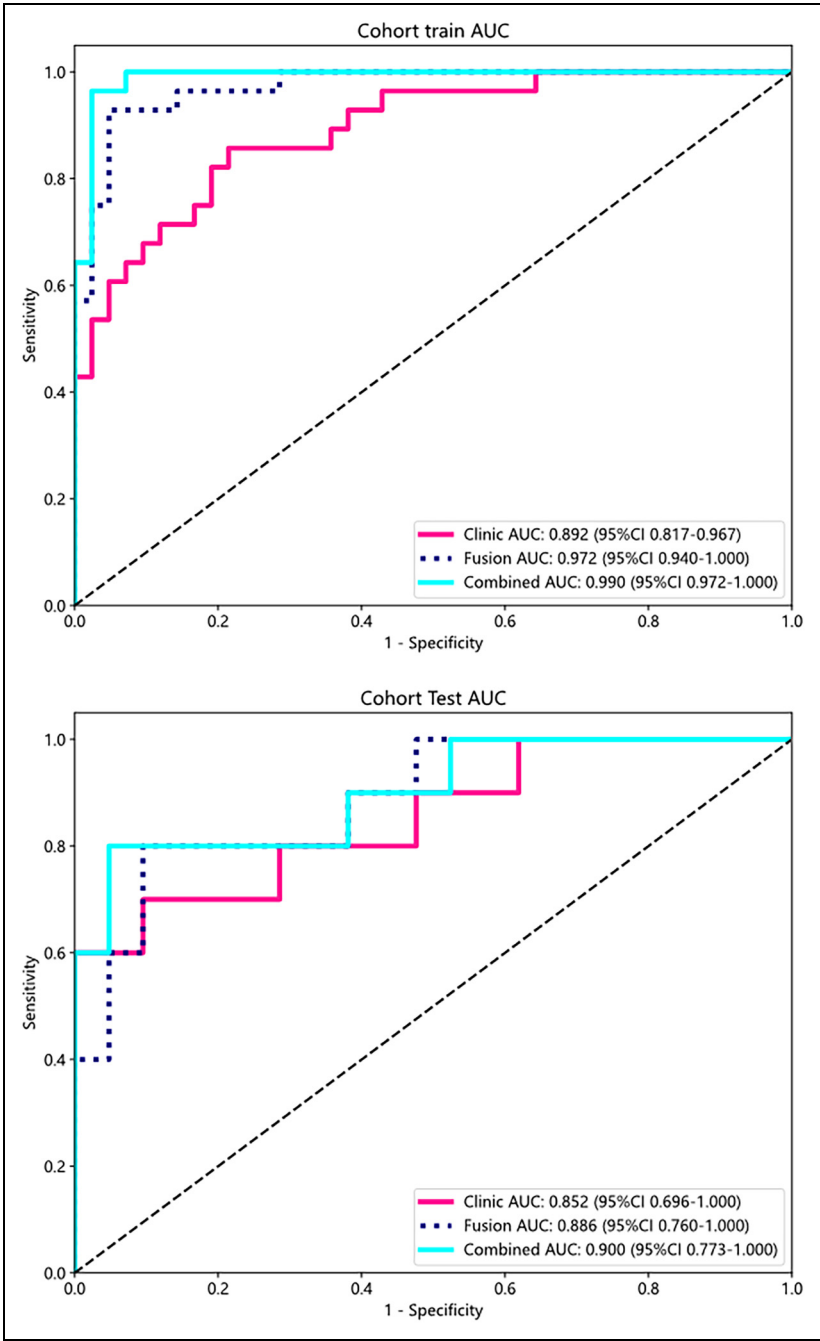


Figure 6. Different models' ROC on train and test cohort.

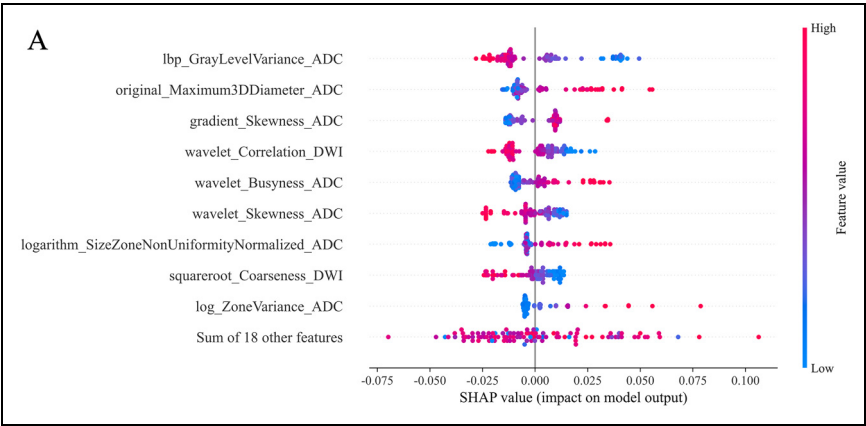


Figure 7A. SHAP beeswarm plot of fusion model. The plot illustrated feature attributions to the model's predictive performance.

density clusters showing common effect ranges. Notably, lbp_GrayLevelVariance_ADC emerges as the most influential predictor for CR/nonCR classification and the coloring showed that the model's output increased with decreasing value of this feature. The feature importance analysis (Figure 7B) quantifies these relationships through mean absolute SHAP values, identifying lbp_GrayLevelVariance_ADC, original_maximum3Ddiameter_ADC, and gradient_Skewness_ADC as key predictive features. Together, these visualizations provide clinicians with transparent insights into the model's decision-making process by quantifying individual feature contributions and revealing effect directionality, thereby aligning with current best practices for interpretable machine learning in medical applications.

Figure 8 presents a heatmap illustrating the distribution of 27 selected radiomic features across CR and non-CR patient groups. The color intensity in the heatmap corresponds to feature value magnitudes, with darker shades typically indicating higher feature values. This visualization enables intuitive comparison of radiomic feature differences between patient groups, facilitating identification of features potentially associated with treatment response.

Clinical use

Supplement Figure 5 presents the DCA curves for both the training and test datasets. The analysis indicates that our combined model offers substantial advantages regarding predicted probabilities. Furthermore, the analysis demonstrates a higher potential for net benefit compared to other models, as evidenced in the DCA results.

Discussion

There is a lack of validation for a radiomic analysis of multi-sequences MR using machine learning to predict early response in LA-NPC after CCRT. In this research, a

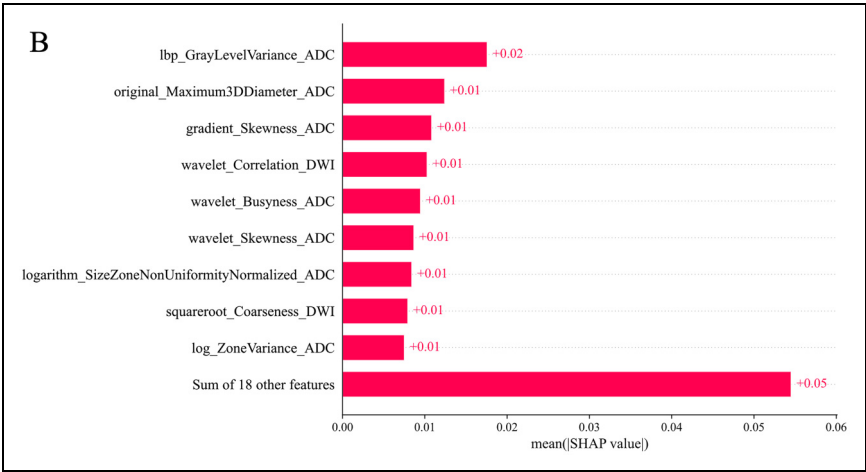


Figure 7B. SHAP feature importance bar chart. The plot illustrates the relative importance ranking of features.

combined model merging clinical and multi-sequences MR radiomics model was constructed to make it come true. This approach demonstrated good performance and proved beneficial for clinicians in assessing the early response in time.

Indeed, numerous studies have examined the efficacy of conventional radiomics models in forecasting early responses or disease progression in patients diagnosed with LA-NPC.^{11,24,36,37} However, it is crucial to note that traditional radiomics models typically rely on a single or a limited number of sequences for feature extraction and analysis. In contrast, our approach stands out due to its comprehensive utilization of multiple sequences.

With the deepening of cancer radiomics research, the current cancer radiomics based on mono-sequence has exposed its shortcomings. Mono-sequence imaging can only partially reflect the tumor information, and inevitably miss some tumor information. Multi-sequences MR radiomic involves the extraction of comprehensive information from each image sequence, followed by the integration of these data for the purpose of model development.³⁸ Consequently, the risk of information loss is mitigated. The segmentation of subregions through the integration of multi-sequence MR facilitates the retention of tumor heterogeneity in the radiomic features derived from these subregions, thereby enhancing the overall efficacy of radiomic analysis.³⁹ Thus, multi-sequences MR models have demonstrated superior performance compared to mono-sequence models.⁴⁰

It is worth noting that our multi-sequences MR model construction process involved a selection-based approach. Through evaluating the performance of different sequences, we observed that T1 (AUC: 0.505) and T2 (AUC: 0.738) sequences were less effective in predicting the outcome. Conversely, ADC (AUC: 0.819) emerged as the most effective sequences in the test dataset. By integrating features from the selected sequences (T1C, DWI, and ADC), the fusion model exhibited enhanced performance with an AUC of

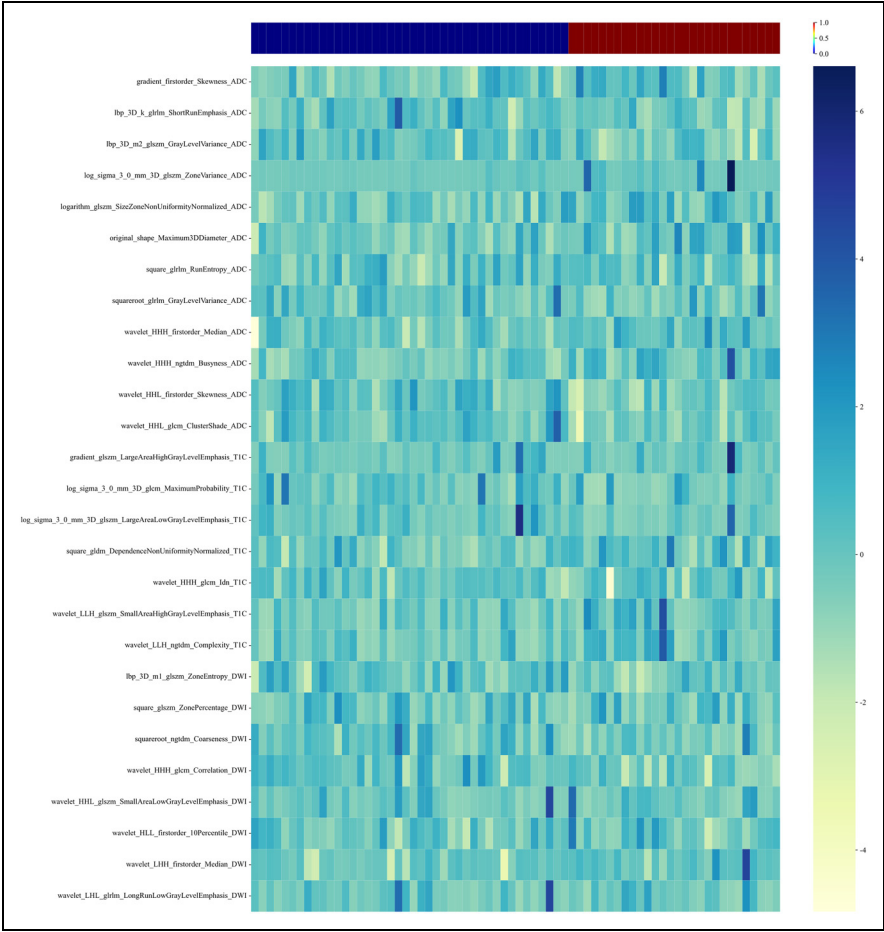


Figure 8. Heatmap visualization of 27 selected radiomic features in the fusion model.

0.886 in the test dataset. This result demonstrates the superiority of fusion model over mono-sequence approaches and validates the decision to exclude T1 and T2 features from the model construction.

Notably, the performance of T1 in the test dataset demonstrated a rapid decline, which is consistent with findings from previous studies. In a study conducted by Shi et al., similar outcome was observed where T1 and T2 sequences showed inferior predictive performance compared to T1C sequences in the test dataset. Among the sequences, T1 exhibited the poorest predictive performance.²⁴ This observation aligns with the clinical perspective, as the T1C sequence provides better information regarding tumor contrast and blood supply due to the enhancement provided by the contrast agent. This enables physicians to more accurately detect and localize tumors and discern differences between

tumor and surrounding normal tissue. Due to the absence of contrast agents, T1 sequences alone may not provide sufficient information for determining the location and extent of nasopharyngeal tumors. Therefore, in clinical practice, T1 sequences are often combined with T1C imaging to accurately assess the position and involvement range of nasopharyngeal lesions. Additionally, T2 images exhibit high sensitivity in detecting effusion or edema. Consequently, T2 sequences facilitate improved visualization of lymph nodes, thereby allowing for a more accurate assessment of lymph node metastasis and its specific localization. Recent research has validated the potential clinical application of T2 sequences in differentiating between benign and metastatic lymph nodes within the retropharyngeal region in cases of NPC.^{41,42} As our study primarily revolved around feature extraction and the establishment of radiomics models focusing on the ROI of nasopharyngeal lesions, the contribution of T2 and T1 sequences to this particular radiomics analysis was relatively minor compared to other sequences.

Hu et al. conducted a study where they found the use of a machine learning model that utilized features extracted from DWI maps as a potential prognosis detection tool for NPC patients.⁴³ In malignant tumors, water molecule diffusion is often restricted by high cell density, resulting in higher signal intensity on DWI and lower values on ADC. DWI technology provides both quantitative and qualitative information, thereby improving the specificity of disease diagnosis.⁴³ This study provides evidence that incorporating DWI and ADC sequences in radiomics has the potential to complement or potentially replace the conventional sequences currently used (T2, T1, and T1C). By doing so, it can offer additional high-specificity information and support for decision-making in the field of radiology.

Radiomics clinical visualization methods are diverse, ranging from nomograms to complex interactive tools.⁴⁴⁻⁴⁷ These methods not only enhance the interpretability of radiomics data but also provide intuitive support for clinical decision-making. A nomogram is an interpretability tool used to visualize predictive models, graphically displaying the contribution of each feature to the prediction outcome. In a study by Shen et al., a nomogram was constructed to visualize a radiological model for identifying early hematoma expansion in spontaneous intracerebral hemorrhage.⁴⁸ Heatmaps are commonly used to display the distribution or patterns of high-dimensional data. A heatmap was utilized to study the association between radiomic features and clinical parameters in patients with early-stage cervical cancer.⁴⁷ SHAP is a feature importance interpretation method that demonstrates the contribution of each feature to the model's prediction. The SHAP tool identified two features of the venous phase as the most significant, effectively differentiating between G1 and G2/3 of pancreatic neuroendocrine tumors (pNETs), demonstrating favorable interpretability.⁴⁶ Other methods include generating attention maps from deep learning models and developing interactive tools, such as web-based visualization platforms,⁴⁵ to enable clinicians to explore radiomics data. With the continuous advancement of technology, radiomics visualization is poised to play an increasingly significant role in precision medicine.

The translation of ML-based medical tools into clinical practice requires careful consideration of regulatory requirements, such as FDA approval or CE marking. Regulatory bodies emphasize the need for robust validation across diverse datasets, transparency in model interpretability, and adherence to data privacy standards. Furthermore, demonstrating clinical utility through prospective trials and ensuring generalizability across different

populations are critical steps in the approval process. Addressing these regulatory hurdles is essential to ensure the safety, efficacy, and reliability of AI-driven tools in real-world healthcare settings.

Our predictive model demonstrates encouraging accuracy ($AUC = 0.90$), suggesting the ability to effectively identify patients at high risk of residual disease. First of all, clinicians can leverage the risk stratification provided by the model to customize treatment plans. For these high-risk patients, treatment intensification should be considered through combination with targeted therapies, immunotherapy, or enhanced maintenance therapy following radiotherapy. Conversely, for low-risk residual disease patients - particularly those with T4 or N3 classification who have undergone induction chemotherapy and concurrent chemoradiotherapy - the use of metronomic chemotherapy may potentially be omitted. Further non-inferiority studies based on these findings will be required to validate the safety of reducing unnecessary interventions in selected patient populations. Secondly, the model's outputs facilitate meaningful discussions between clinicians and patients regarding treatment options, using visual aids to clarify risks and recommendations. Thirdly, accurate predictions allow healthcare facilities to optimize resources, such as scheduling radiotherapy and follow-up care for high-risk patients.

Limitations

However, it is crucial to acknowledge and address the limitations of this study. Firstly, it is important to acknowledge that our study was conducted with a relatively modest sample size and was limited to a single institution, which may affect the generalizability of our findings. The absence of external validation further highlights the need for caution when interpreting the results. In future work, we aim to collaborate with multi-institutional cohorts to validate our findings across diverse populations and settings. Secondly, while standard MRI sequences were used, DCE-MRI and PET-CT offer additional insights in reflecting changes in tissue microcirculation and metabolic activity, providing more accurate diagnostic and treatment recommendations. Although the integration of multi-sequences MR is valuable for advancing radiomics, there are indeed challenges in the implementation process. Variances in spatial structures can be observed among images obtained from different sequences due to factors like temporal differences, patient positioning, device field of view, and the number of slices captured.³⁸ We are optimistic that future advancements in image registration and alignment tools will facilitate smoother integration of multi-sequence data. We remain committed to exploring these advancements in our ongoing efforts to improve the accuracy and reliability of radiological analyses. Thirdly, the biological interpretation of radiomic features remains challenging. In future studies, we will conduct radiogenomic analyses to correlate radiomic features with molecular characteristics, thereby providing biological explanations for currently ambiguous radiomic signatures.

Conclusion

In conclusion, we have successfully developed and validated a combined model that integrated clinical features with multi-sequence MR radiomics model. This combined model

demonstrates exceptional predictive performance, achieves an impressive AUC of 0.900, which significantly outperforms models based on clinical or radiomics features alone. These results underscore the effectiveness of radiomic analysis of multi-sequences MR, supported by machine learning, in enhancing the early response prediction for LA-NPC patients following standard treatment. The superior performance of the combined model highlights its potential as a robust tool for personalized treatment planning and clinical decision-making in LA-NPC management.

Acknowledgements

Thanks to our colleagues and staff at the Department of Radiation Oncology.

ORCID iDs

Kexin Shi  <https://orcid.org/0009-0002-0009-461X>

Shu Zhou  <https://orcid.org/0000-0002-6864-5177>

Ethical considerations

The study received full ethical approval under the identifier 2023-SRFA-552 from The First Affiliated Hospital with Nanjing Medical University Institutional Review Board. The study process was in accordance with the Declaration of Helsinki Ethics statement.

Authors contributions

LQ, YF and YZ contributed equally to this article and share first authorship.

Data management, Lei Qiu, Yinjiao Fei, Mengxing Wu and Yurong Li; Formal analysis, Kexin Shi, Mengxing Wu and Weilin Xu; Funding, Shu Zhou; Survey, Yuchen Zhu, Kexin Shi, Gefei Jiang and Weilin Xu; Methodology, Lei Qiu; Project Management, Shu Zhou; Resources, Yuandong Cao; Software, Yuchen Zhu, Jinling Yuan, Gefei Jiang and Xingjian Sun; Supervision, Yuandong Cao; Validation, Xingjian Sun; Visualization, Yinjiao Fei, Jinling Yuan and Jinyan Luo; Writing - Original draft, Lei Qiu, Yinjiao Fei and Yuchen Zhu; Writing-review and editing, Weilin Xu, Yuandong Cao and Shu Zhou.

All authors contributed to the article and approved the submitted version.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Thanks to the support of Jiangsu Province Entrepreneurship and Innovation Doctoral Talent Program (2019303073386ER19) and Jiangsu Province People's Hospital Clinical Capability Enhancement Project (JSPH-MC-2021-17).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Supplemental material

Supplemental material for this article is available online.

References

1. Chen Y-P, Chan ATC, Le Q-T, et al. Nasopharyngeal carcinoma. *Lancet* 2019; 394: 64–80.
2. Mnejja W, Nouri O, Fourati N, et al. État des lieux et perspectives thérapeutiques pour les carcinomes nasopharyngés localement évolués. *Cancer/Radiothérapie* 2022; 26: 730–735.
3. Juarez-Vignon Whaley JJ, Afkhami M, Sampath S, et al. Early stage and locally advanced nasopharyngeal carcinoma treatment from present to future: where are we and where are we going? *Curr Treat Options Oncol* 2023; 24: 845–866.
4. Al-Sarraf M, LeBlanc M, Giri PG, et al. Chemoradiotherapy versus radiotherapy in patients with advanced nasopharyngeal cancer: phase III randomized intergroup study 0099. *J Clin Oncol* 1998; 16: 1310–1317.
5. al-Sarraf M, Pajak TF, Cooper JS, et al. Chemo-radiotherapy in patients with locally advanced nasopharyngeal carcinoma: a radiation therapy oncology group study. *J Clin Oncol* 1990; 8: 1342–1351. 1990/08/01.
6. Wee J, Tan EH, Tai BC, et al. Randomized trial of radiotherapy versus concurrent chemoradiotherapy followed by adjuvant chemotherapy in patients with American joint committee on cancer/international union against cancer stage III and IV nasopharyngeal cancer of the endemic variety. *J Clin Oncol* 2005; 23: 6730–6738.
7. Fountzilias G, Ciuleanu E, Bobos M, et al. Induction chemotherapy followed by concomitant radiotherapy and weekly cisplatin versus the same concomitant chemoradiotherapy in patients with nasopharyngeal carcinoma: a randomized phase II study conducted by the hellenic cooperative oncology group (HeCOG) with biomarker evaluation. *Ann Oncol* 2012; 23: 427–435. 2011/04/29.
8. Lee SJ, Kim YS, Kay CS, et al. The effect of adjuvant chemotherapy and early tumor regression on the outcome of nasopharyngeal cancer patients treated with concurrent chemoradiotherapy. *Oral Oncol* 2021; 113: 105130. 20210125.
9. He Y, Zhou Q, Shen L, et al. A retrospective study of the prognostic value of MRI-derived residual tumors at the end of intensity-modulated radiotherapy in 358 patients with locally-advanced nasopharyngeal carcinoma. *Radiat Oncol* 2015; 10: 89. 2015/04/17.
10. Huang YY, Cao X, Cai ZC, et al. Short-term efficacy and long-term survival of nasopharyngeal carcinoma patients with radiographically visible residual disease following observation or additional intervention: a real-world study in China. *Laryngoscope Investig Otolaryngol* 2022; 7: 1881–1892. 2022/12/23.
11. Wu M, Xu W, Fei Y, et al. MRI-based clinical radiomics nomogram may predict the early response after concurrent chemoradiotherapy in locally advanced nasopharyngeal carcinoma. *Front Oncol* 2023; 13: 1192953. 2023/05/31.

12. Mayerhoefer ME, Materka A, Langs G, et al. Introduction to radiomics. *J Nucl Med* 2020; 61: 488–495. 2020/02/16.
13. Braman NM, Etesami M, Prasanna P, et al. Intratumoral and peritumoral radiomics for the pre-treatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res* 2017; 19: 57. 20170518.
14. Jiang Y, Zhou K, Sun Z, et al. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. *Cell Rep Med* 2023; 4: 101146. 20230808.
15. Le VH, Kha QH, Hung TNK, et al. Risk score generated from CT-based radiomics signatures for overall survival prediction in non-small cell lung cancer. *Cancers (Basel)* 2021; 13: 20210719.
16. Lee SY, Jeon SI, Jung S, et al. Targeted multimodal imaging modalities. *Adv Drug Delivery Rev* 2014; 76: 60–78.
17. Shin J, Seo N, Baek S-E, et al. MRI Radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy. *Radiology* 2022; 303: 351–358.
18. Zhang X-Y, Wang L, Zhu H-T, et al. Predicting rectal cancer response to neoadjuvant chemoradiotherapy using deep learning of diffusion kurtosis MRI. *Radiology* 2020; 296: 56–64.
19. Zhang L, Li X, Yang L, et al. Multi-Sequence and multi-regional MRI-based radiomics nomogram for the preoperative assessment of muscle invasion in bladder cancer. *J Magn Reson Imaging* 2023; 58: 258–269. 20221027.
20. Xue K, Liu L, Liu Y, et al. Radiomics model based on multi-sequence MR images for predicting preoperative immunoscore in rectal cancer. *Radiol Med* 2022; 127: 702–713. 20220713.
21. Panebianco V, De Berardinis E, Barchetti G, et al. An evaluation of morphological and functional multi-parametric MRI sequences in classifying non-muscle and muscle invasive bladder cancer. *Eur Radiol* 2017; 27: 3759–3766. 20170208.
22. Zhou Y, Gu HL, Zhang XL, et al. Multiparametric magnetic resonance imaging-derived radiomics for the prediction of disease-free survival in early-stage squamous cervical cancer. *Eur Radiol* 2022; 32: 2540–2551. 20211012.
23. Zhuo EH, Zhang WJ, Li HJ, et al. Radiomics on multi-modalities MR sequences can subtype patients with non-metastatic nasopharyngeal carcinoma (NPC) into distinct survival subgroups. *Eur Radiol* 2019; 29: 5590–5599. 2019/03/16.
24. Zhao L, Gong J, Xi Y, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. *Eur Radiol* 2020; 30: 537–546. 2019/08/03.
25. 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: 4259–4269. 2017/03/11.
26. Xu H, Lv W, Zhang H, et al. Multimodality radiomics analysis based on [(18)F]FDG PET/CT imaging and multisequence MRI: application to nasopharyngeal carcinoma prognosis. *Eur Radiol* 2023; 33: 6677–6688. 2023/04/16.
27. Sprinkart AM, Marx C, Träber F, et al. Evaluation of exponential ADC (eADC) and computed DWI (cDWI) for the detection of prostate cancer. *Rofo* 2018; 190: 758–766. 2018/07/26.
28. Filograna L, Magarelli N, Cellini F, et al. Diffusion weighted imaging (DWI) and apparent diffusion coefficient (ADC) values for detection of malignant vertebral bone marrow lesions. *Eur Rev Med Pharmacol Sci* 2018; 22: 590–597. 2018/02/21.

29. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Br Med J* 2007; 335: 806–808.
30. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009; 45: 228–247. 2008/12/23.
31. Viera AJ and Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005; 37: 360–363.
32. Lv JW, Zhou GQ, Li JX, et al. Magnetic resonance imaging-detected tumor residue after intensity-modulated radiation therapy and its association with post-radiation plasma Epstein-Barr virus deoxyribonucleic acid in nasopharyngeal carcinoma. *J Cancer* 2017; 8: 861–869. 2017/04/07.
33. Spadarella G, Stanzione A, Akinci D, Antonoli T, et al. Systematic review of the radiomics quality score applications: an EuSoMII radiomics auditing group initiative. *Eur Radiol* 2023; 33: 1884–1894. 2022/10/26.
34. Wang W, Liang H, Zhang Z, et al. Comparing three-dimensional and two-dimensional deep-learning, radiomics, and fusion models for predicting occult lymph node metastasis in laryngeal squamous cell carcinoma based on CT imaging: a multicentre, retrospective, diagnostic study. *EClinicalMedicine* 2024; 67: 102385. 2024/01/23.
35. Chen K, Huang X, Chen J, et al. Efficacy and safety of combined PD-1 inhibitor with induction chemotherapy followed by IMRT plus nimotuzumab in locally advanced nasopharyngeal carcinoma: a retrospective analysis. *Onco Targets Ther* 2025; 18: 283–296. 20250226.
36. Kang L, Niu Y, Huang R, et al. Predictive value of a combined model based on Pre-treatment and mid-treatment MRI-radiomics for disease progression or death in locally advanced nasopharyngeal carcinoma. *Front Oncol* 2021; 11: 774455. 2021/12/25.
37. Yongfeng P, Chuner J, Lei W, et al. The usefulness of pretreatment MR-based radiomics on early response of neoadjuvant chemotherapy in patients with locally advanced nasopharyngeal carcinoma. *Oncol Res* 2021; 28: 605–613. 20201026.
38. Zhang L, Wang Y, Peng Z, et al. The progress of multimodal imaging combination and sub-region based radiomics research of cancers. *Int J Biol Sci* 2022; 18: 3458–3469. 2022/06/01.
39. Jardim-Perassi BV, Huang S, Dominguez-Viqueira W, et al. Multiparametric MRI and coregistered histology identify tumor habitats in breast cancer mouse models. *Cancer Res* 2019; 79: 3952–3964. 2019/06/13.
40. Wang Q, Li Q, Mi R, et al. Radiomics nomogram building from multiparametric MRI to predict grade in patients with glioma: a cohort study. *J Magn Reson Imaging* 2019; 49: 825–833. 2018/09/28.
41. Wang P, Hu S, Wang X, et al. Synthetic MRI in differentiating benign from metastatic retropharyngeal lymph node: combination with diffusion-weighted imaging. *Eur Radiol* 2023; 33: 152–161.
42. Qu J, Pan B, Su T, et al. T1 and T2 mapping for identifying malignant lymph nodes in head and neck squamous cell carcinoma. *Cancer Imaging* 2023; 23: 25.
43. 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: 3201.

44. Le NQK. Hematoma expansion prediction: still navigating the intersection of deep learning and radiomics. *Eur Radiol* 2024; 34: 2905–2907. 20240122.
45. Ma Y, Yue P, Zhang J, et al. Early prediction of acute gallstone pancreatitis severity: a novel machine learning model based on CT features and open access online prediction platform. *Ann Med* 2024; 56: 2357354. 20240530.
46. Ye JY, Fang P, Peng ZP, et al. A radiomics-based interpretable model to predict the pathological grade of pancreatic neuroendocrine tumors. *Eur Radiol* 2024; 34: 1994–2005. 20230902.
47. Zheng RR, Cai MT, Lan L, et al. An MRI-based radiomics signature and clinical characteristics for survival prediction in early-stage cervical cancer. *Br J Radiol* 2022; 95: 20210838. 20211129.
48. Feng C, Ding Z, Lao Q, et al. Prediction of early hematoma expansion of spontaneous intracerebral hemorrhage based on deep learning radiomics features of noncontrast computed tomography. *Eur Radiol* 2024; 34: 2908–2920. 20231108.