

Longitudinal intratumoral and peritumoral MRI radiomics for predicting pathologic complete response to neoadjuvant chemoimmunotherapy in locally advanced head and neck squamous cell carcinoma: a retrospective two-center study

Shitong Lou

Beijing Tongren Hospital CMU

Ling Duan

Beijing Tongren Hospital CMU

Xinyu Li

Beijing Tongren Hospital CMU

Changyu Zhu

Beijing Tongren Hospital CMU

Yijun Shi

Beijing Tongren Hospital CMU

Weitao Zhang

Beijing Tongren Hospital CMU

Shuyang Zhang

Beijing Tongren Hospital CMU

Heshu Liu

Beijing Tongren Hospital CMU

Wei Li

weili18989@ccmu.edu.cn

Beijing Tongren Hospital CMU

Research Article

Keywords: head and neck squamous cell carcinoma, neoadjuvant chemoimmunotherapy, magnetic resonance imaging, peritumoral region, radiomics, pathologic complete response, longitudinal imaging

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11118822/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Pathologic complete response (pCR) after neoadjuvant chemoimmunotherapy (NACI) is an important marker of treatment sensitivity in locally advanced head and neck squamous cell carcinoma (HNSCC), but reliable preoperative prediction remains challenging. This study investigated whether longitudinal intratumoral and peritumoral MRI radiomics could improve pCR prediction and whether integration with clinicopathologic factors provided additional value.

Methods

This retrospective two-center study included 310 patients with locally advanced HNSCC who underwent NACI followed by definitive surgery, comprising 151 patients in the training cohort, 66 in the internal validation cohort, and 93 in the independent external test cohort. MRI was performed before treatment and after the second NACI cycle. Radiomic features were extracted from contrast-enhanced fat-suppressed T1-weighted images within the primary tumor and 1-, 3-, and 5-mm peritumoral regions. Features were selected using reproducibility assessment, correlation analysis, and least absolute shrinkage and selection operator regression. Pre-treatment intratumoral, on-treatment intratumoral, longitudinal intratumoral, multiscale longitudinal peritumoral, spatiotemporal radiomics, and clinical-spatiotemporal radiomics models were sequentially developed. Performance was evaluated using the area under the receiver operating characteristic curve (AUC), calibration analysis, decision curve analysis, and incremental discrimination metrics.

Results

Overall, 124 of 310 patients (40.0%) achieved pCR. The clinical-spatiotemporal radiomics model (M6), integrating histologic grade, combined positive score, longitudinal intratumoral features, and 3-mm longitudinal peritumoral features, showed the highest observed discrimination. The AUCs were 0.928 (95% CI, 0.882–0.974), 0.856 (95% CI, 0.749–0.963), and 0.864 (95% CI, 0.780–0.948) in the training, internal validation, and external test cohorts, respectively. In the external test cohort, sensitivity was 0.857 and specificity was 0.845. M6 also showed favorable calibration and net benefit across a broad range of threshold probabilities.

Conclusions

Integrating longitudinal intratumoral and peritumoral MRI radiomics with clinicopathologic factors improved preoperative prediction of pCR after NACI in locally advanced HNSCC. The model showed

consistent performance in independent external testing and may support individualized pCR risk stratification, warranting further prospective multicenter validation.

Introduction

Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignancies worldwide, with more than 900,000 new cases diagnosed annually^{1,2}. Approximately 60% of patients present with locally advanced disease at initial diagnosis, and treatment is often associated with substantial impairment in speech, swallowing, and appearance, while long-term survival remains suboptimal^{2,3}. In recent years, immune checkpoint inhibitors have reshaped the perioperative treatment landscape for locally advanced HNSCC. Neoadjuvant chemoimmunotherapy (NACI) shows favorable pathologic response rates across multiple clinical studies^{4–6}. Pathologic response reflects tumor sensitivity to neoadjuvant therapy and is closely associated with long-term outcomes⁷. Among pathologic response endpoints, pathologic complete response (pCR) is an important indicator of the depth of response to neoadjuvant therapy. However, treatment response varies substantially among patients⁸. For patients who show little response early during treatment, continuing neoadjuvant therapy without timely reassessment may not only increase the risk of immune-related adverse events but also compromise subsequent definitive treatment if the tumor progresses⁹. Therefore, identifying patients with different probabilities of achieving pCR before definitive surgical pathology becomes available is important for preoperative response assessment and risk stratification.

At present, confirmation of pCR still relies on histopathologic evaluation after definitive surgery, and preoperative assessment therefore depends on surrogate predictors¹⁰. Although clinicopathologic factors and immune-related biomarkers, such as the programmed death ligand 1 combined positive score (PD-L1 CPS), provide some predictive information, their ability to predict response at the individual-patient level remains limited, and they cannot adequately characterize dynamic tumor changes during treatment^{11,12}. Conventional imaging assessment based on Response Evaluation Criteria in Solid Tumors (RECIST) primarily relies on changes in tumor size. However, treatment response may also involve necrosis, fibrosis, changes in tissue composition, and alterations in internal heterogeneity, which are not fully captured by size-based assessment alone^{4,13}.

Radiomics enables the extraction of quantitative features from routine MRI that characterize tumor heterogeneity, providing a noninvasive approach for evaluating response to neoadjuvant therapy¹⁴. Compared with single-time-point imaging, longitudinal radiomics can integrate complementary phenotypic information before and during treatment, thereby capturing early tumor responses to therapy. Previous studies suggest that both longitudinal imaging features and intratumoral-peritumoral radiomic features have predictive potential^{11,15,16}, but these two dimensions are generally investigated separately. Longitudinal studies commonly model relative changes in imaging features or delta-radiomic features between pre- and post-treatment examinations¹⁷, whereas MRI-based intratumoral-peritumoral studies mainly focus on static spatial phenotypes derived from a single pre-treatment time point¹¹. Dynamic

peritumoral information during treatment therefore remains insufficiently explored. Consequently, systematic evidence is still lacking on how treatment-induced temporal changes, spatial information at the tumor-host interface¹⁸, and clinicopathologic characteristics can be integrated, and on whether these information dimensions provide complementary and incremental value for predicting pCR after NACI.

Accordingly, this study uses multiparametric MRI acquired before treatment and after the second cycle of NACI to develop and validate a multimodal prediction model integrating longitudinal intratumoral radiomic features, longitudinal peritumoral radiomic features at multiple spatial scales, and clinicopathologic factors. We systematically evaluate the incremental value of temporal, spatial, and clinical information for pCR prediction, with the aim of providing a basis for individualized and dynamic preoperative risk assessment of pCR after NACI in patients with locally advanced HNSCC.

Materials and Methods

Study Design and Participants

This retrospective, two-center study was approved by the institutional review boards of Beijing Tongren Hospital, Capital Medical University, and the First Affiliated Hospital of Zhengzhou University. De-identified clinical and imaging data were used, and the requirement for informed consent was waived. The study design and reporting followed relevant recommendations, including CLEAR, IBSI, CLAIM, and TRIPOD+AI^{19–22}.

The study consecutively included patients with locally advanced head and neck squamous cell carcinoma (HNSCC) who received neoadjuvant chemoimmunotherapy (NACI) between January 2023 and December 2025. Eligible patients had pathologically confirmed HNSCC, received 2–3 cycles of platinum-based chemotherapy combined with an immune checkpoint inhibitor, underwent baseline MRI within 30 days before the first NACI cycle and early follow-up MRI after the second NACI cycle, subsequently underwent definitive surgical resection, and had complete postoperative pathologic data available for assessment of the primary outcome. Patients were excluded if they had received previous treatment for a head and neck malignancy, had severe motion or metal artifacts on MRI that interfered with lesion segmentation or feature extraction, had distant metastases, or had another concurrent malignancy.

Patients from the primary center were randomly allocated to the training and internal validation cohorts at an approximate 7:3 ratio, whereas all patients from the second center were assigned to the independent external test cohort. The final study population comprised 310 patients, including 151 in the training cohort, 66 in the internal validation cohort, and 93 in the external test cohort.

Clinical, pathologic, and treatment-related variables were collected from the electronic medical record system, including age, sex, smoking history, alcohol consumption, Eastern Cooperative Oncology Group (ECOG) performance status, clinical T category, N category, overall clinical stage, histologic grade, p16

status, combined positive score (CPS), Ki-67 index, treatment regimen, and type of immune checkpoint inhibitor.

Pathologic Response Assessment

The primary outcome was postoperative pathologic complete response (pCR), defined as no residual viable tumor in the primary tumor bed or regional lymph nodes (ypT0N0). Surgical specimens were independently evaluated by two experienced head and neck pathologists blinded to imaging-derived features and model predictions, with disagreements resolved by consensus.

MRI Acquisition, Tumor Segmentation, and Peritumoral Region Construction

All patients underwent 3.0T MRI. T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), and contrast-enhanced fat-suppressed T1-weighted imaging (CE-T1WI) were acquired at both baseline and on-treatment time points. CE-T1WI served as the primary sequence for primary tumor segmentation and radiomic analysis, whereas T1WI and T2WI were used to assist anatomic localization. Baseline MRI was performed within 30 days before the first NACI cycle, and on-treatment MRI was acquired after the second NACI cycle. Detailed acquisition parameters were provided in Table S1.

The primary tumor was independently segmented at the pre-treatment and on-treatment time points according to the visible tumor boundaries on the corresponding images. A radiologist with 10 years of experience in head and neck MRI manually delineated the three-dimensional region of interest (ROI) slice by slice on CE-T1WI using ITK-SNAP version 3.8, and all segmentations were reviewed by a senior radiologist with 15 years of experience. Both readers were blinded to postoperative pathologic outcomes, radiomic features, and model predictions. Clearly identifiable vessels, airways or luminal spaces, as well as areas of necrosis, calcification, and hemorrhage, were excluded from the tumor ROI.

Based on the primary tumor ROI at each time point, the tumor boundary was expanded outward by 1, 3, and 5 mm in the original physical space to generate three-dimensional pure peritumoral regions, designated Peri1, Peri3, and Peri5, respectively. Air, bone, luminal spaces, large vessels, and other neoplastic lesions that were not contiguous with the primary tumor were excluded, while adjacent contiguous soft tissue was retained. All peritumoral ROIs underwent quality review by the senior radiologist. Detailed procedures for peritumoral region generation and cropping were provided in the Supplementary Methods.

Image Preprocessing and Radiomic Feature Extraction

Radiomic analysis was primarily performed using the Onekey platform (version 20250330), with conventional radiomic features extracted using PyRadiomics (version 3.0.1; Computational Imaging and

Bioinformatics Laboratory, Brigham and Women's Hospital, Boston, MA, USA). A uniform preprocessing pipeline was applied across all centers, imaging time points, and regions of interest (ROIs).

MRI signal intensities were first normalized using z-score normalization and then resampled to isotropic voxels of $3 \times 3 \times 3 \text{ mm}^3$. First-order statistics and texture features, including gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size-zone matrix (GLSZM), gray-level dependence matrix (GLDM), and neighboring gray-tone difference matrix (NGTDM) features, were extracted from the original, Laplacian of Gaussian-filtered, and wavelet-transformed images. Three-dimensional shape features were calculated from the geometric properties of the ROIs. In addition, Hessian-derived features and surface mean curvature features were extracted to provide complementary information on local second-order spatial structure and boundary geometry. Detailed preprocessing parameters and the mathematical definitions and implementations of the Hessian matrix and surface mean curvature were provided in the Supplementary Methods and Table S2.

Feature Stability and Longitudinal Radiomic Feature Integration

A random sample of 50 patients from the training cohort was used to assess segmentation-related feature stability. Two radiologists independently performed ROI segmentation. The radiologist who performed the initial segmentation repeated the segmentation of the same cases after a 1-month interval without access to the initial results. Both pre-treatment and on-treatment MRI examinations were resegmented using the same procedure, and the corresponding peritumoral ROIs were regenerated from each primary tumor segmentation.

Intraclass correlation coefficients (ICCs) were used to assess intraobserver and interobserver reproducibility. Only radiomic features with both intraobserver ICC > 0.75 and interobserver ICC > 0.75 were retained for subsequent feature selection.

Radiomic features underwent sequential standardization, univariable screening, Pearson correlation-based redundancy reduction, and least absolute shrinkage and selection operator (LASSO) logistic regression using the training data. For highly correlated feature pairs with $|r| > 0.90$, the feature with the higher univariable predictive AUC was retained. If the AUCs were identical, the feature with the smaller P value in univariable logistic regression was retained. The LASSO regularization parameter was determined by 10-fold cross-validation, and features with nonzero coefficients were retained. Detailed selection criteria and parameter settings were provided in the Supplementary Methods.

To characterize longitudinal changes in imaging phenotypes during NACI, radiomic features were first extracted separately at the pre-treatment and on-treatment time points. The retained features from the two time points were then concatenated at the patient level to form a combined dual-time-point candidate feature matrix, followed by further correlation-based redundancy reduction and joint dimensionality reduction using LASSO.

To prevent information leakage, all feature selection, dimensionality reduction, and parameter optimization procedures were performed exclusively within the training data or within the training folds of the corresponding cross-validation procedure. Neither the internal validation cohort nor the external test cohort was involved in model development.

Model Definition and Development

A hierarchical modeling strategy was used to progressively integrate temporal, spatial, and clinical information. The model definitions and corresponding input components were summarized in Table 1.

Table 1
Definitions and input components of the hierarchical prediction models

Model	Definition	Input components
M0	Clinical	Clinicopathologic variables selected by univariable and multivariable logistic regression
M1	Pre-treatment-Intra	Pre-treatment intratumoral MRI radiomic features
M2	On-treatment-Intra	Intratumoral MRI radiomic features obtained after the second NACI cycle
M3	Longitudinal-Intra	Intratumoral radiomic features from the pre-treatment and on-treatment
M4 ₁ /M4 ₃ /M4 ₅	Longitudinal-Peri1/3/5	Dual-time-point radiomic features from the corresponding 1-, 3-, and 5-mm peritumoral regions
M5 ₁ /M5 ₃ /M5 ₅	Candidate Spatiotemporal Radiomics	Predicted probabilities from M3 and the corresponding M4 model (M4 ₁ , M4 ₃ , or M4 ₅) (score-level fusion using logistic regression)
M5	Spatiotemporal Radiomics	Predicted probabilities from M3 and M4 ₃ (score-level fusion using logistic regression)
M6	Clinical-Spatiotemporal Radiomics	Predicted probabilities from M3 and M4 ₃ together with independent clinicopathologic predictors (multimodal integration logistic regression)

The clinical model M0 was constructed using clinicopathologic predictors selected through univariable and multivariable logistic regression. Pre-treatment and on-treatment intratumoral features were used to construct M1 and M2, respectively, whereas intratumoral features from the two time points were fused at the feature level to construct M3. Using the same strategy, features from the 1-, 3-, and 5-mm peritumoral regions at both time points were used to construct M4₁, M4₃, and M4₅, respectively. Subsequently, the predicted probability from M3 was fused separately with that from each corresponding M4 model to construct the candidate spatiotemporal radiomics models M5₁, M5₃, and M5₅, among which M5₃ was selected as the final spatiotemporal radiomics model M5. Finally, the

predicted probabilities from M3 and M4₃ were further integrated with the independent clinicopathologic predictors identified for M0 to construct M6, on the basis of which a nomogram was developed.

Five candidate classifiers, including logistic regression, support vector machine, k-nearest neighbors, random forest, and XGBoost, were compared. Hyperparameters were optimized using grid search with 10-fold cross-validation, with the mean cross-validated AUC used as the primary criterion for classifier selection. The selected classifier was then fixed and consistently applied to subsequent radiomic models. Detailed hyperparameter settings were provided in Table S3.

Model Evaluation

Model discrimination was evaluated using receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC), with corresponding 95% confidence intervals. AUCs of different models within the same cohort were compared using the paired DeLong test. Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were also calculated.

The incremental predictive value of each model was assessed using the continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI). Calibration performance was evaluated using calibration curves, and clinical net benefit was assessed using decision curve analysis (DCA). Classification thresholds were determined exclusively in the training cohort according to the Youden index and were subsequently fixed and applied without modification to the internal validation cohort and the independent external test cohort.

Statistical Analysis

Radiomics feature processing and machine-learning analyses were primarily performed using the Onekey platform (version 20250330), whereas the remaining statistical analyses were conducted using Python (version 3.12.7; Python Software Foundation) and R (version 4.5.2; R Foundation for Statistical Computing, Vienna, Austria). The normality of continuous variables was assessed using the Shapiro-Wilk test. Categorical variables were presented as counts and percentages and were compared between groups using the χ^2 test or Fisher exact test, as appropriate. Continuous variables were summarized as mean $\pm$ standard deviation or median (interquartile range), according to their distributions. Comparisons among the three groups were performed using one-way analysis of variance or the Kruskal-Wallis test, as appropriate. All statistical tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Patient Characteristics and Analysis Cohorts

A total of 310 patients were included in the overall analysis and were divided into a training cohort ($n = 151$), an internal validation cohort ($n = 66$), and an external test cohort ($n = 93$). Baseline characteristics

were compared across all three cohorts. Overall, 124 of 310 patients (40.0%) achieved pCR, including 67 of 151 (44.37%) in the training cohort, 22 of 66 (33.33%) in the internal validation cohort, and 35 of 93 (37.63%) in the external test cohort. The pCR rates did not differ significantly among the three cohorts ($P = 0.267$). All 93 patients in the external test cohort were included in the external performance evaluation. Table 2 summarized and compared the baseline clinical characteristics across the three cohorts. The mean age was 61.03 ± 7.66 years, and 301 patients (97.1%) were male. A history of smoking was present in 236 patients (76.1%), a history of alcohol consumption in 242 (78.1%), and an Eastern Cooperative Oncology Group performance status score of 0 in 239 (77.1%).

Table 2

The clinicopathologic characteristics of patients in the training, validation and testing cohorts

Variable	ALL	Training cohort	Validation cohort	Testing cohort	P value
	(n = 310)	(n = 151)	(n = 66)	(n = 93)	
Age*	61.03 ± 7.66	61.17 ± 7.31	60.09 ± 8.38	61.46 ± 7.72	0.515
Gender					0.964
Female	9(2.90)	4(2.65)	2(3.03)	3(3.23)	
Male	301(97.10)	147(97.35)	64(96.97)	90(96.77)	
Tobacco consumption					0.646
No	74(23.87)	38(25.17)	17(25.76)	19(20.43)	
Yes	236(76.13)	113(74.83)	49(74.24)	74(79.57)	
pack-years*	27.11 ± 23.23	27.90 ± 25.03	23.21 ± 21.28	28.61 ± 21.36	0.299
Alcohol consumption					0.426
No	67(21.61)	34(22.52)	14(21.21)	19(20.43)	
Yes	242(78.06)	117(77.48)	51(77.27)	74(79.57)	
Not recorded	1(0.32)	0(0.00)	1(1.52)	0(0.00)	
ECOG					0.472
0	239(77.10)	112(74.17)	52(78.79)	75(80.65)	
1	71(22.90)	39(25.83)	14(21.21)	18(19.35)	
Family history of cancer					0.006
No	255(82.26)	134(88.74)	47(71.21)	74(79.57)	
Yes	55(17.74)	17(11.26)	19(28.79)	19(20.43)	
Primary tumor site					0.400
Hypopharynx	230 (74.19)	111 (73.51)	48 (72.73)	71 (76.34)	
Oropharynx	18 (5.81)	6 (3.97)	3 (4.55)	9 (9.68)	
Oral cavity	31 (10.00)	16 (10.60)	8 (12.12)	7 (7.53)	
Larynx	31 (10.00)	18 (11.92)	7 (10.61)	6 (6.45)	

Variable	ALL	Training cohort	Validation cohort	Testing cohort	P value
	(n = 310)	(n = 151)	(n = 66)	(n = 93)	
T category					0.294
T1	9(2.90)	4(2.65)	0(0.00)	5(5.38)	
T2	66(21.29)	26(17.22)	16(24.24)	24(25.81)	
T3	64(20.65)	32(21.19)	14(21.21)	18(19.35)	
T4	171(55.16)	89(58.94)	36(54.55)	46(49.46)	
N category					0.147
N0	36(11.61)	22(14.57)	10(15.15)	4(4.30)	
N1	52(16.77)	27(17.88)	9(13.64)	16(17.20)	
N2a	16(5.16)	5(3.31)	3(4.55)	8(8.60)	
N2b	91(29.35)	47(31.13)	22(33.33)	22(23.66)	
N2c	96(30.97)	41(27.15)	17(25.76)	38(40.86)	
N3a	2(0.65)	1(0.66)	0(0.00)	1(1.08)	
N3b	17(5.48)	8(5.30)	5(7.58)	4(4.30)	
Clinical stage					0.798
III	40(12.90)	21(13.91)	9(13.64)	10(10.75)	
IVA	241(77.74)	116(76.82)	49(74.24)	76(81.72)	
IVB	29(9.35)	14(9.27)	8(12.12)	7(7.53)	
Histologic grade					0.172
Poorly differentiated	106(34.19)	50(33.11)	22(33.33)	34(36.56)	
Moderately differentiated	184(59.35)	86(56.95)	41(62.12)	57(61.29)	
Well differentiated	20(6.45)	15(9.93)	3(4.55)	2(2.15)	
EGFR					0.645
+	35(11.29)	16(10.60)	9(13.64)	10(10.75)	
++	27(8.71)	10(6.62)	6(9.09)	11(11.83)	
+++	248(80.00)	125(82.78)	51(77.27)	72(77.42)	
p16					0.382

Variable	ALL (n = 310)	Training cohort (n = 151)	Validation cohort (n = 66)	Testing cohort (n = 93)	P value
Negative	285(91.94)	141(93.38)	58(87.88)	86(92.47)	
Positive	25(8.06)	10(6.62)	8(12.12)	7(7.53)	
CPS*	21.86 ± 24.15	21.26 ± 23.56	25.33 ± 27.92	20.35 ± 22.17	0.404
Ki-67*	57.03 ± 17.76	56.82 ± 17.74	56.97 ± 18.23	57.42 ± 17.64	0.968
Neoadjuvant therapy regimen					0.946
nab-PP+anti-PD-1	269(86.77)	132(87.42)	57(86.36)	80(86.02)	
PF+anti-PD-1	41(13.23)	19(12.58)	9(13.64)	13(13.98)	
Immunotherapy drug					0.472
Pembrolizumab	196(63.23)	97(64.24)	35(53.03)	64(68.82)	
Tislelizumab	70(22.58)	32(21.19)	22(33.33)	16(17.20)	
Toripalimab	1(0.32)	1(0.66)	0(0.00)	0(0.00)	
Nivolumab	12(3.87)	6(3.97)	3(4.55)	3(3.23)	
Sintilizumab	31(10.00)	15(9.93)	6(9.09)	10(10.75)	
Neoadjuvant therapy cycles					0.167
2	131(42.26)	72(47.68)	24(36.36)	35(37.63)	
3	179(57.78)	79(52.32)	42(63.64)	58(62.37)	
Treatment response					0.267
pCR	124(40.00)	67(44.37)	22(33.33)	35(37.63)	
Non-pCR	186(60.00)	84(55.63)	44(66.67)	58(62.37)	
*The variable was summarized as mean and SD (mean ± SD). nab-PP, nanoparticle albumin-bound-paclitaxel/platinum (cisplatin or carboplatin); pCR, pathologic complete response; PD-1, programmed death receptor-1; PF, platinum (cisplatin or carboplatin)/5-Fluorouracil (5-FU).					

All patients had stage III–IV disease, including 241 patients (77.7%) with stage IVA disease and 171 (55.2%) with T4 disease. Moderate differentiation was the most common histologic grade and was observed in 184 patients (59.4%). A total of 269 patients (86.8%) received nab-paclitaxel combined with

platinum-based chemotherapy and an anti-PD-1 antibody. Pembrolizumab was the most frequently used immune checkpoint inhibitor and was administered to 196 patients (63.2%).

Except for family history of cancer ($P = 0.006$), no significant differences were observed in the distributions of the remaining baseline variables among the three cohorts. The patient selection and cohort allocation process was shown in Fig. 1.

Clinical Model and Radiomics Feature Selection

As shown in Table 3, univariable and multivariable logistic regression analyses identified histologic grade and combined positive score (CPS) as independent factors associated with pCR after neoadjuvant chemoimmunotherapy (NACI) (both $P < 0.05$). These two variables were used to construct the clinical model M0. The AUCs of M0 were 0.707 (95% CI, 0.624–0.790), 0.793 (95% CI, 0.660–0.926), and 0.599 (95% CI, 0.480–0.718) in the training, internal validation, and external test cohorts, respectively.

Table 3
Univariable and multivariable logistic regression analyses of clinicopathologic factors in the training cohort.

Variable	Univariate logistic regression		Multivariate logistic regression	
	OR(95%CI)	P value	OR(95%CI)	P value
Age	0.996(0.992,1.001)	0.166		
Gender	0.815(0.620,1.070)	0.217		
Tobacco consumption	0.794(0.581,1.083)	0.222		
pack-years	0.994(0.982,1.007)	0.364		
Alcohol consumption	0.773(0.569,1.050)	0.167		
ECOG	1.167(0.688,1.978)	0.631		
Family history of cancer	1.313 (0.730,2.363)	0.363		
Primary tumor site		0.785		
Hypopharynx	Reference			
Oropharynx	1.527 (0.584,3.993)	0.382		
Oral cavity	1.103 (0.516,2.361)	0.800		
Larynx	0.840 (0.384,1.836)	0.662		
T category		0.321		
T1	Reference			
T2	1.176 (0.290,4.774)	0.820		
T3	0.912 (0.224,3.719)	0.898		
T4	0.693 (0.179,2.678)	0.595		
N category		0.704		
N0	Reference			
N1	0.958 (0.409,2.245)	0.921		
N2a	0.671 (0.201,2.238)	0.516		
N2b	0.801 (0.369,1.741)	0.576		
N2c	0.559 (0.256,1.219)	0.144		
N3a	1.118 (0.065,19.283)	0.939		
N3b	0.610 (0.185,2.005)	0.415		

Variable	Univariate logistic regression		Multivariate logistic regression	
	OR(95%CI)	P value	OR(95%CI)	P value
Clinical stage		0.087		
III	Reference			
IVA	0.479 (0.244,0.941)	0.035		
IVB	0.665 (0.254,1.738)	0.405		
Histologic grade		0.045		0.012
Poorly differentiated	Reference		Reference	
Moderately differentiated	0.688 (0.424,1.116)	0.129	0.570 (0.356,0.912)	0.019
Well differentiated	0.280 (0.088,0.893)	0.032	0.260 (0.088,0.764)	0.014
EGFR expression		0.665		
+	Reference			
++	1.393 (0.506,3.838)	0.522		
+++	0.964 (0.468,1.985)	0.920		
p16	1.500(0.519,4.336)	0.530		
CPS	1.014(1.004,1.023)	0.016	1.025(1.012,1.038)	< 0.001
Ki-67	0.998(0.993,1.002)	0.353		
Neoadjuvant therapy regimen	0.727(0.339,1.562)	0.493		
Neoadjuvant therapy cycles				
2	Reference			
3	0.893 (0.469,1.698)	0.730		
Immunotherapy drug		0.680		
Pembrolizumab	Reference			
Tislelizumab	0.840 (0.485,1.455)	0.533		
Nivolumab	0.900 (0.278,2.916)	0.861		
Sintilimab	0.720 (0.342,1.515)	0.387		
Toripalimab	NE	NE*		

Variable	Univariate logistic regression		Multivariate logistic regression	
	OR(95%CI)	P value	OR(95%CI)	P value
*NE: Not estimable due to insufficient sample size (n = 1) and complete separation in logistic regression analysis. pCR, pathologic complete response; ECOG, Eastern Cooperative Oncology Group; CPS, combined positive score. <i>P</i> < 0.05 was considered statistically significant.				

Based on contrast-enhanced T1-weighted imaging (CE-T1WI), 1,605 radiomic features were initially extracted from the intratumoral and peritumoral regions at each time point. After sequential feature selection based on stability assessment, univariable analysis, correlation-based redundancy reduction, and least absolute shrinkage and selection operator (LASSO) regression, 7 pre-treatment and 18 on-treatment intratumoral features were retained in the training cohort to construct the single-time-point baseline model M1 and on-treatment model M2, respectively. For construction of the longitudinal model M3, the candidate feature pools retained at an earlier stage of the pre-treatment and on-treatment pipelines were concatenated at the feature level and subjected to additional correlation-based redundancy reduction and joint LASSO selection, resulting in 29 longitudinal intratumoral features (Fig. 2). For the 1-, 3-, and 5-mm peritumoral regions, 9/16, 11/23, and 8/11 features were retained at the pre-treatment/on-treatment time points, respectively. The joint dimensionality-reduction procedures after fusion of the two time-point feature sets at each peritumoral scale, together with the final selected features, were shown in Supplementary Figures S1–S3.

Integration of Temporal Information from Pre-treatment and On-treatment Intratumoral MRI

The pre-treatment intratumoral model M1 (Pre-treatment-Intra) achieved AUCs of 0.806 (95% CI, 0.733–0.880), 0.524 (95% CI, 0.367–0.682), and 0.738 (95% CI, 0.637–0.840) in the training, internal validation, and external test cohorts, respectively. The on-treatment intratumoral model M2 (On-treatment-Intra) achieved corresponding AUCs of 0.823 (95% CI, 0.755–0.890), 0.723 (95% CI, 0.590–0.856), and 0.700 (95% CI, 0.580–0.821) (Fig. 3A–C).

After integrating intratumoral information from both time points, M3 (Longitudinal-Intra) achieved AUCs of 0.843 (95% CI, 0.779–0.908), 0.716 (95% CI, 0.581–0.851), and 0.752 (95% CI, 0.641–0.864) in the training, internal validation, and external test cohorts, respectively (Fig. 3A–C). The sensitivity and specificity of M3 were 0.806 and 0.762 in the training cohort, 0.545 and 0.818 in the internal validation cohort, and 0.743 and 0.845 in the external test cohort, respectively. M3 achieved higher AUCs than both M1 and M2 in the training and external test cohorts, whereas its performance was comparable to that of M2 in the internal validation cohort (0.716 vs 0.723). Detailed model performance was provided in Table S4. These results indicated that the longitudinal intratumoral phenotype derived from two time points integrated the pre-treatment baseline state with early imaging changes during treatment and served as the core imaging phenotype for subsequent integration of spatial information.

Selection of Peritumoral Spatial Scale and Development of the Spatiotemporal Radiomics Model

To evaluate spatial information in the peritumoral region, longitudinal peritumoral models M4 were constructed at 1-, 3-, and 5-mm scales using imaging information from both time points and were subsequently integrated with M3 to generate the candidate spatiotemporal models M5₁, M5₃, and M5₅. As shown in Fig. 4, stratified 10-fold cross-validation within the training cohort shown that M5₃ achieved the highest mean AUC (0.872 ± 0.036), compared with M5₁ (0.823 ± 0.046) and M5₅ (0.847 ± 0.041). The 3-mm peritumoral scale was therefore selected for further model development.

As shown in Table S5, the 3-mm longitudinal peritumoral model M4₃ (Longitudinal-Peri3) achieved AUCs of 0.839 (95% CI, 0.772–0.906), 0.735 (95% CI, 0.600–0.870), and 0.777 (95% CI, 0.673–0.881) in the training, internal validation, and external test cohorts, respectively, demonstrating consistent discriminative performance of the longitudinal peritumoral phenotype across all three cohorts.

The performance of the three candidate spatiotemporal models across the three cohorts further supported selection of the 3-mm scale. In the training cohort, M5₁, M5₃, and M5₅ achieved AUCs of 0.843 (95% CI, 0.778–0.908), 0.906 (95% CI, 0.861–0.952), and 0.869 (95% CI, 0.812–0.925), respectively. The corresponding AUCs were 0.715 (95% CI, 0.573–0.856), 0.740 (95% CI, 0.605–0.875), and 0.730 (95% CI, 0.588–0.873) in the internal validation cohort and 0.796 (95% CI, 0.700–0.892), 0.818 (95% CI, 0.730–0.906), and 0.795 (95% CI, 0.702–0.889) in the external test cohort (Fig. 3D–F and Table S6). M5₃ achieved the highest AUC among the three candidate models in all three cohorts and was therefore defined as the final spatiotemporal radiomics model M5.

Furthermore, compared with M3, M5 increased the AUC from 0.843 to 0.906 in the training cohort ($\Delta\text{AUC} = 0.063$), from 0.716 to 0.740 in the internal validation cohort ($\Delta\text{AUC} = 0.024$), and from 0.752 to 0.818 in the external test cohort ($\Delta\text{AUC} = 0.066$) (Fig. 5A–C and Table S7). Paired DeLong tests shown that the improvements were statistically significant in the training cohort ($P = 0.014$) and external test cohort ($P = 0.047$), but not in the internal validation cohort ($P = 0.610$). These findings indicated that incorporating longitudinal information from the 3-mm peritumoral region into the longitudinal intratumoral phenotype further improved model discrimination in the training and external test cohorts.

Integration of Clinical Information and Development of the Final Clinical-Spatiotemporal Model

Histologic grade and combined positive score (CPS) were further integrated with M5 to construct the Clinical-Spatiotemporal Radiomics model, M6. As shown in Fig. 5A–C, M6 achieved AUCs of 0.928 (95% CI, 0.882–0.974), 0.856 (95% CI, 0.749–0.963), and 0.864 (95% CI, 0.780–0.948) in the training, internal validation, and external test sets, respectively, representing the highest observed AUCs among the major candidate models in all three cohorts.

As shown in Figure S4 and Table S7, compared with M5, M6 increased the AUC by 0.022, 0.116, and 0.046 in the training, internal validation, and external test cohorts, respectively. Paired DeLong tests yielded *P* values of 0.043, 0.018, and 0.041, respectively, indicating that integration of clinicopathologic information with spatiotemporal radiomics provided a further improvement in discrimination across all three cohorts.

Incremental predictive analyses were consistent with the direction of the AUC comparisons. In the training cohort, M6 yielded a continuous net reclassification improvement (NRI) of 0.180 (95% CI, 0.030–0.330; *P* = 0.020) and an integrated discrimination improvement (IDI) of 0.049 (95% CI, 0.006–0.092; *P* = 0.026) compared with M5. In the internal validation cohort, the continuous NRI was 0.410 (95% CI, 0.090–0.730; *P* = 0.012) and the IDI was 0.092 (95% CI, 0.014–0.170; *P* = 0.021). In the external test cohort, the continuous NRI was 0.240 (95% CI, 0.018–0.462; *P* = 0.034) and the IDI was 0.049 (95% CI, 0.005–0.093; *P* = 0.029).

Calibration, Clinical Net Benefit, and Nomogram

The calibration curves in Fig. 5D–F illustrated the agreement between predicted probabilities and observed event rates. Among the evaluated models, the Clinical-Spatiotemporal Radiomics model showed the most consistent calibration across cohorts. Its calibration curve closely follows the 45° reference line in the training cohort and maintained a similar overall pattern in the internal validation and external test cohorts, without evident systematic overestimation or underestimation.

In comparison, the single-modality models showed greater calibration variation across cohorts. The Clinical model exhibited marked non-monotonic fluctuations in the external test cohort, indicating poor agreement between predicted probabilities and observed event rates. The Longitudinal-Intra model showed slope attenuation and plateauing in both the internal validation and external test cohorts, suggesting reduced sensitivity in distinguishing patients at the lower and higher ends of the predicted risk spectrum. Although the Spatiotemporal Radiomics model generally preserved the expected diagonal trend, its cross-cohort calibration consistency remained lower than that of the model incorporating clinical variables.

Decision curve analysis (DCA) evaluated the clinical net benefit of each model across a range of threshold probabilities (Fig. 5G–I). Across the three cohorts, the Clinical-Spatiotemporal Radiomics model generally provided the greatest net benefit over the broadest threshold range and outperformed the treat-all and treat-none strategies across clinically relevant thresholds. In the training cohort, the combined model maintained a favorable net benefit over threshold probabilities extending to approximately 0.85.

This clinical utility was also preserved in the validation datasets. In the internal validation and external test cohorts, the Clinical-Spatiotemporal Radiomics model consistently outperformed the single-modality models across the clinically relevant threshold range of approximately 0.10 to 0.70. Although the Spatiotemporal Radiomics model approached the net benefit of the combined model in the external

test cohort, its benefit declined more clearly at higher thresholds in the internal validation cohort. The Clinical model showed a rapid reduction in net benefit in the external test cohort once the threshold probability exceeded approximately 0.35, approaching the zero-net-benefit reference line. These differences in decision-analytic performance suggested that incorporating longitudinal spatiotemporal imaging information helped overcome the limitations of clinical variables alone and provided additional value for individualized risk stratification.

Finally, the nomogram shown in Fig. 6 translated the multimodal prediction model into a clinically interpretable scoring system. The nomogram incorporated the selected baseline clinicopathologic factors, including histologic grade and CPS, together with longitudinal intratumoral and peritumoral MRI radiomics features. By summing the corresponding point values, clinicians could obtain an individualized quantitative estimate of the probability of pathologic complete response (pCR).

Discussion

By progressively integrating temporal information before and during treatment, spatial phenotypes at the tumor–host interface, and baseline clinicopathologic characteristics, this study develops and validates a preoperative multimodal framework for predicting pathologic complete response (pCR) after neoadjuvant chemoimmunotherapy (NACI) in locally advanced head and neck squamous cell carcinoma (HNSCC). A single imaging time point or an isolated anatomical region may not fully characterize the complexity of treatment response. In contrast, longitudinal intratumoral imaging phenotypes preserve both baseline tumor heterogeneity and early treatment-induced phenotypic changes. The subsequent incorporation of longitudinal features from a 3-mm peritumoral region provides additional predictive gain, suggesting that the spatial tumor–host interface contains treatment-response information that is complementary to that captured within the tumor. The final clinicospacial-temporal radiomics model (M6), which further integrates histologic grade and combined positive score (CPS), achieved the best performance among all compared models, with AUCs of 0.928, 0.856, and 0.864 in the training, internal validation, and independent external test cohorts, respectively. In the external test cohort, M6 achieved a sensitivity of 0.857 and a specificity of 0.845. These findings indicate that longitudinal imaging changes during treatment, peritumoral spatial phenotypes, and clinicopathologic characteristics provide complementary predictive information. Their joint modeling more comprehensively reflects the multilayered biological state underlying pathologic response to NACI than any single information dimension, resulting in a pCR prediction model with good external predictive performance.

Single-time-point imaging primarily reflects the tumor state at a specific moment, whereas the joint analysis of pre-treatment and on-treatment imaging captures both baseline heterogeneity and early treatment-induced phenotypic changes. Previous studies provide initial evidence for the value of this temporal dimension. Liu et al. use changes in histogram features derived from multisequence MRI before and after treatment to predict pCR following neoadjuvant chemoimmunotherapy in HNSCC²³, while Huang et al. demonstrate that multiphase MRI fusion outperforms single-time-point models in the setting of neoadjuvant therapy for breast cancer¹⁵. Accordingly, this study integrates MRI features

obtained before treatment and after the second treatment cycle and constructs longitudinal phenotypes through feature-level fusion to retain complementary information from both time points. The longitudinal intratumoral model M3 achieves AUCs of 0.843 in the training cohort and 0.752 in the external test cohort, both higher than those of the corresponding pre-treatment model M1 (0.806 and 0.738) and on-treatment model M2 (0.823 and 0.700). These results suggest that longitudinal MRI more closely reflects the dynamic process of treatment response by integrating baseline tumor status with early response-related changes. Biologically, treatment-induced tumor cell death, stromal remodeling, inflammatory responses, and local microvascular alterations may manifest as changes in MRI signal intensity and texture distributions^{24,25}. Longitudinal radiomics therefore captures treatment-related heterogeneity that is difficult to characterize adequately using conventional RECIST-based diameter measurements or a single imaging examination, thereby providing an additional dynamic temporal dimension for response assessment.

The peritumoral region lies at the invasive tumor front, where tumor cells, stroma, vasculature, and immune components interact, and therefore provides complementary information beyond the intratumoral phenotype²⁶. Lin et al. demonstrated that combining intratumoral and peritumoral MRI features with CPS significantly improved the prediction of pCR after neoadjuvant chemoimmunotherapy in HNSCC¹¹ suggesting that tissues extending beyond the tumor boundary contain underused phenotypic information. An imaging-pathology correlation study in oral squamous cell carcinoma further showed that specific peritumoral texture features were directly associated with stromal CD45 + immune cell density and PD-L1 CPS, providing in situ pathologic evidence for these macroscopic imaging phenotypes²⁷. The findings of the present study are consistent with this spatial mechanism. In the independent external test cohort, incorporating longitudinal peritumoral features increases the AUC of the longitudinal intratumoral model from 0.752 to 0.818 ($\Delta\text{AUC} = 0.066$, $P = 0.047$), indicating that the peritumoral region provides radiomic information that is complementary to that derived from the intratumoral lesion. With respect to spatial scale, the 3-mm expansion achieves the highest discrimination in cross-validation of the training cohort compared with the 1-mm and 5-mm expansions, with mean AUCs of 0.823, 0.872, and 0.847 for the 1-mm, 3-mm, and 5-mm regions, respectively, and this advantage remains consistent across the validation cohorts. These findings suggest that, within the current anatomical and pathologic setting, a 3-mm expansion provides a favorable spatial scale for enriching biological information from the peritumoral microenvironment while limiting contamination from adjacent normal tissues.

Histologic grade and CPS are independently associated with pCR in multivariable analysis (both $P < 0.05$), providing important clinicopathologic context and baseline information for the imaging-based model. Their incorporation into M5 yields M6, which increases the AUC by 0.022 ($P = 0.043$), 0.116 ($P = 0.018$), and 0.046 ($P = 0.041$) in the training, internal validation, and external test cohorts, respectively, while also producing significant improvements in both NRI and IDI. Previous prospective studies demonstrated an association between PD-L1 CPS and response to neoadjuvant immunotherapy in HNSCC^{11,28,29}. The present study further shows that, even after longitudinal temporal information and

peritumoral spatial information are integrated into the higher-order imaging model M5, the predictive contributions of CPS and histologic grade remain nonredundant. This finding likely reflects the different biological scales represented by imaging features and histopathologic or molecular markers. MRI signals primarily capture macroscopic tissue properties and indirect manifestations of the tumor microenvironment, whereas CPS quantifies the intensity of tumor-immune interactions mediated by PD-L1 expression, and histologic grade reflects the degree of tumor dedifferentiation and intrinsic invasive potential at the cellular level^{30–32}. Such molecular and cellular information cannot be fully represented by macroscopic MRI signals and texture features alone.

The clinical value of the model first lies in the timing of prediction within a critical preoperative decision window. A previous study in oral squamous cell carcinoma³³ performed definitive surgery within 2–4 weeks after two cycles of neoadjuvant chemoimmunotherapy and reported no treatment-related surgical delays, which is broadly consistent with the MRI assessment schedule used in the present study. Setting the prediction time point after completion of the second NACI cycle and before definitive surgery enables the model to incorporate early treatment-induced imaging changes while preserving sufficient time for reassessment of the surgical plan or adjustment of the treatment strategy before surgery^{34,35}. In the independent external test cohort, M6 maintains strong discrimination, with an AUC of 0.864, a sensitivity of 0.857, and a specificity of 0.845, while also demonstrating favorable calibration in the external test cohort, supporting the reliability of its individualized risk estimates. Decision curve analysis further demonstrated favorable net benefit across all three cohorts, supporting its potential utility for preoperative risk stratification.

Building on these findings, this study integrates histologic grade, CPS, and longitudinal intratumoral and peritumoral radiomic features into a visual nomogram that directly translates multimodal information into an individualized probability of pCR. As risk-adapted treatment and function-preserving strategies after NACI receive increasing attention, this quantitative tool provides an interpretable and objective basis for multidisciplinary decision-making and may also support risk stratification and treatment design in future prospective clinical trials.

Several limitations should be considered. First, this study is based on a retrospective two-center cohort in which most participants are male and p16-negative. The transportability of the model across populations with different sex distributions, HPV/p16 status, and more heterogeneous HNSCC characteristics therefore requires further validation in prospective multicenter cohorts. Second, although independent external testing supports the cross-center generalizability of the model, its stability and reproducibility remain to be evaluated across MRI scanners from different vendors, magnetic field strengths, and acquisition protocols. Third, the current radiomic features primarily serve a predictive role, whereas their underlying histologic, immune microenvironmental, and molecular correlates remain to be clarified through spatially matched imaging-pathology analyses and multiomics studies. Future work should also evaluate the associations of the model with event-free survival, disease-free survival, and overall survival using long-term follow-up. Prospective decision studies based on prespecified risk

thresholds and defined clinical management pathways are further needed to determine the practical value of the model for risk-adapted treatment and patient selection in clinical trials.

Conclusion

In conclusion, through stepwise multimodal modeling and independent external testing, this study demonstrates within a unified framework the complementary value of longitudinal intratumoral, peritumoral, and clinicopathologic information for predicting pCR and develops an integrated prediction tool with good discrimination, calibration, and clinical net benefit. These findings support the progressive integration of temporal, spatial, and clinical information as an effective strategy for individualized preoperative stratification of pCR probability in patients with locally advanced head and neck squamous cell carcinoma undergoing neoadjuvant chemoimmunotherapy. By overcoming the limitations of conventional single-time-point imaging assessment, this model provides a quantitative reference for treatment response evaluation throughout the neoadjuvant treatment course. For patients with a high predicted probability of pCR after neoadjuvant therapy, the model may also provide a basis for future prospective investigation of individualized surgical de-escalation and function-preserving strategies, consistent with the broader direction toward precision and individualized treatment.

Abbreviations

AJCC

American Joint Committee on Cancer

AUC

area under the curve

CI

confidence interval

CPS

combined positive score

DCA

decision curve analysis

ECOG

Eastern Cooperative Oncology Group

HNSCC

head and neck squamous cell carcinoma

ICC

intraclass correlation coefficient

IDI

integrated discrimination improvement

NRI

net reclassification improvement

MRI

magnetic resonance imaging
T1WI
T1-weighted imaging
T2WI
T2-weighted imaging
CE-T1WI
contrast-enhanced fat-suppressed T1-weighted imaging
NACI
neoadjuvant chemoimmunotherapy
pCR
pathologic complete response
PD-1
programmed cell death protein 1
PPV
positive predictive value
NPV
negative predictive value
ROI
region of interest
ROC
receiver operating characteristic

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. This retrospective study was approved by the Institutional Review Board of Beijing Tongren Hospital, Capital Medical University, and the Ethics Committee of the First Affiliated Hospital of Zhengzhou University. The requirement for informed consent was waived because of the retrospective nature of the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific funding.

Author Contribution

SL and WL conceived and designed the study and contributed to the analysis strategy. XL, CZ, YS, and WZ collected the clinical and imaging data. SL, LD, SZ, and HL performed the data analysis. SL drafted the manuscript. WL supervised the study and critically revised the manuscript.

Acknowledgement

We acknowledge the technical support provided by the Onekey platform for radiomics feature extraction and analysis.

Availability of data and materials

The data generated in this study are available upon request from the corresponding author.

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. *CA Cancer J Clin*. 2025;75(1):10–45. 10.3322/caac.21871.
2. Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. *CA Cancer J Clin*. 2026;76(4):e70090. 10.3322/caac.70090.
3. Marur S, Forastiere AA. Head and neck squamous cell carcinoma: Update on epidemiology, diagnosis, and treatment. *Mayo Clin Proc*. 2016;91(3):386–396. 10.1016/j.mayocp.2015.12.017
4. Zhang Z, Wu B, Peng G, et al. Neoadjuvant chemoimmunotherapy for the treatment of locally advanced head and neck squamous cell carcinoma: A single-arm phase 2 clinical trial. *Clin Cancer Res*. 2022;28(15):3268–76. 10.1158/1078-0432.CCR-22-0666.
5. Schoenfeld JD, Hanna GJ, Jo VY, et al. Neoadjuvant nivolumab or nivolumab plus ipilimumab in untreated oral cavity squamous cell carcinoma: A phase 2 open-label randomized clinical trial. *JAMA Oncol*. 2020;6(10):1563–70. 10.1001/jamaoncol.2020.2955.
6. Wu D, Li Y, Xu P, et al. Neoadjuvant chemo-immunotherapy with camrelizumab plus nab-paclitaxel and cisplatin in resectable locally advanced squamous cell carcinoma of the head and neck: A pilot phase II trial. *Nat Commun*. 2024;15(1):2177. 10.1038/s41467-024-46444-z.
7. Uppaluri R, Haddad RI, Tao Y, et al. Neoadjuvant and adjuvant pembrolizumab in locally advanced head and neck cancer. *N Engl J Med*. 2025;393(1):37–50. 10.1056/NEJMoa2415434.

8. Baratz HQ, Hidalgo C, Price DL, et al. Neoadjuvant immunotherapy and chemoimmunotherapy regimens in head and neck cancer: A systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg.* 2026;152(5):490–501. 10.1001/jamaoto.2026.0080.
9. Oncology Group of Chinese Medical Association, Head and Neck Tumor Specialty Committee of Chinese Pharmaceutical Education Association, Head and Neck Surgery Group of the Otorhinolaryngology Society of the Chinese Medical Doctor Association. [expert consensus on neoadjuvant therapy for locally advanced head and neck squamous cell carcinoma]. *Zhonghua Yi Xue Za Zhi.* 2025;105(20):1578–95. 10.3760/cma.j.cn112137-20241230-02960.
10. Smussi D, Mattavelli D, Paderno A, et al. Revisiting the concept of neoadjuvant and induction therapy in head and neck cancer with the advent of immunotherapy. *Cancer Treat Rev.* 2023;121:102644. 10.1016/j.ctrv.2023.102644.
11. Lin P, Xie W, Li Y, et al. Intratumoral and peritumoral radiomics of MRIs predicts pathologic complete response to neoadjuvant chemoimmunotherapy in patients with head and neck squamous cell carcinoma. *J Immunother Cancer.* 2024;12(11):e009616. 10.1136/jitc-2024-009616.
12. Ren S, Lan T, Wu F, et al. Intratumoral CD103 + CD8+ T cells predict response to neoadjuvant chemoimmunotherapy in advanced head and neck squamous cell carcinoma. *Cancer Commun.* 2023;43(10):1143–63. 10.1002/cac2.12480.
13. Shankar LK, Schöder H, Sharon E, et al. Harnessing imaging tools to guide immunotherapy trials: Summary from the national cancer institute cancer imaging steering committee workshop. *Lancet Oncol.* 2023;24(3):e133–43. 10.1016/S1470-2045(22)00742-2.
14. Bera K, Braman N, Gupta A, Velcheti V, Madabhushi A. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol.* 2022;19(2):132–46. 10.1038/s41571-021-00560-7.
15. Huang Y, Shi Z, Zhu T, et al. Longitudinal MRI-driven multi-modality approach for predicting pathological complete response and B cell infiltration in breast cancer. *Adv Sci (Weinh).* 2025;12(12):2413702. 10.1002/advs.202413702.
16. Chen Y, Huang Y, Li W, et al. Intratumoral microbiota-aided fusion radiomics model for predicting tumor response to neoadjuvant chemoimmunotherapy in triple-negative breast cancer. *J Transl Med.* 2025;23:352. 10.1186/s12967-025-06369-7.
17. Huang Y, Zhu T, Zhang X, et al. Longitudinal MRI-based fusion novel model predicts pathological complete response in breast cancer treated with neoadjuvant chemotherapy: A multicenter, retrospective study. *EClinicalMedicine.* 2023;58:101899. 10.1016/j.eclinm.2023.101899.
18. Ruffin AT, Li H, Vujanovic L, Zandberg DP, Ferris RL, Bruno TC. Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment. *Nat Rev Cancer.* 2023;23(3):173–88. 10.1038/s41568-022-00531-9.
19. Kocak B, Baessler B, Bakas S, et al. CheckList for EvaluAtion of radiomics research (CLEAR): A step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. *Insights Imaging.* 2023;14(1):75. 10.1186/s13244-023-01415-8.

20. 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(2):328–38. 10.1148/radiol.2020191145.
21. Tejani AS, Klontzas ME, Gatti AA, et al. Checklist for artificial intelligence in medical imaging (CLAIM): 2024 update. *Radiology: Artif Intell*. 2024;6(4):e240300. 10.1148/ryai.240300.
22. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD + AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. 10.1136/bmj-2023-078378.
23. Liu H, Zhu C, Wang X, Chen X, Li Z, Xian J. Prediction of pathological complete response in locally advanced head and neck squamous cell carcinoma treated with neoadjuvant chemo-immunotherapy using volumetric multisequence MRI histogram analysis. *Neuroradiology*. 2024;66(6):919–29. 10.1007/s00234-024-03339-6.
24. Teng X, Ma J, Zhang J, et al. Longitudinal DCE MRI vascular textures: Radiologic and biologic insights for pCR prediction in HER2-negative breast cancer. *Radiology: Artif Intell*. 2026;8(3):e250734. 10.1148/ryai.250734.
25. Huang Y, Gong H, Hu W, et al. Fractal analysis of longitudinal MRI for predicting response to neoadjuvant chemotherapy in breast cancer. *NPJ Precis Oncol*. 2025;10:25. 10.1038/s41698-025-01220-0.
26. Arora R, Cao C, Kumar M, et al. Spatial transcriptomics reveals distinct and conserved tumor core and edge architectures that predict survival and targeted therapy response. *Nat Commun*. 2023;14(1):5029. 10.1038/s41467-023-40271-4.
27. Yuan Z, Ai S, He Q, et al. Intratumoral and peritumoral habitat radiomics of MRI predicts pathologic complete response to neoadjuvant chemoimmunotherapy in oral squamous cell carcinoma. *Int J Surg*. 2025;111(9):6232–44. 10.1097/JS9.0000000000002715.
28. Rosenberg AJ, Juloori A, Jelinek MJ, et al. Neoadjuvant nivolumab plus chemotherapy followed by response-stratified chemoradiation therapy in HPV-negative head and neck cancer: The DEPEND phase 2 nonrandomized clinical trial. *JAMA Oncol*. 2025;11(5):492–501. 10.1001/jamaoncol.2025.0081.
29. Wise-Draper TM, Gulati S, Palackdharry S, et al. Phase II clinical trial of neoadjuvant and adjuvant pembrolizumab in resectable local–regionally advanced head and neck squamous cell carcinoma. *Clin Cancer Res*. 2022;28(7):1345–52. 10.1158/1078-0432.CCR-21-3351.
30. Mints M, Sample RA, Parikh AS, et al. Single-cell analysis highlights the significance of malignant cell IFN/MHC-II for immunotherapy response in head and neck squamous cell carcinoma. *Cell Rep Med*. 2026;7(4):102715. 10.1016/j.xcrm.2026.102715.
31. Liu N, Wu J, Deng E, et al. Immunotherapy and senolytics in head and neck squamous cell carcinoma: Phase 2 trial results. *Nat Med*. 2025;31(9):3047–61. 10.1038/s41591-025-03873-7.
32. Li H, Zandberg DP, Kulkarni A, et al. Distinct CD8 + T cell dynamics associate with response to neoadjuvant cancer immunotherapies. *Cancer Cell*. 2025;43(4):757–e7758.

33. Xiang Z, Wei X, Zhang Z, et al. Efficacy, safety and single-cell analysis of neoadjuvant immunochemotherapy in locally advanced oral squamous cell carcinoma: A phase II trial. Nat Commun. 2025;16(1):3968. 10.1038/s41467-025-59004-w.

34. Zhang X, Zhou Y, Huang J, et al. Response-adapted surgery after neoadjuvant immunochemotherapy in locally advanced head and neck squamous cell carcinoma (REMATCH): A single-arm phase II trial. Clin Cancer Res Published online September. 2026;10:OF1–9. 10.1158/1078-0432.CCR-26-1337.

35. Mastrolonardo EV, Nunes KL, Llerena P, et al. Response-adaptive surgical timing in neoadjuvant immunotherapy demonstrates enhanced pathologic treatment response in head and neck squamous cell carcinoma. Clin Cancer Res. 2025;31(3):515–28. 10.1158/1078-0432.CCR-24-0037.

Figures

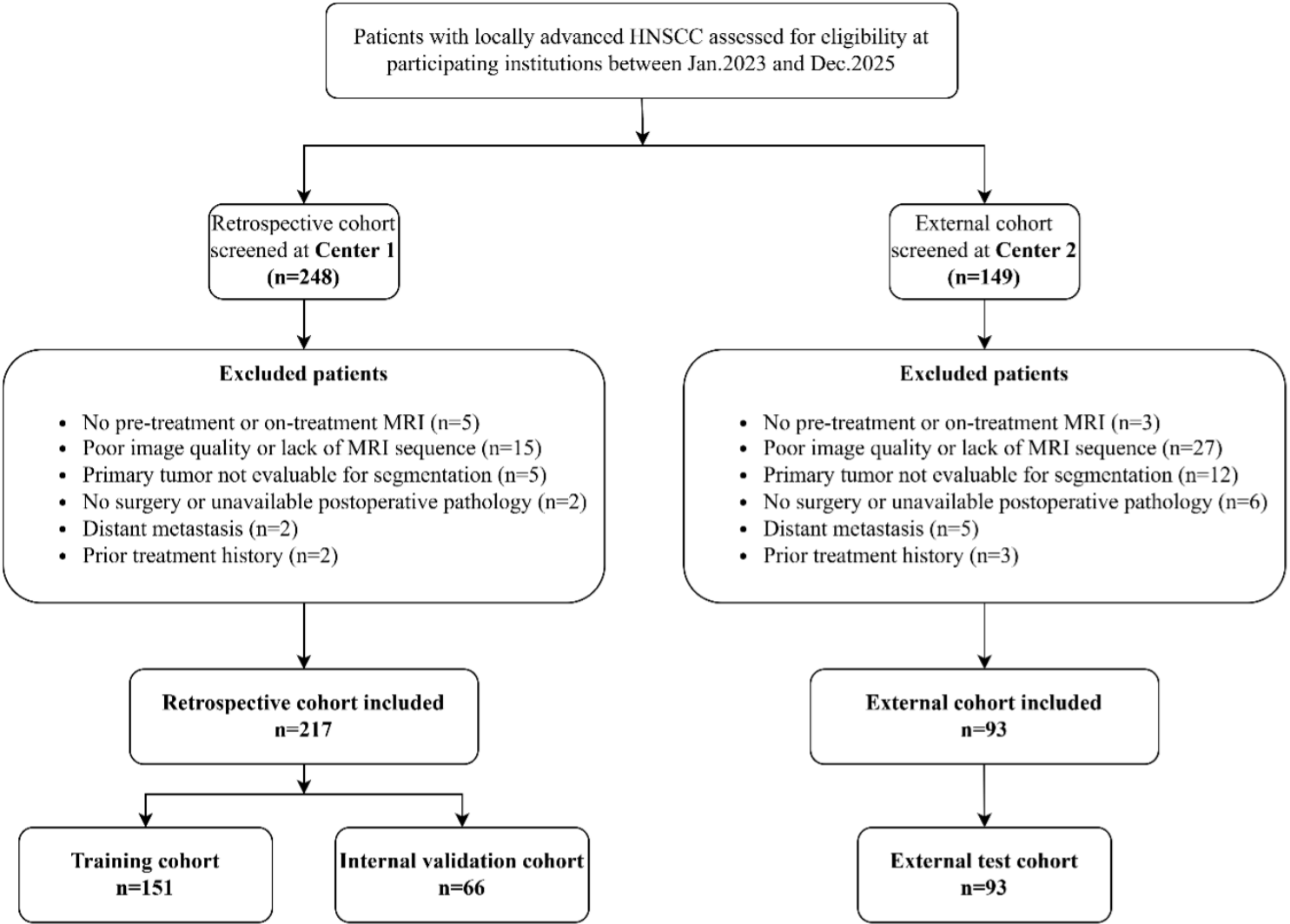


Figure 1

Flowchart of patient selection and cohort allocation.

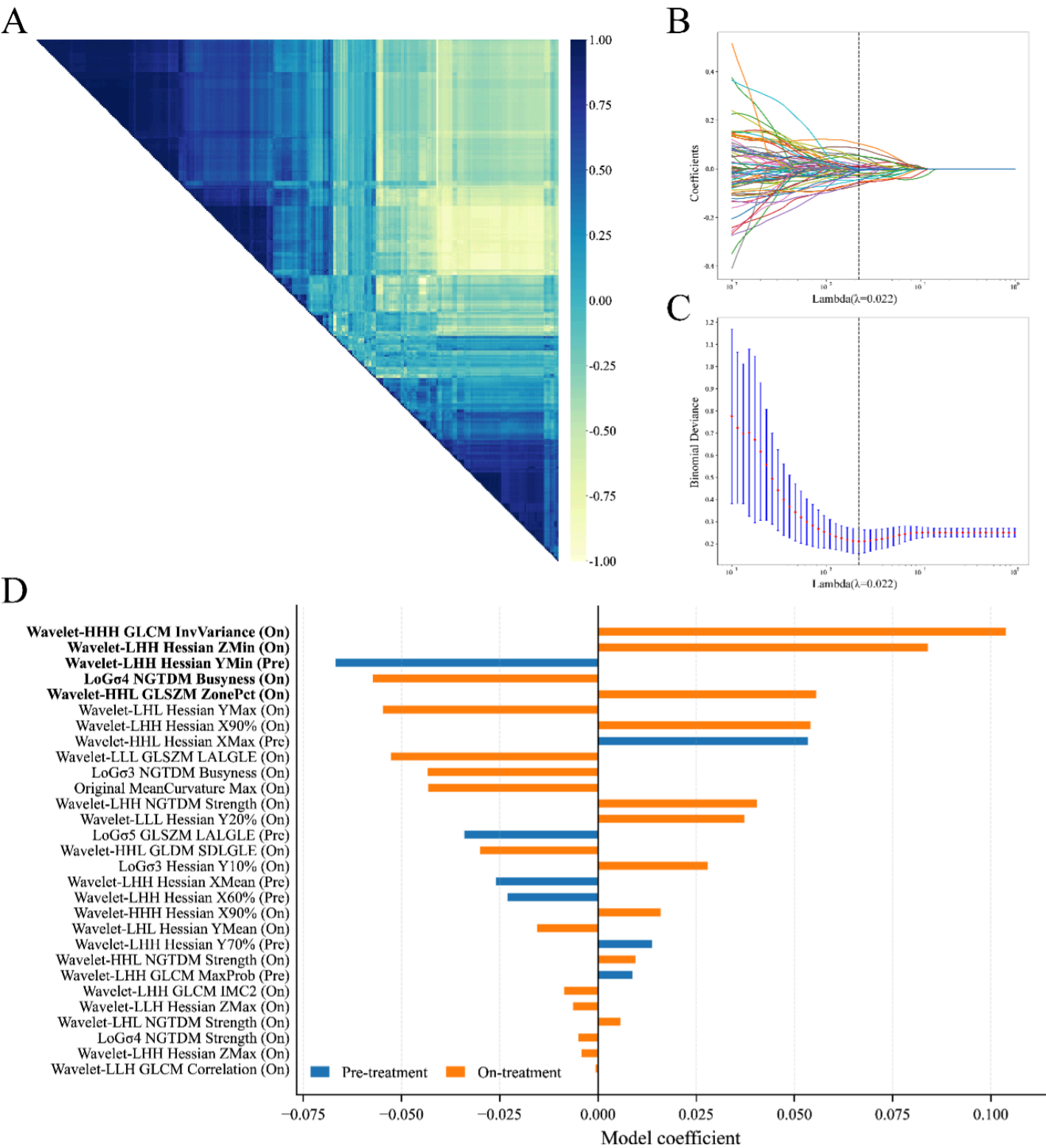


Figure 2

Longitudinal radiomic feature selection for predicting neoadjuvant chemoimmunotherapy response in HNSCC. (A) Pearson correlation matrix of candidate radiomic features used to identify and remove highly collinear redundant features ($|r| > 0.90$). (B) Coefficient profiles from LASSO logistic regression as

a function of the penalty parameter λ . The vertical dashed line indicated the optimal regularization parameter ($\lambda = 0.022$). (C) Binomial deviance as a function of λ based on 10-fold cross-validation. The vertical dashed line indicated the optimal cutoff corresponding to $\lambda = 0.022$. (D) Distribution of the 29 nonzero longitudinal intratumoral features and their regression coefficients retained at the optimal λ . Positive coefficients indicated that higher feature values were associated with achieving pCR, whereas negative coefficients indicated an association with non-pCR. Blue bars represented pre-treatment baseline features, and orange bars represented on-treatment features obtained during the second treatment cycle.

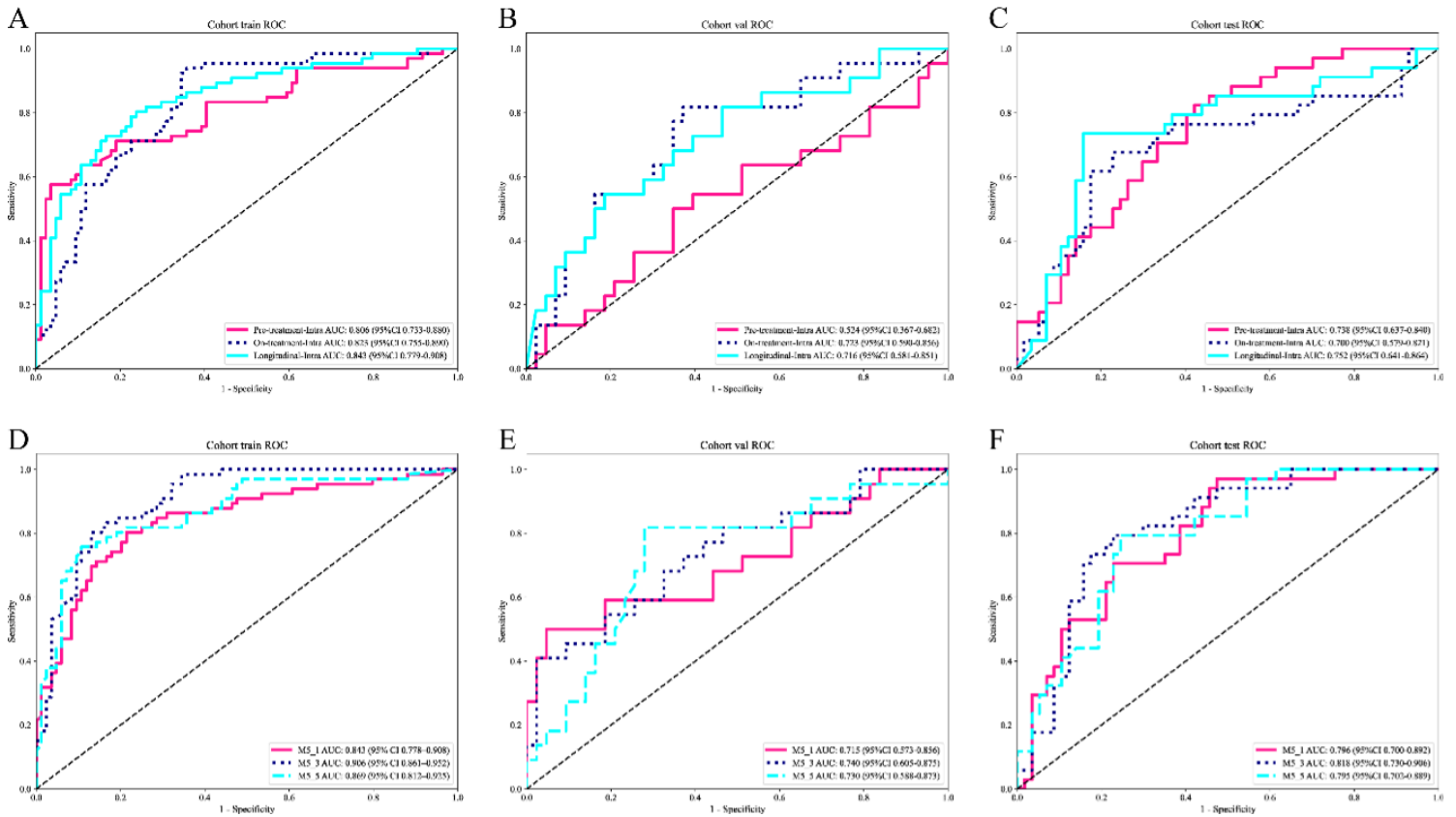


Figure 3

ROC curves of longitudinal intra- and multi-scale peritumoral spatiotemporal radiomics predicting pCR. (A–C) Comparison of ROC curves and AUCs (95% CIs) for the pre-treatment intratumoral model (Pre-treatment-Intra, M1), on-treatment intratumoral model (On-treatment-Intra, M2), and longitudinal intratumoral model (Longitudinal-Intra, M3) in the training cohort (A), internal validation cohort (B), and independent external test cohort (C). (D–F) Comparison of ROC curves and AUCs (95% CIs) for candidate spatiotemporal radiomics models integrating longitudinal intratumoral features with longitudinal peritumoral features at different spatial scales (M5₁, 1 mm; M5₃, 3 mm; M5₅, 5 mm) in the training cohort (D), internal validation cohort (E), and independent external test cohort (F). The black diagonal dashed line represented the reference line for no discriminative ability.

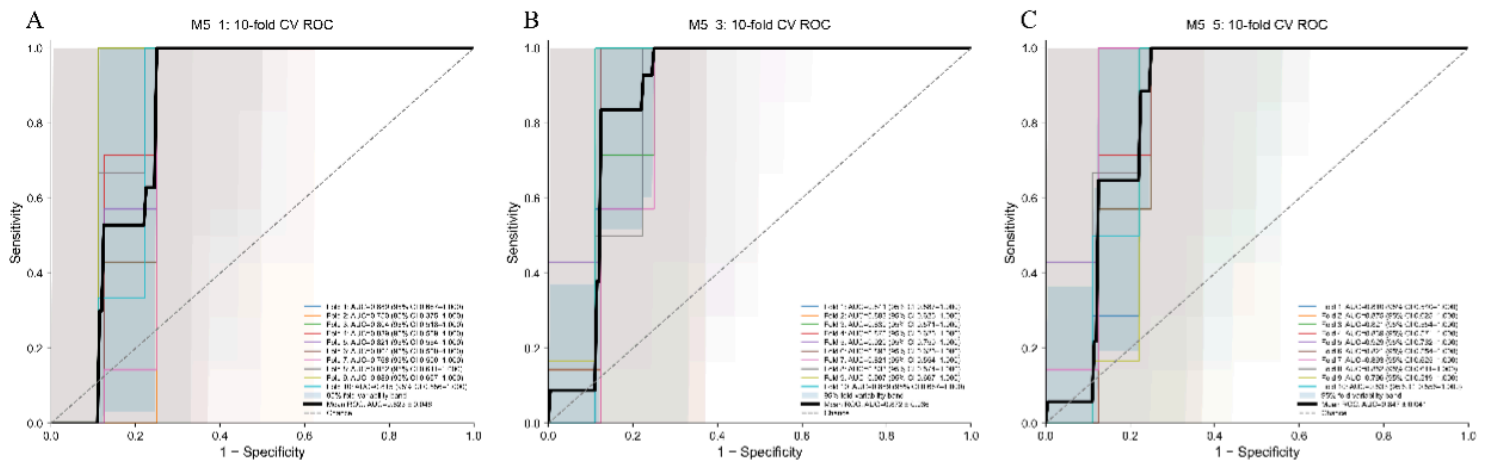


Figure 4

Internal validation of multiscale spatiotemporal radiomics models across 10-fold cross-validation in training cohort.

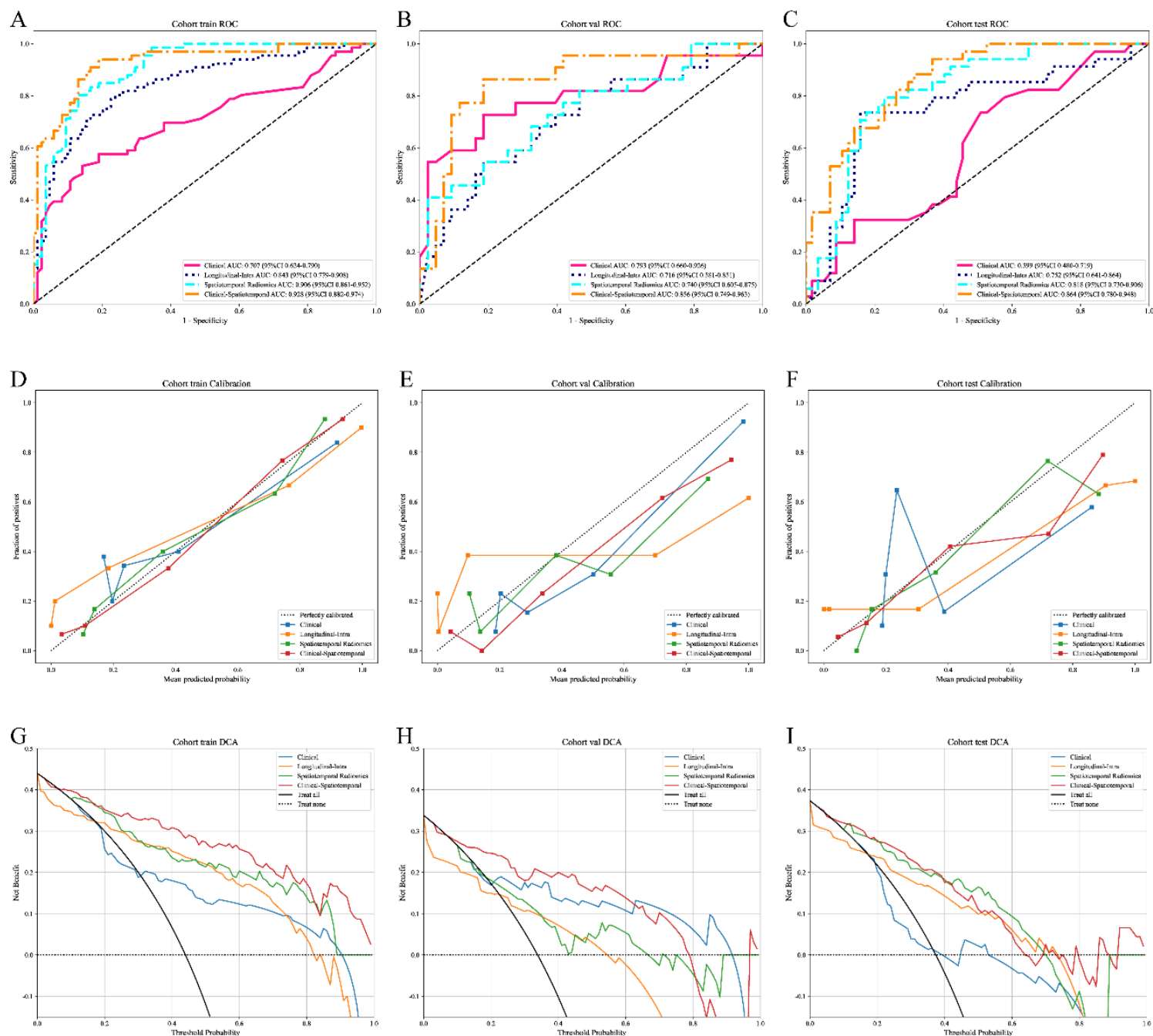


Figure 5

Comprehensive multicohort validation of major predictive models for pCR in HNSCC. (A–C) Receiver operating characteristic curves of the Clinical (M0), Longitudinal-Intra (M3), Spatiotemporal Radiomics (M5), and Clinical-Spatiotemporal Radiomics (M6) models in the training cohort (A), internal validation cohort (B), and external test cohort (C). (D–F) Calibration curves of the four models in the corresponding cohorts. (G–I) Decision curve analyses of the four models in the corresponding cohorts.

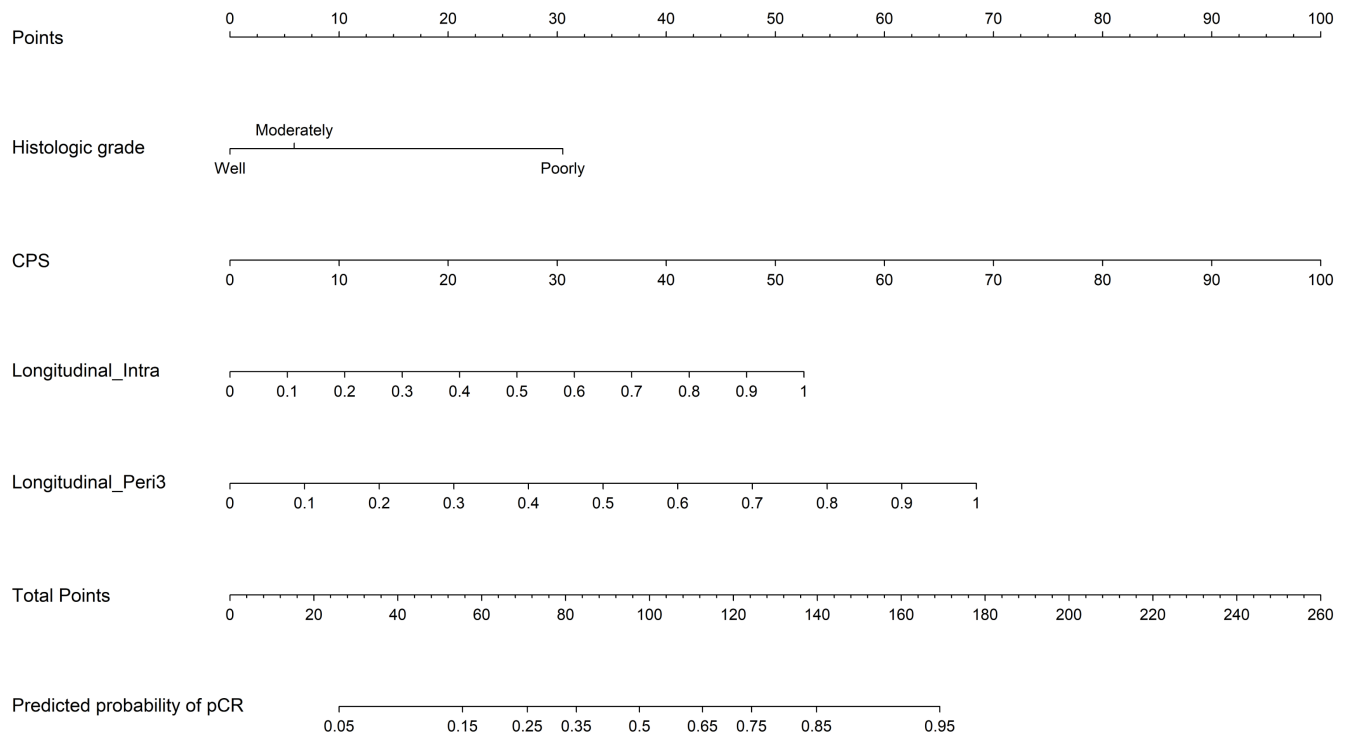


Figure 6

Nomogram of the Clinical-Spatiotemporal Radiomics model for pCR prediction.

Supplementary Files

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

- [Supplementarymaterial1.docx](#)
- [Supplementarymaterial2.docx](#)