



Intratumoral and peritumoral heterogeneity based on CT to predict the pathological response after neoadjuvant chemoimmunotherapy in esophageal squamous cell carcinoma

Xiaodong Ling, MD^b, Xiaolong Yang, MS^a, Ping Wang, MD^c, Yingjie Li, MS^a, Zhubin Wen, MS^a, Jiayang Wang, MS^a, Kaige Chen, MS^a, Yanhong Yu, MS^d, Aoyu Liu, MS^a, Jianqun Ma, MD^{b,*}, Wei Meng, MD^{a,*}

Background: Neoadjuvant chemoimmunotherapy (NACI) regimen (camrelizumab plus paclitaxel and nedaplatin) has shown promising potential in patients with esophageal squamous cell carcinoma (ESCC), but accurately predicting the therapeutic response remains a challenge.

Objective: To develop and validate a CT-based machine learning model that incorporates both intratumoral and peritumoral heterogeneity for predicting the pathological response of ESCC patients after NACI.

Methods: Patients with ESCC who underwent surgery following NACI between June 2020 and July 2024 were included retrospectively and prospectively. Univariate and multivariate logistic regression analyses were performed to identify clinical variables associated with pathological response. Traditional radiomics features and habitat radiomics features from the intratumoral and peritumoral regions were extracted from posttreatment CT images, and 6 predictive models were established using 14 machine learning algorithms. The combined model was developed by integrating intratumoral and peritumoral habitat radiomics features with clinical variables. The performance of the models was evaluated using the area under the receiver operating characteristic curve (AUC).

Results: A total of 157 patients (mean [SD] age, 59.6 [6.5] years) were enrolled in our study, of whom 60 (38.2%) achieved major pathological response (MPR) and 40 (25.5%) achieved pathological complete response (pCR). The combined model demonstrated excellent predictive ability for MPR after NACI, with an AUC of 0.915 (95% CI: 0.844–0.981), accuracy of 0.872, sensitivity of 0.733, and specificity of 0.938 in the test set. In sensitivity analysis focusing on pCR, the combined model exhibited robust performance, with an AUC of 0.895 (95% CI: 0.782 – 0.980) in the test set.

Conclusion: The combined model integrating intratumoral and peritumoral habitat radiomics features with clinical variables can accurately predict MPR in ESCC patients after NACI and shows promising potential in predicting pCR.

Keywords: chemoimmunotherapy, esophageal squamous cell carcinoma, habitat radiomics, pathological response

Introduction

Esophageal cancer ranks as the seventh most common cancer worldwide and the sixth leading cause of cancer-related deaths^[1]. Globally, more than half of new esophageal cancer cases occur in China, with more than 90% of these cases are esophageal squamous

cell carcinoma (ESCC)^[2]. Despite advancements in therapeutic approaches, the prognosis for ESCC patients remains unsatisfactory, particularly in those with locally advanced cases, where the 5-year survival rate is less than 20%^[3]. A previous study developed a novel treatment regimen consisting of neoadjuvant camrelizumab in combination with paclitaxel and nedaplatin^[4]. The results demonstrated that this neoadjuvant chemoimmunotherapy (NACI) regimen showed satisfactory efficacy in patients with locally advanced ESCC, achieving a major pathological response

^aDepartment of Radiology, Harbin Medical University, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang, China, ^bDepartment of Thoracic Surgery, Harbin Medical University, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang, China, ^cDepartment of Radiology, Beijing Jishuitan Hospital, Capital Medical University, Beijing, China and ^dMagnetic Resonance Imaging Department, The Second People's Hospital, Yangming District, Mudanjiang City, Heilongjiang Province, China

Xiaodong Ling and Xiaolong Yang authors are contributed equally to this work and share the first authorship.

Jianqun Ma and Wei Meng are co-corresponding authors.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding authors. Address: Radiology Department, Harbin Medical University, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin, Heilongjiang 150081, China. Tel.: +86 010 69159590. E-mail: articlemengwei@163.com (W. Meng); Department of Thoracic Surgery, Harbin Medical University, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin 150081, Heilongjiang, China. E-mail: jianqunma@hrbmu.edu.cn (J. Ma).

Copyright © 2025 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

International Journal of Surgery (2026) 112:314–324

Received 24 March 2025; Accepted 26 August 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website www.ijsof.com.

Published online 19 September 2025

<https://dx.doi.org/10.1097/JS9.0000000000003422>

(MPR) rate of 45.2% and a pathological complete response (pCR) rate of 27.4%.

Pathological response is an important indicator for evaluating the efficacy of neoadjuvant therapy. MPR is defined as the percentage of residual tumor cells in the primary tumor and lymph nodes not exceeding 10%, while no residual tumor is observed is classified as a pCR^[5]. MPR and pCR have been shown to correlate with higher long term survival and a lower risk of recurrence^[6–8]. The image-based response evaluation criteria in solid tumors (RECIST) scoring system is commonly used to assess treatment response in ESCC. However, due to the unique mechanisms of immunotherapy, discrepancies often arise between RECIST scores and pathological results, such as delayed response, pseudoprogression, and hyperprogression^[9,10]. These discrepancies may lead to an underestimation of neoadjuvant efficacy, potentially resulting in high-response patients receiving unnecessary prolonged treatment. Several studies suggest that blood biomarkers show promise in predicting pathological response to neoadjuvant therapy, although their reliability can be affected by factors such as inflammation or infection^[11,12]. Therefore, there is an urgent need for a more reliable method to accurately predict the pathological response of ESCC patients after NACI.

Radiomics extracts quantitative features from medical images and transforms image data into structured formats, offering a promising method for comprehensive tumor phenotype quantification^[13]. However, traditional radiomics does not fully capture the spatial heterogeneity within the tumor^[14]. This heterogeneity is a known factor influencing treatment response and prognosis^[15]. Habitat radiomics is an innovative approach that divides the tumor into different subregions based on similar imaging features, providing a better representation of tumor heterogeneity^[16]. Machine learning algorithms excel at extracting key information from high-dimensional data, overcoming the limitations and biases associated with manual feature extraction^[17]. Previous studies developed a CT-based habitat model to predict pathological response after neoadjuvant therapy in ESCC patients, but they only analyzed the tumor, neglecting the peritumoral region^[18]. Hu *et al*^[19] showed that peritumoral regions provide additional predictive value for pathological response after neoadjuvant chemoradiotherapy in ESCC patients.

Therefore, the aim of this study was to develop a machine learning model based on CT images, integrating intratumoral and peritumoral heterogeneity features along with clinical variables, to predict the pathological response of ESCC patients after receiving NACI, in order to provide a basis for personalized neoadjuvant therapy and avoid overtreatment of well-responders. Additionally, the study employed SHapley Additive exPlanations (SHAP) analysis for the interpretability analysis of the model.

Materials and methods

Patients

This study was approved by the institutional ethics committee. Written informed consent was obtained for the prospective cohort, while it was waived for the retrospective cohort. The work has been reported in line with the Standards for the Reporting of Diagnostic accuracy studies criteria.

Between June 2020 and June 2023, we retrospectively reviewed data from 141 consecutive patients with ESCC who received NACI

HIGHLIGHTS

- The habitat radiomics model demonstrates superior predictive performance compared to the traditional radiomics model.
- The combined model integrating intratumoral and peritumoral habitat features with clinical variables exhibits the best predictive performance for pathological response, with an AUC of 0.915.
- SHapley Additive exPlanations (SHAP) analysis enhances the interpretability of the model.

treatment at our institution. The inclusion criteria were as follows: (1) pathology-confirmed ESCC, (2) receipt of at least two cycles of NACI, and (3) availability of contrast-enhanced CT images within 2 weeks prior to surgery. The exclusion criteria were: (1) failure to undergo surgery, (2) poor-quality CT scans, (3) prior antitumor therapy, and (4) incomplete clinicopathological data. Ultimately, 110 patients (109 men and 1 woman; mean age, 60 years) were included in the study and used for training the predictive model.

Between June 2023 and May 2024, ESCC patients who received NACI treatment were prospectively recruited from our institution. The enrollment criteria were identical to those for the retrospective cohort. Finally, 47 eligible ESCC patients were included in the prospective test set to evaluate the model's performance. The details of patient inclusion are shown in Fig. 1A.

In this study, the NACI regimen consisted of camrelizumab (200 mg), paclitaxel (155 mg/m²), and nedaplatin (80 mg/m²). All patients received the medications via intravenous injection on the first day of each treatment cycle, with each cycle lasting 3 weeks. Patients underwent two to four cycles of NACI treatment based on individual condition. After completing treatment, patients underwent radical surgical resection.

Clinical information collection and pathologic evaluation

Clinical data were collected through the electronic health record, with details provided in Supplemental Digital Content A1, available at: <http://links.lww.com/JS9/F149>. Two pathologists, with 10 and 15 years of experience, evaluated the postoperative pathology specimens and determined the MPR and pCR status based on the proportion of residual tumor cells in both the primary tumor and lymph nodes. A detailed description of the evaluation process is provided in Supplemental Digital Content A2, available at: <http://links.lww.com/JS9/F149>.

CT image acquisition and segmentation

The posttreatment contrast-enhanced CT images were retrieved from the picture archiving and communication system. The details of acquisition are provided in Supplemental Digital Content A3, available at: <http://links.lww.com/JS9/F149>. Two radiologists, with 8 and 5 years of experience, respectively, were blinded to the pathological response and manually delineated the intratumoral and peritumoral regions of interest (ROIs) using ITK-SNAP software (version 3.8.0, <https://www.itksnap.org>). Intratumoral ROIs were defined by outlining the tumor boundaries on each slice of the transverse CT images, ensuring the exclusion of any gas present in the esophageal lumen. The peritumoral ROIs included adjacent tissues and lymph nodes,

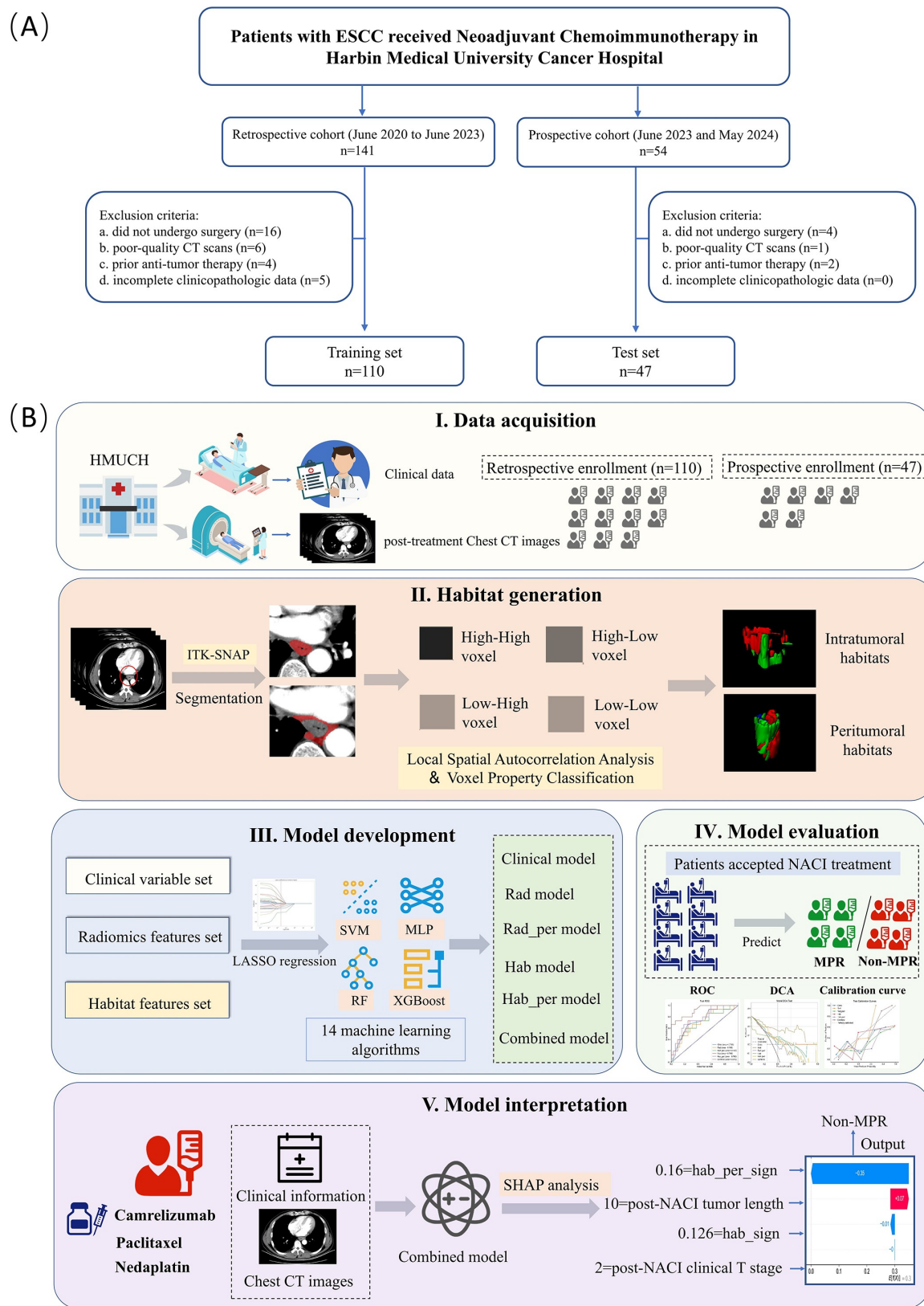


Figure 1. Flowchart of patient enrollment (A) and the overall flowchart of this study (B). ESCC, esophageal squamous cell carcinoma; AUC, area under the receiver operating characteristic curve; DCA, decision curves analysis. Rad model, intratumoral radiomics model; Rad_per model, peritumoral radiomics model; Hab model, intratumoral habitat radiomics model; Hab_per model, peritumoral habitat radiomics model.

excluding the airway, aorta, vertebrae, and azygos vein. Disagreements were resolved through mutual discussion and consensus among the radiologists.

Intratumoral and peritumoral habitat generation

To objectively extract intratumoral and peritumoral subregions, a method based on local spatial autocorrelation and regional attribute values was employed. CT images and corresponding mask files were inputted, spatial attributes were analyzed, and voxel sets for each ROI were extracted. A neighborhood weight matrix was constructed based on Euclidean distance, and the local Moran's index for each voxel was calculated to capture the spatial correlations between voxels. Based on the magnitude and direction of the Moran's index, as well as the relationship between voxel property values and the regional mean, each voxel was reclassified into one of five categories: high-high, low-low, high-low, low-high, or non-significant. The generation of the above habitat was carried out using the PixelMed AI platform. Additional details of methodology for habitat generation can be found in the Supplemental Digital Content A4, available at: <http://links.lww.com/JS9/F149>.

Feature extraction and selection

We used the open-source Python package, Pyradiomics (version 3.0.1; <https://pyradiomics.readthedocs.io>) to extract radiomics features. All images were standardized preprocessing prior to feature extraction. Details of preprocessing and feature extraction are described in Supplemental Digital Content A5, available at: <http://links.lww.com/JS9/F149>. To select reproducible and nonredundant radiomic features, we first calculated the mean of the radiomic features from different subregions and performed an independent t-test, removing features with $P > 0.05$. Next, Pearson correlation analysis was used to eliminate highly correlated features, with a threshold set at 0.9. Finally, the Least Absolute Shrinkage and Selection Operator algorithm was applied for feature dimensionality reduction in the multivariate analysis. The optimal regularization parameter λ was determined through 10-fold cross-validation to identify independent and stable radiomics features. For clinical variables, univariate and multivariate logistic regression (LR) were used to identify independent predictors associated with pathological response ($P < 0.05$).

Model development and evaluation

The workflow of the study is illustrated in Fig. 1B. We developed six models using the selected features: (1) clinical features (clinical model), (2) intratumoral radiomics features (Rad model), (3) peritumoral radiomics features (Rad_per model), (4) intratumoral habitat radiomics features (Hab model), (5) peritumoral habitat radiomics features (Hab_per model), and (6) a combined model integrating clinical variables, intratumoral habitat radiomics features, and peritumoral habitat radiomics features. The model development utilized 14 machine learning algorithms supported by the scikit-learn library, including LR, naive Bayes, linear support vector machine (linear_SVM), polynomial support vector machine (poly_SVM), sigmoid support vector machine (sigmoid_SVM), radial basis function support vector machine (rbf_SVM), decision tree, Random Forest, Extremely Randomized Trees, eXtreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost), multilayer perceptron, Gradient Boosting Machine, and LightGBM. The hyperparameters for each

model are shown in Supplementary Digital Content A6, available at: <http://links.lww.com/JS9/F149>. The algorithm with the highest area under the receiver operating characteristic curve (AUC) and minimal evidence of overfitting was selected for further investigation. In our study, the performance of the models was assessed by the AUC and its 95% confidence interval (95% CI). Calibration curves and decision curve analysis (DCA) were used to evaluate the usefulness of the predictive models.

Model explanation

SHAP analysis is a model interpretation method based on game theory, which quantifies the importance of features and their roles in predictions by calculating each feature's contribution to the model's output. In this study, we employed SHAP analysis to perform an interpretability assessment of the developed combined model. The results were visualized using SHAP bar plots and beeswarm plots. Additionally, we selected representative cases to create SHAP waterfall plots, allowing for an in-depth examination of the model's prediction process for specific individuals.

Development of web-based calculator

To facilitate the application of the model in clinical settings, we integrated the optimal prediction model into a web platform developed using the Shiny framework. After inputting the relevant feature values and clicking the "Calculate" button, the application automatically computes the probability of ESCC patients achieving MPR after receiving NACI treatment.

Statistical analyses

Continuous variables following a normal distribution were analyzed using an independent samples t-test and expressed as mean and standard deviation. The Mann-Whitney U test was used to analyze nonnormally distributed data, which were expressed as median (interquartile range). Levene's test was used to assess the homogeneity of variances. Categorical data were compared using the Chi-square test or Fisher's exact test and expressed as frequencies with percentages. The differences in AUC between models were compared using the DeLong test. Two-tailed $P < 0.05$ was considered indicative of a statistically significant difference. The sample size estimation indicates that a minimum of 100 participants are required (Supplemental Digital Content A7, available at: <http://links.lww.com/JS9/F149>). All statistical analyses were performed using Python (version 3.8.2; <https://www.python.org>).

Results

Patient characteristics

The baseline characteristics are summarized in Table 1. This study included 157 patients, of whom 60 (38.2%) achieved MPR and 40 (25.5%) achieved pCR. The average age of the entire cohort was 59.6 years (standard deviation, 6.5 years), with 98.1% ($n = 154$) being male. The majority of tumors were located in the lower thoracic esophagus ($n = 94$, 59.9%). Regarding pretreatment staging, most patients were in the T3 ($n = 119$, 75.8%) and N1 ($n = 70$, 44.6%) stages, with the highest proportion in stage III ($n = 83$, 52.9%). Compared to the training set, patients in the test set had larger tumor thickness, larger maximum short-axis lymph node diameters after NACI, and a higher proportion of patients

Table 1
The baseline characteristics of the enrolled patients

Characteristics	Total (n = 157)	Training set (n = 110)	Test set (n = 47)	P value
Age (year) ^a	59.6 ± 6.5	60.0 ± 6.0	58.5 ± 7.4	0.228
Sex	—	—	—	0.214
Female	154 (98.1)	109 (99.1)	45 (95.7)	—
Male	3 (1.9)	1 (0.9)	2 (4.3)	—
Smoking	—	—	—	0.257
Yes	81 (51.6)	53 (48.2)	28 (59.6)	—
No	76 (48.4)	57 (51.8)	19 (40.4)	—
Drinking	—	—	—	0.315
Yes	89 (56.7)	59 (53.6)	30 (63.8)	—
No	68 (43.3)	51 (46.4)	17 (36.2)	—
Tumor location	—	—	—	0.615
Upper thoracic	12 (7.6)	8 (7.3)	4 (8.5)	—
Middle thoracic	94 (59.9)	64 (58.2)	30 (63.8)	—
Lower thoracic	42 (26.8)	30 (27.3)	12 (25.5)	—
Cardia	9 (5.7)	8 (7.3)	1 (2.1)	—
Pre-NACI clinical T stage	—	—	—	0.115
T2	38 (24.2)	31 (28.2)	7 (14.9)	—
T3	119 (75.8)	79 (71.8)	40 (85.1)	—
Pre-NACI clinical N stage	—	—	—	0.253
N0	55 (35.0)	34 (30.9)	21 (44.7)	—
N1	70 (44.6)	52 (47.3)	18 (38.3)	—
N2	32 (20.4)	24 (21.8)	8 (17.0)	—
Pre-NACI clinical stage	—	—	—	0.904
II	74 (47.1)	51 (46.4)	23 (48.9)	—
III	83 (52.9)	59 (53.6)	24 (51.1)	—
Post-NACI clinical T stage	—	—	—	0.035 ^c
T0	14 (8.9)	14 (12.7)	—	—
T2	107 (68.2)	71 (64.5)	36 (76.6)	—
T3	36 (22.9)	25 (22.7)	11 (23.4)	—
Post-NACI clinical N stage	—	—	—	0.612
N0	98 (62.4)	67 (60.9)	31 (66.0)	—
N1	52 (33.1)	37 (33.6)	15 (31.9)	—
N2	7 (4.5)	6 (5.5)	1 (2.1)	—
Baseline tumor length (mm) ^b	45.0 [35.0, 60.0]	45.0 [35.0, 58.8]	45.0 [30.0, 57.5]	0.699
Baseline tumor thickness (mm) ^b	16.0 [13.0, 20.0]	16.0 [13.0, 19.8]	18.0 [14.0, 22.5]	0.030 ^c
Baseline maximum short diameter of lymph nodes (mm) ^b	9.0 [7.0, 13.0]	9.0 [7.0, 13.0]	9.0 [8.0, 12.5]	0.532
Post-NACI tumor length (mm) ^b	15.0 [10.0, 25.0]	15.0 [10.0, 25.0]	15.0 [10.0, 20.0]	0.592
Post-NACI tumor thickness (mm) ^b	10.0 [7.0, 12.0]	9.0 [6.0, 11.0]	11.0 [10.0, 13.0]	<0.00 ^c
Post-NACI maximum short diameter of lymph nodes (mm) ^b	7.0 [5.0, 9.0]	6.0 [5.0, 9.0]	8.0 [6.0, 9.0]	0.040 ^c
Pathological response	—	—	—	0.291
MPR without pCR	20 (12.7)	13 (11.8)	7 (14.9)	—
pCR	40 (25.5)	32 (29.1)	8 (17.0)	—
Non-MPR	97 (61.8)	65 (59.1)	32 (68.1)	—

Except where indicated, data are numbers of patients, with percentages in parentheses. NACI, neoadjuvant chemoimmunotherapy; n, number of patients; P values correspond to the comparison of baseline characteristics between the training set and the test set; MPR, major pathological response; pCR, pathological complete response.

^aData are mean, with SD in parentheses.

^bData are medians, with IQRs in parentheses.

^cIndicates statistical significance; $P < 0.05$.

with higher T stages. No significant differences were observed between the two groups for the other characteristics.

Feature extraction and selection

Based on the analysis results of local spatial autocorrelation and regional attribute values, the intratumoral and peritumoral regions were each divided into four subregions. We extracted 1678 radiomics features from the tumor, peritumor, and each subregion, respectively. After feature selection, 17 traditional radiomics features and 15 habitat radiomics features from the tumor region were selected to develop the Rad model and Hab

model, respectively. In the peritumoral region, 8 traditional radiomics features and 10 habitat radiomics features were selected to develop the Rad_per model and Hab_per model, respectively.

After univariate and multivariate LR, post-NACI tumor length and post-NACI clinical T stage remained independent predictors of MPR and were used to develop the clinical model (Table 2).

Model performance and sensitivity analysis

The linear-SVM algorithm outperformed other machine learning algorithms in constructing the clinical and Rad models. For the

Table 2
Logistic regression analysis for the association between MPR to NACI and patient characteristics

Characteristics	Univariable analysis		Multivariable analysis	
	OR (95% CI)	P value	OR (95% CI)	P value
Age (year)	1.024 (0.973, 1.078)	0.351	—	—
Sex	3.310 (0.294, 37.320)	0.333	—	—
Smoking	2.137 (1.109, 4.119)	0.023 ^a	2.266 (0.709, 7.249)	0.168
Drinking	1.937 (1.008, 3.722)	0.047 ^a	1.314 (0.410, 4.211)	0.646
Tumor location	0.977 (0.683, 1.399)	0.900	—	—
Pre-NACI clinical T stage	1.080 (0.508, 2.299)	0.841	—	—
Pre-NACI clinical N stage	0.849 (0.545, 1.324)	0.470	—	—
Pre-NACI clinical stage	1.205 (0.632, 2.296)	0.571	—	—
Post-NACI clinical T stage	0.293 (0.160, 0.536)	<0.001 ^a	0.392 (0.202, 0.760)	0.005 ^a
Post-NACI clinical N stage	0.586 (0.323, 1.065)	0.079	—	—
Baseline tumor length (mm)	1.006 (0.990, 1.021)	0.463	—	—
Baseline tumor thickness (mm)	0.991 (0.939, 1.046)	0.746	—	—
Baseline maximum short diameter of lymph nodes (mm)	1.013 (0.960, 1.069)	0.631	—	—
Post-NACI tumor length (mm)	0.826 (0.771, 0.884)	<0.001 ^a	0.852 (0.793, 0.916)	0.547
Post-NACI tumor thickness (mm)	0.827 (0.740, 0.924)	<0.001 ^a	0.958 (0.833, 1.102)	<0.001 ^a
Post-NACI maximum short diameter of lymph nodes (mm)	0.909 (0.798, 1.036)	0.152	—	—

P values represent the comparison between patients with MPR and non-MPR after NACI in the training set. NACI, neoadjuvant chemioimmunotherapy; MPR, major pathological response.

^aIndicates statistical significance; $P < 0.05$.

Rad-per and Hab-per models, the LR algorithm achieved the best performance. The NB and XGBoost algorithms showed superior performance in constructing the Hab and combined models, respectively. Detailed data are provided in Supplemental Digital Content Table S1-6, available at: <http://links.lww.com/JS9/F149>.

For predicting MPR in ESCC patients after NACI, the AUCs of the Rad and Rad_per models were 0.738 (95% CI: 0.572–0.835) and 0.733 (95% CI: 0.512–0.862) in the test set, respectively. The habitat radiomics models outperformed the traditional models, with the Hab model achieving an AUC of 0.756 (95% CI: 0.609–0.887) and the Hab_per model an AUC of 0.769 (95% CI: 0.632–0.898) in the test set. The combined model, integrating clinical variables, intratumoral and peritumoral habitat radiomics features, achieved

the highest predictive performance, with AUCs of 0.987 (95% CI: 0.967–0.999) and 0.915 (95% CI: 0.844–0.981) in the training and test sets, respectively (Fig. 2). Calibration curves showed high consistency between the combined model's predictions and actual observations (Fig. 3A and B). Moreover, DCA demonstrated that the combined model offered greater net benefits across a wider range of threshold probabilities (Fig. 3C and D). Other diagnostic metrics for the six models are detailed in Table 3.

In the sensitivity analysis focusing on the pCR, the combined model based on the AdaBoost algorithm performed the best, with AUCs of 0.938 (95% CI: 0.948, 0.95) in the training set and 0.895 (95% CI: 0.782, 0.980) in the test

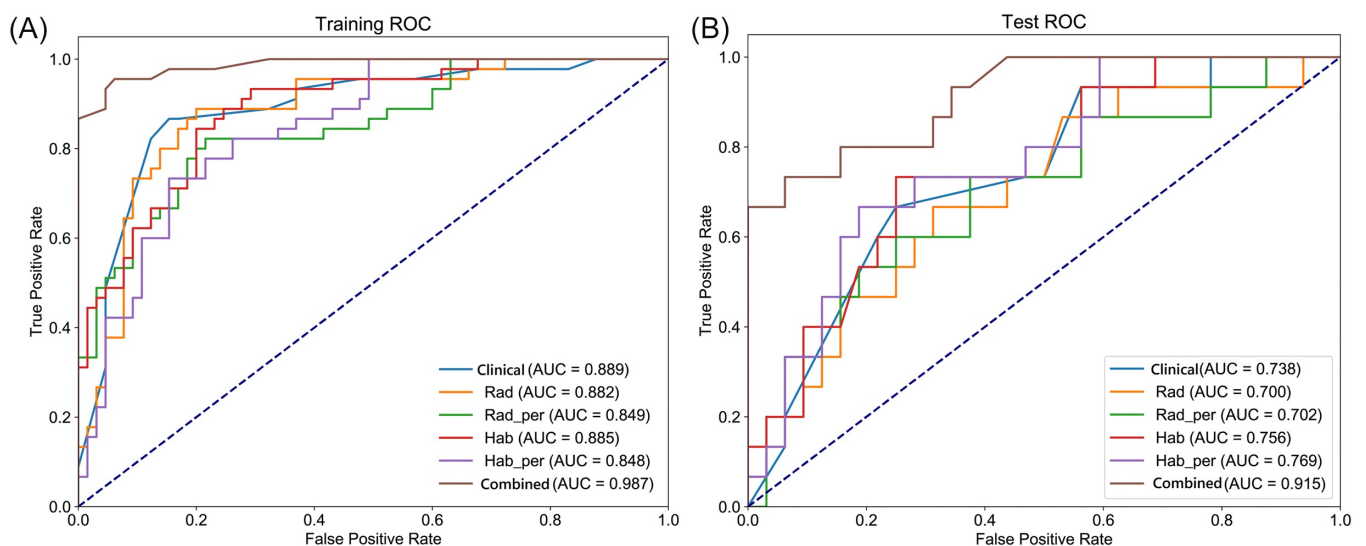


Figure 2. ROC of models in predicting MPR in the training set (A) and the test set (B). ROC, receiver operating characteristic curve; AUC, area under the receiver operating characteristic curve.

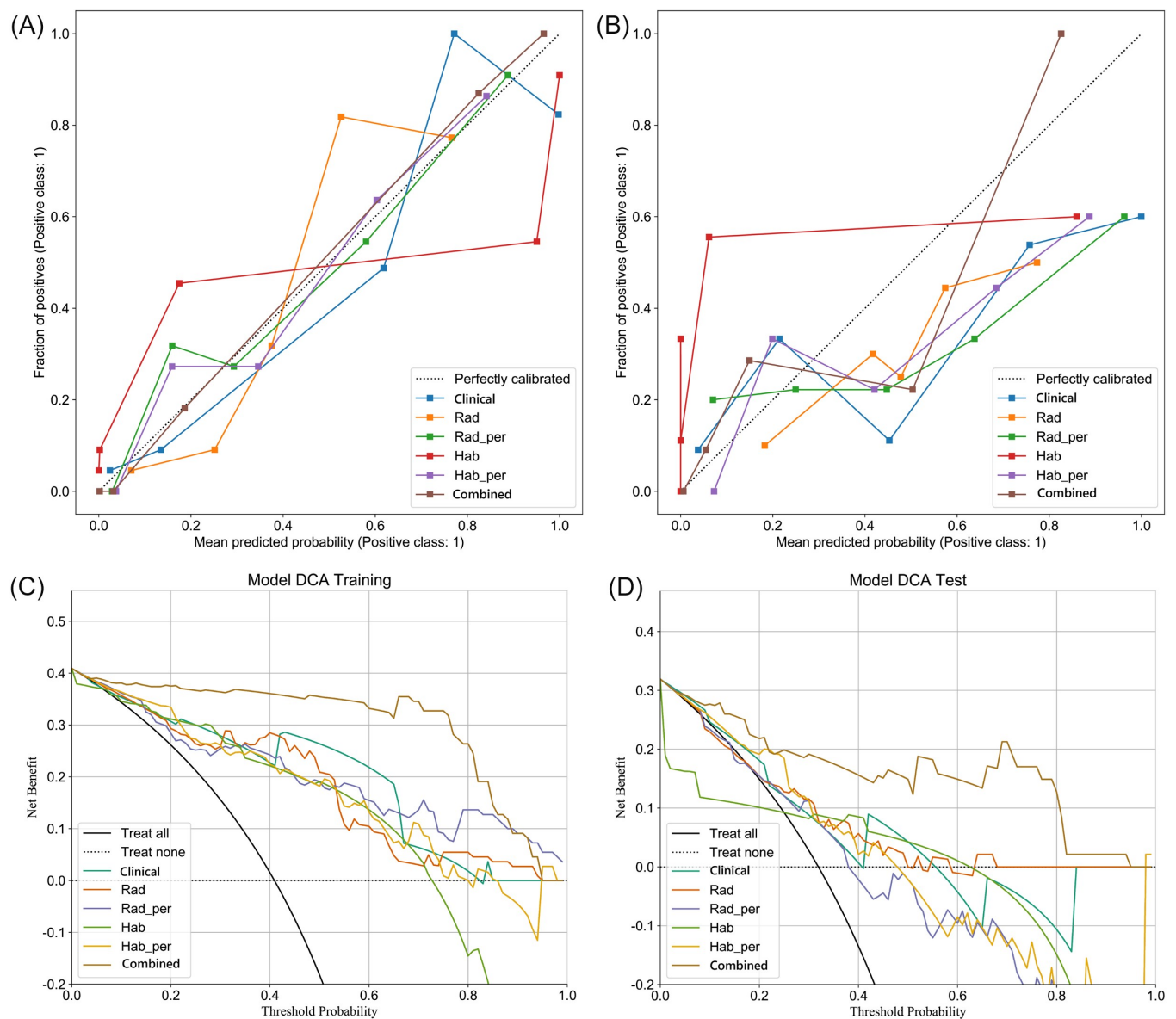


Figure 3. Decision curve analysis and calibration curves. Decision curve analysis (DAC) of models in predicting MPR in the training set (A) and the test set (B); calibration curves of models in predicting MPR in the training set (A) and the test set (B).

set, respectively (Supplemental Digital Content Table S7, available at: <http://links.lww.com/JS9/F149>, Supplemental

Digital Content Fig. S1A, available at: <http://links.lww.com/JS9/F149>). Calibration curve and DCA also showed good

Table 3
Predictive performance of the established models in the test set

Model	AUC	ACC	Sensitivity	Specificity	NPV	PPV	F1
Clinical model	0.738	0.723	0.667	0.750	0.828	0.556	0.606
Rad model	0.681	0.359	0.700	0.667	0.688	0.815	0.500
Rad_per model	0.702	0.66	0.733	0.625	0.833	0.478	0.579
Hab model	0.756	0.745	0.733	0.75	0.857	0.579	0.647
Hab_per model	0.769	0.766	0.667	0.812	0.839	0.625	0.645
Combined model	0.915	0.872	0.733	0.938	0.882	0.846	0.786

Rad model, intratumoral radiomics model; Rad_per model, peritumoral radiomics model; Hab model, intratumoral habitat radiomics model; Hab_per model, peritumoral habitat radiomics model; ACC, accuracy; AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value; F1, F1 score.

calibration and clinical application performance of the model (Supplemental Digital Content Fig. S1B-D, available at: <http://links.lww.com/JS9/F149>).

Model explanation

To address the problem of the black box of the machine learning model, we performed SHAP analysis on the combined model to visualize the contribution of each variable to the output of the model. The SHAP summary bar plot (Fig. 4A) offers a global explanation of the combined model, evaluating the contribution of each feature to the model using the mean absolute SHAP values, displayed in descending order: post-NACI tumor length, hab_per_sign, hab_sign, and post-NACI T stage. The SHAP beeswarm plot (Fig. 5B) visually displays the direction and intensity of each feature’s influence on the model’s prediction. For example, shorter post-NACI tumor length, higher hab_per_sign, and higher hab_sign significantly increased the model’s output probability, that is, the probability of MPR.

The SHAP waterfall plot (Fig. 5) enables the analysis of individual samples at the local level and explains how the model predicts whether a patient has achieved MPR after NACI. The plot details each feature’s specific value and its corresponding SHAP value, visually reflecting the impact of features on the prediction result in both positive and negative directions. Figure 5D demonstrates the prediction process for a patient achieving MPR. Figure 5H demonstrates the

prediction process for a patient with non-MPR. The waterfall plot vividly reconstructs the prediction process for a specific patient, providing intuitive and powerful support for a deeper understanding of the model’s decision-making mechanism.

Web-based calculator

The combined model was integrated into a web application to facilitate its clinical use. By inputting the actual values of the four features required by the model, the application automatically predicts the probability of MPR in ESCC patients after NACI treatment. The web application can be accessed online via the following link: <https://predict13.shinyapps.io/predict/>.

Discussion

NACI regimens consisting of camrelizumab plus paclitaxel and nedaplatin offer significant advantages over conventional neoadjuvant regimens, including higher chances of achieving radical resection and improved prognosis, making them a promising treatment for ESCC^[4]. Tumor heterogeneity leads to highly variable responses to treatment among patients. Accurate prediction of pathological response after NACI in ESCC patients is crucial, as it enables clinicians to adjust subsequent treatment plans in a timely manner and avoid overtreatment. In this study, we developed a combined model integrating clinical variables and intratumoral and peritumoral habitat features, which demonstrated excellent performance in predicting MPR after NACI in ESCC patients, with AUCs of 0.987 in the

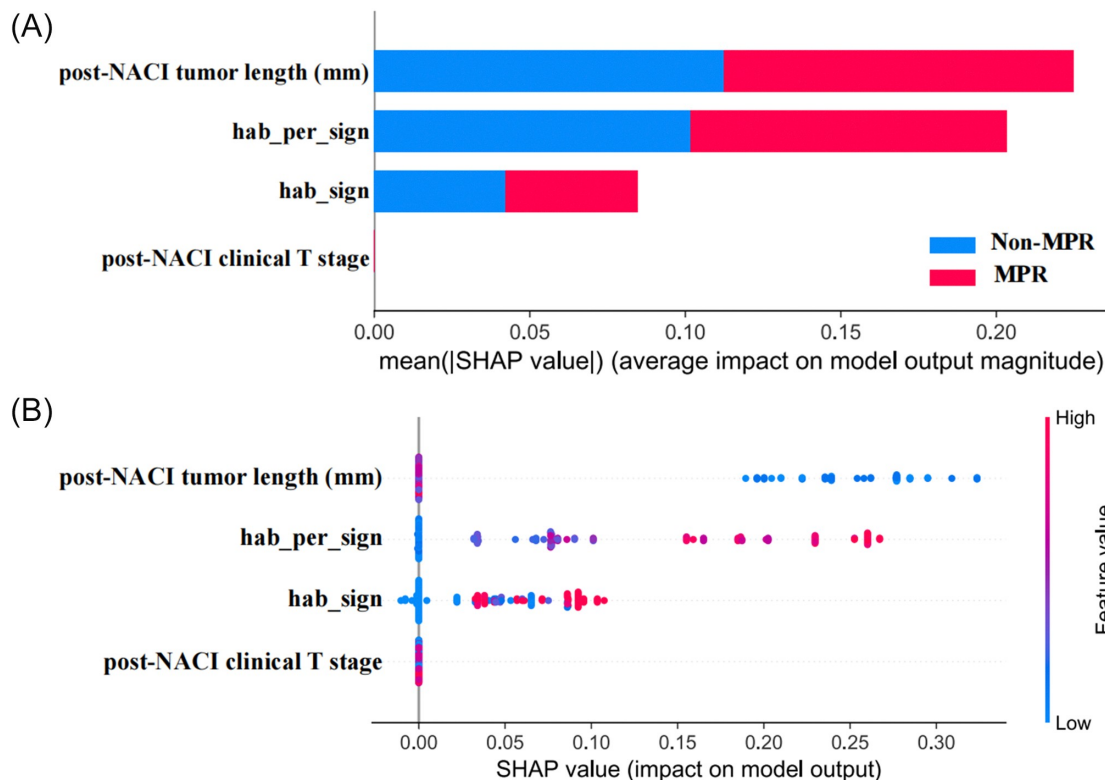


Figure 4. Global model explanation by the SHapley Additive exPlanations (SHAP) analysis method. (A) SHAP summary bar plot. (B) The SHAP beeswarm plot.

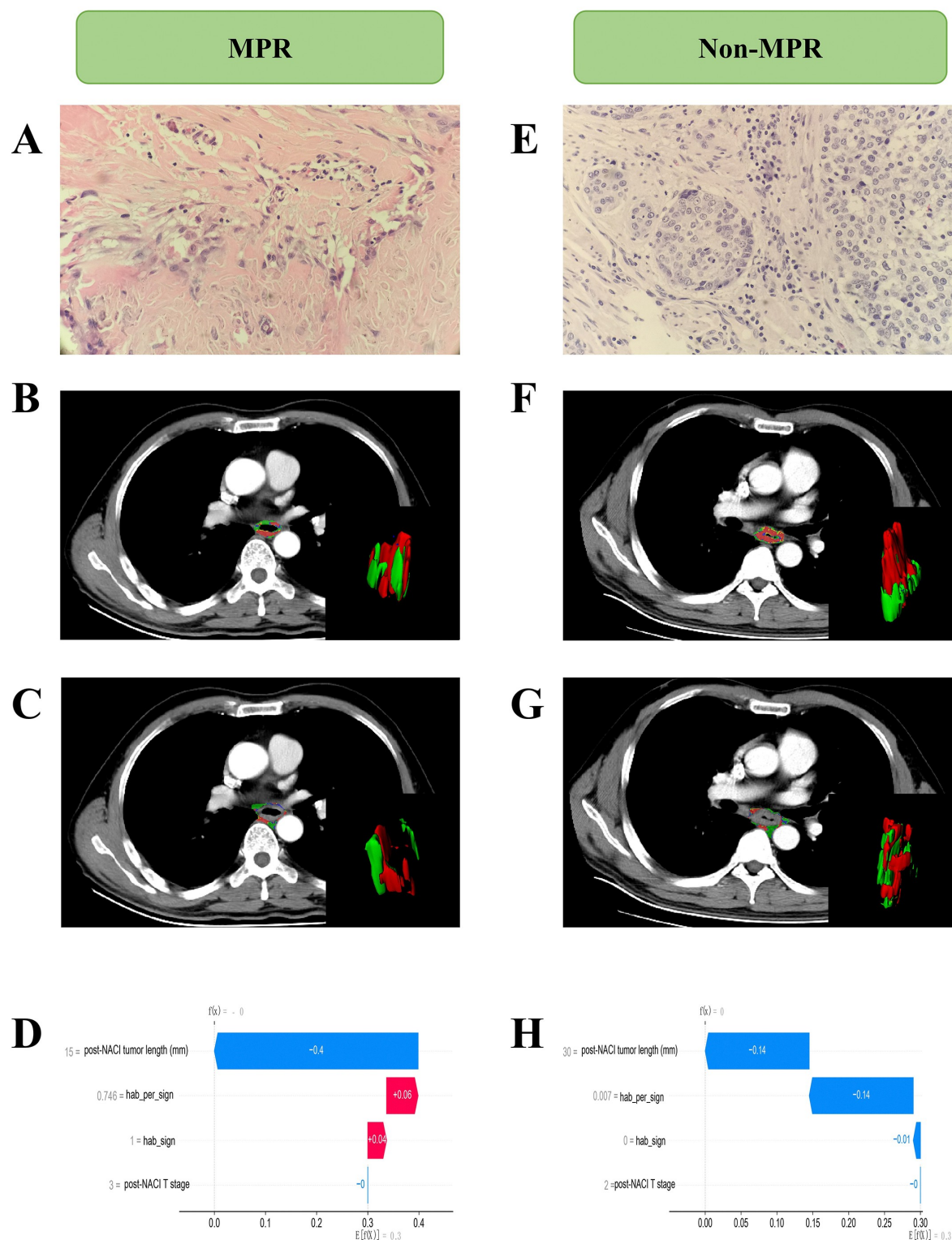


Figure 5. Visualizes representative cases correctly predicted as MPR and Non-MPR. (A) and (E) show tumor tissue sections for the MPR and Non-MPR patients, respectively. (B) and (F) illustrate examples of intratumoral habitat visualization. (C) and (G) provide examples of peritumoral habitat visualization. (D) and (H) present SHAP waterfall plots, reflecting the contribution of relevant clinical and radiomic features to the predicted outcomes.

training set and 0.915 in the test set. Furthermore, sensitivity analysis showed that the combined model also effectively predicted pCR after NACI in ESCC patients.

In this study, we included 157 ESCC patients who underwent NACI, with 38.2% achieving MPR and 25.5% achieving pCR, which aligns with the probabilities observed in real-world clinical

settings. Previous studies have investigated the clinicopathologic factors associated with treatment response to NACI in ESCC patients and reported that pCR was related to tumor thickness, length, and area^[20,21]. In our study, we identified post-NACI tumor length and post-NACI T stage as independent predictors of MPR. These findings suggest that tumor characteristics are important indicators of pathological response, providing valuable insights for guiding clinical decisions and treatment planning.

Several previous studies have used radiomics model to predict pathological responses after neoadjuvant therapy in ESCC patients. Wang *et al*^[22] developed a prediction model that integrates clinical data and radiomics features to predict pCR in ESCC patients undergoing neoadjuvant immunotherapy, with an AUC of 0.84. Zheng *et al*^[23] developed CT-based delta-radiomics to predict pathological response to neoadjuvant immunochemotherapy in ESCC, achieving AUCs of 0.847 and 0.719 in the training and test sets, respectively. However, these studies primarily relied on traditional radiomics, which assume that tumors are homogeneous or uniformly heterogeneous, thus overlooking the impact of intratumoral heterogeneity on treatment response. Habitat radiomics has shown promise in characterizing tumor heterogeneity and predicting pathological responses to neoadjuvant therapy in various cancers, including breast cancer^[24], nonsmall cell lung cancer^[25], serous ovarian cancer^[26], and nasopharyngeal carcinoma^[27].

Previous studies have primarily focused on predicting pCR. Notably, our prior research found that the pCR rate after NACI in ESCC patients was relatively low at 27.4%, while the probability of achieving MPR was 45.2%^[4]. Furthermore, MPR has been shown to serve as an independent prognostic factor and a surrogate for survival endpoints in ESCC patients^[8,28]. Therefore, we chose MPR as the primary predictive endpoint in this study. In this study, we used habitat radiomics to analyze the heterogeneity of both intratumoral and peritumoral regions. The results demonstrated that both the intratumoral and peritumoral habitat models exhibited good predictive performance for MPR, with AUCs of 0.756 and 0.769, respectively. When intratumoral and peritumoral habitat radiomics features were combined with clinical variables, the performance of the combined model improved significantly, achieving an AUC of 0.915. This finding highlights the importance of peritumoral tissue in predicting pathological responses, suggesting that peritumoral radiomic features may offer valuable insights for optimizing treatment response assessment and individualized treatment planning.

In ESCC patients after NACI treatment, pCR and MPR are clinically correlated but do not fully overlap. This suggests that although both MPR and pCR are closely related to treatment response, they represent distinct clinical endpoints. Sensitivity analysis further revealed that the combined prediction model demonstrated strong performance in predicting pCR with high accuracy. This finding underscores the potential of the combined model in assessing treatment response and guiding clinical decisions, offering more precise treatment strategies, particularly for personalized treatment planning.

Machine learning models face limitations in clinical applications due to the opacity of their decision-making processes^[29]. To improve the interpretability and transparency of our model, we performed SHAP analysis to assess the contribution of each feature to the predicted outcome, thereby enhancing clinicians' confidence in the model. Additionally, we integrated the model into a user-friendly, web-based system, providing clinicians with a convenient and efficient tool to assess a patient's response to

treatment. Given the wide range of machine learning algorithms available for radiomics modeling, selecting the appropriate algorithm is crucial for model development. In our study, we employed 14 mainstream machine learning algorithms, which helps mitigate the risk of overfitting or underfitting and improves the robustness and reliability of our results.

This study has several limitations. First, we included patients from a single center. Future studies will aim to complete multi-center validation. However, the inclusion of a prospective cohort enhances the reliability and validity of our findings. Second, male patients were disproportionately represented in this study. While it is well-established that around 70% of esophageal cancer cases occur in men, the higher proportion of male patients in our study may be attributed to the greater incidence of esophageal cancer in men at our research center. Third, tumor and peritumoral segmentation relied on manual delineation by experienced radiologists, which is time-consuming and labor-intensive. While this presents a challenge, ongoing research into automated volume segmentation techniques offers the potential to streamline the process and reduce human error.

Conclusion

In conclusion, we established and evaluated a model that combines clinical variables and intratumoral and peritumoral habitat features to accurately predict pathological response in patients with ESCC receiving NACI. This approach has the potential to aid clinical decision-making in ESCC patients, prevent overtreatment, and facilitate personalized, precise cancer treatment.

Ethical approval

This study was approved by the Medical Ethics Committee of the Affiliated Tumor Hospital of Harbin Medical University (KY2024-104).

Consent

Written informed consent was obtained for the prospective cohort, while it was waived for the retrospective cohort.

Source of funding

This work was supported by National Natural Science Foundation of China [grant number 82371922].

Author contributions

X.L.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, writing—original draft, writing—review and editing. X. Y.: conceptualization, data curation, formal analysis, methodology, validation, visualization, writing—original draft, writing—review and editing. P.W.: data curation, formal analysis, investigation, methodology, writing—review and editing. Y.L.: data curation, investigation, software, writing—review and editing. Z. W.: data curation, investigation, software, writing—review and editing. J.W.: conceptualization, data curation, writing—review and editing. K.C.: data curation, writing—review and editing. Y. Y.: data curation, writing—review and editing. A.L.: data curation, writing—review and editing. J.M.: conceptualization, data curation, investigation, project administration, resources, supervision, writing—review and editing. W.M.: conceptualization,

data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualization, writing–review and editing. All authors have read and agreed to publish the version of the manuscript.

Conflicts of interest disclosure

The authors declare no competing interests.

Guarantor

The guarantor of this publication is Wei Meng.

Research registration unique identifying number (UIN)

This study has been registered in the Chinese Clinical Trial Registry, ID: ChiCTR2000033761 (<https://www.chictr.org.cn/showproj.html?proj=55022>).

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

Data are available from the corresponding author upon reasonable request.

Acknowledgements

We sincerely appreciate the efforts and contributions of every team member in this study. We also extend our gratitude to all the patients and their families for their support of this research.

Assistance with the study

None.

Presentation

None.

References

- [1] Sung H, Ferlay J, Siegel RL, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209–49.
- [2] Morgan E, Soerjomataram I, Rumgay H, *et al.* The global landscape of esophageal squamous cell carcinoma and esophageal adenocarcinoma incidence and mortality in 2020 and projections to 2040: new estimates from GLOBOCAN 2020. *Gastroenterology* 2022;163:649–58.e2.
- [3] Thrift AP. Global burden and epidemiology of Barrett oesophagus and oesophageal cancer. *Nat Rev Gastroenterol Hepatol* 2021;18:432–43.
- [4] Yang Y, Zhang J, Meng H, *et al.* Neoadjuvant camrelizumab combined with paclitaxel and nedaplatin for locally advanced esophageal squamous cell carcinoma: a single-arm phase 2 study (cohort study). *Int J Surg* 2024;110:1430–40.
- [5] Wu Z, Zheng Q, Chen H, *et al.* Efficacy and safety of neoadjuvant chemotherapy and immunotherapy in locally resectable advanced esophageal squamous cell carcinoma. *J Thorac Dis* 2021;13:3518–28.
- [6] Blum Murphy M, Xiao L, Patel VR, *et al.* Pathological complete response in patients with esophageal cancer after the trimodality approach: the association with baseline variables and survival-The University of Texas MD Anderson Cancer Center experience. *Cancer* 2017;123:4106–13.
- [7] Nie R, Chen F, Provencio M, *et al.* Predictive value of radiological response, pathological response and relapse-free survival for overall survival in neoadjuvant immunotherapy trials: pooled analysis of 29 clinical trials. *Eur J Cancer* 2023;186:211–21.
- [8] Yang Y, Liu J, Liu Z, *et al.* Two-year outcomes of clinical N2-3 esophageal squamous cell carcinoma after neoadjuvant chemotherapy and immunotherapy from the phase 2 NICE study. *J Thorac Cardiovasc Surg* 2024;167:838–847.e1.
- [9] Derclé L, Sun S, Seban RD, *et al.* Emerging and evolving concepts in cancer immunotherapy imaging. *Radiology* 2023;306:e239003.
- [10] Wang JL, Tang LS, Zhong X, *et al.* A machine learning radiomics based on enhanced computed tomography to predict neoadjuvant immunotherapy for resectable esophageal squamous cell carcinoma. *Front Immunol* 2024;15:1405146.
- [11] Sun X, Feng Y, Zhang B, *et al.* The role of neutrophil-to-lymphocyte ratio in predicting pathological response for resectable non-small cell lung cancer treated with neoadjuvant chemotherapy combined with PD-1 checkpoint inhibitors. *Cancer Res Treat* 2022;54:1017–29.
- [12] Yang Y, Xin D, Wang H, *et al.* A novel predictor of pathological complete response for neoadjuvant immunotherapy in resectable locally advanced esophageal squamous cell carcinoma. *J Inflamm Res* 2023;16:1443–55.
- [13] Lambin P, Leijenaar RTH, Deist TM, *et al.* Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749–62.
- [14] Zhang L, Wang Y, Peng Z, *et al.* The progress of multimodal imaging combination and subregion based radiomics research of cancers. *Int J Biol Sci* 2022;18:3458–69.
- [15] El-Sayes N, Vito A, Mossman K. Tumor heterogeneity: a great barrier in the age of cancer immunotherapy. *Cancers (Basel)* 2021;13:806.
- [16] Zhang X, Su GH, Chen Y, Gu YJ, You C. Decoding intratumoral heterogeneity: clinical potential of habitat imaging based on radiomics. *Radiology* 2023;309:e232047.
- [17] Cuocolo R, Caruso M, Perillo T, Ugga L, Petretta M. Machine learning in oncology: a clinical appraisal. *Cancer Lett* 2020;481:55–62.
- [18] Kong W, Xu J, Huang Y, *et al.* CT-based habitat radiomics for predicting treatment response to neoadjuvant chemioimmunotherapy in esophageal cancer patients. *Front Oncol* 2024;14:1418252.
- [19] Hu Y, Xie C, Yang H, *et al.* Assessment of intratumoral and peritumoral computed tomography radiomics for predicting pathological complete response to neoadjuvant chemoradiation in patients with esophageal squamous cell carcinoma. *JAMA Netw Open* 2020;3:e2015927.
- [20] Qi WX, Li S, Xiao J, Li H, Chen J, Zhao S. A machine learning approach using (18)F-FDG PET and enhanced CT scan-based radiomics combined with clinical model to predict pathological complete response in ESCC patients after neoadjuvant chemoradiotherapy and anti-PD-1 inhibitors. *Front Immunol* 2024;15:1351750.
- [21] Ruan Y, Ma Y, Ma M, *et al.* Dynamic radiological features predict pathological response after neoadjuvant immunotherapy in esophageal squamous cell carcinoma. *J Transl Med* 2024;22:471.
- [22] Wang X, Gong G, Sun Q, Meng X. Prediction of pCR based on clinical-radiomic model in patients with locally advanced ESCC treated with neoadjuvant immunotherapy plus chemoradiotherapy. *Front Oncol* 2024;14:1350914.
- [23] Zheng Y, Mei P, Wang M, *et al.* CT-based delta-radiomics for predicting pathological response to neoadjuvant immunotherapy in esophageal squamous cell carcinoma: a multicenter study. *BMC Med Imaging* 2024;24:329.
- [24] Shi Z, Huang X, Cheng Z, *et al.* MRI-based quantification of intratumoral heterogeneity for predicting treatment response to neoadjuvant chemotherapy in breast cancer. *Radiology* 2023;308:e222830.
- [25] Caii W, Wu X, Guo K, Chen Y, Shi Y, Chen J. Integration of deep learning and habitat radiomics for predicting the response to immunotherapy in NSCLC patients. *Cancer Immunol Immunother* 2024;73:153.
- [26] Bi Q, Miao K, Xu N, *et al.* Habitat radiomics based on MRI for predicting platinum resistance in patients with high-grade serous ovarian carcinoma: a multicenter study. *Acad Radiol* 2024;31:2367–80.
- [27] Zhu Y, Zheng D, Xu S, *et al.* Intratumoral habitat radiomics based on magnetic resonance imaging for preoperative prediction treatment response to neoadjuvant chemotherapy in nasopharyngeal carcinoma. *Jpn J Radiol* 2024;42:1413–24.
- [28] Hong Z, Xie S, Xu H, *et al.* Major pathologic response as a prognostic surrogate in esophageal squamous cell carcinoma patients receiving neoadjuvant chemotherapy/chemoimmunotherapy: a multi-center cohort study. *Eur J Surg Oncol* 2025;51:109500.
- [29] Adlung L, Cohen Y, Mor U, Elinav E. Machine learning in clinical decision making. *Med* 2021;2:642–65.