

# Supplementary material

## Study design and patients

This study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) reporting guidelines. To construct a predictive model of TLSs, we retrospectively reviewed data for 611 patients with gastric cancer who underwent surgery at Xijing Hospital (Xi'an, China), one of the largest centers for gastric cancer in China, between August 1, 2015, and June 30, 2018. The inclusion criteria were histologically confirmed diagnosis of gastric adenocarcinoma; pathological stage II–III (according to the 8<sup>th</sup> Edition of the American Joint Committee on Cancer TNM Staging Manual); curative resection (R0); available preoperative contrast-enhanced CT scans of the abdomen; and complete clinicopathological and follow-up data. The exclusion criteria included prior partial gastrectomy, other synchronous or metachronous malignant tumors, or previous neoadjuvant chemotherapy. We randomly allocated the 611 patients with gastric adenocarcinoma into the training cohort ( $n = 304$ ) or validation cohort 1 ( $n = 307$ ) according to a random assignment function. This study was approved by the Ethical Committee of Xijing Hospital (KY20222241-X-1).

To rigorously evaluate the predictive and prognostic performance of the TLS classification model, we enrolled 393 patients who received preoperative neoadjuvant therapy at 2 clinical centers. The inclusion criteria were (1) histologically confirmed gastric adenocarcinoma; (2) clinical stage II–III, determined through contrast-enhanced CT; (3) availability of pre-treatment contrast-enhanced abdominal CT performed before neoadjuvant therapy; and (4) complete clinicopathological and follow-up data. The exclusion criteria included prior partial gastrectomy and the presence of synchronous or metachronous malignancies. The internal evaluation cohorts came from the Xijing Hospital, where 202 patients received neoadjuvant chemotherapy, and 101 patients underwent neoadjuvant immunotherapy combined with chemotherapy. The external evaluation cohort comprised 90 additional patients treated with neoadjuvant chemotherapy at the First Hospital of China Medical University in Shenyang, China. This multi-center design enabled comprehensive validation of the model across distinct therapeutic settings.

## CT acquisition and image processing

All patients in the internal training and validation cohorts underwent contrast-enhanced abdominal CT scans before radical surgery. All patients in the model evaluation cohorts

underwent contrast-enhanced abdominal CT scans before neoadjuvant therapy, to rule out the effect of neoadjuvant therapy on the formation of TLSs.

CT images at the portal venous phase were retrieved for training the deep learning model, because of their superior contrast. The primary tumor was outlined on these CT scans by 2 collaborating radiologists, YJ and CL, in ITK-SNAP software (version 3.4.0). Both specialists worked together and achieved a consensus in defining the tumor boundaries.

## Development of an imaging model to assess the TLS classifier

### Data processing

To mitigate potential inter-center variability in CT acquisition protocols, we resampled CT scans to a uniform in-plane resolution of  $1 \text{ mm}^2$  per pixel with linear interpolation. Voxel intensities were clipped to the range  $[-1,024, 2,048]$  and linearly normalized to  $[0,1]$ . Fixed-size  $50 \times 50$  pixel patches centered on predefined tumor seed points were then extracted for deep feature learning. Radiomic features were extracted from segmented tumor regions and subsequently harmonized with the ComBat method to adjust for potential center-related batch effects. Harmonization was applied after feature extraction and before model training.

### Deep feature extraction

Deep imaging features were extracted with FMCIB, a large-scale foundation model pretrained through self-supervised learning on 11,467 radiographic lesions encompassing diverse anatomical sites<sup>1</sup>. This model captures comprehensive visual semantics relevant to lesion morphology, texture, and contextual tissue patterns. For each patient, preprocessed CT image patches centered on the primary tumor were input into FMCIB to generate high-dimensional feature embeddings, which served as deep representations of the tumor phenotype and microenvironmental characteristics. These features were subsequently integrated with the handcrafted radiomic features in the predictive modeling framework. Importantly, the pretrained FMCIB foundation model was frozen during training to decrease the number of trainable parameters and improve model stability.

## Radiomic feature extraction

With *PyRadiomics*<sup>2</sup>, 525 radiomic features were extracted from the tumor masks, including 9 shape, 18 first-order, 22 GLCM, 16 GLRLM, 16 GLSZM, 14 GLDM, 86 LBP, and 344 wavelet features. To decrease annotation costs, we calculated features from the slice with the largest tumor area.

## Feature fusion and prediction

The deep and radiomic features were standardized *via* Z score normalization and concatenated into a 4,621-dimensional vector. This fused representation was processed by a regression head consisting of (i) a fully connected layer (256 units, ReLU, dropout  $P = 0.5$ ), (ii) a second fully connected layer (128 units, ReLU, dropout), and (iii) a final linear layer outputting a scalar prediction.

## Training

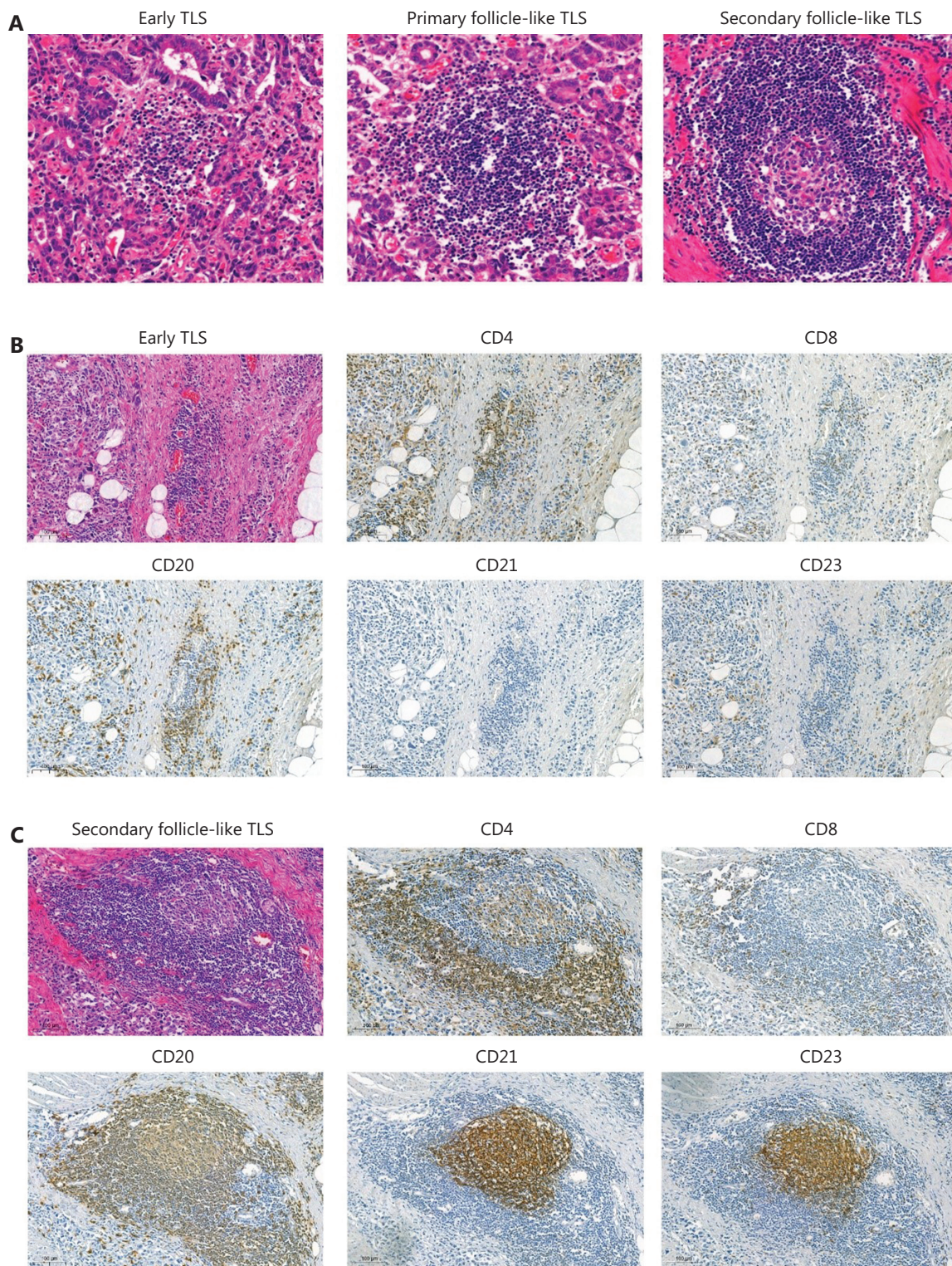
Models were trained with a batch size of 8 and an initial learning rate of  $1 \times 10^{-5}$  with exponential decay and were optimized with Adam. Data augmentation included reflections and random rotations between  $-15^\circ$  and  $+15^\circ$ .

## Statistical analysis

Statistical analysis was performed between January 1, 2025, and June 30, 2025. Categorical variables were compared with the  $\chi^2$  test (or Fisher's exact test for groups with small sample sizes), and continuous variables were compared with Student's  $t$  test, as appropriate. Survival curves were generated with the Kaplan–Meier method and compared with the log-rank test. Multiplicative interactions between the TLS classifier results and the response to adjuvant chemotherapy were assessed with a Cox regression model. To explore the relationship between the TLS classifier results and clinical outcomes, we performed univariate and multivariable analyses with the Cox proportional hazards model. A two-tailed  $P < 0.05$  was considered to indicate statistical significance. Statistical analyses were performed in R software (version 4.1.1) and SPSS software (version 24.0).

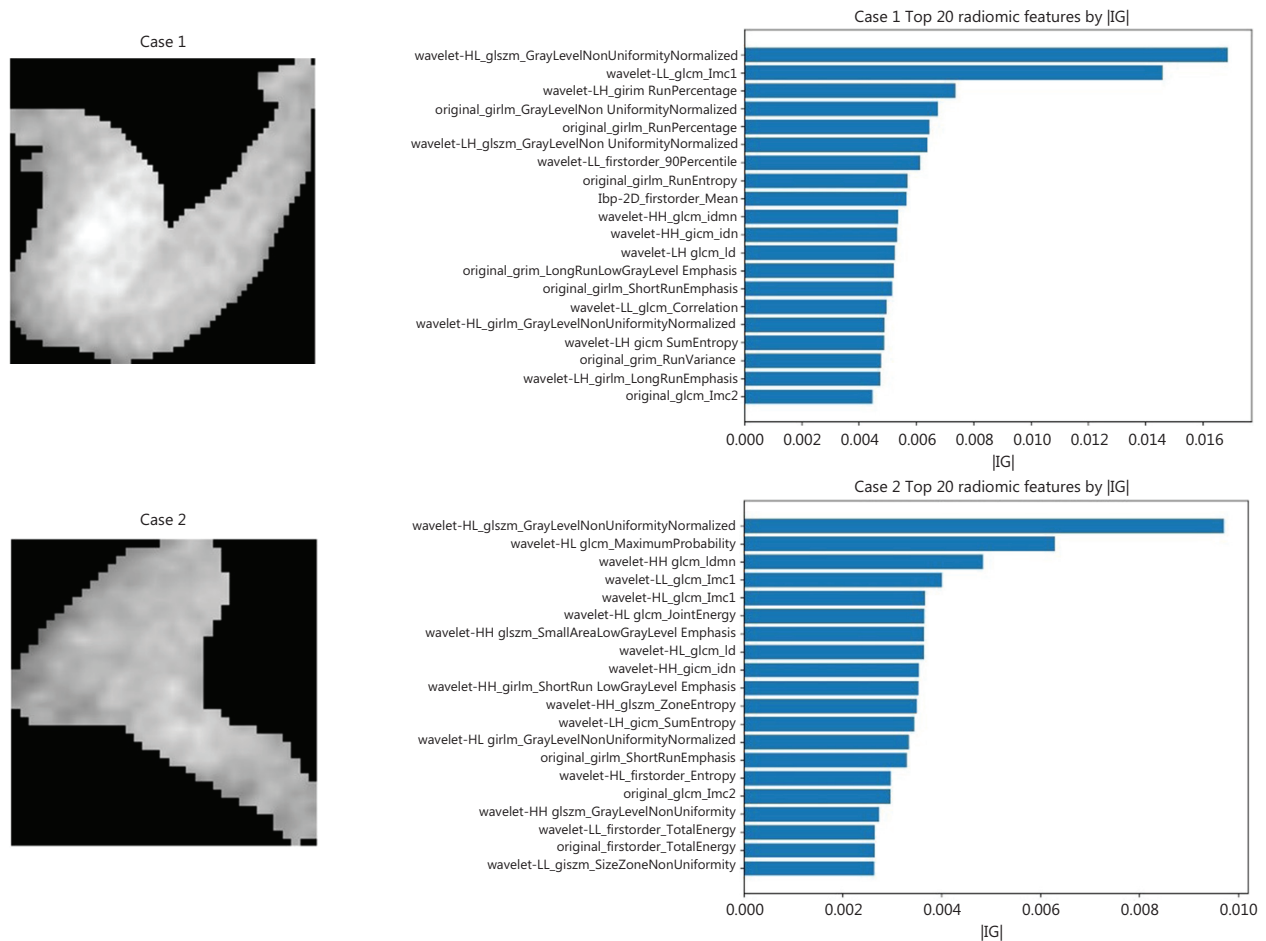
## References

1. Pai S, Bontempi D, Hadzic I, Prudente V, Sokac M, Chaunzwa TL, et al. Foundation model for cancer imaging biomarkers. *Nat Mach Intell.* 2024; 6: 354-67.
2. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017; 77: e104-e7.

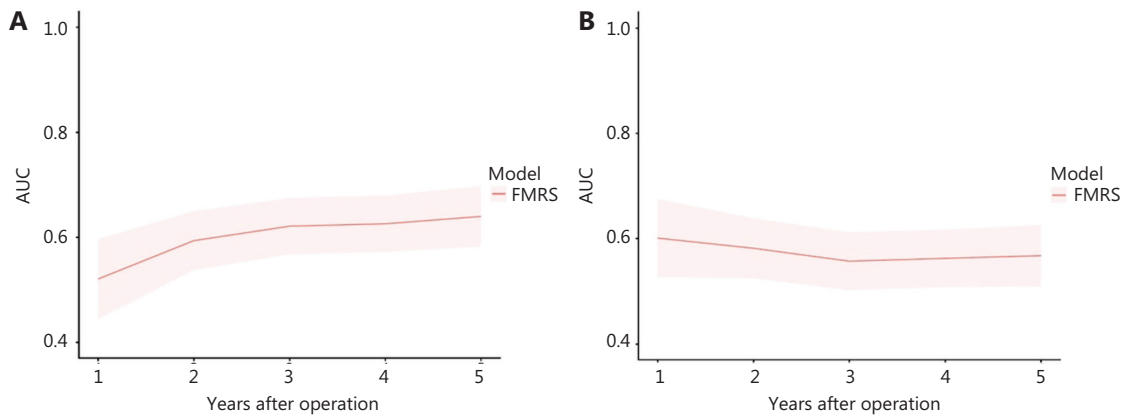


**Figure S1** Tertiary lymphoid structures (TLSs) in gastric cancer. (A) Representative hematoxylin–eosin-stained images of TLSs. TLSs were categorized into 3 structural types: (i) aggregates (Agg) (early TLS), which represent loose and poorly organized lymphocyte clusters; (ii) primary follicle-like (PFL) structures, which are compact, round clusters lacking germinal centers; and (iii) secondary follicle-like (SFL) structures, which are well-formed follicles containing germinal centers. In further assessment of maturation status, tumors were considered immature if only

Agg TLSs or no TLSs were present, or were considered mature if PFL and/or SFL TLSs were detected. (B) Representative immunohistochemistry images showing CD4, CD8, CD20, CD21, and CD23 staining ( $\times 100$  magnification). TLSs were subsequently classified according to B-cell organization and the presence of follicular dendritic cell (FDC) networks into 3 categories: (i) Agg, consisting of scattered CD20<sup>+</sup> lymphocyte clusters without discernible FDC structures; (ii) PFL structures, consisting of compact CD20<sup>+</sup> B-cell clusters associated with CD21<sup>+</sup> FDC networks but lacking CD23 expression, indicating the absence of germinal centers; and (iii) SFL structures, characterized by well-organized CD20<sup>+</sup> B-cell follicles containing both CD21<sup>+</sup> and CD23<sup>+</sup> FDCs, consistently with germinal center formation.



**Figure S2** Feature attribution analysis. Top 20 radiomic features contributing to model predictions identified with Integrated Gradients (IG) in 2 representative cases. Bar plots show the absolute attribution values (|IG|) indicating the relative importance of individual radiomic features. Representative tumor regions from CT images are also shown.



**Figure S3** Time-dependent ROC analysis of FMRS. Time-dependent AUC curves of FMRS for predicting postoperative overall survival at 1–5 years in the training cohort (A) and internal validation cohort (B).

**Table S1** Association of the foundation-model-enhanced CT radiomics signature (FMRS) with overall survival (OS) in the training cohort

Variables	Univariate		Multivariate	P value
	HR (95% CI)	P value	HR (95% CI)	
<b>Sex</b>				
Male	1			
Female	0.985 (0.691–1.403)	0.931		
<b>Age, years</b>				
<60	1			
≥60	1.107 (0.812–1.508)	0.520		
<b>Primary tumor location</b>				
Proximal	1			
Body	0.917 (0.611–1.377)	0.676		
Antrum	0.805 (0.567–1.143)	0.226		
<b>CEA, ng/mL</b>				
<5	1		1	
≥5	1.732 (1.233–2.433)	0.002	1.690 (1.190–2.400)	0.003
<b>CA 19-9, U/mL</b>				
<37	1		1	
≥37	1.454 (1.038–2.037)	0.029	1.061 (0.746–1.508)	0.742
<b>Pathologic TNM stage</b>				
pStage II	1		1	
pStage III	3.306 (2.107–5.186)	<0.001	3.341 (2.123–5.258)	<0.001
<b>Adjuvant chemotherapy</b>				
No	1		1	
Yes	0.47 (0.328–0.675)	<0.001	0.505 (0.351–0.728)	<0.001

**Table S1** Continued

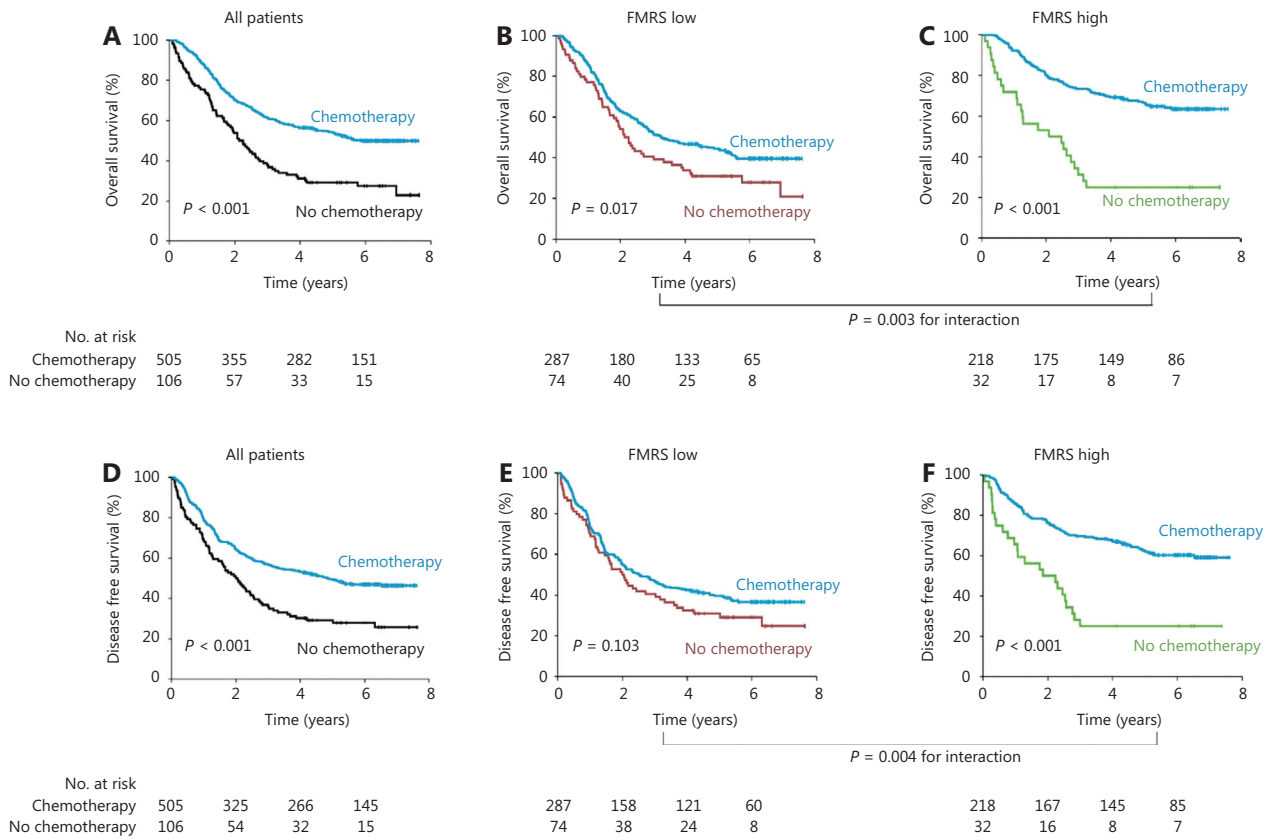
Variables	Univariate		Multivariate	P value
	HR (95% CI)	P value	HR (95% CI)	
FMRS class				
Low	1		1	
High	0.475 (0.339–0.666)	<0.001	0.512 (0.364–0.721)	<0.001

CI, confidence interval; FMRS, foundation-model-enhanced CT radiomics signature; HR, hazard ratio.

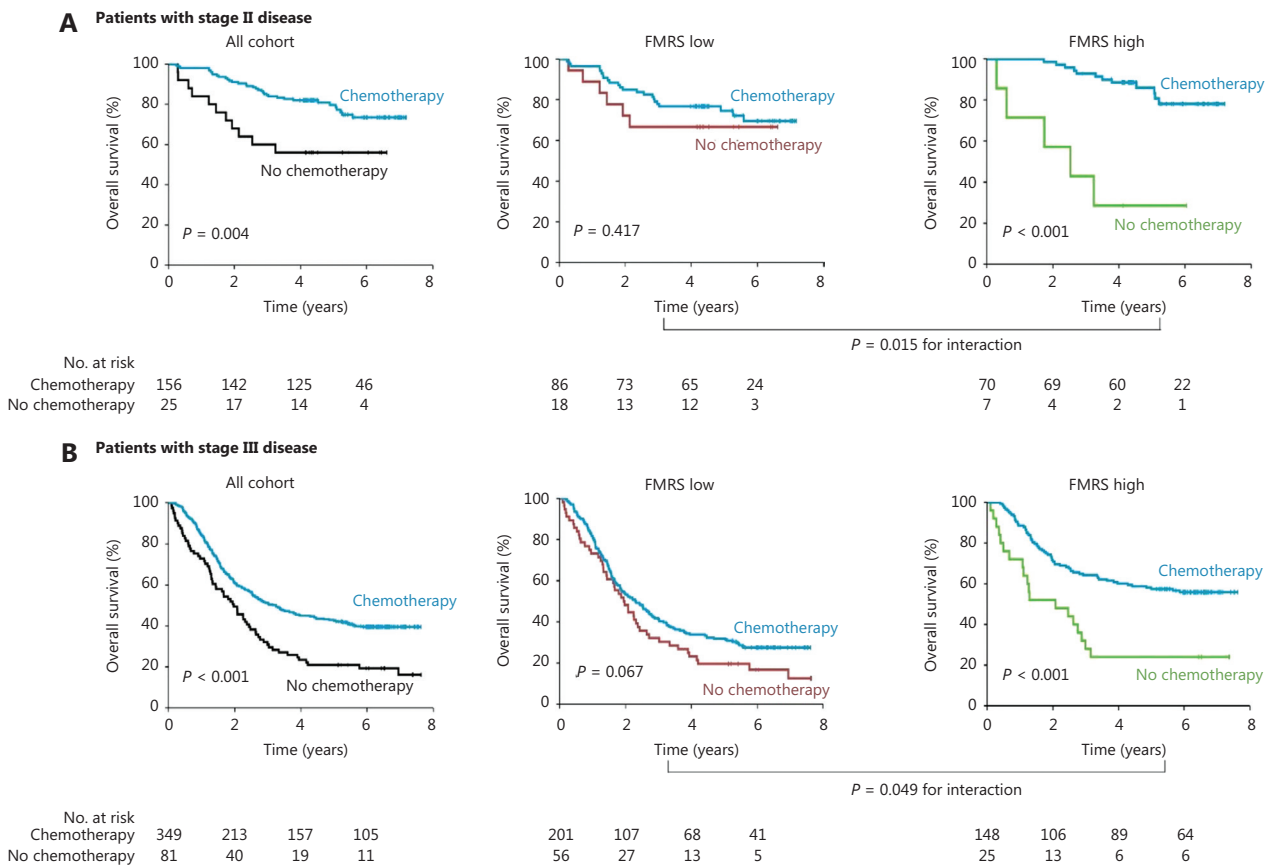
**Table S2** Association of the foundation-model-enhanced CT radiomics signature (FMRS) with overall survival (OS) in validation cohort

Variables	Univariate		Multivariate	P value
	HR (95% CI)	P value	HR (95% CI)	
Sex				
Male	1			
Female	0.921 (0.616–1.378)	0.690		
Age, years				
<60	1			
≥60	1.219 (0.891–1.668)	0.216		
Primary tumor location				
Proximal	1			
Body	0.777 (0.516–1.169)	0.226		
Antrum	0.682 (0.478–0.972)	0.035		
CEA, ng/mL				
<5	1		1	
≥5	1.648 (1.14–2.382)	0.008	1.669 (1.140–2.443)	0.008
CA 19-9, U/mL				
<37	1		1	
≥37	1.432 (1.011–2.027)	0.043	1.143 (0.793–1.649)	0.473
Pathologic TNM stage				
pStage II	1		1	
pStage III	3.398 (2.199–5.25)	<0.001	3.441 (2.219–5.334)	<0.001
Adjuvant chemotherapy				
No	1		1	
Yes	0.534 (0.37–0.769)	0.001	0.666 (0.46–0.966)	0.032
FMRS class				
Low	1		1	
High	0.629 (0.451–0.876)	0.006	0.699 (0.495–0.988)	0.043

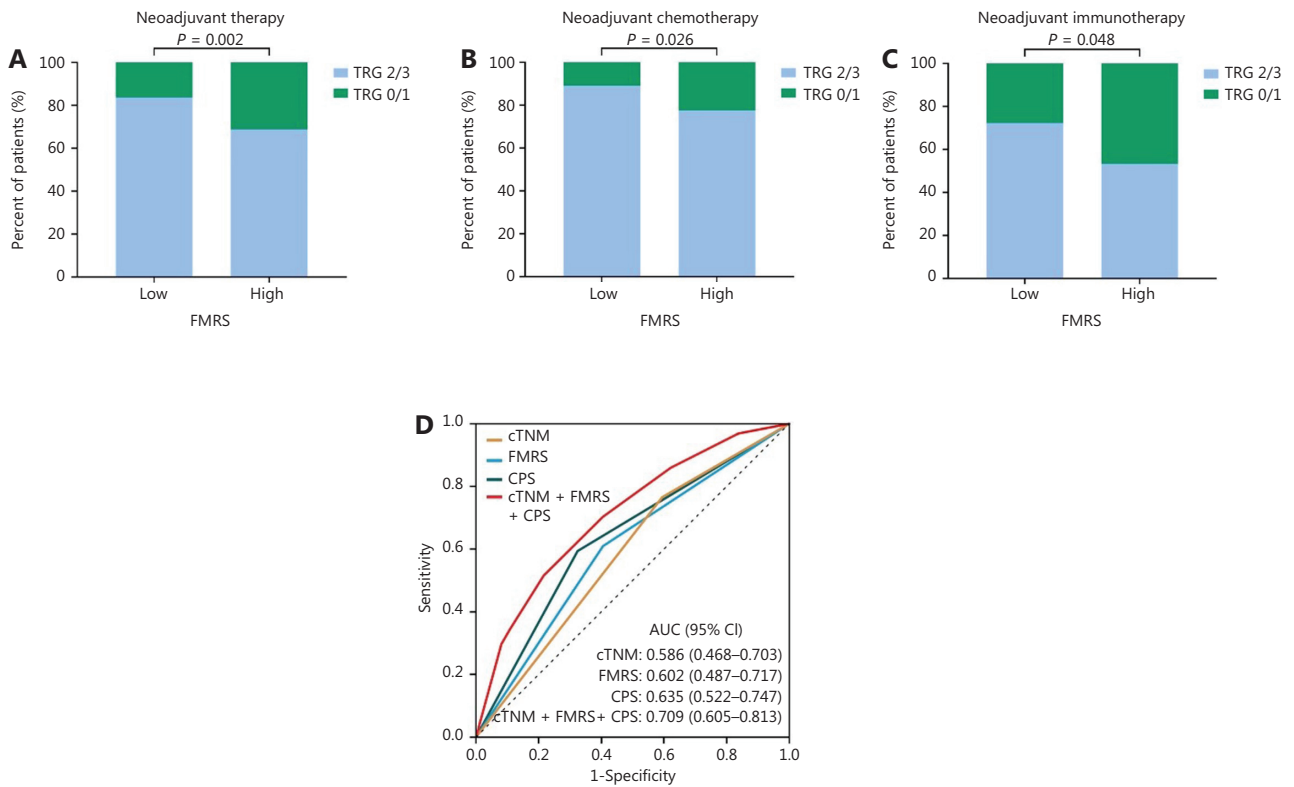
CI, confidence interval; FMRS, foundation-model-enhanced CT radiomics signature; HR, hazard ratio.



**Figure S4** Association between FMRS status and survival benefit from adjuvant chemotherapy. (A–C) Kaplan–Meier curves for overall survival (OS), stratified by adjuvant chemotherapy in the overall cohort (A), low-FMRS group (B), and high-FMRS group (C). (D–F) Kaplan–Meier curves for disease-free survival (DFS), stratified by adjuvant chemotherapy in the overall cohort (D), low-FMRS group (E), and high-FMRS group (F). Interaction analyses demonstrated that FMRS status significantly modified the survival benefit associated with adjuvant chemotherapy. Patients in the high-FMRS group derived greater survival benefit from adjuvant chemotherapy, whereas those in the low-FMRS group showed limited benefit. DFS, disease-free survival; FMRS, foundation-model-enhanced CT radiomics signature; OS, overall survival.



**Figure S5** Association between FMRS status and survival benefit from adjuvant chemotherapy, stratified by pathologic stage. (A) Kaplan–Meier curves for overall survival (OS) according to adjuvant chemotherapy in patients with stage II gastric cancer. (B) Kaplan–Meier curves for OS according to adjuvant chemotherapy in patients with stage III gastric cancer. Stage-stratified analyses demonstrated consistent chemotherapy benefit and a significant interaction between FMRS status and treatment. FMRS, foundation-model-enhanced CT radiomics signature; OS, overall survival.



**Figure S6** Association between FMRS status and response to neoadjuvant therapy. (A) Comparison of pathologic response rates (TRG 0/1) between FMRS-high and FMRS-low groups in the overall internal neoadjuvant therapy cohort. (B) Pathologic response rates, stratified by FMRS status in patients receiving neoadjuvant chemotherapy alone. (C) Pathologic response rates, stratified by FMRS status in patients receiving neoadjuvant immunotherapy plus chemotherapy. (D) Receiver operating characteristic (ROC) curves comparing the predictive performance of FMRS, PD-L1 combined positive score (CPS), clinical TNM (cTNM) stage, and their composite models for predicting pathological response to neoadjuvant immunotherapy. The combined cTNM + CPS + FMRS model demonstrated more favorable discrimination than individual predictors. AUC, area under the curve; CPS, combined positive score; cTNM, clinical tumor–node–metastasis stage; FMRS, foundation-model-enhanced CT radiomics signature; TRG, tumor regression grade; ROC, receiver operating characteristic.