BMC Medicine

RESEARCH

Open Access

Radiomics-based ensemble model predicts postoperative recurrence of gastric cancer

Ping'an Ding[1,2,3†], Sheng Chen[1,2,3,4†], Honghai Guo[1,2,3†], Jiaxuan Yang[1,2,3†], Qiang Fu[5], Baoquan Wang[5], Renjun Gu[6,7], Lilong Zhang[8], Jiaxiang Wu[1,2,3], Haotian Wu[1,2,3], Ning Meng[9], Xiaolong Li[10], Zhenjiang Guo[11], Yuan Tian[1,2,3], Lingjiao Meng[12*†] and Qun Zhao[1,2,3*†]

Abstract

Background Early postoperative recurrence in locally advanced gastric cancer (LAGC) severely compromises patient outcomes, yet current predictive models are inadequate due to limited generalizability and insufficient validation. This study aimed to establish and robustly validate a multimodal Radiomics-Clinical Integrated Risk Stratification Assessment (RSA) model to accurately predict early recurrence.

Methods We retrospectively analyzed 2,516 LAGC patients from six hospitals across northern and southern China, divided into training, internal validation, and external validation cohorts. Radiomics features were extracted from portal venous-phase abdominal CT images using nnU-Net-based segmentation, followed by a hierarchical feature-selection framework integrating the mRMR algorithm and LASSO regression. A predictive RSA model combining clinical factors and radiomics features was developed. Model robustness was validated internally, externally, prospectively (clinical trial cohort, NCT01516944, n = 569), and further independently tested using a publicly available dataset from The Cancer Imaging Archive (TCIA, n = 41).

Results The RSA model exhibited superior performance across all validation cohorts, achieving AUC values of 0.873 (training), 0.871–0.872 (internal validation), 0.870–0.873 (external validation), 0.857 (prospective validation), and 0.850 (TCIA validation). Decision curve analyses confirmed the RSA model provided significant clinical benefits beyond traditional models. Transcriptomic analyses revealed that low-risk patients exhibited enhanced immune infiltration and significant activation of immune-related pathways, including IL6/JAK/STAT3 and interferon signaling. Multivariate Cox regression demonstrated the RSA model independently predicted five-year overall survival across validation cohorts (HR range: 1.830–2.166, all $P < 0.001$), surpassing standard clinical staging.

Conclusions Our RSA model, integrating robust radiomics methodologies and critical clinical parameters, consistently demonstrated high accuracy and clinical applicability in diverse populations, significantly improving prediction

†Ping'an Ding, Sheng Chen, Honghai Guo and Jiaxuan Yang contributed equally to this work.

†Lingjiao Meng and Qun Zhao contributed equally as co-corresponding authors.

*Correspondence:
Lingjiao Meng
ljmeng@hebmu.edu.cn
Qun Zhao
zhaoqun@hebmu.edu.cn
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

of early postoperative recurrence in LAGC. This approach provides novel insights into tumor-immune microenvironment interactions, paving the way toward personalized postoperative management strategies.

Keywords Radiomics, Gastric Cancer, Early Recurrence, Multimodal Prediction, Tumor Immune Microenvironment

Background

Gastric cancer remains one of the most lethal malignancies worldwide, especially in East Asia, where nearly half of the global cases occur [1, 2]. Locally advanced gastric cancer (LAGC) poses significant clinical challenges due to its aggressive nature, frequent recurrence, and poor prognosis despite standard surgical and adjuvant therapies [3–5]. Postoperative recurrence within the early follow-up period profoundly impacts patient survival, emphasizing the critical need for precise predictive tools that surpass traditional clinical and pathological assessments [6, 7]. While current staging systems, such as the TNM classification, offer foundational prognostic guidance, their reliance solely on clinicopathological features limits their ability to capture the complex biological heterogeneity inherent in gastric cancer, underscoring the urgent necessity for integrating multidimensional biomarkers [8–10].

Emerging evidence suggests that radiomics, which extracts high-dimensional quantitative features from medical imaging, provides promising opportunities to noninvasively decode tumor biology and predict clinical outcomes [11–15]. Recent advances in artificial intelligence (AI) and machine learning methodologies have significantly expanded the potential of radiomics, enabling more nuanced and robust predictive modeling [16–18]. However, despite its promise, the clinical translation of radiomics-based predictive models remains hindered by inconsistent feature selection, lack of reproducibility across different populations, and poor interpretability, which limit their widespread adoption in clinical practice [19, 20]. Moreover, existing radiomics-based predictive models for gastric cancer recurrence frequently lack rigorous validation, particularly across diverse geographic regions and clinical scenarios, further highlighting the imperative for robustly validated multimodal approaches.

Motivated by these unmet clinical needs and methodological gaps, we designed this study to develop and validate an integrative predictive framework for early postoperative recurrence in LAGC, harnessing advanced radiomics, rigorous feature selection algorithms, and comprehensive multimodal data fusion approaches. Our primary objective was to establish a robust and clinically actionable Radiomics-Clinical Integrated Risk Stratification Assessment (RSA) model capable of accurately predicting early recurrence, thereby facilitating precise patient stratification and personalized clinical decision-making. Importantly, our study leverages a large-scale multicenter cohort encompassing geographically diverse patient populations from both northern and southern China, supplemented by rigorous internal and external validations, including an additional prospective multicenter clinical trial dataset. Furthermore, we conducted comprehensive transcriptomic profiling to elucidate the biological underpinnings of radiomic signatures, providing mechanistic insights into immune microenvironment characteristics associated with recurrence risks.

Thus, this study stands apart from previous literature by combining methodological rigor, robust multicenter validation, interpretability, and biological relevance, addressing critical limitations of existing prognostic tools in gastric cancer management. Our findings not only promise substantial improvements in predictive accuracy but also provide novel insights into the tumor immune microenvironment, paving the way toward precision oncology approaches that may transform postoperative care and clinical outcomes in patients with LAGC.

Methods

Study cohort

This study leveraged a multicenter, prospectively curated database established by the Hebei Gastric Cancer Collaborative Network (HGCCN), encompassing gastric cancer cases reported from major healthcare institutions in Hebei Province, China (http://hbss.suvalue.com/). We retrospectively analyzed a cohort of patients with LAGC diagnosed between 2012 and 2019, drawn from four hospitals in Hebei [the Fourth Hospital of Hebei Medical University (FHHMU), Shijiazhuang People's Hospital (SJZPH), Baoding Central Hospital (BDCH), and Hengshui People's Hospital (HSPH)] and two academic centers in Southern China [Wuhan University People's Hospital (WHPH) and Nanjing Jinling Hospital (NJJLH)]. After applying rigorous selection criteria outlined in the Additional file 3: Supplementary Methods, 2,516 eligible cases were included in the analysis.

Cohorts for model training and validation were delineated according to institution and treatment period. The training dataset consisted of 926 patients treated at FHHMU from 2014 to 2017. The internal validation cohort, also from FHHMU, was divided into two temporal subsets: internal set I (2012–2013; n = 231) and internal set II (2018–2019; n = 206). For external validation,

patients from the other five hospitals were grouped by geographic region: external validation set I included 314 patients from SJZPH, BDCH, and HSPH (Northern China), and external validation set II comprised 229 cases from WHPH and NJJLH (Southern China).

To further assess clinical applicability, we incorporated an independent cohort of 569 LAGC patients from a registered prospective trial (NCT01516944), comprising both surgical and neoadjuvant chemotherapy cases. In addition, we collected fresh tumor specimens from 114 patients for transcriptomic profiling to investigate the model's biological underpinnings. Technical details on RNA sequencing and downstream analysis are available in the Additional file 3: Supplementary Methods

This investigation was conducted in accordance with the ethical standards of the Declaration of Helsinki, with approval from local ethics committees. While informed consent was waived for retrospective clinical data use, all participants contributing biospecimens for RNA sequencing provided written informed consent.

Clinical outcomes

This study adhered to postoperative treatment principles outlined in national and international guidelines, specifically those from the Chinese Society of Clinical Oncology [21] and the National Comprehensive Cancer Network [22]. Patients with satisfactory physical performance were advised to receive adjuvant chemotherapy regimens incorporating 5-fluorouracil. Postoperative surveillance followed a predefined schedule, starting with clinical evaluations such as history-taking and physical examination, and incorporating laboratory testing, diagnostic imaging, or endoscopy based on individual clinical needs.

The follow-up intervals varied with time since surgery, occurring every 3 to 6 months during the initial two years, then every 6 months in the third through fifth years, and annually thereafter. Early recurrence was defined as tumor relapse identified within 24 months after curative-intent gastrectomy for LAGC, based on radiological evidence (e.g., ultrasound, CT, or PET/CT), clinical deterioration, or findings at reoperation [23–25]. Overall survival (OS) was calculated from the date of surgery to death from any cause. In this analysis, early recurrence was the primary clinical endpoint, with OS considered as a secondary measure. The final follow-up was conducted in June 2024, and the median follow-up period was 66 months (range: 55–162 months).

Image acquisition and segmentation

All patients underwent contrast-enhanced abdominal CT scans within two weeks prior to surgery, with portal venous phase images retrieved from the picture archiving and communication system (PACS). Automated segmentation was performed using nnU-Net, a deep-learning-based model designed for efficient segmentation through automatic configuration of preprocessing, network architecture, and training parameters according to dataset properties. In total, 200 CT scans with corresponding manual segmentation labels were included, divided into a training set ($n=160$) and a test set ($n=40$). The segmentation pipeline consisted of: (1) resampling images to a voxel size of 1.25 mm × 0.93 mm × 0.93 mm^3; (2) CT intensity normalization using clipping at the 0.5–99.5 percentile and Z-score standardization; and (3) training a six-level 3D U-Net with channel dimensions ranging from 32 to 320, employing Dice and cross-entropy loss functions (Adam optimizer, learning rate $=3\times10^{-4}$, batch size $=2$) over 1,000 epochs. Data augmentation and five-fold cross-validation were also conducted. The entire training and inference processes were implemented using the PyTorch framework on a Windows operating system, achieving a Dice coefficient of 0.890. The nnU-Net model was subsequently used to automatically segment the remaining patients' CT images, thereby generating three-dimensional regions of interest (3D-ROI). Detailed CT acquisition parameters for each participating center, including scanner models, contrast protocols, and reconstruction settings, are provided in the Additional file 3: Supplementary Methods.

Radiomics feature selection

Radiomics features were extracted using 3D Slicer software. To standardize CT images from different scanners, all scans were resampled to isotropic voxels of 1 × 1 × 1 mm, and voxel intensity discretization was set to a fixed bin width of 25. Initially, a total of 1,130 radiomics features were extracted from each patient's ROI. Prior to feature selection, the ComBat harmonization algorithm was applied to pool and harmonize imaging data from multiple institutions, followed by Z-score normalization to eliminate dimensional differences between features. To evaluate the effectiveness of ComBat harmonization, we performed principal component analysis (PCA) before and after correction. PCA plots based on the first two principal components were generated to assess clustering by imaging center and dataset. As shown in Additional file 1: Fig. S1, center-specific clustering was substantially reduced following correction, indicating improved feature alignment.

A hierarchical feature-selection framework was developed to efficiently reduce dimensionality and retain prognostically relevant information. Initially, 400 candidate features were selected using the minimum redundancy maximum relevance (mRMR) algorithm to optimize information gain while minimizing redundancy (Additional file 1: Fig. S2A). Subsequently, candidate features

underwent normality and homogeneity of variance assessments, followed by statistical significance testing (two-sample t-test or Mann–Whitney U test, two-tailed, significance threshold $p < 0.05$), resulting in 329 statistically significant features. A Pearson correlation matrix was then constructed, and highly collinear features (correlation coefficient > 0.9) were excluded, leaving 249 non-redundant features. Finally, the least absolute shrinkage and selection operator (LASSO) regression with tenfold cross-validation was applied, and an optimal regularization parameter (λ) was identified using the minimum standard error criterion, yielding eight features with non-zero coefficients for further analysis (Additional file 1: Fig. S2B-C).

Development and validation of the radiomics-driven ensemble model

An ensemble learning strategy was employed to predict early postoperative recurrence in patients with LAGC. Multiple machine learning algorithms including logistic regression, decision trees, ridge regression, support vector machines, and a voting regressor were integrated. The voting regressor served as the meta-estimator, combining the predictions of the base models by averaging their outputs to construct a robust radiomics-based predictive model. To further evaluate the risk of overfitting commonly associated with high-dimensional radiomics features, we conducted two additional analyses to assess model complexity and generalizability. First, a learning curve was generated by progressively increasing the size of the training set and plotting AUC values for both the training and validation cohorts. As shown in Additional file 1: Fig. S3, the training AUC demonstrated a gradual decline with increasing data, while the validation AUC remained stable around 0.81–0.82, suggesting limited overfitting and good model generalization. Second, we examined the relationship between model performance and the number of selected radiomic features. Using a stepwise approach, features were ranked by importance, and models were incrementally built with increasing feature subset sizes. As illustrated in Additional file 1: Fig. S4, model AUC improved initially and then plateaued after approximately 8 features, indicating that a small number of informative features contributed most to predictive performance. These findings confirm that our feature selection strategy effectively balanced model complexity and performance while minimizing overfitting risk.

Development and evaluation of the multimodal RSA model

To enhance predictive performance, we developed a multimodal RSA model by integrating clinical factors with radiomic features. Univariate logistic regression was first used to assess the association between clinical variables and early recurrence in the training cohort. Subsequently, a multivariate logistic regression model was constructed using independent clinical predictors of recurrence. These clinical variables were then combined with radiomic features to build a nomogram-based RSA model.

Risk scores generated by the model were dichotomized into high-risk and low-risk groups, using the Youden index to determine the optimal cutoff that maximized sensitivity and specificity. Model performance was evaluated using multiple metrics, including the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR), negative likelihood ratio (NLR), and F1-score.

To quantify the incremental value of the RSA model over clinical-only or radiomics-only models, we calculated the net reclassification improvement (NRI) and integrated discrimination improvement (IDI). Differences in AUC values between models were statistically compared using the DeLong test. Calibration curves were plotted to assess the agreement between predicted and observed recurrence rates, and predictive performance was further evaluated using prediction error curves and the Brier score.

Biological characteristics and immune infiltration

To explore the biological basis of the radiomic features, RNA sequencing was performed on 114 tumor tissue samples from the study cohort. Functional enrichment analysis was conducted to identify the biological roles of relevant features. P-values were adjusted using the Benjamini–Hochberg method to control the false discovery rate (FDR), and gene sets with FDR-adjusted p-values < 0.05 were considered statistically significant. Immune cell composition and tumor ecosystem structure were inferred using CIBERSORTx and the EcoTyper framework, which enabled characterization of cell types, cell states, and multicellular communities. For immune score and cell composition comparisons between groups, Wilcoxon rank-sum tests were performed, and p-values were corrected using the Benjamini–Hochberg method. Immune features with FDR-adjusted p-values < 0.05 were considered statistically significant.

To support external validation, an additional 38 transcriptomic samples with matched CT imaging from the TCIA (derived from TCGA-STAD) were incorporated. While limited in scale, these publicly available samples enabled cross-cohort consistency checks and multimodal integration. Detailed methods for RNA processing, sequencing, and bioinformatics analysis are provided in the Additional file 3: Supplementary Methods [26–35].

Statistical analysis

All statistical analyses were performed using SPSS version 28.0 (IBM) and R version 4.3.3 (http://www.r-project.org). Continuous variables were analyzed using unpaired two-tailed t-tests or Mann–Whitney U tests, as appropriate. Categorical variables were compared using chi-square or Fisher's exact tests. Cox proportional hazards regression models (both univariate and multivariate) were applied to evaluate the prognostic significance of variables on survival outcomes. A two-sided p-value of less than 0.05 was considered statistically significant. Detailed statistical algorithms and procedures are described in the Additional file 3: Supplementary Methods.

Results

Patient characteristics

A total of 1,906 patients with LAGC were included, distributed across the training ($n=926$), internal validation ($n=437$), and external validation ($n=543$) cohorts. Additional file 2: Table. S1 presents the detailed baseline characteristics. Male patients accounted for 65.8% of the total cohort, and 53.6% were older than 65 years. More than three-quarters (75.7%) had tumors classified as pT4, while 72.2% exhibited lymph node metastasis. The distribution of ECOG performance status, tumor location, and histological differentiation was comparable among cohorts, with no substantial imbalance observed. Systemic inflammation and nutritional indices were stratified by cohort using median cut-offs, and their distributions were generally consistent. Early postoperative recurrence occurred in 52.1%, 54.0%, and 55.8% of patients in the training, internal, and external validation cohorts, respectively. A comparison of clinicopathological characteristics between patients with and without recurrence in each cohort is provided in Additional file 2: Table. S2.

Development and validation of the radiomics-based ensemble model

An ensemble model based on radiomic features was constructed to predict early recurrence following curative surgery in LAGC patients (Fig. 1B). This integrative approach outperformed each single machine learning classifier in the training dataset, achieving an area under the ROC curve (AUC) of 0.837 (95% CI: 0.812–0.863), as presented in Additional file 1: Fig. S5. The model also yielded robust sensitivity (0.774) and specificity (0.736), detailed in Additional file 2: Table. S3.

Following successful development in the training set, the ensemble model maintained high predictive accuracy in the internal validation sets, with AUCs of 0.826 (95% CI: 0.774–0.878) for validation set I and 0.827 (95% CI: 0.772–0.880) for validation set II (Additional file 1: Fig. S6 and Additional file 2: Table. S3). Moreover, the model consistently outperformed individual algorithms in both external validation cohorts, achieving an AUC of 0.827 (95% CI: 0.780–0.874) in external validation set I and 0.820 (95% CI: 0.766–0.874) in external validation set II (Additional file 1: Fig. S7 and Additional file 2: Table. S3).

Development and validation of the RSA prediction model

Multivariate logistic regression identified pathological type, pT stage, pN stage, and radiomic features as independent predictors of early recurrence (Additional file 2: Table. S4). These variables were incorporated into a combined nomogram, referred to as the Radiomics based RSA model, which integrates clinical and radiomics data (Fig. 2A). The RSA model demonstrated strong predictive power, with an AUC of 0.873 (95% CI: 0.850–0.895) in the training cohort (Figs. 2B–E; Additional file 2: Table. S5).

Compared with single-modality models, the RSA model provided significant incremental value. It showed improved reclassification and discrimination over the clinical model [NRI = 0.416 (95%CI:0.298–0.534), IDI = 0.230 (95%CI:0.204–0.258); both $P<0.001$) and the radiomics model (Additional file 2: Table. S6). DeLong test further confirmed the superiority of the RSA model (AUC: 0.873 vs. 0.751, $P<0.001$). The model demonstrated excellent calibration (Brier score = 0.144; Fig. 2F) and greater net clinical benefit in decision curve analysis (Fig. 2G).

(See figure on next page.)

Fig. 1 Schematic Overview of Study Design and Workflow. **A** Radiomics analysis pipeline: abdominal CT images were segmented using the nnU-Net algorithm, followed by extraction of radiomic features and construction of radiomic signatures. **B** Development and validation of the RSA: an integrated predictive model incorporating radiomic and clinicopathological features was established to assess early recurrence risk in LAGC patients. **C** Bioinformatics analysis: transcriptomic profiling of 38 TCIA patients and 114 prospectively collected tumor samples was performed to explore the biological underpinnings and immune infiltration patterns associated with radiomic features and RSA-defined risk groups. **D** Patient cohort distribution: a total of 2,516 LAGC patients with preoperative CT imaging were included from six medical centers and public databases. Contributing institutions: Center 1 – The Fourth Hospital of Hebei Medical University; Center 2 – Shijiazhuang People's Hospital; Center 3 – Baoding First Central Hospital; Center 4 – Hengshui People's Hospital; Center 5 – Nanjing Jinling Hospital; Center 6 – Renmin Hospital of Wuhan University

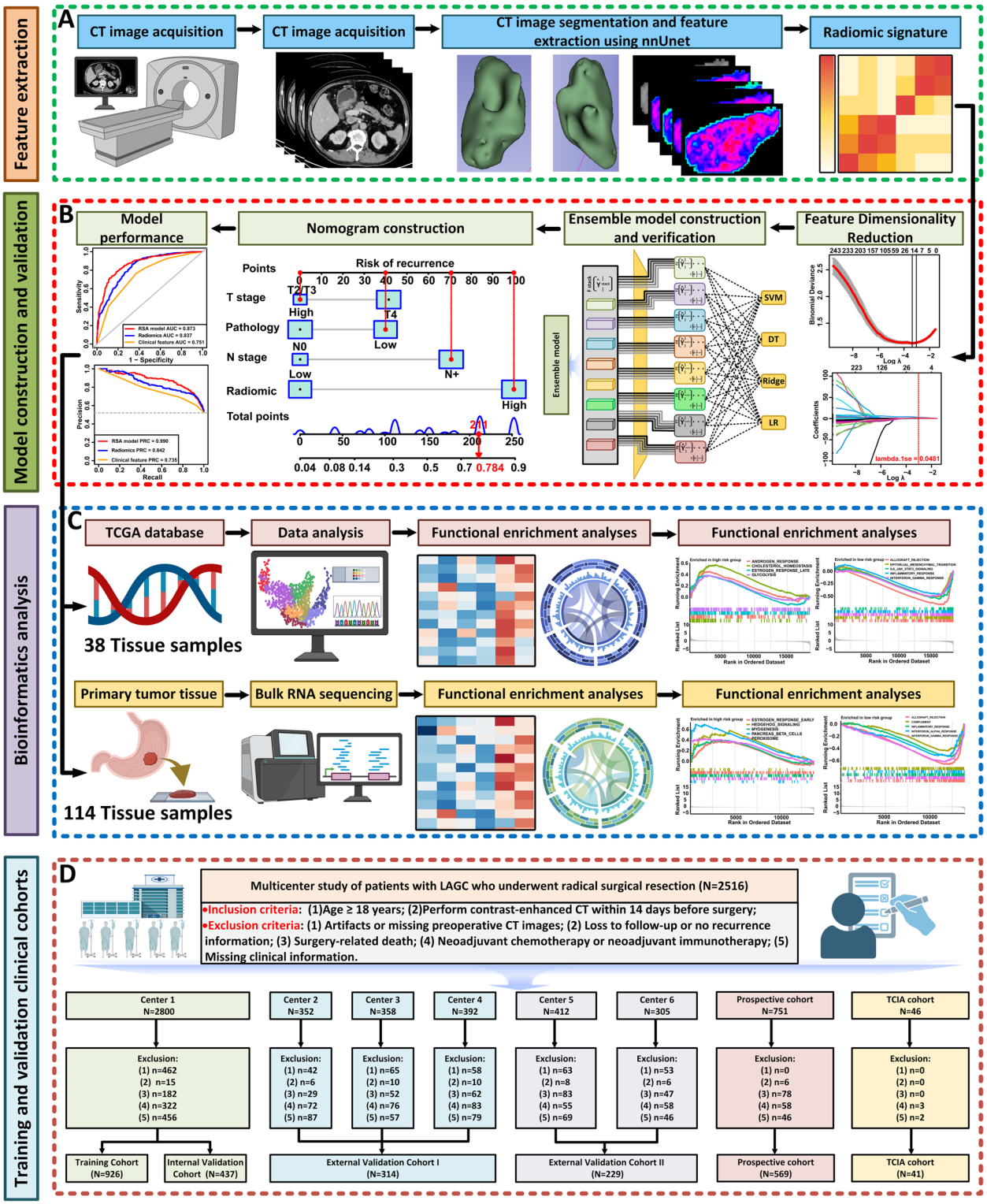


Fig. 1 (See legend on previous page.)

Patients were stratified into high- and low-risk groups using the Youden index. The RSA model identified more recurrence cases in the high-risk group than the clinical model (40.8% vs. 39.8%; Fig. 2H), and five-year OS was significantly lower in high-risk patients (43.4% vs. 57.4%; $P<0.001$; F ig. 2I). Multivariate Cox regression confirmed the RSA model as an independent prognostic factor for OS (HR = 1.416, 95% CI: 1.011–1.983; $P=0.043$; Additional file 2: Table. S7).

Subgroup analyses based on blood tumor markers and molecular features (HER2 and PD-L1 expression) also supported the RSA model's superior performance in recurrence prediction (Figs. 2J–K, Additional file 1: Fig. S8-9, Additional file 2: Table. S8). Across all subgroups, high-risk patients had significantly worse five-year OS (Additional file 1: Fig. S10). Ten-fold cross-validation and bootstrap resampling (1,000 iterations) confirmed the model's robustness, with a consistently higher AUC (mean = 0.8724, 95%CI: 0.8717–0.8732) than clinical-only (0.7501) and radiomics-only models (0.8374) (Figs. 2L–M).

Internal cohort-based evaluation of the radiomics-clinical RSA model

To further validate the RSA model, we retrospectively analyzed 437 LAGC patients treated at FHHMU. Based on treatment periods, these patients were categorized into two internal validation subsets: set I (January 2012–December 2013, $n=231$) and set II (January 2018–December 2019, $n=206$) (Fig. 1D). Clinical profiles were well-balanced between the two groups (all $P>0.05$; Additional file 2: Table. S9).

In validation set I, the RSA model demonstrated clear advantages over the clinical-only and radiomics-only models, with an AUC of 0.871 (95% CI: 0.826–0.915; Figs. 3A–C, 3E; Additional file 2: Table. S10). The superiority was statistically supported by the DeLong test ($P<0.001$; Additional file 2: Table. S6). Compared with the clinical model, RSA yielded significant gains in reclassification and discrimination metrics (NRI = 0.559, 95% CI: 0.275–0.832; IDI = 0.232, 95% CI: 0.177–0.285; both $P<0.001$), along with good calibration (Brier score = 0.149; Fig. 3D). Similar model performance was observed in validation set II (F igs. 3I–M; Additional file 2: Table. S10).

The two-layer concentric circle plots illustrated a greater proportion of early recurrence cases correctly identified as high risk by the RSA model compared to the clinical model (set I: 43.7% vs. 37.7%, Fig. 3G; set II: 44.2% vs. 39.3%, Fig. 3N). Decision curve analysis also demonstrated that the RSA model provided higher net clinical benefit across both internal subsets (Figs. 3F and 3O). Sensitivity analysis further showed that the RSA model captured a larger percentage of high-risk recurrence cases (82.7% vs. 69.0%), while reducing low-risk misclassification by 8.9% (Fig. 3Q).

Kaplan–Meier survival analysis revealed significantly better five-year OS in the RSA-defined low-risk group compared to high-risk patients (set I: 62.3% vs. 44.0%, $P=0.002$; set II: 52.5% vs. 36.2%, $P=0.005$; Figs. 3H and 3P). Moreover, multivariate Cox regression confirmed that the RSA classification remained an independent prognostic factor for five-year OS in both validation sets (set I: HR = 2.166, 95% CI: 1.417–3.312, $P<0.001$; set II: HR = 1.830, 95% CI: 1.258–2.660, $P=0.002$; Additional file 2: Table. S11-12).

External validation of the RSA model across multicenter cohorts

To further assess the robustness and generalizability of the RSA model, we conducted validation in two independent multicenter cohorts. External validation set I included 314 LAGC patients treated at SJZPH, BDCH, and HSPH (all in northern China), while set II comprised 229 patients from WHPH and NJJLH in the south (Fig. 1D). The baseline characteristics for both cohorts are summarized in Additional file 2: Table. S13.

In validation set I, the RSA model achieved strong predictive performance for early recurrence, with an AUC of 0.870 (95% CI: 0.831–0.911; Figs. 4A–C, 4E; Additional file 2: Table. S14). Compared to the clinical model, RSA yielded significant gains in both net reclassification

(See figure on next page.)

Fig. 2 Construction and Performance Evaluation of the RSA Model in the Training Dataset. **A** The integrative nomogram was established by incorporating selected radiomic and clinicopathological predictors to estimate early recurrence risk in LAGC. **B**–**C** Discrimination performance of all candidate models was assessed using ROC and PR curves. **D** Comparative summary of key performance metrics across models is visualized through a radar diagram. **E** Confusion matrices depict classification performance for each model within the training set. **F** Calibration assessment of the RSA model indicates strong concordance between predicted probabilities and actual outcomes. **G** Clinical benefit analysis was conducted using decision curve analysis for each model. **H** Concentric circle visualization highlights relative clinical applicability of competing models. **I** Kaplan–Meier curves based on RSA-defined risk stratification demonstrate significant survival differences; threshold derived from Youden index. **J–K** Stratified performance evaluations across patient subgroups defined by serum tumor markers and HER2 status, presented via bar and radar plots. **L** Boxplots display tenfold cross-validation results for model metrics including AUC, sensitivity, specificity, and accuracy. **M** Model robustness evaluated through bootstrap-based ROC curve comparisons from 1,000 iterations

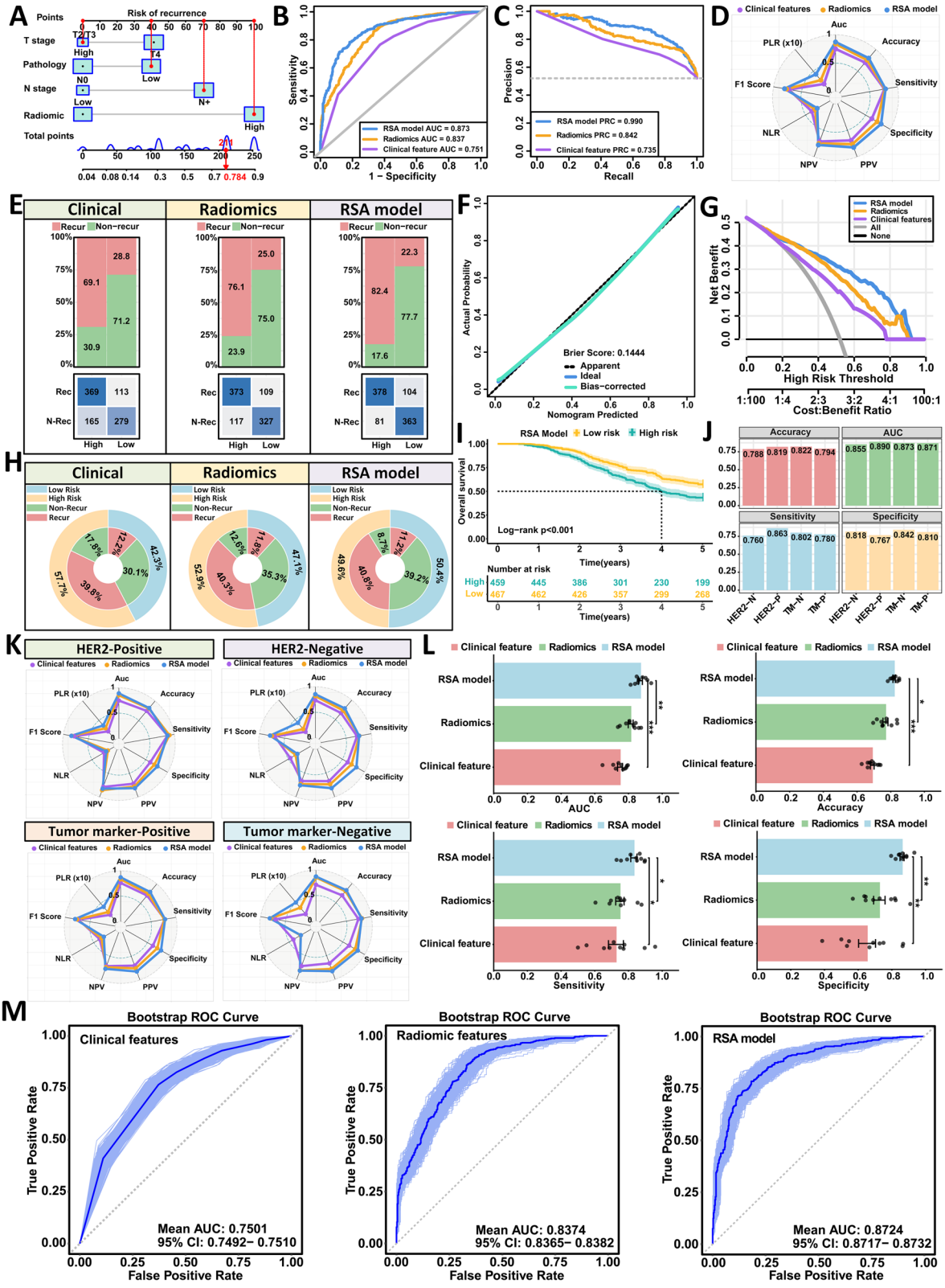


Fig. 2 (See legend on previous page.)

(NRI = 0.616, 95% CI: 0.403–0.823; $P < 0.001$) and discrimination (IDI = 0.237, 95% CI: 0.188–0.283; $P < 0.001$; Additional file 2: Table. S6), and demonstrated good calibration (Brier score = 0.139; Fig. 4D). Validation set II showed comparable results (F igs. 4I–M; Additional file 2: Table. S14).

Decision curve analysis indicated that RSA offered the greatest clinical net benefit across both external cohorts (Figs. 4F and 4O). Concentric circle plots showed that RSA identified a larger share of recurrence cases in the high-risk group compared to the clinical model (set I: 49.0% vs. 40.8%, Fig. 4G; set II: 42.8% vs. 40.2%, Fig. 4N). Sensitivity analysis combining both sets revealed a 14.3% improvement in identifying high-risk patients, along with a 6.4% reduction in low-risk misclassification (Fig. 4Q).

Five-year overall survival was significantly lower among high-risk patients than among those classified as low risk by the RSA model (set I: 42.4% vs. 65.4%, $P < 0.001$; Fig. 4H; set II: 38.5% vs. 67.0%, $P < 0.001$; Fig. 4P). Multivariate Cox regression confirmed RSA as an independent predictor of long-term prognosis in both validation cohorts (set I: HR = 1.998, 95% CI: 1.403–2.845; set II: HR = 2.144, 95% CI: 1.402–3.279; both $P < 0.001$; Additional file 2: Table. S15-16).

Evaluation of the RSA model using a multicenter prospective cohort

To assess the RSA model in a prospective clinical context, we analyzed data from a multicenter trial (NCT01516944) enrolling LAGC patients. After applying strict inclusion and exclusion criteria, 569 patients were included: 193 underwent upfront surgery (33.9%), while 376 received neoadjuvant chemotherapy, including 182 treated with the XELOX regimen (32.0%) and 194 with the SOX regimen (34.1%) (Fig. 5A). Baseline clinical characteristics are detailed in Additional file 2: Table. S17, and trial outcomes have been reported previously [36].

In the primary surgery subgroup, the RSA model demonstrated excellent discrimination for early recurrence, outperforming clinical and radiomics-only models (AUC = 0.857, 95% CI: 0.804–0.908; Figs. 5B–F; Additional file 2: Table. S18). Compared to the clinical model, RSA provided meaningful improvements in classification and discrimination (NRI = 0.597, 95% CI: 0.282–0.901; IDI = 0.244, 95% CI: 0.183–0.306; both P < 0.001; Additional file 2: Table. S6). The calibration was acceptable (Brier score = 0.153; Fig. 5F), and decision curve analysis confirmed superior clinical utility (Fig. 5G). The concentric circle plot also showed a higher proportion of recurrence-free patients in the RSA-defined low-risk group than in the clinical model (39.4% vs. 26.9%; Fig. 5H).

In patients receiving neoadjuvant chemotherapy, model performance declined (AUC = 0.751, 95% CI: 0.698–0.797; Figs. 5K–O; Additional file 2: Table. S6, S18), and no clear link was observed between RSA classification and treatment response (Figs. 5P–Q). However, in both neoadjuvant and surgery-only groups, high-risk patients consistently exhibited worse five-year OS compared to low-risk patients (F igs. 5I–J). Subgroup analysis further indicated limited predictive value of the RSA model within both XELOX and SOX regimens (Additional file 1: Fig. S11).

Stage-specific stratification analysis of the RSA model

To examine whether the RSA model could improve survival prediction within the same TNM stage, we conducted stratified survival analyses across stages I to III in the full LAGC cohort. Kaplan–Meier curves showed that the RSA model effectively distinguished high- and low-risk patients within each TNM stage, with significant survival differences (all log-rank $P < 0.05$; Fig. 6A).

To address potential confounding, we applied 1:1 propensity score matching (PSM) within each TNM stage. Matched baseline characteristics are provided in Additional file 2: Table. S19. Before PSM, RSA-defined high-risk patients had consistently poorer overall survival than low-risk patients (Stage I: 17.9% vs. 91.4%, $P < 0.001$; Stage II: 49.4% vs. 67.4%, $P < 0.001$; Stage III: 36.4% vs. 47.0%, $P < 0.001$). After matching, the survival gap remained statistically significant across all stages (Stage I: 54.2%

(See figure on next page.)

Fig. 3 Validation of RSA Model Performance in Internal Cohorts. **A–B** ROC and PR curves depict model discrimination ability in Internal Validation Cohort I. **C** Radar diagram summarizes model-level AUC, sensitivity, specificity, and accuracy in Cohort I. **D** Calibration curve illustrates the agreement between predicted and observed recurrence probabilities in Cohort I using the RSA model. **E** Classification results in Cohort I are shown using confusion matrices for each model. **F**–**G** Decision curve and concentric circle plots assess and compare the net clinical benefit of models in Cohort I. **H** Survival outcomes of Cohort I stratified by RSA-based risk groups (cutoff: Youden index) are visualized via Kaplan–Meier plots, with log-rank test applied for significance. **I**–**J** Discriminative performance of predictive models in Internal Validation Cohort II represented by ROC and PR curves. **K** Radar plot presents model accuracy indicators in Cohort II. **L** RSA model calibration in Cohort II is illustrated to assess predictive consistency. **M** Cohort II classification accuracy is depicted using model-wise confusion matrices. **N**–**O** Model utility in Cohort II evaluated by double-layer circle visualization and DCA curve, respectively. **P** Kaplan–Meier analysis compares recurrence-free survival in high- vs. low-risk patients in Cohort II, based on RSA nomogram cutoff. **Q** Comparative analysis of RSA versus clinical-only predictors across both internal cohorts for early recurrence prediction

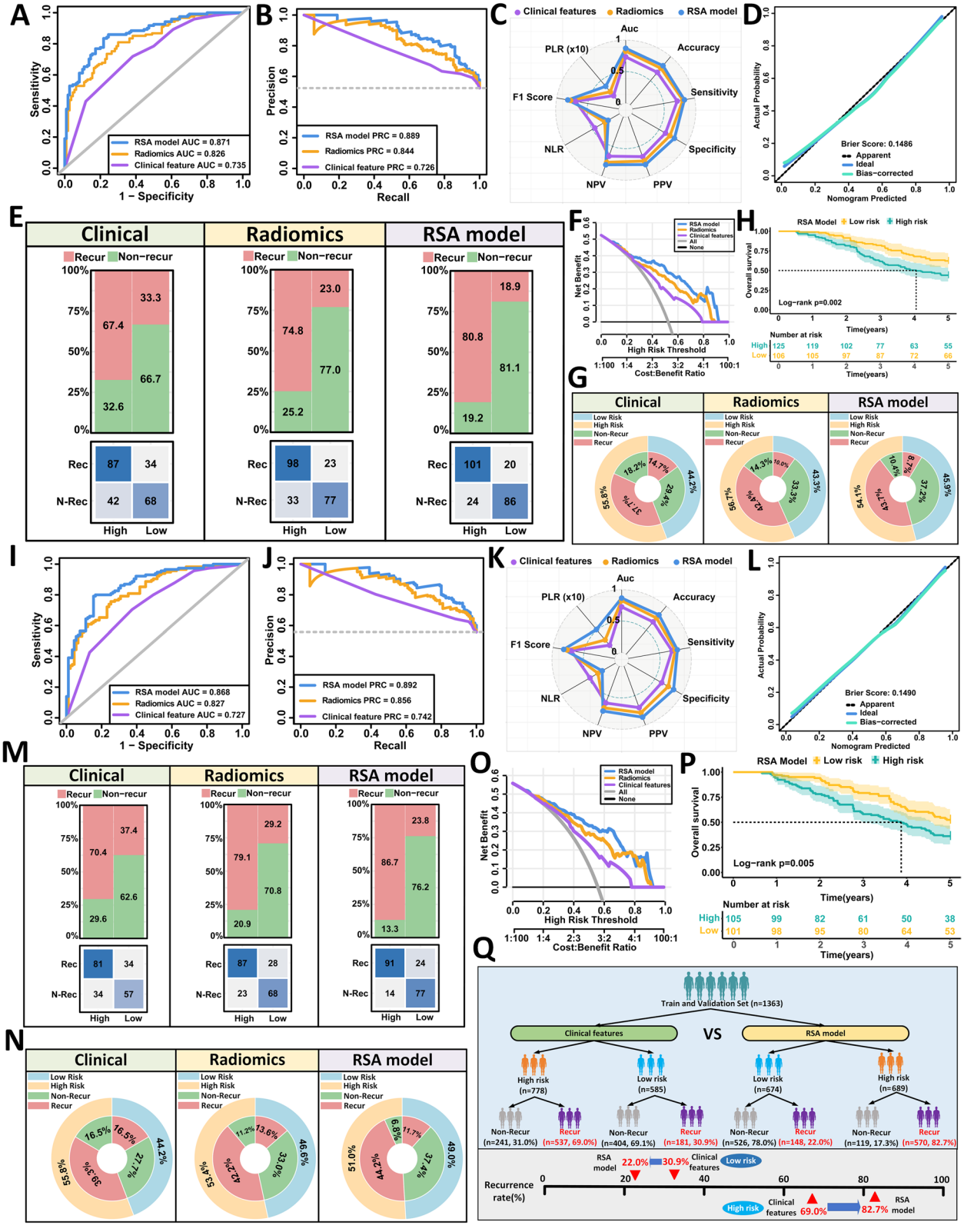


Fig. 3 (See legend on previous page.)

vs. 79.2%, $P=0.038$; Stage II: 49.7% vs. 65.1%, $P=0.003$; Stage III: 34.1% vs. 47.1%, $P<0.001$; Fig. 6B).

Given the complexity of ensemble models, direct interpretation of their decision processes is often limited. To address this, we employed a simplified interpretability framework known as SHAP (SHapley Additive exPlanations), which attributes a model's output to individual input features. SHAP values quantify each feature's contribution to the model's prediction, with higher absolute values indicating greater influence. We computed SHAP values to assess feature importance and ranked the top eight radiomic features most strongly associated with recurrence prediction. These were visualized using swarm plots (Fig. 6C) and importance ranking charts (Additional file 1: Fig. S12). In addition, SHAP dependence plots were generated for each of the eight selected features to visualize their value distributions and directional influence on recurrence risk prediction (Additional file 1: Fig. S13). For each plot, we identified the approximate threshold where the SHAP value crosses zero. This point indicates a potential inflection in the direction of the feature's contribution to the model output and may serve as a clinically meaningful cutoff.

Figure 6D presents a representative comparison between two patients with identical clinical characteristics and treatment regimens. Despite their clinical similarity, differences in radiomic profiles led to divergent outcomes: the patient with a higher radiomics-derived risk score developed liver metastasis within five months post-surgery (Additional file 1: Fig. S14). Collectively, these findings suggest that the RSA model provides significant added prognostic value beyond the traditional TNM staging system.

External validation and immune characterization of the RSA model in gastric cancer

To externally validate the RSA model, we evaluated 41 gastric cancer patients from the TCIA cohort who met the inclusion criteria (Additional file 1: Fig. S15A). The RSA model exhibited excellent predictive accuracy for early postoperative recurrence, outperforming single-feature models with an AUC of 0.850 (95% CI: 0.703–0.962; Fig. 7A, Additional file 1: Fig. S15B-F, Additional file 2: Table. S20), supported by favorable calibration (Brier score = 0.149; Additional file 1: Fig. S15G).

We further explored the biological differences underlying the RSA-defined high- and low-risk groups through transcriptomic analyses using 38 samples from TCIA and an additional prospective cohort of 114 patients. Gene set enrichment analysis revealed significant upregulation of immune-inflammatory pathways, notably IL6/JAK/STAT3 signaling and interferon-alpha/gamma responses, in low-risk patients (Figs. 7B–C). Correspondingly, ESTIMATE analysis confirmed higher stromal and immune scores in this group (Fig. 7D, Additional file 1: Fig. S16A).

Immune activity assessed by single-sample gene set enrichment analysis in TCGA gastric cancer samples indicated significantly greater activity of T cells, APC co-stimulation, immune checkpoints, and inflammatory pathways within the low-risk group (Fig. 7E, Additional file 1: Fig. S16B). Analyses using CIBERSORT and MCPcounter consistently showed higher infiltration of immune cells, especially T cells, in low-risk patients (Figs. 7F–G, Additional file 1: Fig. S16C-D). Higher expression of immune checkpoint genes related to T-cell activation further characterized this group (Additional file 1: Fig. S16E-F).

These observations were validated in the prospective cohort, reinforcing findings of elevated immune-inflammatory pathway activities, increased immune infiltration, and higher immune checkpoint expression in the low-risk group (Figs. 7H–L, Additional file 1: Fig. S17). Finally, expression signatures predictive of immune checkpoint therapy response, specifically interferon-gamma and T-cell effector function signatures, were significantly enhanced in the low-risk group, suggesting potential greater benefit from immunotherapy (Figs. 7M–N).

(See figure on next page.)

Fig. 4 External Validation Results of the RSA Model in Independent Cohorts. **A–B** ROC and PR curves evaluate model discrimination in External Validation Cohort I. **C** Radar chart summarizes the classification metrics including AUC, sensitivity, specificity, and accuracy for models in Cohort I. **D** The calibration curve of the RSA model in Cohort I reflects alignment between predictions and actual outcomes. **E** Confusion matrix visualization highlights model-level classification performance in Cohort I. **F–G** Decision curve analysis and concentric circle plot illustrate net benefit and comparative utility of predictive models in Cohort I. **H** Kaplan–Meier curves depict survival outcomes stratified by RSA risk groups in Cohort I using the nomogram-based Youden threshold; statistical significance determined by the log-rank test. **I–J** External Validation Cohort II model performance is illustrated using ROC and PR curves. **K** Radar diagram conveys accuracy and discrimination metrics for each model in Cohort II. **L** The calibration analysis demonstrates predictive reliability of the RSA model in Cohort II. **M** Confusion matrices display model prediction performance in Cohort II. **N–O** Clinical applicability of models in Cohort II evaluated using concentric circle visualization and DCA, respectively. **P** Recurrence-free survival in RSA-defined subgroups of Cohort II presented using Kaplan–Meier curves; log-rank test applied for comparison. **Q** RSA model's prognostic advantage over clinicopathological variables alone analyzed across both external cohorts

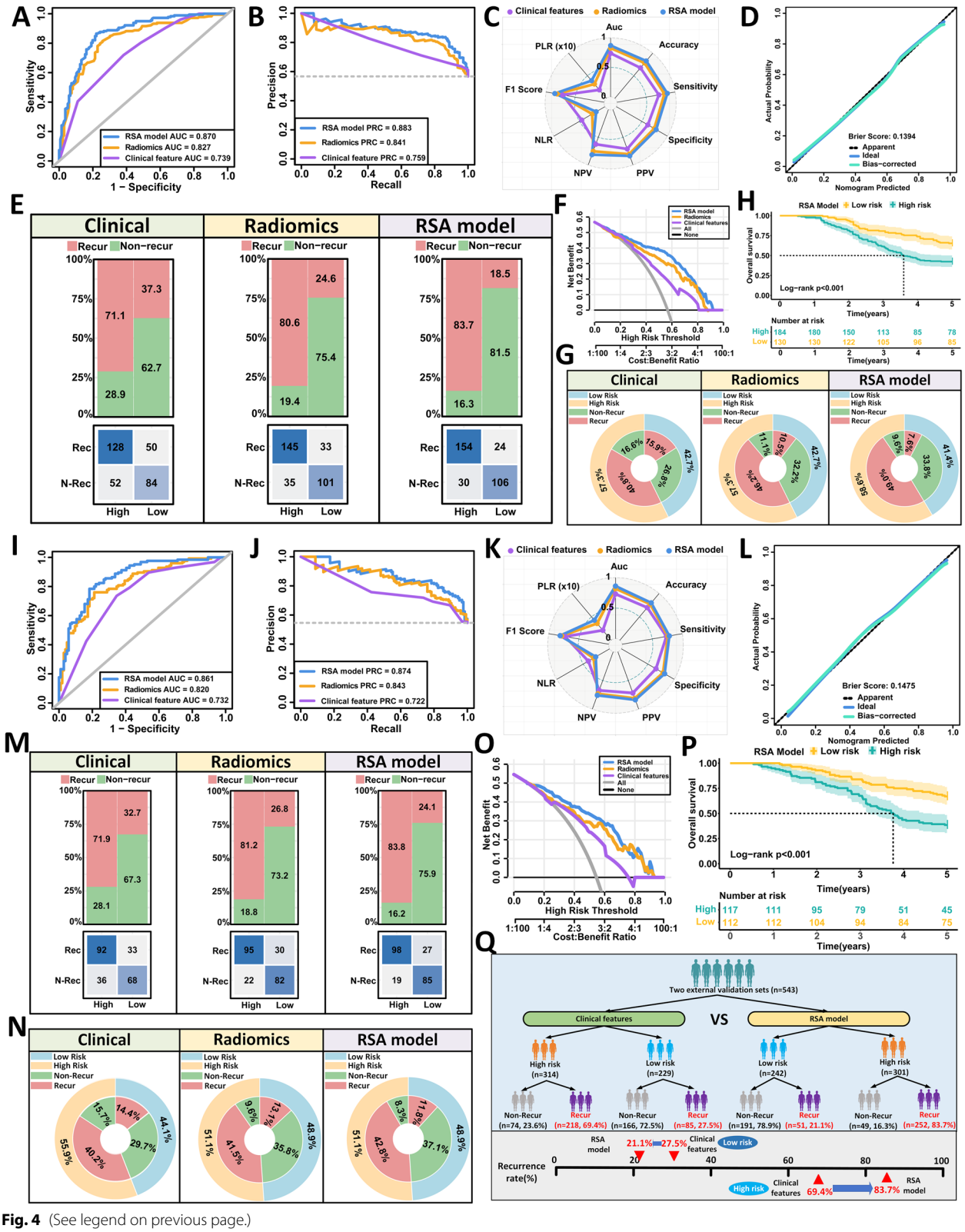


Fig. 4 (See legend on previous page.)

Discussion

In this multicenter study, we successfully developed and validated an integrative Radiomics-Clinical RSA model to accurately predict early postoperative recurrence in LAGC patients. The RSA model uniquely integrates critical clinical parameters (pathological type, tumor invasion depth, and lymph node metastasis status) with carefully selected radiomics features, significantly enhancing predictive accuracy. Across extensive internal and external validations, including geographically diverse cohorts and a prospective multicenter clinical trial, our model consistently outperformed both clinical-only and radiomics-only models, as evidenced by significantly higher AUC values and substantial improvements in NRI and IDI. Such robust predictive power underscores our model's potential utility for personalized patient stratification, enabling clinicians to optimize postoperative surveillance strategies, ultimately improving survival outcomes.

To position the current study within the existing Literature context, we conducted a comprehensive comparative analysis with previously published radiomics-based recurrence models. Huang et al. developed a radiomics nomogram based on CT images for gastric cancer recurrence prediction, reporting an AUC of approximately 0.830 [37]. However, their model was primarily validated in a single-center retrospective cohort, lacking rigorous external validation and integration of clinical parameters. Similarly, Zhang et al. constructed a radiomics signature using enhanced CT imaging, achieving an AUC of 0.831; yet, their validation was limited to hospitals in one region with limited generalizability [38]. By contrast, our RSA model addresses these limitations by combining a broader set of critical clinical parameters with a refined selection of radiomics features validated through rigorous internal, external, and prospective datasets, achieving notably higher AUC values (0.873 in training, 0.870 external validation). Furthermore, our model underwent comprehensive performance evaluation using decision curve analysis, demonstrating substantial clinical benefits across various levels of clinical decision thresholds, an important aspect of validation that was not addressed in previous studies.

Moreover, comparing our findings with those from Cao et al., who also integrated radiomics with clinical features, reveals distinct methodological advantages of our approach [39]. Guo et al. reported an AUC of 0.814 and focused primarily on short-term retrospective validation [40]. Our study not only achieved significantly higher predictive performance but also utilized advanced methodologies, including nnU-Net-based automated segmentation, robust feature selection via LASSO with cross-validation, and harmonization of multi-institutional imaging data using ComBat methods, thus effectively minimizing variability and enhancing generalizability. Importantly, our RSA model was further validated in geographically diverse external cohorts from both northern and southern regions of China, demonstrating consistent accuracy and reliability across varied clinical settings, which represents a significant improvement over previous models. Collectively, these methodological innovations substantially advance the current radiomics modeling landscape for postoperative recurrence prediction in gastric cancer.

In addition to the predictive performance, we also explored the potential biological relevance of the selected radiomic features to enhance interpretability. The radiomic signatures retained in our model are likely to capture fundamental tumor characteristics, including intratumoral heterogeneity, morphology, and intensity variations. Texture features may reflect genomic instability and treatment resistance, shape features describe morphological invasiveness, and first-order intensity features are associated with cellular density, necrosis, or stromal composition. These imaging-derived traits are consistent with our transcriptomic analyses, which demonstrated that low-risk patients exhibited stronger immune activation and greater immune infiltration. Such concordance between imaging phenotypes and molecular profiles supports

(See figure on next page.)

Fig. 5 Prospective Cohort-Based Evaluation of the RSA Model for Early Recurrence Prediction. **A** Flow diagram outlining patient inclusion and risk stratification in the prospective validation dataset. **B**–**C** ROC and PR curves assess model discrimination in the surgery-alone subgroup. **D** Radar chart compares AUC, sensitivity, specificity, and accuracy among different models in this subgroup. **E** Model-wise confusion matrices visualize classification results in the surgery-alone group. **F** Calibration plot shows agreement between predicted and observed outcomes for the RSA model. **G**–**H** Decision curve analysis and concentric circle visualization demonstrate net clinical benefit and comparative utility for each model. **I**–**J** Kaplan–Meier survival analyses for the surgery-alone (I) and neoadjuvant chemotherapy (J) subgroups, stratified by the RSA-defined Youden index threshold; significance tested with log-rank method. **K**–**L** ROC and PR curve assessments of predictive performance in the neoadjuvant chemotherapy subgroup. **M** Radar diagram summarizes accuracy metrics across models for patients receiving neoadjuvant chemotherapy. **N** Calibration analysis evaluates prediction accuracy of the RSA model in this subgroup. **O** Confusion matrices detail classification outcomes across models. **P**–**Q** Comparisons of radiologic treatment response (P) and postoperative pathological grading (Q) among different neoadjuvant regimens, stratified by RSA risk categories

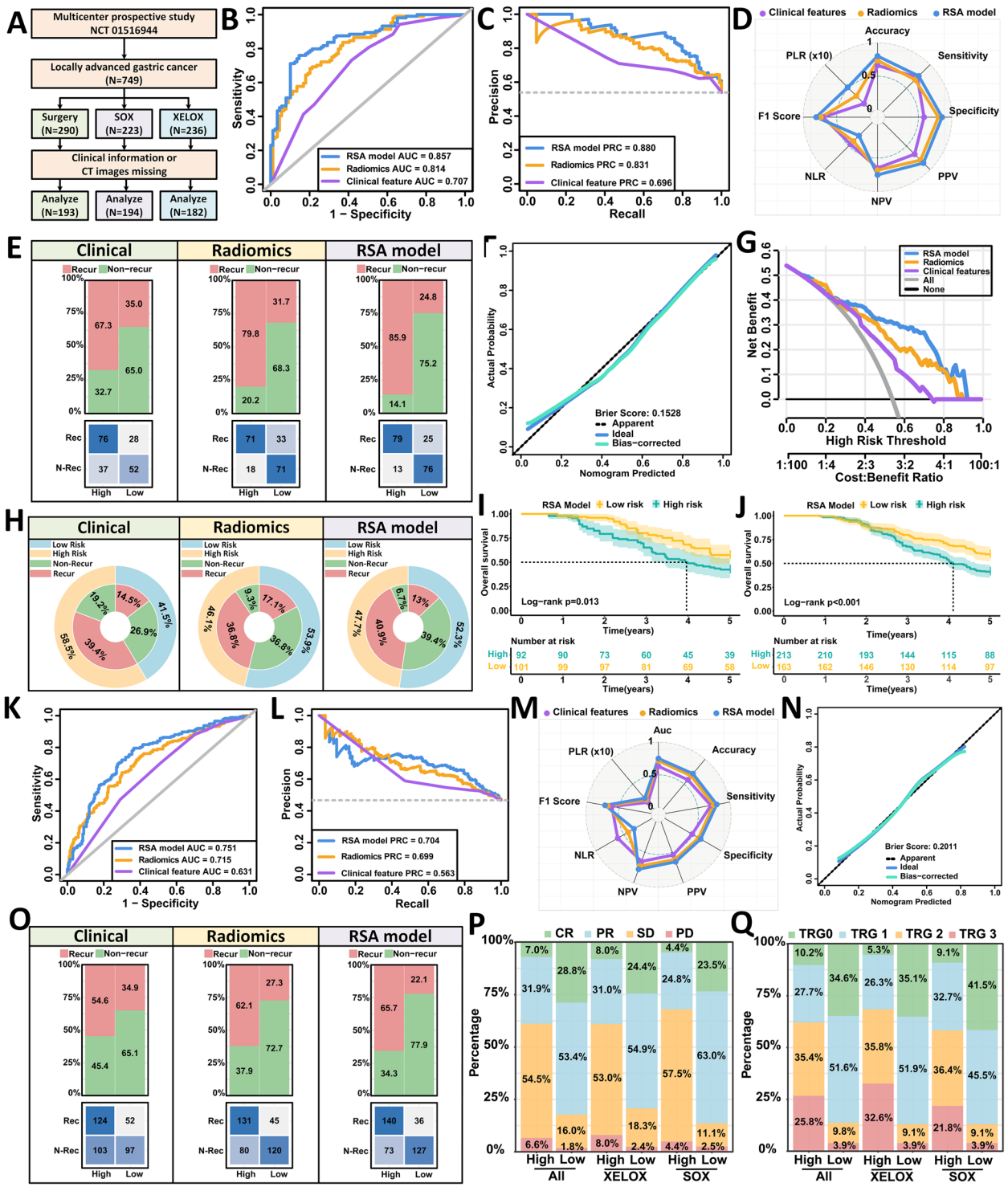


Fig. 5 (See legend on previous page.)

the biological plausibility of our RSA model and highlights its translational potential in clinical practice.

At the molecular level, our study provides significant biological insights into the radiomics-derived risk stratification. Transcriptomic analyses of patients classified as high risk and low risk revealed distinct characteristics of the immune microenvironment, with the low-risk subgroup showing significant enrichment of

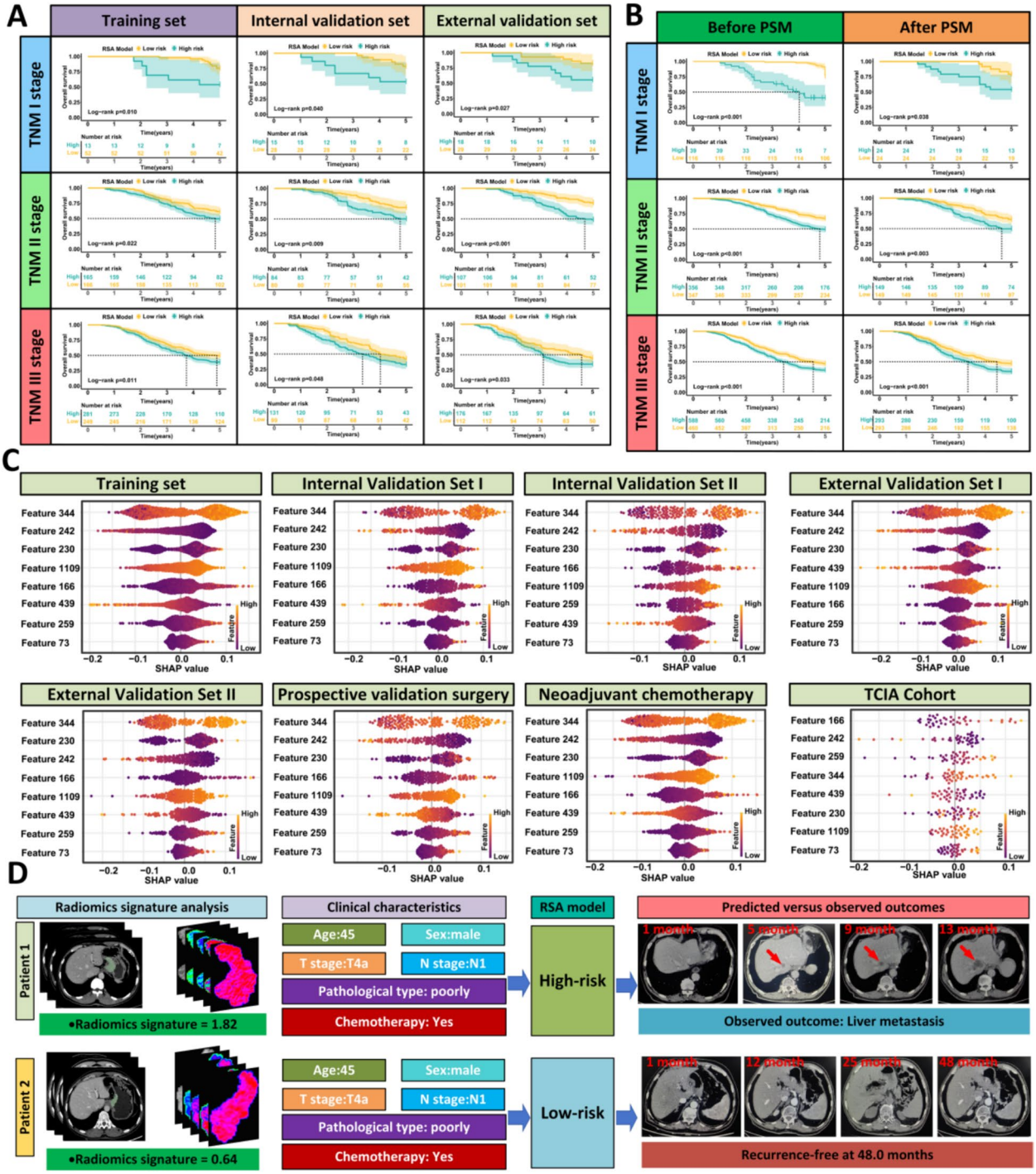


Fig. 6 Added Prognostic Discrimination and Model Interpretability of the RSA Beyond Conventional TNM Stratification. **A** Kaplan–Meier curves compare survival between RSA-based high- and low-risk subgroups within each TNM stage, across the training and validation cohorts, using the Youden index as cutoff. **B** PSM-adjusted survival comparisons highlight prognostic value of RSA-defined risk levels across TNM stages, reducing potential bias. **C** SHAP summary plots depict the relative contribution and directionality of top radiomic and clinical features across datasets. Dots represent individual patients, with color indicating the original feature value and horizontal spread denoting SHAP value magnitude. **D** Clinical trajectory of two example patients from distinct RSA-defined risk groups, illustrating differences in prognosis post-surgery

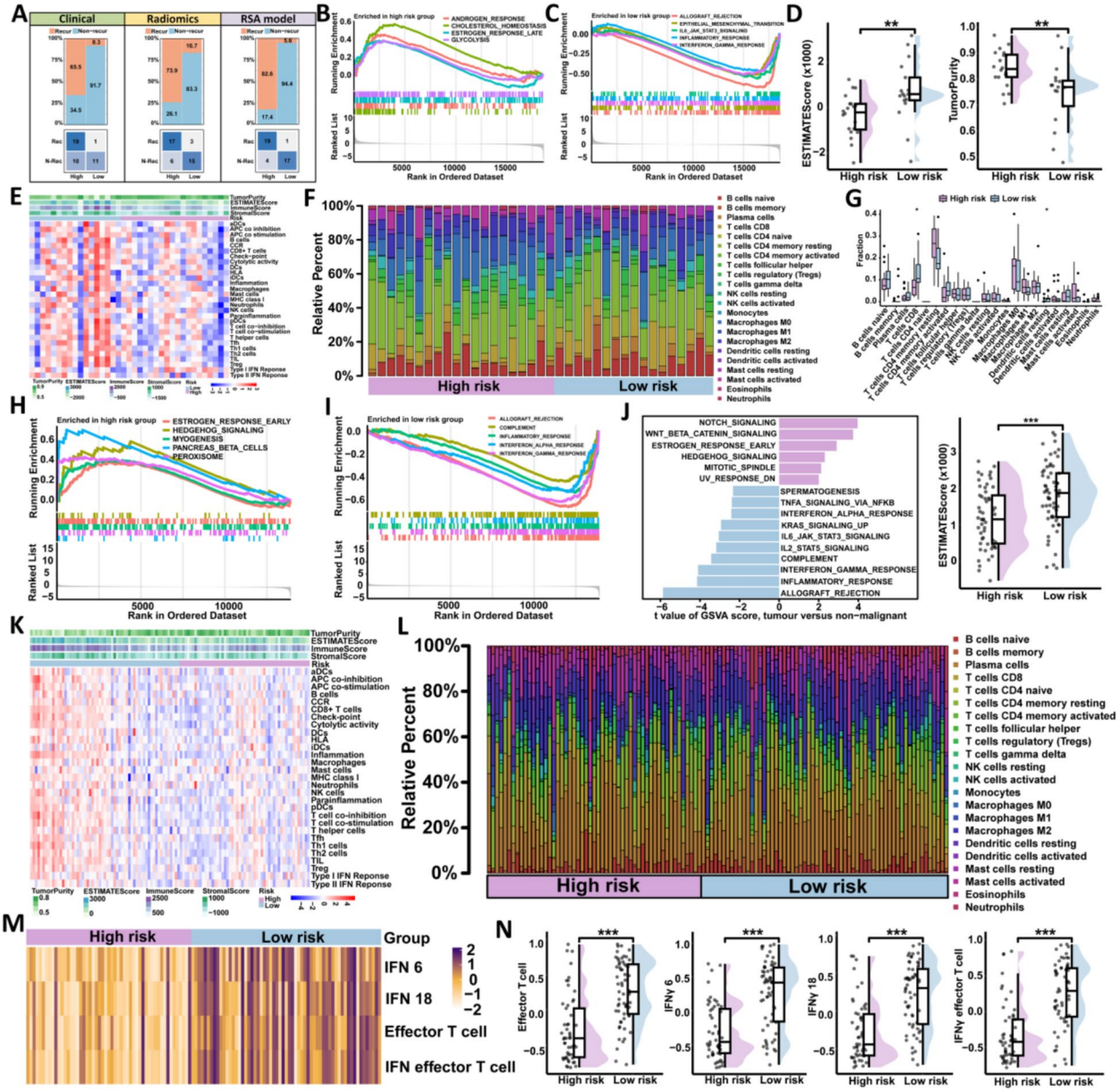


Fig. 7 Molecular and Immune Profiling Between RSA-Derived Risk Subgroups Across Gastric Cancer Cohorts. **A** Confusion matrices illustrate classification accuracy of predictive models in the TCIA cohort. **B**–**C** GSEA highlights enriched pathways specific to the high-risk (B) and low-risk (C) groups. **D** Violin plots compare immune scores and tumor purity, revealing distinct immunological profiles between risk strata. **E** ssGSEA heatmap displays scores for 29 immune-related gene sets across individual TCGA samples. **F** Immune cell composition inferred by CIBERSORT in the TCGA cohort shows differential subtype distribution between groups. **G** Boxplots quantify immune infiltration differences across high- and low-risk patients. **H**–**J** Enrichment analyses using GSEA (H–I) and GSVA (J) identify significant signaling pathways associated with risk levels. **K** ssGSEA-based heatmap of immune gene sets from the internal sequencing dataset. **L** Immune cell landscape derived from CIBERSORT in the sequencing cohort highlights group-specific infiltration patterns. **M** Heatmaps visualizing IFN-γ–related and T cell–associated gene signatures in both risk groups. **N** Violin plots comparing the expression of IFN-γ and T cell–signature scores in the internal dataset. Statistical significance is denoted as follows: ns = not significant; $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$

immune-related signaling pathways, including IL6/JAK/STAT3 signaling, interferon alpha response, and interferon gamma response. These findings indicate an association between radiomics signatures and underlying immunological heterogeneity. Further immune cell infiltration analysis confirmed that the low-risk group exhibited increased immune cell infiltration, particularly T-cell populations, and higher expression of immune

checkpoint molecules. These molecular findings align with contemporary literature, suggesting that robust immunosurveillance correlates with favorable clinical outcomes, providing a compelling biological basis for the superior prognostic accuracy demonstrated by our RSA model [41–43]. In particular, we observed that the RSA model showed reduced predictive performance in patients receiving neoadjuvant chemotherapy (AUC = 0.751), likely due to treatment-related changes in tumor imaging characteristics such as necrosis, fibrosis, or volume reduction. These alterations may disrupt radiomic feature stability and diminish the effectiveness of models trained on pre-treatment data. Therefore, future studies should consider model adaptation or retraining strategies specific to this subgroup, as well as the development of radiomic signatures that dynamically capture post-treatment tumor biology.

Despite our study's strengths, several limitations remain. Firstly, inherent biases associated with retrospective data collection in parts of our cohort may have introduced selection bias. Although mitigated by our rigorous multicenter design and prospective validation, future randomized controlled trials could further strengthen evidence supporting the RSA model's clinical application. Secondly, variations in CT imaging protocols across participating institutions, despite being harmonized using the ComBat algorithm, might still introduce subtle heterogeneity affecting radiomic feature extraction. To evaluate the effectiveness of this harmonization, we conducted principal component analysis before and after ComBat correction, which demonstrated visibly reduced inter-center and inter-dataset clustering, thereby supporting the validity of the feature alignment strategy. Future studies employing standardized imaging protocols could further enhance reproducibility. Additionally, the RSA model exhibited relatively lower performance in patients receiving neoadjuvant chemotherapy, indicating that treatment-induced biological and radiomic feature alterations warrant further exploration to refine predictions in this clinical subgroup. Prospective validation involving dynamic longitudinal imaging and novel deep learning techniques could further optimize the predictive potential and applicability of our RSA model.

Conclusions

In conclusion, this study presents a clinically robust, biologically interpretable, and methodologically innovative RSA model for predicting early postoperative recurrence in LAGC patients. Through extensive multicenter and prospective validations combined with advanced radiomic feature extraction, robust modeling methodologies, and comprehensive biological analyses, we substantially enhance existing predictive capabilities. Our findings not only establish a strong foundation for integrating radiomics into precision oncology but also significantly contribute to the broader implementation of personalized postoperative management strategies in gastric cancer, ultimately improving clinical outcomes.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-04393-4.

Additional file 1: Figure S1. PCA before and after ComBat harmonization. Figure S2. Radiomic feature selection using LASSO in the training cohort. Figure S3. Learning curve analysis of the ensemble model. Figure S4. Model performance (AUC) across different radiomic feature subset sizes. Figure S5. Performance of radiomic models in the training cohort. Figure S6. Performance of radiomic models in internal validation cohorts. Figure S7. Performance of radiomic models in external validation cohorts. Figure S8. RSA model performance in subgroup analyses based on tumor markers and HER2. Figure S9. Confusion matrices across subgroups stratified by tumor markers and HER2. Figure S10. Kaplan–Meier survival by RSA-defined risk groups across subgroups. Figure S11. RSA model performance across NAC regimens in the prospective cohort (NCT01516944). Figure S12. Feature importance of radiomic predictors across datasets (SHAP analysis). Figure S13. SHAP dependence plots of eight radiomic features selected by mRMR and LASSO. Figure S14. SHAP feature contribution visualization for two representative patients. Figure S15. RSA model performance in the TCIA gastric cancer validation cohort. Figure S16. Immune microenvironment differences between RSA-defined groups in TCGA cohort. Figure S17. Immune microenvironment differences between RSA-defined groups in transcriptome cohort.

Additional file 2: Table S1. Characteristics of patients in training, internal, and external cohorts. Table S2. Characteristics by recurrence status in training/internal/external cohorts. Table S3. Performance of different ML and ensemble models. Table S4. Logistic regression of recurrence factors (training set). Table S5. Model performance in predicting recurrence (training set). Table S6. Incremental predictive value of RSA model across validation cohorts. Table S7. Cox regression of OS factors (training set). Table S8. Model performance by HER2/tumor markers. Table S9. Characteristics of patients by admission period (internal validation). Table S10. Model performance in predicting recurrence (internal validation). Table S11. Cox regression of OS factors (internal validation I: 2012–2014). Table S12. Cox regression of OS factors (internal validation II: 2017–2019). Table S13. Characteristics of patients (external validation set). Table S14. Model performance in predicting recurrence (external validation set). Table S15. Cox regression of OS factors (external validation I: Northern). Table S16. Cox regression of OS factors (external validation II: Southern). Table S17. Characteristics of patients (prospective cohort, NCT01516944). Table S18. Model performance (prospective cohort, NCT01516944). Table S19. Baseline characteristics before/after PSM by TNM stage. Table S20. Model performance in predicting recurrence (TCIA set).

Additional file 3: Supplementary Methods. This file provides detailed descriptions of the supplementary methods, including inclusion and exclusion criteria, clinical data collection, RNA sequencing procedures, differential expression analysis, functional enrichment analyses, evaluation of immune activity, immune infiltration analyses (CIBERSORT, MCPcounter, EcoTyper), and statistical analysis strategies. Full references for the methods are included.

Acknowledgements

We sincerely thank the pathologists who participated in this study, namely Li Yang, Xiaosheng Xu and Min Li.

Code availability

The code used for computation analysis in this study can be found at https://github.com/hebeidpa/DeepComp/tree/main/Gastric-nnUNet. For any additional questions, pleasecon-tact the corresponding author.

Authors' contributions
(I)Conception and design: QZ and LM; (II) Administrative support: QZ; (III) Provision of study materials or patients: PD, HG, JY, SC, JW, HW, YT, RG, LZ, NM, XL, ZG, LM, QZ; (IV) Collection and assembly of data: HG, JY, SC, JW, HW; (V) Data analysis and interpretation: PD, HG, JY, SC; (VI) Manuscript writing: PD, SC, HG, JY; and (VII) all authors read and approved the final manuscript.

Funding
Hebei Provincial Major Science and Technology Special Project (23297701Z).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
The study protocol was approved by the Ethics Committee of the Fourth Hospital of Hebei Medical University (approval number: 2024KY199), and informed consent was obtained from all the study participants. All authors followed applicable ethical standards to maintain research integrity without duplication, fraud, or plagiarism.

Ethics statement
All authors certify that they comply with the ethical guidelines for authorship.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]The Third Department of Surgery, the Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei, China. [2]Hebei Key Laboratory of Precision Diagnosis and Comprehensive Treatment of Gastric Cancer, Shijiazhuang 050011, Hebei, China. [3]Big Data Analysis and Mining Application for Precise Diagnosis and Treatment of Gastric Cancer Hebei Provincial Engineering Research Center, Shijiazhuang 050011, Hebei, China. [4]Clinical College of Hebei University, Baoding, China. [5]Hebei Institute of Advanced Computing Technology, Langfang 065000, Hebei, China. [6]School of Chinese Medicine & School of Integrated Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing 210023, Jiangsu, China. [7]Department of Gastroenterology and Hepatology, Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, Jiangsu, China. [8]Department of General Surgery, Renmin Hospital of Wuhan University, Wuhan 430065, Hubei, China. [9]Department of General Surgery, Shijiazhuang People's Hospital, Shijiazhuang 050050, Hebei, China. [10]Department of General Surgery, Baoding Central Hospital, Baoding 071030, Hebei, China. [11]Department of General Surgery, Hengshui People's Hospital, Hengshui 053099, Hebei, China. [12]Research Center and Tumor Research Institute of the Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei, China.

Received: 12 May 2025 Accepted: 11 September 2025
Published online: 25 November 2025

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12–49.
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
3. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. CA Cancer J Clin. 2021;71(3):264–79.
4. Sexton RE, Al Hallak MN, Diab M, Azmi AS. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev. 2020;39(4):1179–203.
5. Li GZ, Doherty GM, Wang J. Surgical management of gastric cancer: a review. JAMA Surg. 2022;157(5):446–54.
6. Chen Y, Jia K, Xie Y, Yuan J, Liu D, Jiang L, et al. The current landscape of gastric cancer and gastroesophageal junction cancer diagnosis and treatment in China: a comprehensive nationwide cohort analysis. J Hematol Oncol. 2025;18(1):42.
7. Sundar R, Nakayama I, Markar SR, Shitara K, van Laarhoven HWM, Janjigian YY, et al. Gastric cancer. Lancet. 2025;S0140–6736(25):00052–62.
8. van Laarhoven H. Management of locally advanced gastric cancer. Lancet Oncol. 2025;26(3):267–9.
9. Svrcek M, Voron T, André T, Smyth EC, de la Fouchardière C. Improving individualised therapies in localised gastro-oesophageal adenocarcinoma. Lancet Oncol. 2024;25(9):e452–63.
10. Gao YX, Guo XJ, Lin B, Huang XB, Tu RH, Lin M, et al. Targeting LHPP in neoadjuvant chemotherapy resistance of gastric cancer: insights from single-cell and multi-omics data on tumor immune microenvironment and stemness characteristics. Cell Death Dis. 2025;16(1):306.
11. Jiang Y, Zhou K, Sun Z, Wang H, Xie J, Zhang T, et al. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. Cell Rep Med. 2023;4(8):101146.
12. Wang W, Peng Y, Feng X, Zhao Y, Seeruttun SR, Zhang J, et al. Development and validation of a computed tomography-based radiomics signature to predict response to neoadjuvant chemotherapy for locally advanced gastric cancer. JAMA Netw Open. 2021;4(8):e2121143.
13. Sun Z, Zhang T, Ahmad MU, Zhou Z, Qiu L, Zhou K, et al. Comprehensive assessment of immune context and immunotherapy response via noninvasive imaging in gastric cancer. J Clin Invest. 2024;134(6):e175834.
14. Huang W, Jiang Y, Xiong W, Sun Z, Chen C, Yuan Q, et al. Noninvasive imaging of the tumor immune microenvironment correlates with response to immunotherapy in gastric cancer. Nat Commun. 2022;13(1):5095.
15. Feng L, Liu Z, Li C, Li Z, Lou X, Shao L, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicentre observational study. Lancet Digit Health. 2022;4(1):e8-17.
16. Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol. 2020;31(7):912–20.
17. Adili D, Mohetaer A, Zhang W. Diagnostic accuracy of radiomics-based machine learning for neoadjuvant chemotherapy response and survival prediction in gastric cancer patients: a systematic review and meta-analysis. Eur J Radiol. 2024;173:111249.
18. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. Ann Oncol. 2019;30(3):431–8.
19. Chen Q, Zhang L, Liu S, You J, Chen L, Jin Z, et al. Radiomics in precision medicine for gastric cancer: opportunities and challenges. Eur Radiol. 2022;32(9):5852–68.
20. Spadarella G, Stanzione A, Akinci D'Antonoli T, Andreychenko A, Fanni SC, et al. Systematic review of the radiomics quality score applications: an EuSoMII Radiomics Auditing Group Initiative. Eur Radiol. 2023;33(3):1884–94.
21. Wang FH, Zhang XT, Tang L, Wu Q, Cai MY, Li YF, et al. The Chinese society of clinical oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer, 2023. Cancer Commun. 2024;44(1):127–72.
22. Ajani JA, D'Amico TA, Bentrem DJ, Chao J, Cooke D, Corvera C, et al. Gastric cancer, version 2.2022, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw. 2022;20(2):167–92.
23. Zhang XQ, Huang ZN, Wu J, Liu XD, Xie RZ, Wu YX, et al. Machine learning prediction of early recurrence in gastric cancer: a nationwide real-world study. Ann Surg Oncol. 2025;32(4):2637–50.
24. Takahashi R, Ohashi M, Kano Y, Ida S, Kumagai K, Nunobe S, et al. Timing and site-specific trends of recurrence in patients with pathological stage II or III gastric cancer after curative gastrectomy followed by adjuvant S-1 monotherapy. Gastric Cancer. 2019;22(6):1256–62.
25. Chiang CY, Huang KH, Fang WL, Wu CW, Chen JH, Lo SS, et al. Factors associated with recurrence within 2 years after curative surgery for gastric adenocarcinoma. World J Surg. 2011;35(11):2472–8.

26. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9.
27. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12(4):357–60.
28. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. Stringtie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol. 2015;33(3):290–5.
29. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.
30. 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.
31. Wang Z, Fang M, Zhang J, Tang L, Zhong L, Li H, et al. Radiomics and deep learning in nasopharyngeal carcinoma: a review. IEEE Rev Biomed Eng. 2024;17:118–35.
32. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
33. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–7.
34. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218.
35. Zhong L, Dong D, Fang X, Zhang F, Zhang N, Zhang L, et al. A deep learning-based radiomic nomogram for prognosis and treatment decision in advanced nasopharyngeal carcinoma: a multicentre study. EBioMedicine. 2021;70:103522.
36. Zhao Q, Lian C, Huo Z, Li M, Liu Y, Fan L, et al. The efficacy and safety of neoadjuvant chemotherapy on patients with advanced gastric cancer: a multicenter randomized clinical trial. Cancer Med. 2020;9(16):5731–45.
37. Huang L, Feng B, Li Y, Liu Y, Chen Y, Chen Q, et al. Computed tomography-based radiomics nomogram: potential to predict local recurrence of gastric cancer after radical resection. Front Oncol. 2021;11:638362.
38. Zhang W, Fang M, Dong D, Wang X, Ke X, Zhang L, et al. Development and validation of a CT-based radiomic nomogram for preoperative prediction of early recurrence in advanced gastric cancer. Radiother Oncol. 2020;145:13–20.
39. Cao M, Hu C, Li F, He J, Li E, Zhang R, et al. Development and validation of a deep learning model for predicting gastric cancer recurrence based on CT imaging: a multicenter study. Int J Surg. 2024;110(12):7598–606.
40. Guo X, Chen M, Zhou L, Zhu L, Liu S, Zheng L, et al. Predicting early recurrence in locally advanced gastric cancer after gastrectomy using CT-based deep learning model: a multicenter study. Int J Surg. 2025;111(2):2089–104.
41. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017;14(12):717–34.
42. Bailey C, Black JRM, Reading JL, Litchfield K, Turajlic S, McGranahan N, et al. Tracking cancer evolution through the disease course. Cancer Discov. 2021;11(4):916–32.
43. Huang B, Lang X, Li X. The role of IL-6/JAK2/STAT3 signaling pathway in cancers. Front Oncol. 2022;12:1023177.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.