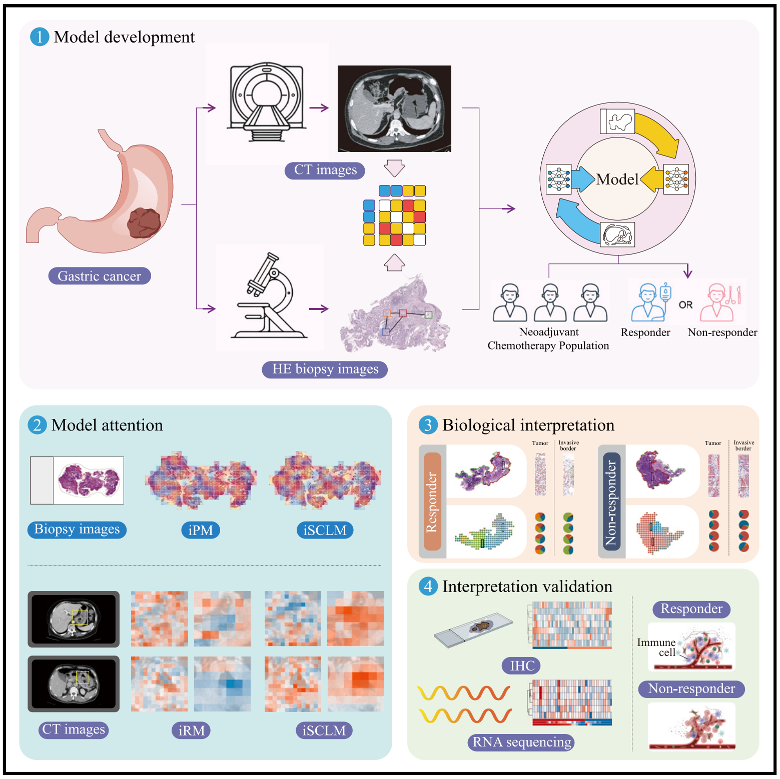
Cell Reports
Medicine

Article

Interpretable multi-modal artificial intelligence model for predicting gastric cancer response to neoadjuvant chemotherapy

Graphical abstract



Highlights

- iSCLM is a multi-modal framework to predict neoadjuvant chemotherapy response
- iSCLM enables a focus on tumor-invasive borders with multi-modal data
- iSCLM is interpreted with increased inflammatory cell infiltration

Authors

Peng Gao, Qiong Xiao, Hui Tan, ...,
Ting Deng, Dinggang Shen,
Zhenning Wang

Correspondence

cmuwangguan@sina.com (G.W.),
zhycmu@163.com (Y.Z.),
xymcdengting@126.com (T.D.),
dinggang.shen@gmail.com (D.S.),
znwang@cmu.edu.cn (Z.W.)

In brief

Gao et al. develop an interpretable AI model (iSCLM) integrating CT scans and biopsy images to predict the response of neoadjuvant chemotherapy in gastric cancer. Validated with a multicenter cohort, iSCLM shows interpretable pathology changes in responders, contributing to the advancement of clinical practices in screening patients for neoadjuvant chemotherapy administration.

Gao et al., 2024, Cell Reports Medicine *5*, 101848
December 17, 2024 © 2024 The Authors. Published by Elsevier Inc.
https://doi.org/10.1016/j.xcrm.2024.101848

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Interpretable multi-modal artificial intelligence model for predicting gastric cancer response to neoadjuvant chemotherapy

Peng Gao,[1,26] Qiong Xiao,[1,26] Hui Tan,[1,26] Jiangdian Song,[2,26] Yu Fu,[1,26] Jingao Xu,[3,26] Junhua Zhao,[1] Yuan Miao,[4] Xiaoyan Li,[5] Yi Jing,[3] Yingying Feng,[6] Zitong Wang,[1] Yingjie Zhang,[1] Enbo Yao,[1] Tongjia Xu,[1] Jipeng Mei,[1]

(Author list continued on next page)

[1]Department of Surgical Oncology and General Surgery, The First Hospital of China Medical University, Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors (China Medical University), Ministry of Education, Shenyang 110001, China
[2]School of Health Management, China Medical University, Shenyang 110122, China
[3]Neusoft Research of Intelligent Healthcare Technology, Co. Ltd., Shenyang 110169, China
[4]Department of Pathology, The First Hospital of China Medical University, Shenyang 110001, China
[5]Department of Pathology, Cancer Hospital of China Medical University, Liaoning Cancer Hospital & Institute, Cancer Hospital of Dalian University of Technology, Shenyang 110042, China
[6]The School of Computer Science and Engineering, Northeastern University, Shenyang 110167, China
[7]Department of Radiology, The First Hospital of China Medical University, Shenyang 110001, China
[8]Department of GI Medical Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin Key Laboratory of Digestive Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin 300202, China
[9]Department of Pathology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450000, China
[10]State Key Laboratory of Holistic Integrative Management of Gastrointestinal Cancers, National Clinical Research Center for Digestive Diseases, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, China
[11]Department of Radiology, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, China
[12]Department of Pathology, Xijing Hospital, Fourth Military Medical University, Xi'an 710032, China
[13]Department of Oncology, Shengjing Hospital of China Medical University, Shenyang 110004, China
[14]Department of Pathology, Shenyang Fifth People Hospital, Shenyang 110001, China
[15]Department of Oncology, People's Hospital of Liaoning Province, People's Hospital of China Medical University, Shenyang 110000, China
[16]Department of Medical Oncology, The Second Hospital of Dalian Medical University, Dalian 116021, China
[17]The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, China

(Affiliations continued on next page)

SUMMARY

Neoadjuvant chemotherapy assessment is imperative for prognostication and clinical management of locally advanced gastric cancer. We propose an incremental supervised contrastive learning model (iSCLM), an interpretable artificial intelligence framework integrating pretreatment CT scans and H&E-stained biopsy images, for improved decision-making regarding neoadjuvant chemotherapy. We have constructed and tested iSCLM using retrospective data from 2,387 patients across 10 medical centers and evaluated its discriminative ability in a prospective cohort (132 patients; ChiCTR2300068917). iSCLM achieves areas under receiver operating characteristic curves of 0.846–0.876 across different test cohorts. Computed tomography (CT) and pathological attention heatmaps from Shapley additive explanations and global sort pooling illustrate additional benefits for capturing morphological features through supervised contrastive learning. Specifically, pathological top-ranked tiles exhibit decreased distances to tumor-invasive borders and increased inflammatory cell infiltration in responders compared with non-responders. Moreover, CD11c expression is elevated in responders. The developed interpretable model at the molecular pathology level accurately predicts chemotherapy efficacy.

INTRODUCTION

Despite substantial efforts to improve treatment for gastric cancer via complete tumor resection and adjacent lymph node removal,[1,2] the overall patient prognosis remains dismal.[3,4] Neoadjuvant chemotherapy has been used for patients with locally advanced gastric cancer to reduce tumor size prior to curative resectioning.[5,6] However, outcomes vary substantially among patients with locally advanced gastric cancer receiving neoadjuvant chemotherapy,[7–9] thereby leading to potential adverse

Cell Reports Medicine 5, 101848, December 17, 2024

Hanyu Chen,[1] Xue Jiang,[7] Yuchong Yang,[8] Zhengyang Wang,[9] Xianchun Gao,[10] Minwen Zheng,[11] Liying Zhang,[12] Min Jiang,[13] Yuying Long,[14] Lijie He,[15] Jinghua Sun,[16] Yanhong Deng,[17] Bin Wang,[18,19,20] Yan Zhao,[21] Yi Ba,[22] Guan Wang,[7,*] Yong Zhang,[5,*] Ting Deng,[8,*] Dinggang Shen,[23,24,25,*] and Zhenning Wang[1,27,*]

[18]Department of Gastroenterology & Chongqing Key Laboratory of Digestive Malignancies, Daping Hospital, Army Medical University (Third Military Medical University), Chongqing 400042, China

[19]School of Medicine, Chongqing University, Chongqing 400000, China

[20]Institute of Pathology and Southwest Cancer Center, and Key Laboratory of Tumor Immunopathology of Ministry of Education of China, Southwest Hospital, Army Medical University (Third Military Medical University), Chongqing 400038, China

[21]Department of Stomach Surgery, Cancer Hospital of China Medical University, Liaoning Cancer Hospital & Institute, Cancer Hospital of Dalian University of Technology, The Liaoning Provincial Key Laboratory of Interdisciplinary Research on Gastrointestinal Tumor Combining Medicine with Engineering, Shenyang 110042, China

[22]Cancer Medical Center & Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China

[23]School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China

[24]Shanghai United Imaging Intelligence Co., Ltd., Shanghai 201807, China

[25]Shanghai Clinical Research and Trial Center, Shanghai 201210, China

[26]These authors contributed equally

[27]Lead contact

*Correspondence: cmuwangguan@sina.com (G.W.), zhycmu@163.com (Y.Z.), xymcdengting@126.com (T.D.), dinggang.shen@gmail.com (D.S.), znwang@cmu.edu.cn (Z.W.)

https://doi.org/10.1016/j.xcrm.2024.101848

effects and increased expenses for non-responders.[10] Therefore, identifying patient response to neoadjuvant chemotherapy is crucial for optimizing therapeutic decision-making.

Currently, there is no global consensus concerning optimal selection criteria.[11–14] National Comprehensive Cancer Network (NCCN) guidelines recommend neoadjuvant chemotherapy followed by surgery for patients with resectable diseases $\geq$cT2.[11] Meanwhile, the Japanese Gastric Cancer Association recommends this approach for cases of bulky lymph node metastasis.[14] Significant discrepancies among different guidelines highlight the current clinical challenges in accurately predicting patients who would benefit from the treatment.

The application of artificial intelligence (AI) for extracting information and identifying new patterns in multi-modal biomedical data is increasing, particularly through machine learning and deep learning.[15–18] Deep learning-based radiology and pathology models have demonstrated promise in predicting neoadjuvant chemotherapy responses and survival prognosis for patients with locally advanced gastric cancer.[19–22] However, major image-based studies exhibit limited accuracy and lack prospective validation using single-image modality data. The exploration of multi-modal feature representation, such as aligning microscopic whole-slide images (WSIs) with macroscopic computed tomography (CT), to facilitate clinical decision-making, remains insufficient. Additionally, studies on deep learning methods aimed at analyzing cell components in WSI to interpret variations in chemotherapy response among patients with locally advanced gastric cancer are rare.

In this study, we developed an interpretable AI framework to predict the response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer. Supervised contrastive learning was employed to integrate features from pretreatment contrast-enhanced CT scans and hematoxylin and eosin (H&E)-stained biopsy slides. The AI-enabled model was developed to interpret how underlying cellular level features influence the extraction of CT features for predicting chemotherapy response. Top-ranked pathology tiles enriched from diverse geographic locations were analyzed to identify cellular differences. These aforementioned findings were further validated using RNA sequencing (RNA-seq) and immunohistochemistry (IHC).

RESULTS

Overview

The overall study design is illustrated in Figure 1. First, pretreatment CT scans and gastric endoscopic biopsy pathological WSI were collected from patients with locally advanced gastric cancer, and regions of interest (ROIs) were manually annotated. A supervised contrastive learning model (SCLM) was proposed by incorporating CT and histopathological features extracted from these ROIs. The interpretability of the model was evaluated across multiple levels, including CT, pathology, cellular, and molecular dimensions. Notably, we evaluated changes in attention patterns after employing our training strategy to map CT and pathological features and further investigated the contribution of cellular and molecular information within the focus areas to the model's predictions. Finally, the performance of the model to guide neoadjuvant chemotherapy selection for improving survival outcomes in patients with locally advanced gastric cancer was evaluated.

After applying inclusion and exclusion criteria, 2,003 patients from Eastern and Northern China, sourced from the First Hospital of China Medical University, Liaoning Cancer Hospital and Institute, Tianjin Medical University Cancer Institute and Hospital, Second Hospital of Dalian Medical University, Shengjing Hospital of China Medical University, and People's Hospital of Liaoning Province, were enrolled in the development cohort (n = 1,208) and incremental cohort (n = 795) for model construction. Two independent cohorts were used for external testing; external test cohort 1 comprised 162 patients from the First Affiliated Hospital of Zhengzhou University in Central China, and

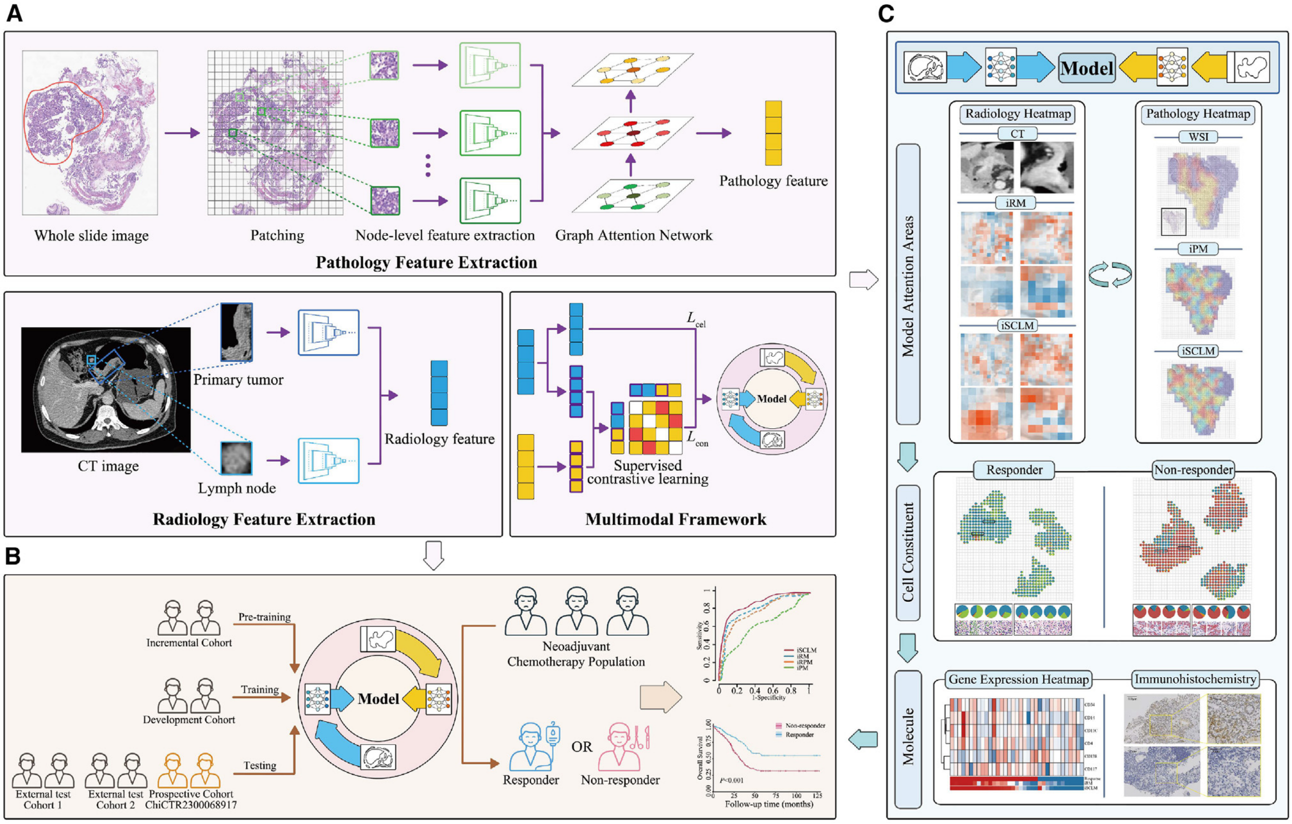


Figure 1. Workflow of the study

(A) Feature extraction and model development. L_{con}, the supervised contrastive learning loss; L_{cel}, the cross-entropy loss.

(B) External validation and patient risk stratification.

(C) Biological interpretation and molecular verification of the model. iRM, incremental radiology model; iPM, incremental pathology model; iSCLM, incremental supervised contrastive learning model.

external test cohort 2 comprised 222 patients from the Sixth Affiliated Hospital of Sun Yat-sen University in Southern China, Xijing Hospital, and Daping Hospital in Western China. In addition, a prospective cohort (132 eligible patients; ChiCTR2300068917) from the First Hospital of China Medical University and Tianjin Medical University Cancer Institute and Hospital (treated between March 1, 2023, and February 29, 2024) was enrolled and used to assess the applicability of the model (Figure S1).

The clinicopathological characteristics of the enrolled patients are detailed in Table 1. Response rates to chemotherapy were 63.3% in the development cohort (765/1,208), 57.4% in external test cohort 1 (93/162), 73.0% in external test cohort 2 (162/222), and 45.3% in the incremental cohort (360/795). Among the prospectively enrolled patients, 67.4% (89/132 individuals) responded to neoadjuvant chemotherapy. In the development cohort, we observed significant differences in the distribution proportions of cT ($p = 0.011$) and cN ($p = 0.001$) categories between responders and non-responders.

Integration of multi-modal features in a SCLM improves predictive performance

A deep residual neural network (ResNet-34)[23] to extract CT features of the primary lesion and lymph nodes was used to predict the response to neoadjuvant chemotherapy. We concatenated the features and used a fully connected layer with 256 neurons. Moreover, inspired by the spatial capability of tumor slides, we converted WSIs into spatially connected graph representations, where each image tile represented a graph node. Pretrained ResNet-18 was used to extract node features as input, and the graph attention network (GAT)[24] was subsequently employed to learn the attention score of each edge by aggregating neighborhood features to differentiate node weights. A SCLM was then proposed for effective learning of feature representations from both CT and pathological images, facilitating cross-modal understanding.

We developed six robust SCLM variants using different loss combinations. Among these models, the SCLM-d variant exhibited the best performance, achieving an area under the receiver operating characteristic curve (AUROC) of 0.880 (95% confidence interval [95% CI] 0.778–0.982; Table S1) during validation. Accordingly, SCLM-d was used for further development, yielding AUROC values of 0.817 (95% CI 0.752–0.882) in external test cohort 1 and 0.847 (95% CI 0.783–0.910) in external test cohort 2.

To enhance the training of feature extractors, an incremental learning[25] strategy integrated data from patients with unresected

Table 1. Characteristics of patients

	Development cohort (*n* = 1,208)			Incremental cohort (*n* = 795)[a]			External test cohort 1 (*n* = 162)			External test cohort 2 (*n* = 222)			Prospective test cohort (*n* = 132)		
	Non-responder	Responder	*p* value	Non-responder	Responder	*p* value	Non-responder	Responder	*p* value	Non-responder	Responder	*p* value	Non-responder	Responder	*p* value
Sex	–	–	0.320	–	–	0.844	–	–	0.049	–	–	0.238	–	–	0.081
Male	324	539	–	309	258	–	49	78	–	47	114	–	38	67	–
Female	119	226	–	126	102	–	20	15	–	13	48	–	5	22	–
Age	–	–	0.974	–	–	0.334	–	–	0.691	–	–	0.379	–	–	0.381
<60	206	355	–	163	123	–	29	42	–	22	70	–	14	36	–
≥60	237	410	–	272	237	–	40	51	–	38	92	–	29	53	–
cT	–	–	0.011	–	–	0.973	–	–	0.215	–	–	0.162	–	–	0.729
T2	12	18	–	6	5	–	2	4	–	2	15	–	1	2	–
T3	122	275	–	77	66	–	27	48	–	29	87	–	20	35	–
T4	309	472	–	352	289	–	40	41	–	29	60	–	22	52	–
cN	–	–	0.001	–	–	<0.001	–	–	0.097	–	–	0.063	–	–	0.061
N0	32	97	–	35	13	–	4	17	–	2	15	–	2	5	–
N1	101	193	–	95	58	–	33	40	–	14	56	–	8	29	–
N2	176	305	–	171	132	–	26	32	–	31	55	–	17	40	–
N3	134	170	–	134	157	–	6	4	–	13	36	–	16	15	–
CA19-9 (U/mL)	11.54 (6.12–37.67)	13.64 (6.77–38.14)	0.128	18.73 (7.91–126.10)	18.29 (7.88–120.00)	0.737	10.44 (5.84–36.92)	10.30 (3.76–22.05)	0.255	8.83 (4.08–27.09)	10.17 (4.14–35.81)	0.616	12.41 (6.22–47.39)	13.24 (6.68–41.17)	0.830
CEA (ng/mL)	2.68 (1.43–6.87)	2.64 (1.43–5.76)	0.525	3.81 (1.80–12.21)	5.06 (2.23–23.52)	0.001	2.64 (1.39–5.07)	2.77 (1.59–13.20)	0.307	2.62 (1.73–4.61)	3.13 (1.73–8.43)	0.306	2.67 (1.52–5.23)	3.25 (1.33–6.70)	0.990

Data are represented as *n* or median (IQR). cT, clinical primary tumor category; cN, clinical regional lymph node category; IQR, interquartile range.

[a] The incremental cohort included 211 responders and 235 non-responders with cM1.

Table 2. Prediction performance of iSCLM compared with other models

Model	External test cohort 1	p value	External test cohort 2	p value	Prospective test cohort	p value
iSCLM						
AUROC	0.866 (0.804–0.927)	reference	0.876 (0.820–0.933)	reference	0.846 (0.780–0.911)	reference
Accuracy	0.802 (0.734–0.856)	reference	0.847 (0.794–0.888)	reference	0.788 (0.711–0.849)	reference
Sensitivity	0.946 (0.880–0.977)	–	0.901 (0.846–0.938)	–	0.775 (0.678–0.850)	–
Specificity	0.609 (0.491–0.715)	–	0.700 (0.575–0.801)	–	0.814 (0.674–0.903)	–
PPV	0.765 (0.680–0.833)	–	0.890 (0.833–0.929)	–	0.896 (0.808–0.946)	–
NPV	0.894 (0.774–0.954)	–	0.724 (0.598–0.822)	–	0.636 (0.504–0.751)	–
iRM						
AUROC	0.802 (0.734–0.870)	0.005	0.822 (0.759–0.885)	0.007	0.781 (0.700–0.861)	0.007
Accuracy	0.722 (0.649–0.785)	0.117	0.631 (0.565–0.691)	<0.001	0.674 (0.590–0.748)	0.052
Sensitivity	0.624 (0.522–0.715)	–	0.525 (0.448–0.600)	–	0.607 (0.503–0.702)	–
Specificity	0.855 (0.753–0.919)	–	0.917 (0.819–0.964)	–	0.814 (0.674–0.903)	–
PPV	0.853 (0.750–0.918)	–	0.944 (0.876–0.976)	–	0.871 (0.766–0.933)	–
NPV	0.628 (0.527–0.719)	–	0.417 (0.336–0.502)	–	0.500 (0.386–0.614)	–
iPM						
AUROC	0.550 (0.459–0.641)	<0.001	0.596 (0.511–0.682)	<0.001	0.546 (0.441–0.650)	<0.001
Accuracy	0.531 (0.454–0.606)	<0.001	0.631 (0.565–0.691)	<0.001	0.538 (0.453–0.621)	<0.001
Sensitivity	0.495 (0.395–0.594)	–	0.698 (0.623–0.763)	–	0.584 (0.480–0.681)	–
Specificity	0.580 (0.462–0.689)	–	0.450 (0.331–0.575)	–	0.442 (0.304–0.589)	–
PPV	0.613 (0.500–0.715)	–	0.774 (0.700–0.834)	–	0.684 (0.573–0.778)	–
NPV	0.460 (0.359–0.564)	–	0.355 (0.257–0.467)	–	0.339 (0.229–0.470)	–
iRPM						
AUROC	0.784 (0.714–0.855)	0.002	0.833 (0.772–0.894)	0.082	0.698 (0.603–0.794)	<0.001
Accuracy	0.722 (0.649–0.785)	0.117	0.725 (0.663–0.780)	0.003	0.659 (0.575–0.734)	0.028
Sensitivity	0.785 (0.691–0.856)	–	0.691 (0.616–0.757)	–	0.663 (0.560–0.753)	–
Specificity	0.638 (0.520–0.741)	–	0.817 (0.701–0.894)	–	0.651 (0.502–0.776)	–
PPV	0.745 (0.650–0.821)	–	0.911 (0.847–0.949)	–	0.797 (0.692–0.873)	–
NPV	0.688 (0.566–0.788)	–	0.495 (0.399–0.592)	–	0.483 (0.359–0.608)	–

Data are metric value (95% CI). AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; iRM, incremental radiology model; iPM, incremental pathology model; iRPM, model constructed by direct concatenation of radiological and pathological features with incremental learning; iSCLM, incremental supervised contrastive learning model.

tumors based on the incremental cohort. The performance of the incremental SCLM (iSCLM) was enhanced, resulting in AUROC scores of 0.866 (95% CI 0.804–0.927) in external test cohort 1 and 0.876 (95% CI 0.820–0.933) in external test cohort 2 (Table 2; Figures 2A and 2B). Moreover, the results of stratified analysis based on various histological types, chemotherapy regimens, and tumor locations also demonstrated consistent performance across two external test datasets (Figure S2).

iSCLM outperformed unimodal models and the model constructed by directly concatenating features. Notably, the incremental radiology-based model (iRM), focusing on single-CT image modality, exhibited moderately lower performance compared with iSCLM, with AUROC values of 0.802 (95% CI 0.734–0.870) in external test cohort 1 ($p = 0.005$) and 0.822 (95% CI 0.759–0.885) in external test cohort 2 ($p = 0.007$). Similarly, the incremental pathology model (iPM) demonstrated a substantially lower AUROC compared with iSCLM in both the external test cohorts, with values of 0.550 (95% CI 0.459–0.641) in external test cohort 1 ($p < 0.001$) and 0.596 (95% CI 0.511–0.682) in external test cohort 2 ($p < 0.001$). The model constructed by direct concatenation of radiological and pathological features (iRPM) showed no significant improvement over single modality, with AUROC values of 0.784 (95% CI 0.714–0.855) in external test cohort 1 and 0.833 (95% CI 0.772–0.894) in external test cohort 2. Furthermore, iSCLM demonstrated superior performance in the prospective cohort, with an AUROC of 0.846 (95% CI 0.780–0.911, Figure 2C), which was better than that of iRM ($p = 0.007$), iPM ($p < 0.001$), and iRPM ($p < 0.001$).

To evaluate the reproducibility of iSCLM, a trainee resident with 2 years of experience in oncology, blinded to both pathological results and previous ROIs, was asked to reannotate tumor regions and retest the results of the external test set. iSCLM exhibited stable performance, with AUROC values of 0.834 (95% CI 0.764–0.904) in external test cohort 1 and 0.851 (95% CI 0.789–0.912) in external test cohort 2 (both $p > 0.05$). Corresponding intraclass correlation coefficient values were 0.890 and 0.848, respectively (both $p < 0.001$).

Additionally, we evaluated the enhancement in iSCLM performance using the Shapley additive explanations (SHAP)[26] technique. We observed substantial changes attributable to

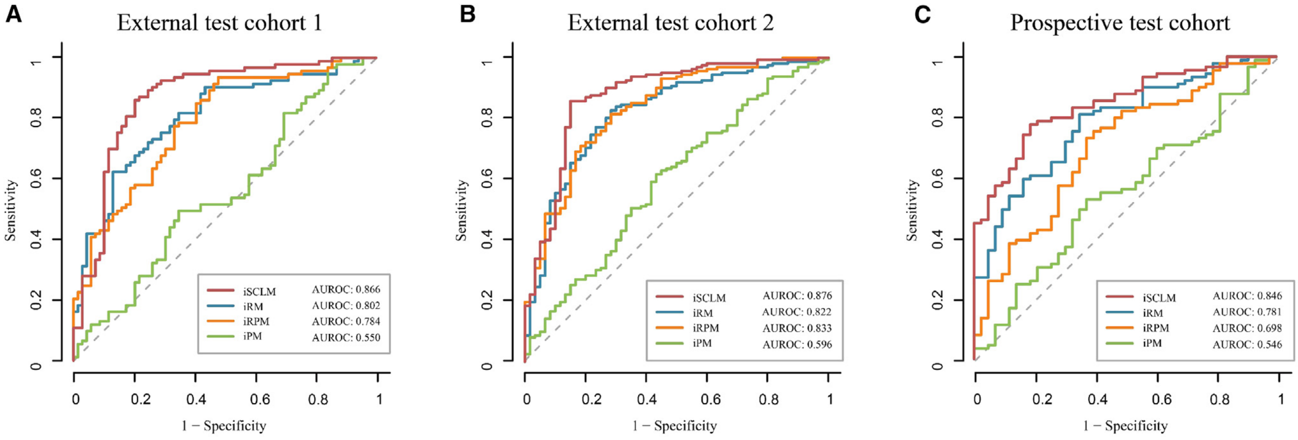


Figure 2. Prediction performance in the external and prospective test cohorts

Receiver operating characteristic curve (ROC) of iSCLM, iRM, iRPM and iPM for predicting tumor response to neoadjuvant chemotherapy. (A) External test cohort 1, (B) external test cohort 2, (C) prospective test cohort. iSCLM, incremental supervised contrastive learning model; iRM, incremental radiology model; iRPM, model constructed by direct concatenation of radiological and pathological features with incremental learning; iPM, incremental pathology model.

supervised contrastive learning in the feature representation of CT images (Figure S3) based on Moran's Index ($p < 0.001$). The SHAP values illustrated on the CT scans, represented as blue (negative) regions and red (positive) regions on neoadjuvant chemotherapy response prediction, demonstrated that supervised contrastive learning utilization enhanced the model's ability to identify critical features in these regions. As shown in Figure S3, the SHAP values for tumor regions in the CT scans of patients A and B shifted from negative in iRM predictions to positive in iSCLM predictions following supervised contrastive learning application.

iSCLM focuses on the tumor-invasive border and inflammatory cell proportions

We examined the top-ranked tiles identified by the GAT cluster from the tumor-invasive border with supervised contrastive learning (Figure 3A). As shown in Table S2, the top iSCLM regions in the tumor side exhibited significantly shorter distances to the tumor-invasive border relative to the average distances of all tiles. Although the top 5 tiles from the iPM exhibited attention to the tumor-invasive border ($p = 0.013$), iSCLM consistently focused more on the tumor-invasive border across the top 5 to top 20 tiles ($p < 0.05$). Furthermore, iSCLM exhibited greater attention compared with iPM among the top 5 tiles ($p < 0.001$).

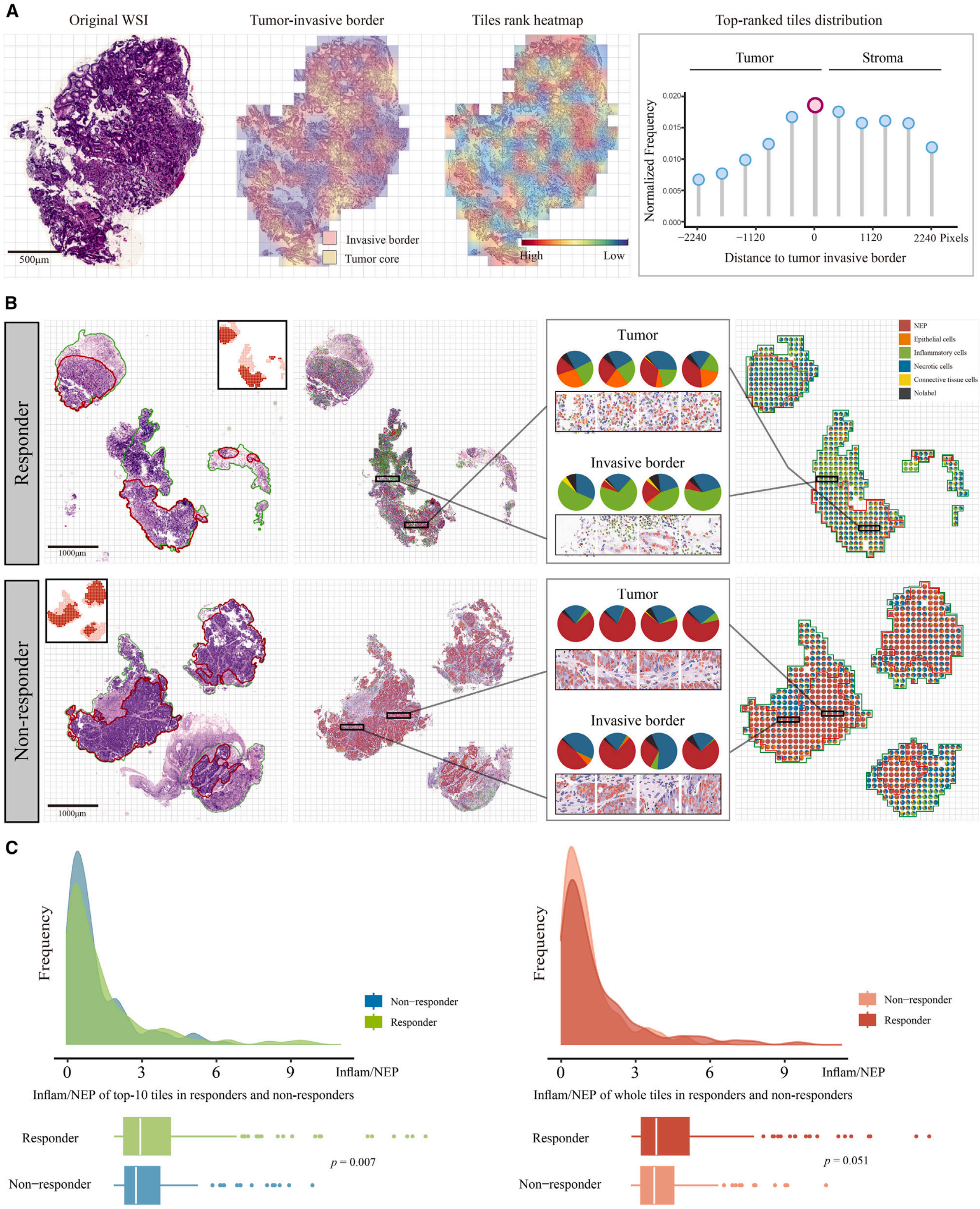
To further explore the biological variances in pathology images within the tumor-invasive border between responder and non-responder groups, we employed the HoVer-Net[27–29] model to identify the cellular content of tiles (Figure 3B). Following standardization of these contents based on the amount of neoplastic epithelial cells (NEPs), we observed significant differences in the ratio of inflammatory cells/NEP ($p = 0.007$, Figure 3C), connective tissue cells/NEP ($p = 0.013$), and necrotic cells/NEP ($p = 0.027$) from the top 10 ranked tiles by iSCLM between patient subgroups. Meanwhile, no significant distinctions were found in the ratio of non-neoplastic epithelial cells/NEP ($p = 0.300$, Table S3). A validation experiment aimed at evaluating the reproducibility of the HoVer-Net results suggested strong consistency with that of manual counting and another automatic cell classification method (TSFD-Net)[30] (NEP: HoVer-Net vs. manual $r = 0.753$; HoVer-Net vs. TSFD-Net $r = 0.713$; inflammatory cells: HoVer-Net vs. manual $r = 0.749$; HoVer-Net vs. TSFD-Net $r = 0.660$; Figure S4). Moreover, HoVer-Net demonstrated greater consistency with manual annotations than TSFD-Net (Figure S4).

To validate cell differences identified by iSCLM associated with the response to neoadjuvant chemotherapy, we analyzed 35 RNA-seq samples from 23 responders and 12 non-responders, using single-sample gene set enrichment analysis (ssGSEA) to assess cell type contents. The findings indicated increased dendritic cell infiltration ($p = 0.005$, Figures 4A and 4D; Table S4). The results of gene set variation analysis (GSVA) based on Gene Ontology showed significant activation of dendritic cell-related pathways (Figure 4B). To enhance interpretability, we compared the expression levels of inflammatory cell-related biomarkers and observed elevated CD11c expression in responders ($p = 0.040$, Figures 4C, 4E, and S5). Additionally, IHC for seven biomarkers of inflammatory cells (i.e., CD4, CD8, CD11c, CD20, CD56, CD68, and CD163) from 50 endoscopic biopsy samples (from 28 responders and 22 non-responders) who underwent neoadjuvant chemotherapy (Figures 4F and S6) revealed significantly increased CD11c expression ($p = 3.49\text{e}{-03}$, Figure 4G) and decreased CD163 expression ($p = 0.031$, Figure 4H) in responders; these are the primary markers for dendritic cells and M2 macrophages, respectively. Specifically, in the tumor-invasive border (Figure S6B), the differences in CD11c expression were pronounced ($p = 4.72\text{e}{-04}$, Figures 4G and S6C). Multiplex IHC corroborated these observations through the verification of CD4, CD8, CD11c, and CD163 staining patterns in the prospective cohort (Figures 4I and 4J).

iSCLM corrects misclassifications made by iRM and iPM in the prospective cohort

In the prospective cohort, iSCLM corrected the predictions made by iRM and iPM, improving performance of neoadjuvant

CellPress
OPEN ACCESS



(legend on next page)

chemotherapy response prediction (both $p < 0.05$, Table 2). As illustrated in Figure 5A for patient 1, the prominent CT features of the tumor region included significant gastric wall thickening and an irregular mucosal surface. However, the iRM incorrectly classified this patient as a non-responder. As shown on the attention map of iRM, when not reinforced by pathological WSI features, it failed to accurately identify regions that positively contribute to predicting the patient's response to neoadjuvant chemotherapy (blue regions on the attention map). The use of iSCLM corrected the prediction, accurately identifying the patient as a responder. After incorporating pathological knowledge of the tumor-invasive border and differences in inflammatory cells, the model highlighted abnormal regions of the gastric wall in red on the attention map, which may explain the corrected prediction and improved model performance. Similarly, in Figure 5B, patient 2 was initially incorrectly misclassified as a non-responder by iPM. The use of iSCLM corrected the prediction. iSCLM identified lesion characteristics on CT (indicated by red regions on the attention map of iSCLM), and the average distance between the top 10 ranked pathology tiles and the tumor-invasive border was significantly reduced (from 756 pixels using iPM to 597 pixels using iSCLM).

Prognostic value of iSCLM and subsequent real-world impact

Next, we assessed the prognostic value of the iSCLM for patient treatment outcomes. Univariate and multivariate analyses indicated that patients predicted as responders to neoadjuvant chemotherapy based on iSCLM demonstrated improved overall survival compared with those predicted as non-responders ($p < 0.001$, Figure S7B; Table S5).

To simulate the real-world impact of model adoption on survival, an evaluation dataset was compiled including an additional 394 patients with $\geq$cT2 locally advanced gastric cancer recommended for neoadjuvant chemotherapy based on NCCN guidelines, who had proceeded directly to surgery without receiving neoadjuvant chemotherapy treatment (untreated group). iSCLM was used to perform simulated predictions for the untreated group. Along with the patients who had received neoadjuvant chemotherapy in the present study (treated group), all patients were grouped into four categories: patients with locally advanced gastric cancer who were predicted to respond favorably to neoadjuvant chemotherapy and subsequently received neoadjuvant chemotherapy (predicted responder-treated group), patients who were predicted to benefit from neoadjuvant chemotherapy but did not receive it (predicted responder-untreated group), patients who were predicted to be non-responsive to neoadjuvant chemotherapy and did not receive the treatment (predicted non-responder-untreated group), and patients who were predicted to be non-responsive to neoadjuvant chemotherapy but received the treatment (predicted non-responder-treated group). Owing to differences in baseline characteristics between patients who received neoadjuvant chemotherapy and those who did not, propensity score matching was performed to adjust for potential biases. Matching variables included patient age, T stage, and N stage. After propensity score matching, the predicted responder-treated group demonstrated the most favorable outcome (5-year survival rate: 53.6%, Figure S7C). The predicted non-responder-treated group exhibited worse outcomes (5-year survival rate: 25.9% vs. 43.4%, $p = 0.027$ compared with the predicted non-responder-untreated group, Figure S7C).

Subsequently, we evaluated survival differences between patients who chose whether to undergo neoadjuvant chemotherapy based on the model's recommendation and those who made this decision without the model's guidance. The aforementioned four groups were further stratified into two distinct categories: those who adhered to the iSCLM-recommended neoadjuvant chemotherapy (including the predicted responder-treated and predicted non-responder-untreated groups) and those who did not follow the iSCLM-recommended neoadjuvant chemotherapy (including the predicted responder-untreated and predicted non-responder-treated groups). Accepting the treatment decisions based on iSCLM recommendations demonstrated improved overall survival (5-year survival rate for those accepting vs. those rejecting iSCLM recommendations: 48.6% vs. 35.5%, $p = 0.005$, Figure S7D). Results of the stratified survival analysis by clinical stage are presented in Figures S7E and S7F.

DISCUSSION

Developing interpretable AI-enabled predictions for neoadjuvant chemotherapy in patients with locally advanced gastric cancer is crucial for enhancing clinical decision-making concerning patients. In this study, we developed a multi-modal approach to predict neoadjuvant chemotherapy responses and survival outcomes based on a large-scale comprehensive dataset in China. Employing a multi-modal iSCLM strategy mutually reinforced feature representation of CT and pathology images. Specifically, top-ranked pathology tiles used to determine responses to neoadjuvant chemotherapy were more enriched at the tumor-invasive border. Moreover, biological interpretations of these borders revealed significant differences in inflammatory cell components, especially in dendritic cells, among patients with distinct neoadjuvant chemotherapy responses.

Although CT images have been reported to be useful in predicting neoadjuvant chemotherapy responses in locally advanced gastric cancer, with AUROCs ranging from 0.679 to 0.860 in external test cohorts,[19–21] a persistent need for improved accuracy and validation in prospective cohorts remains. Similarly, while previous pathology image prediction models have shown promise,[22] potential performance

Figure 3. Spatial analysis of model attention and cell component analysis

(A) Original whole-slide image (WSI), manually annotated tumor-invasive border, tiles rank heatmap (left), and distribution of top-ranked tiles (right).

(B) Manual annotation (left), cell classification generated by HoVer-Net (middle), and pie chart of cell classification in each tile (right) for two typical patients, one responder and one non-responder. Examples of the cell components in the tumor region and invasive border are demonstrated (middle).

(C) Inflam/NEP ratio of the top-10 tiles (left) and the whole tiles (right) in responders and non-responders. The boxplot summarizes data using the median and interquartile range to display its distribution and variability. Inflam, inflammatory cells; NEP, neoplastic epithelial cells.

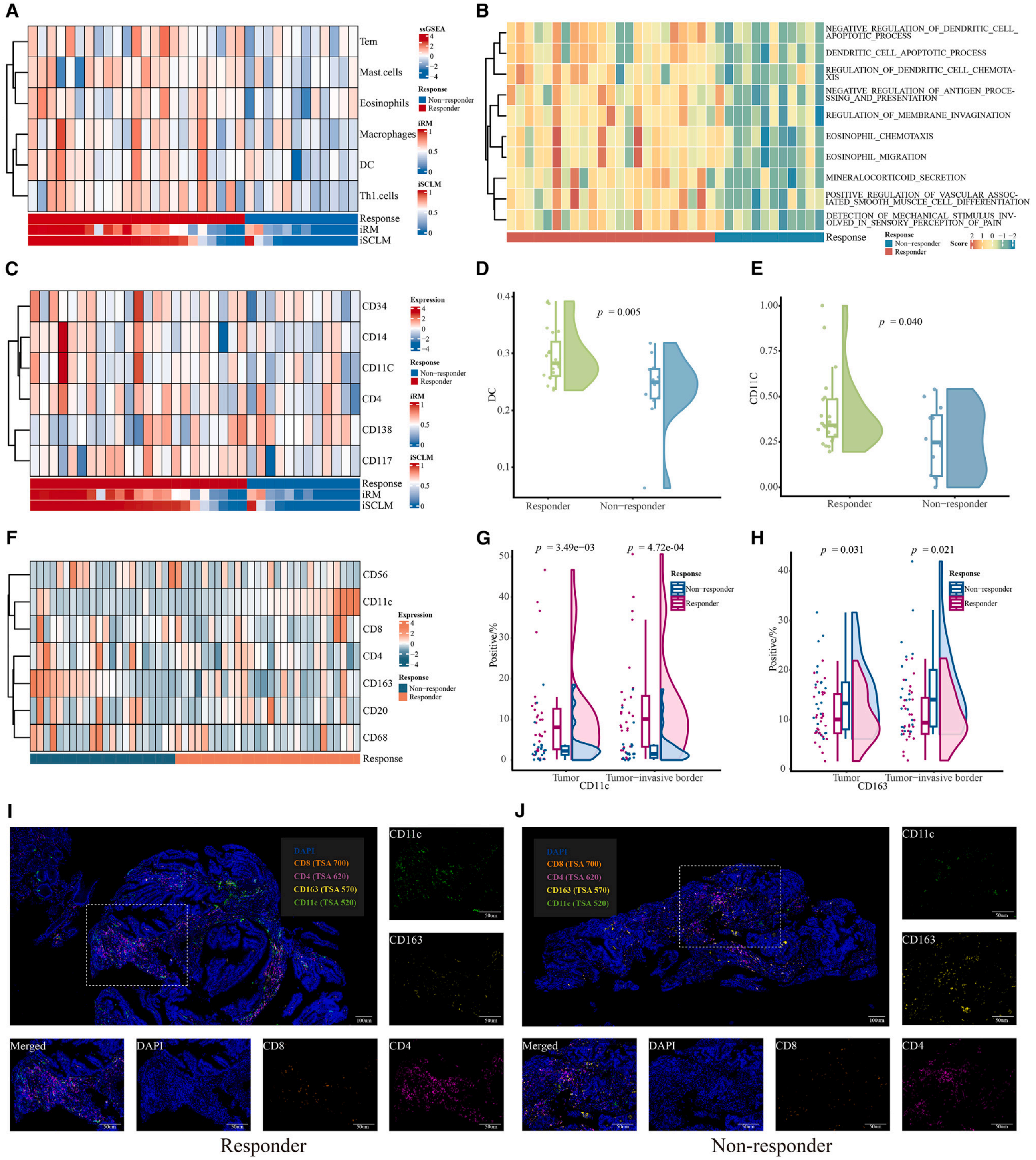


Figure 4. RNA-seq and IHC results

(A) Heatmap plot presenting cell composition based on ssGSEA of RNA-seq.

(B) Heatmap plot presenting the activity of pathways identified using RNA-seq. The colors in the heatmap correspond to the activity scores of each pathway, as quantified through gene set variation analysis.

(C) Heatmap plot depicting the expression levels of inflammatory cell-related biomarkers derived from RNA-seq.

(D) Raincloud plot illustrating the differential enrichment of dendritic cells between responders and non-responders as analyzed using ssGSEA of RNA-seq.

(E) Raincloud plot displaying the variation in CD11c expression between responders and non-responders, based on RNA-seq data.

(F) Heatmap plot illustrating variations in cell marker expression between responders and non-responders, evaluated using IHC.

(legend continued on next page)

enhancements from integrating information from multiple modalities remain underexplored. Therefore, we employed a supervised contrastive learning approach that integrated knowledge from radiology and pathology to enhance image feature representation. According to the validation procedures conducted on retrospective and prospective multicenter cohorts across research centers from various regions of China, our model demonstrated significant performance improvements over single-modality models and the model constructed by directly concatenating features. In addition, compared with NCCN guidelines,[11] which recommend neoadjuvant chemotherapy followed by surgery for patients with resectable diseases ≥cT2, our model indicated a 13.1% divergence in 5-year survival rate when accepting treatment decisions (higher than those who did not comply with the treatment decisions) according to predictive results. These results imply that employing iSCLM-based personalized guidance can potentially confer significant survival benefits to patients.

Through the application of SCLMs that utilize multi-modal images, our proposed approach synergistically integrated complementary information from both radiology and pathology. This collaborative strategy enhanced the extraction of features related to the neoadjuvant chemotherapy response, thereby improving accuracy. The distinct manifestations of macroscopic and microscopic tumors in radiology and pathology, respectively, highlight the inherent correlations between the two modalities. This integration ensures that radiology and pathology mutually reinforce each other, enhancing feature extraction. Hence, supervised contrastive learning can be used to align features between modalities when extracting global representations and characterizing those associated with neoadjuvant chemotherapy. With the assistance of supervised contrastive learning, our results demonstrate a significant improvement in the focus of the model on key pathological features with the help of CT feature representations. Specifically, the GAT model exhibited a propensity for top-ranked tiles to predominantly focus on the tumor-invasive border, which is crucial for predicting responses to neoadjuvant therapy.

Higher levels of inflammatory cell contents were observed within the tumor-invasive border in patients who exhibited a better response to neoadjuvant chemotherapy than in others, which explains the underlying biological processes. Additional RNA-seq and IHC analyses demonstrated that a significant difference in the composition of immune cells, particularly in CD11c expression, was observed between patients with distinct responses to neoadjuvant chemotherapy. These findings are consistent with previously reported results showing that tumor-infiltrated immune cells were significantly related to neoadjuvant chemotherapy responses in tumors (such as those in breast and colorectal cancers).[31–34] Specifically, in human colorectal cancer with liver metastases, cases with immune cell infiltration in the invasive margin indicated better responses to neoadjuvant chemotherapy.[35] These findings are also corroborated by additional studies, which similarly indicated the distinct impact of spatial distribution on the prognosis and chemotherapy outcomes.[36,37] Notably, Kim et al. observed alterations in CD11c$^+$ dendritic cell populations before and after neoadjuvant chemotherapy in patients with gastric cancer.[38] Moreover, Zhao et al. reported that CD163$^+$ macrophages decreased after neoadjuvant chemotherapy in a responder group with gastric cancer.[39] Our comparative experiments showed that the proposed model exhibited reduced attention on the tumor-invasive border found in pathology images without supervised contrastive learning. The aforementioned findings showed that iSCLM can identify human-interpretable differences in inflammatory cells at the tumor-invasive border, potentially improving the accuracy of clinical decision-making.

In conclusion, we developed an iSCLM to accurately predict the efficacy of neoadjuvant chemotherapy, with the overarching aim of improving survival outcomes. We discovered that CT and pathology feature representations were mutually reinforced using the proposed supervised contrastive learning strategy, and intrinsic pathological features contributing to improved model prediction could be attributed to inflammatory cell contents at the tumor-invasive border. Collectively, our findings potentially contribute to the advancement of clinical practices in screening patients with locally advanced gastric cancer for neoadjuvant chemotherapy administration.

Limitations of the study

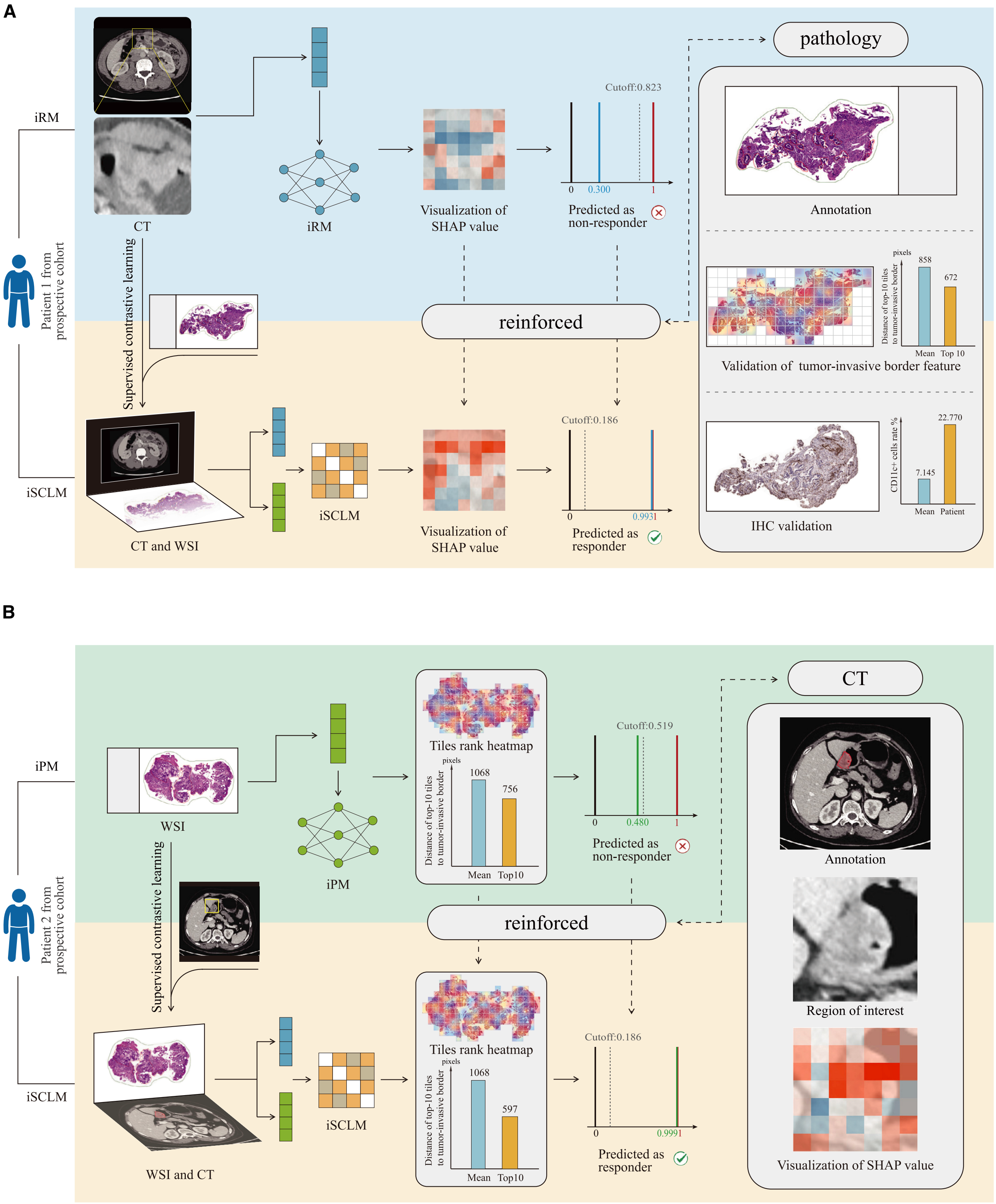
This study has several limitations: Firstly, while efforts were made to elucidate the features associated with neoadjuvant chemotherapy responses in this study, the clinical effectiveness of tumor treatment is influenced by various factors, necessitating future multi-omics analyses. Secondly, due to the lack of publicly available datasets for validation, further validation across multinational populations is essential to assess the generalizability of our findings and methodology. Thirdly, the study utilized manual annotation instead of automatic segmentation strategies due to current technical limitations that hinder the effectiveness of automatic segmentation, thereby restricting clinical applications. As segmentation methods advance, we will attempt to incorporate automated segmentation. Fourthly, due to the high heterogeneity of biopsy samples, such as biopsy site, tissue size, and staining procedures, the performance of the single-modality model of pathology should be improved in the future. Lastly, emerging large language models have manifested potential for mining pathological features; thus, future studies should consider combining such technologies to enhance performance in future applications.

(G) Raincloud plot indicating the proportion of CD11c-positive cells in responders versus non-responders, as determined using IHC.

(H) Raincloud plot illustrating the proportion of CD163-positive cells in responders and non-responders based on IHC. The raincloud plot combines an illustration of data distribution (the "cloud"), with jittered raw data (the "rain"). The boxplot summarizes data using the median and interquartile range to display its distribution and variability.

(I) Multiplex immunohistochemistry (mIHC) staining for CD4 (purple), CD8 (orange), CD163 (yellow), CD11c (green), and DAPI (blue) of a responder example.

(J) mIHC staining for CD4 (purple), CD8 (orange), CD163 (yellow), CD11c (green), and DAPI (blue) of a non-responder example. RNA-seq, RNA sequencing; IHC, immunohistochemistry; ssGSEA, single-sample gene set enrichment analysis; DC, dendritic cell.



(legend on next page)

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents and resources should be directed to and will be fulfilled by the lead contact, Dr. Zhenning Wang (znwang@cmu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The original CT images, H&E-stained WSIs, and clinical data that support the findings of this article are available upon reasonable request from the lead contact (Zhenning Wang: znwang@cmu.edu.cn). The sequencing data have been deposited at Genome Sequence Archive for Human (GSA-Human) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Owing to the presence of sensitive personal information of patients within the dataset, the data will not be publicly available. Applications submitted by researchers will be reviewed by the corresponding author, who will then convene with all authors to make a final decision. The code is available at https://github.com/PengGao-cmu/iSCLM. Any additional information required to reanalyze the data reported in this work paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This study was supported by the National Science and Technology Major Project (2023ZD0501500), the National Key Research and Development Program of China (2022YFA1105300), the National Natural Science Foundation of China (U23A20457, 92259104, 82303846, 82303372, and 82373110), the International Science and Technology Cooperation Project of Liaoning Province (2023JH2/10700005), the Liaoning Provincial Department of Science and Technology supporting high-quality development of China Medical University funding projects (2022JH2/20200072 and 2023JH2/20200061), Liaoning Provincial Department of Education Basic Research Project (LJ232410159010), the Shenyang Public Health Research and Development Medical Engineering Integration Project (22-321-32-20), National High Level Hospital Clinical Research Funding (2023-PUMCH-E-006), and the Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-009A).

AUTHOR CONTRIBUTIONS

Zhenning Wang, D.S., P.G., T.D., and Q.X. conceived the study. P.G., Q.X., H.T., J. Song, Y. Fu, and J.X. contributed to the methodology design. H.T., Zitong Wang, Yingjie Zhang, E.Y., T.X., J.M., H.C., and G.W. were involved in clinical data collection. P.G., G.W., X.J., Y.M., X.L., and Yong Zhang contributed to data annotation and response assessment. P.G., Q.X., H.T., J. Song, Y. Fu, and J.X. constructed and applied the artificial intelligence model. P.G., Q.X., J.Z., M.J., T.D., Y.J., Y. Feng, and Y. Zhao contributed to data analysis. Q.X., Zitong Wang, Yingjie Zhang, E.Y., T.X., and J.M. designed the figures and tables. Y. Fu, Yingjie Zhang, E.Y., and J.Z. performed the experiments. X.G., M.J., Y.L., L.H., J. Sun, Y.D., Y.B., and B.W. contributed to expert review. Zhenning Wang, X.G., M.J., Y.L., L.H., J. Sun, Y.D., Y.Y., Zhengyang Wang, M.Z., L.Z., B.W., Y. Zhao, and T.D. provided patient data and tissue used in this study. P.G., Q.X., H.T., and J. Song edited the first draft of the manuscript. Zhenning Wang, D.S., T.D., Yong Zhang, and Y. Zhao revised the manuscript. All authors contributed to manuscript preparation and data interpretation. All authors read and approved the final version of the manuscript. All authors had full access to the data in the study and undertake final responsibility for the decision to submit this manuscript for publication.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
    - ○ Ethical compliance
    - ○ Study design and participants
- METHOD DETAILS
    - ○ CT annotation
    - ○ Biopsy WSI annotation
    - ○ Model development
    - ○ Model evaluation
    - ○ Visualization of CT attention analysis
    - ○ Biological interpretations of iSCLM
    - ○ Biological validation with RNA-seq
    - ○ Biological validation with IHC
    - ○ Biological validation with multiplex IHC
    - ○ Survival analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2024.101848.

Received: August 28, 2024
Revised: October 15, 2024
Accepted: November 11, 2024
Published: December 4, 2024

REFERENCES

1. Smyth, E.C., Nilsson, M., Grabsch, H.I., van Grieken, N.C., and Lordick, F. (2020). Gastric cancer. Lancet (London, England) *396*, 635–648. https://doi.org/10.1016/s0140-6736(20)31288-5.
2. Joshi, S.S., and Badgwell, B.D. (2021). Current treatment and recent progress in gastric cancer. CA. Cancer J. Clin. *71*, 264–279. https://doi.org/10.3322/caac.21657.
3. Verschoor, Y.L., van de Haar, J., van den Berg, J.G., van Sandick, J.W., Kodach, L.L., van Dieren, J.M., Balduzzi, S., Grootscholten, C., IJsselsteijn, M.E., Veenhof, A.A.F.A., et al. (2024). Neoadjuvant atezolizumab plus chemotherapy in gastric and gastroesophageal junction adenocarcinoma: the phase 2 PANDA trial. Nat. Med. *30*, 519–530. https://doi.org/10.1038/s41591-023-02758-x.

Figure 5. Examples of correction attributed to changes in the focus areas of CT imaging and pathology in the prospective cohort

(A) A responder, with a true label of 1. iRM incorrectly provides a predicted value output of 0.300, while iSCLM corrects this to 0.993 after being reinforced by pathological feature representation. The CT attention map of representative images generated by Shapley additive explanations, which presents the last layer of Convn_x of ResNet-34 and helps explain this correction. IHC shows this patient has a high CD11c expression.

(B) A responder, with a true label of 1. iPM incorrectly provides a predicted value output of 0.480, whereas iSCLM predicts 0.999, reinforced by CT feature representation. The attention heatmap of iPM generated by global sort pooling helps explain this correction. The optimal threshold for prediction in different cohorts was determined using the Youden index in the development cohort. iSCLM, incremental supervised contrastive learning model; iRM, incremental radiology model; iPM, incremental pathology model; CT, computed tomography; WSI, whole-slide images.

4. Alsina, M., Arrazubi, V., Diez, M., and Tabernero, J. (2023). Current developments in gastric cancer: from molecular profiling to treatment strategy. Nat. Rev. Gastroenterol. Hepatol. *20*, 155–170. https://doi.org/10.1038/s41575-022-00703-w.

5. Cunningham, D., Allum, W.H., Stenning, S.P., Thompson, J.N., Van de Velde, C.J.H., Nicolson, M., Scarffe, J.H., Lofts, F.J., Falk, S.J., Iveson, T.J., et al. (2006). Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N. Engl. J. Med. *355*, 11–20. https://doi.org/10.1056/NEJMoa055531.

6. Das, M. (2017). Neoadjuvant chemotherapy: survival benefit in gastric cancer. Lancet Oncol. *18*, e307. https://doi.org/10.1016/s1470-2045(17)30321-2.

7. Sah, B.K., Zhang, B., Zhang, H., Li, J., Yuan, F., Ma, T., Shi, M., Xu, W., Zhu, Z., Liu, W., et al. (2020). Neoadjuvant FLOT versus SOX phase II randomized clinical trial for patients with locally advanced gastric cancer. Nat. Commun. *11*, 6093. https://doi.org/10.1038/s41467-020-19965-6.

8. Schuhmacher, C., Gretschel, S., Lordick, F., Reichardt, P., Hohenberger, W., Eisenberger, C.F., Haag, C., Mauer, M.E., Hasan, B., Welch, J., et al. (2010). Neoadjuvant chemotherapy compared with surgery alone for locally advanced cancer of the stomach and cardia: European Organisation for Research and Treatment of Cancer randomized trial 40954. J. Clin. Oncol. *28*, 5210–5218. https://doi.org/10.1200/jco.2009.26.6114.

9. Aoyama, T., Nishikawa, K., Fujitani, K., Tanabe, K., Ito, S., Matsui, T., Miki, A., Nemoto, H., Sakamaki, K., Fukunaga, T., et al. (2017). Early results of a randomized two-by-two factorial phase II trial comparing neoadjuvant chemotherapy with two and four courses of cisplatin/S-1 and docetaxel/cisplatin/S-1 as neoadjuvant chemotherapy for locally advanced gastric cancer. Ann. Oncol. *28*, 1876–1881. https://doi.org/10.1093/annonc/mdx236.

10. Guo, T., Tang, X.H., Gao, X.Y., Zhou, Y., Jin, B., Deng, Z.Q., Hu, Y., Xing, X.F., Li, Z.Y., and Ji, J.F. (2022). A liquid biopsy signature of circulating exosome-derived mRNAs, miRNAs and lncRNAs predict therapeutic efficacy to neoadjuvant chemotherapy in patients with advanced gastric cancer. Mol. Cancer *21*, 216. https://doi.org/10.1186/s12943-022-01684-9.

11. Ajani, J.A., D'Amico, T.A., Bentrem, D.J., Chao, J., Cooke, D., Corvera, C., Das, P., Enzinger, P.C., Enzler, T., Fanta, P., et al. (2022). Gastric Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. J. Natl. Compr. Canc. Netw. *20*, 167–192. https://doi.org/10.6004/jnccn.2022.0008.

12. Lordick, F., Carneiro, F., Cascinu, S., Fleitas, T., Haustermans, K., Piessen, G., Vogel, A., and Smyth, E.C.; ESMO Guidelines Committee. (2022). Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann. Oncol. *33*, 1005–1020. https://doi.org/10.1016/j.annonc.2022.07.004.

13. Wang, F.H., Zhang, X.T., Tang, L., Wu, Q., Cai, M.Y., Li, Y.F., Qu, X.J., Qiu, H., Zhang, Y.J., Ying, J.E., et al. (2023). The Chinese Society of Clinical Oncology (CSCO): Clinical Guidelines for the Diagnosis and Treatment of Gastric Cancer, 2023 (London, England: Cancer communications). https://doi.org/10.1002/cac2.12516.

14. Japanese Gastric Cancer Association (2023). Japanese Gastric Cancer Treatment Guidelines 2021 (6th edition). Gastric Cancer *26*, 1–25. https://doi.org/10.1007/s10120-022-01331-8.

15. Lu, M.Y., Chen, B., Williamson, D.F.K., Chen, R.J., Zhao, M., Chow, A.K., Ikemura, K., Kim, A., Pouli, D., Patel, A., et al. (2024). A Multimodal Generative AI Copilot for Human Pathology. Nature *634*, 466–473. https://doi.org/10.1038/s41586-024-07618-3.

16. Moor, M., Banerjee, O., Abad, Z.S.H., Krumholz, H.M., Leskovec, J., Topol, E.J., and Rajpurkar, P. (2023). Foundation models for generalist medical artificial intelligence. Nature *616*, 259–265. https://doi.org/10.1038/s41586-023-05881-4.

17. Rajpurkar, P., and Lungren, M.P. (2023). The Current and Future State of AI Interpretation of Medical Images. N. Engl. J. Med. *388*, 1981–1990. https://doi.org/10.1056/NEJMra2301725.

18. Zhang, G.J., and Zhou, Y.B. (2023). Artificial Intelligence and Machine Learning in Clinical Medicine. New England J. Med. *388*, 2398. https://doi.org/10.1056/NEJMc2305287.

19. Cui, Y., Zhang, J., Li, Z., Wei, K., Lei, Y., Ren, J., Wu, L., Shi, Z., Meng, X., Yang, X., and Gao, X. (2022). A CT-based deep learning radiomics nomogram for predicting the response to neoadjuvant chemotherapy in patients with locally advanced gastric cancer: A multicenter cohort study. EClinicalMedicine *46*, 101348. https://doi.org/10.1016/j.eclinm.2022.101348.

20. Hu, C., Chen, W., Li, F., Zhang, Y., Yu, P., Yang, L., Huang, L., Sun, J., Chen, S., Shi, C., et al. (2023). Deep learning radio-clinical signatures for predicting neoadjuvant chemotherapy response and prognosis from pretreatment CT images of locally advanced gastric cancer patients. Int. J. Surg. *109*, 1980–1992. https://doi.org/10.1097/js9.0000000000000432.

21. Wang, W., Peng, Y., Feng, X., Zhao, Y., Seeruttun, S.R., Zhang, J., Cheng, Z., Li, Y., Liu, Z., and Zhou, Z. (2021). Development and Validation of a Computed Tomography-Based Radiomics Signature to Predict Response to Neoadjuvant Chemotherapy for Locally Advanced Gastric Cancer. JAMA Netw. Open *4*, e2121143. https://doi.org/10.1001/jamanetworkopen.2021.21143.

22. Zhou, Z., Ren, Y., Zhang, Z., Guan, T., Wang, Z., Chen, W., Luo, T., and Li, G. (2023). Digital histopathological images of biopsy predict response to neoadjuvant chemotherapy for locally advanced gastric cancer. Gastric Cancer *26*, 734–742. https://doi.org/10.1007/s10120-023-01407-z.

23. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

24. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y.J.I. (2018). Graph attention networks. Stat *1050*, 10–48550. https://doi.org/10.48550/arXiv.1710.10903.

25. Li, Z., and Hoiem, D. (2018). Learning without Forgetting. IEEE Trans. Pattern Anal. Mach. Intell. *40*, 2935–2947. https://doi.org/10.1109/TPAMI.2017.2773081.

26. Lundberg, S.M., and Lee, S.-I. (2017). A unified approach to interpreting model predictions. Adv. Neural Inform. Process. Syst. *30*, 4765–4774. https://doi.org/10.48550/arXiv.1705.07874.

27. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., and Rajpoot, N. (2019). Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med. Image Anal. *58*, 101563. https://doi.org/10.1016/j.media.2019.101563.

28. Lu, M.Y., Chen, T.Y., Williamson, D.F.K., Zhao, M., Shady, M., Lipkova, J., and Mahmood, F. (2021). AI-based pathology predicts origins for cancers of unknown primary. Nature *594*, 106–110. https://doi.org/10.1038/s41586-021-03512-4.

29. Chen, R.J., Lu, M.Y., Williamson, D.F.K., Chen, T.Y., Lipkova, J., Noor, Z., Shaban, M., Shady, M., Williams, M., Joo, B., and Mahmood, F. (2022). Pan-cancer integrative histology-genomic analysis via multimodal deep learning. Cancer Cell *40*, 865–878.e6. https://doi.org/10.1016/j.ccell.2022.07.004.

30. Ilyas, T., Mannan, Z.I., Khan, A., Azam, S., Kim, H., and De Boer, F. (2022). TSFD-Net: Tissue specific feature distillation network for nuclei segmentation and classification. Neural Netw. *151*, 1–15. https://doi.org/10.1016/j.neunet.2022.02.020.

31. Fridman, W.H., Pagès, F., Sautès-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. Nat. Rev. Cancer *12*, 298–306. https://doi.org/10.1038/nrc3245.

32. Denkert, C., Loibl, S., Noske, A., Roller, M., Müller, B.M., Komor, M., Budczies, J., Darb-Esfahani, S., Kronenwett, R., Hanusch, C., et al. (2010). Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. J. Clin. Oncol. *28*, 105–113. https://doi.org/10.1200/jco.2009.23.7370.

33. Denkert, C., von Minckwitz, G., Darb-Esfahani, S., Lederer, B., Heppner, B.I., Weber, K.E., Budczies, J., Huober, J., Klauschen, F., Furlanetto, J.,

et al. (2018). Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. Lancet Oncol. *19*, 40–50. https://doi.org/10.1016/s1470-2045(17)30904-x.

34. Park, Y.H., Lal, S., Lee, J.E., Choi, Y.L., Wen, J., Ram, S., Ding, Y., Lee, S.H., Powell, E., Lee, S.K., et al. (2020). Chemotherapy induces dynamic immune responses in breast cancers that impact treatment outcome. Nat. Commun. *11*, 6175. https://doi.org/10.1038/s41467-020-19933-0.
35. Halama, N., Michel, S., Kloor, M., Zoernig, I., Benner, A., Spille, A., Pommerencke, T., von Knebel, D.M., Folprecht, G., Luber, B., et al. (2011). Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. Cancer Res. *71*, 5670–5677. https://doi.org/10.1158/0008-5472.can-11-0268.
36. Pagès, F., Mlecnik, B., Marliot, F., Bindea, G., Ou, F.S., Bifulco, C., Lugli, A., Zlobec, I., Rau, T.T., Berger, M.D., et al. (2018). International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet (London, England) *391*, 2128–2139. https://doi.org/10.1016/s0140-6736(18)30789-x.
37. Amgad, M., Hodge, J.M., Elsebaie, M.A.T., Bodelon, C., Puvanesarajah, S., Gutman, D.A., Siziopikou, K.P., Goldstein, J.A., Gaudet, M.M., Teras, L.R., and Cooper, L.A.D. (2024). A population-level digital histologic biomarker for enhanced prognosis of invasive breast cancer. Nat. Med. *30*, 85–97. https://doi.org/10.1038/s41591-023-02643-7.
38. Kim, R., An, M., Lee, H., Mehta, A., Heo, Y.J., Kim, K.M., Lee, S.Y., Moon, J., Kim, S.T., Min, B.H., et al. (2022). Early Tumor-Immune Microenvironmental Remodeling and Response to First-Line Fluoropyrimidine and Platinum Chemotherapy in Advanced Gastric Cancer. Cancer Discov. *12*, 984–1001. https://doi.org/10.1158/2159-8290.cd-21-0888.
39. Zhao, S., Liu, Y., Ding, L., Zhang, C., Ye, J., Sun, K., Song, W., Cai, S., He, Y., Peng, J., and Xu, J. (2024). Gastric cancer immune microenvironment score predicts neoadjuvant chemotherapy efficacy and prognosis. J. Pathol. Clin. Res. *10*, e12378. https://doi.org/10.1002/2056-4538.12378.
40. Fey, M., and Lenssen, J.E. (2019). Fast graph representation learning with PyTorch Geometric. arXiv preprint arXiv:1903.02428. https://doi.org/10.48550/arXiv.1903.02428.
41. Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., and Krishnan, D. (2020). Supervised contrastive learning. Adv, Neural Inform. Process. Syst. *33*, 18661–18673. https://doi.org/10.48550/arXiv.2004.11362.
42. Zhang, M., Cui, Z., Neumann, M., and Chen, Y. (2018). An End-To-End Deep Learning Architecture for Graph Classification. In Proceedings of the AAAI conference on artificial intelligence, *32*. https://doi.org/10.1609/aaai.v32i1.11782.
43. Bankhead, P., Loughrey, M.B., Fernández, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., et al. (2017). QuPath: Open source software for digital pathology image analysis. Sci. Rep. *7*, 16878. https://doi.org/10.1038/s41598-017-17204-5.
44. Amin, M., Edge, S., Greene, F., Byrd, D., Brookland, R., Washington, M., Gershenwald, J., Compton, C., Hess, K., Sullivan, D., et al. (2017). AJCC Cancer Staging Manual, 8th ed. (New York).
45. Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., et al. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur. J. Cancer *45*, 228–247. https://doi.org/10.1016/j.ejca.2008.10.026.
46. Borrelli, C., Roberts, M., Eletto, D., Hussherr, M.D., Fazilaty, H., Valenta, T., Lafzi, A., Kretz, J.A., Guido Vinzoni, E., Karakatsani, A., et al. (2024). In vivo interaction screening reveals liver-derived constraints to metastasis. Nature *632*, 411–418. https://doi.org/10.1038/s41586-024-07715-3.
47. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., and Thomas, N.E. (2009). A Method for Normalizing Histology Slides for Quantitative Analysis (IEEE), pp. 1107–1110.
48. Chen, R.J., Ding, T., Lu, M.Y., Williamson, D.F.K., Jaume, G., Song, A.H., Chen, B., Zhang, A., Shao, D., Shaban, M., et al. (2024). Towards a general-purpose foundation model for computational pathology. Nat. Med. *30*, 850–862. https://doi.org/10.1038/s41591-024-02857-3.
49. Ding, K., Zhou, M., Wang, H., Zhang, S., and Metaxas, D.N. (2022). Spatially aware graph neural networks and cross-level molecular profile prediction in colon cancer histopathology: a retrospective multi-cohort study. Lancet. Digit. Health *4*, e787–e795. https://doi.org/10.1016/s2589-7500(22)00168-6.
50. Lee, Y., Park, J.H., Oh, S., Shin, K., Sun, J., Jung, M., Lee, C., Kim, H., Chung, J.H., Moon, K.C., and Kwon, S. (2022). Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning. Nat. Biomed. Eng. *1*, 1. https://doi.org/10.1038/s41551-022-00923-0.
51. Drexler, R., Khatri, R., Sauvigny, T., Mohme, M., Maire, C.L., Ryba, A., Zghaibeh, Y., Dührsen, L., Salviano-Silva, A., Lamszus, K., et al. (2024). A prognostic neural epigenetic signature in high-grade glioma. Nat. Med. *30*, 1622–1635. https://doi.org/10.1038/s41591-024-02969-w.
52. Eastwood, M., Sailem, H., Marc, S.T., Gao, X., Offman, J., Karteris, E., Fernandez, A.M., Jonigk, D., Cookson, W., Moffatt, M., et al. (2023). MesoGraph: Automatic profiling of mesothelioma subtypes from histological images. Cell Rep. Med. *4*, 101226. https://doi.org/10.1016/j.xcrm.2023.101226.
53. Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2018). How powerful are graph neural networks? arXiv preprint arXiv:1810.00826. https://doi.org/10.48550/arXiv.1810.00826.
54. Kipf, T.N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. https://doi.org/10.48550/arXiv.1609.02907.
55. Moran, P.A. (1950). Notes on continuous stochastic phenomena. Biometrika *37*, 17–23. https://doi.org/10.2307/2332142.
56. Zormpas, E., Queen, R., Comber, A., and Cockell, S.J. (2023). Mapping the transcriptome: Realizing the full potential of spatial data analysis. Cell *186*, 5677–5689. https://doi.org/10.1016/j.cell.2023.11.003.
57. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinform. *14*, 7. https://doi.org/10.1186/1471-2105-14-7.
58. D'Andrea, E., Wexler, D.J., Kim, S.C., Paik, J.M., Alt, E., and Patorno, E. (2023). Comparing Effectiveness and Safety of SGLT2 Inhibitors vs DPP-4 Inhibitors in Patients With Type 2 Diabetes and Varying Baseline HbA1c Levels. JAMA Intern. Med. *183*, 242–254. https://doi.org/10.1001/jamainternmed.2022.6664.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD11c	Abcam	Cat# ab52632; RRID:AB_2129793
CD4	Abcam	Cat# ab133616; RRID:AB_2750883
CD68	Abcam	Cat# ab955; RRID:AB_307338
CD163	Abcam	Cat# ab182422; RRID:AB_2753196
CD56	Cell Signaling Technology	Cat# 99746; RRID:AB_2868490
CD8	Cell Signaling Technology	Cat# 85336; RRID:AB_2800052
CD20	Cell Signaling Technology	Cat# 48750; RRID:AB_3107071
Deposited data		
Human RNA sequencing raw data	This paper	HRA008712 in the China National Center for Bioinformation (https://ngdc.cncb.ac.cn/gsa-human/)
Software and algorithms		
iSCLM	This paper; Github	https://github.com/PengGao-cmu/iSCLM
Python 3.7	Python	https://www.python.org
Pytorch 1.11.0+cu113	Pytorch	https://pytorch.org
Pytorch-geometric 2.0.4	Fey et al.[40]	https://github.com/pyg-team/pytorch_geometric/
Incremental learning	Li et al.[25]	https://github.com/ngailapdi/LWF
Supervised contrastive learning	Khosla et al.[41]	https://github.com/google-research/google-research/tree/master/supcon
Shapley additive explanations	Lundberg et al.[26]	https://github.com/shap/shap
Graph attention network	Veličković et al.[24]	https://github.com/PetarV-/GAT
Global sort pooling	Zhang et al.[42]	https://github.com/muhanzhang/pytorch_DGCNN
ResNet	He et al.[23]	https://github.com/pytorch/vision/blob/main/torchvision/models
HoVer-Net	Graham et al.[27]	https://github.com/simongraham/hovernet_inference
TSFD-Net	Ilyas et al.[30]	https://github.com/Mr-TalhaIlyas/TSFD
R 4.3.1	The R Project for Statistical Computing	https://www.r-project.org
ITK-SNAP 3.6	ITK-SNAP Software	https://www.itksnap.org
Qupath v.0.4.3	Bankhead et al.[43]	https://qupath.github.io/
PASS Version 15.0.5	NCSS Statistical Software	https://www.ncss.com/
Other		
NVIDIA Tesla A100 GPU (80GB)	Nvidia Corp., Santa Clara, California.	https://resources.nvidia.com/en-us-gpu-resources/nvidia-a100-datashee-1?lx=CPwSfP

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethical compliance

The study was conducted in accordance with the Declaration of Helsinki and was approved by the First Hospital of China Medical University Ethics Committee (retrospective study approvals [2022]367 and [2022]393, as well as prospective study approval [2022]122). The requirement for informed consent was waived for patients in the retrospective cohorts, since the study posed no more than minimal risk to patients and only pre-existing medical data were used. On the other hand, signed informed consent was obtained from all participants in the prospective cohort and from those who provided biological samples for RNA-seq, IHC, and multiplex IHC.

Study design and participants

We retrospectively collected data from 4284 patients with gastric cancer between August 2012 and December 2022 at 10 medical centers in the Southern, Northern, Western, Eastern, and Central regions of China (Figure S1; Table S6). Patients in retrospective cohorts received standard fluoropyrimidine-based chemotherapy regimens (e.g., XELOX, SOX, FOLFOX, and FLOT). The patients were divided into four distinct cohorts: the development cohort, external test cohort 1, external test cohort 2, and incremental cohort. Patients in the development and external test cohorts received neoadjuvant chemotherapy followed by curative surgery. Inclusion criteria were as follows: (1) Confirmed gastric adenocarcinoma via gastroscopy and histopathological biopsy before treatment initiation; (2) underwent standardized baseline contrast-enhanced CT on portal venous phase sequences before treatment initiation; and (3) locally advanced invasion beyond the submucosal layer and absence of distant metastasis, as determined by pretreatment CT. Exclusion criteria were as follows: (1) Not receiving standard fluoropyrimidine-based chemotherapy regimens; (2) poor CT image quality (e.g., inadequate gastric distension or artifacts) or blurry WSI; (3) absence of tumor regression grade (TRG) of postoperative pathology; (4) underwent previous gastric surgery; (5) presence of other malignancies before neoadjuvant therapy; or (6) receipt of anticancer therapy before baseline CT scans. The development cohort included 1208 patients. Within the development cohort, a subset of 774 patients were exclusively characterized based on CT images. These distinct subsets were systematically used for preliminary model pretraining. In addition, a subset of 434 patients with comprehensive profiles encompassing both CT and pathology WSI was used to build a supervised contrastive learning framework.

The incremental cohort included patients who received chemotherapy and presented with unresected disease ineligible for surgery, including conditions such as distant metastasis or locally unresected disease, in whom chemotherapy response was also evaluated. The inclusion criteria were as follows: (1) histologically confirmed gastric adenocarcinoma via gastroscopy before treatment initiation; (2) underwent standardized baseline contrast-enhanced CT before treatment initiation; and (3) distant metastases or locally unresectable lesions, as determined by pretreatment CT. The exclusion criteria were as follows: (1) not receiving standard fluoropyrimidine-based chemotherapy regimens; (2) poor CT image quality (e.g., inadequate gastric distension or artifacts) or blurry WSIs; (3) absence of post-chemotherapy contrast-enhanced CT scans for Response Evaluation Criteria in Solid Tumors (RECIST) evaluation; (4) underwent curative surgery after chemotherapy completion; (5) underwent previous gastric surgery; (6) presence of other malignancies before neoadjuvant therapy; or (7) receipt of anticancer therapy before baseline CT scans.

Subsequently, we conducted a multicenter, prospective, real-world study on the prediction of neoadjuvant chemotherapy response for locally advanced gastric cancer (Method S1). In total, 164 patients were recruited from the First Hospital of China Medical University and Tianjin Medical University Cancer Institute and Hospital between March 1, 2023, and February 29, 2024. The registration number is ChiCTR2300068917. As this was a non-intervention study, there was no need for randomization. All patients underwent CT imaging and pathological biopsy within 1 month before neoadjuvant chemotherapy administration. The standard XELOX or SOX regimens were administered. The inclusion criteria for the prospective study were as follows: (1) provision of informed consent; (2) age $\geq$18 years; (3) confirmed gastric adenocarcinoma via histopathological biopsy and available WSI; (4) underwent standardized baseline contrast-enhanced CT; (5) locally advanced invasion beyond the submucosal layer and absence of distant metastasis, as determined by pretreatment CT; (6) absence of previous gastric surgery; (7) absence of other malignancies before neoadjuvant therapy; (8) absence of anticancer therapy before the baseline CT scans; and (9) absence of mental conditions or impairment. The exclusion criteria for the prospective study were as follows: (1) receipt of neoadjuvant chemotherapy regimens other than XELOX and SOX; (2) failure of surgery after neoadjuvant chemotherapy; (3) absence of TRG of postoperative pathology; (4) poor CT image quality (e.g., inadequate gastric distension or artifacts) or blurry WSIs; or (5) death during the perioperative period.

Sample size calculations were performed before the prospective study. In clinical practice, no method has historically been accepted based on various guidelines[11–14] for predicting the response to neoadjuvant chemotherapy. A value of 0.750 is generally considered to indicate a model with certain predictive ability. In our previous retrospective study, the AUROC of the iSCLM was between 0.866 and 0.876, whereas the response rate after neoadjuvant chemotherapy was 66.4%. A sample of 76 from the responder group and 38 from the non-responder group demonstrated 80% power to detect a difference of 0.116 between the AUROC under the null hypothesis of 0.750 and an AUROC under the alternative hypothesis of 0.866 using a two-sided z-test at a significance level of 0.050. The calculation was performed using PASS Version 15.0.5 with the Tests for One ROC Curve module.

The response to neoadjuvant therapy was generally evaluated according to the TRG, with evaluation criteria based on the eighth edition of the American Joint Committee on Cancer (AJCC) Staging Manual.[44] Briefly, TRG 0, defined as a complete response, refers to the absence of viable cancer cells, including lymph nodes. TRG 1, defined as near-complete response, is characterized by the presence of single cells or small groups of cancer cells. TRG 2, defined as partial response, is defined as the presence of residual cancer cells, with evident tumor regression beyond single cells or small groups of cancer cells. Finally, TRG 3, defined as poor or no response, refers to the presence of extensive residual cancer with no evident tumor regression (Figure S8A). Pathology experts reconfirmed the TRG criteria during the study. The TRG was assessed by two pathologists (with 20 and 10 years of experience in gastrointestinal pathology diagnosis, respectively) who were blinded to the study data. The results were examined by a pathologist (with 25 years of experience in gastrointestinal pathology diagnosis). Patients with TRG 0–2 were considered responders, whereas those with TRG 3 were defined as non-responders.

For patients who were not initially eligible for surgery, the response to chemotherapy was evaluated after completing at least two cycles of first-line chemotherapy within the initial 3 months of treatment initiation based on RECIST v.1.1[45] (Figure S8B). RECIST was evaluated by a senior radiologist (with 8 years of experience in abdominal imaging) and an oncologist (with 15 years of experience in abdominal imaging). If any discrepancies were found in the two researchers' judgments, a third expert (with 15 years of experience in abdominal imaging) re-evaluated the CT scan to make a final decision. Patients with complete response (CR) and partial response (PR) were considered responders, whereas those with progressive disease (PD) and stable disease (SD) were considered non-responders.

METHOD DETAILS

CT annotation

All patients underwent contrast-enhanced CT examinations within 1 month before starting chemotherapy. Patients enrolled at the 10 centers underwent contrast-enhanced CT scans with similar settings but using different systems and parameters (Table S7). To specify the ROIs, three experts with 8–15 years of experience manually delineated the primary lesions and lymph nodes on the pretreatment CT images of portal venous phases for all patients by using ITK-SNAP (www.itksnap.org, version 3.6, University of Pennsylvania, PA, USA). Initially, a senior radiologist (with 8 years of experience in abdominal imaging) and an oncologist (with 15 years of experience in abdominal imaging) reviewed all CT slices. The ROIs were manually labeled along the border of the primary lesion on the center image slice with the largest tumor area. The largest short-axis slice was chosen for annotating the lymph node, and no delineation was performed if no foci were observed in lymph nodes. Figure S9A provides details on manual segmentation. The experts were blinded to histopathological results. If the lesions remained indeterminate after the initial evaluation or if discrepancies occurred in the two researchers' judgments, a third expert (with 15 years of experience in abdominal imaging) re-evaluated the CT scan to make a final decision. To verify the model's reproducibility, another trainee resident (with 2 years of experience in oncology) used the same method to independently annotate the CT scans of the external test set to assess the effects of interobserver bias.

Biopsy WSI annotation

Each patient underwent gastroscopy and gastric biopsy before chemotherapy. Formalin-fixed paraffin-embedded tissues from the biopsies were used to prepare H&E-stained slides, which were subsequently digitized into WSIs. Pathologists assessed all digitized slides to check their quality and appropriateness for analysis. Images containing blurry regions were rescanned, and the slides with insufficient tumor tissue (<5 mm^2) were excluded. To explore the spatial characteristics of the microenvironment, the tumor region and tumor-related stroma region were annotated separately by two experienced pathologists (with 20 and 10 years of experience in gastrointestinal pathology diagnosis, respectively) using Qupath v.0.4.3 (https://qupath.github.io/).[43,46] The tumor regions were outlined independently with small, closed circles, whereas the surrounding areas of the tumor-related stroma were outlined with different colors (Figure S9B). A third senior professor (with 25 years of experience in gastrointestinal pathology diagnosis) examined these annotations to ensure accuracy.

Model development

Image data preprocessing

We utilized the ROIs of the primary lesion and lymph nodes in a CT scan of the portal venous phase after normalization with a soft tissue window of 350 HU and a level of 50 HU. A rectangular frame was constructed based on the exact contour of the lesion and enlarged in both height and width with an interval of 10 pixels. The rectangular frames were resized to 224 × 224 pixels and normalized to serve as the network input. To address the variation in image intensity, we adopted a data augmentation strategy (i.e., random horizontal flip) during the training procedure.

Using annotations from the pathologists, the WSI was divided into tumor and stromal regions. At a magnification of 40× (0.25 μm/pixel), the WSIs were segmented into nonoverlapping square tiles with an edge length of 448 × 448 pixels. Tiles with blank areas exceeding 90% were excluded. All tiles underwent color normalization using Macenko's[47,48] technique to mitigate color bias and streamline model training and assessment. Furthermore, tiles were resized to 224 × 224 pixels to serve as the model input. All tumor and stromal tiles from the WSI were used for tile-connected graph development.

SCLM construction

Our model was designed to predict patient responses to neoadjuvant chemotherapy by employing supervised contrastive learning with features extracted from CT and pathology images. We employed a deep residual neural network (ResNet-34[23]) to extract 1000 CT features of the primary lesion and lymph nodes, respectively. The features were then concatenated, and a fully connected layer with 256 neurons was used. Inspired by the spatial capability of tumor slides,[49–52] we converted WSIs into spatially connected graph representations G (V, E), where each image tile represented a graph vertex set ($V = \{v_i, i \in N\}$). Edges ($E = \{e_{ij}, i, j \in N\}$) connecting adjacent nodes were created based on Euclidean distances, and the GAT[24] was subsequently employed to learn the attention score of each edge by aggregating neighborhood features to differentiate node weights. Pretrained ResNet-18[23] networks were used to extract the pathological features as node features. A two-layer GAT with two attention heads (64 dimensions per head in the first layer and 32 dimensions per head in the last layer) was used as the basic structure for tiles of all the patients. Using global average pooling as the READout function, node and edge information were integrated to represent the entire graph, and 64 graph-wise pathological features were extracted as the GAT output. Subsequently, these features were input into fully connected layers comprising 256 neurons. Constructing the graph in this way makes it possible to analyze the spatial relationships between different tiles in the WSI, which can be useful for identifying patterns and features associated with chemotherapy response.

For feature integration, we employed supervised contrastive learning[41] to map domain knowledge from pathology to CT (Figure S10). A framework combining supervised contrastive loss and cross-entropy loss was used to effectively learn feature representations from both CT and pathological images, achieving cross-modal understanding and reasoning capabilities. Contrastive learning allows intraclass features (i.e., those extracted for patients in the same class) to be clustered together, whereas inter-class features are pushed further away for patients from different classes. The aim of the model was to align the representations of radiological and pathological features in the feature space, aiding the model in the effective acquisition of feature representations from the above two image modals and serving as input to the fully connected layer for subsequent classification.

We delineated four distinct loss functions: L_{PR}, L_{RP}, L_{PP} and L_{RR}. In this context, L_{PR} and L_{RP} represent the domain knowledge mapping between pathology WSI and CT scans. L_{PP} signifies the knowledge mapping within the same category of pathology WSI, and L_{RR} indicates the knowledge mapping within the same category of CT scans. To provide a detailed description of the four isoform loss functions, we defined the following functions:

$$L_{\mathrm{AB}} = f(A,B) = \sum_{i \in \mathrm{A}} \frac{-1}{|g(B,i)|} \sum_{p \in g(B,i)} \log \frac{\exp\left(z_i \cdot z_p / \tau\right)}{\sum_{\alpha \in \mathrm{C}(i)} \exp\left(z_i \cdot z_\alpha / \tau\right)}$$

Here, A and B are sets of samples, respectively; $g(B,i) = \{p \in \mathrm{B(i)}: \tilde{y}_p = \tilde{y}_i\}$, $\mathrm{B(i)} \equiv \mathrm{B} \setminus \{\mathrm{i}\}$, $g(B,i)$ is the set of samples within B that share the same label as i, excluding i itself; $\mathrm{C(i)} \equiv \mathrm{C} \setminus \{\mathrm{i}\}$, C is the set of all multi-modal samples; z_x represents features extracted based on the encoder and sample x, the • symbol denotes the inner (dot) product, and τ is a scalar temperature parameter ($\tau = 0.1$).

Utilizing the aforementioned formula, the loss function was computed as follows: $L_{\mathrm{RP}} = f(R,P)$, $L_{\mathrm{PR}} = f(P,R)$, $L_{\mathrm{RR}} = f(R,R)$, $L_{\mathrm{PP}} = f(P,P)$. Here, CT imaging samples are denoted by R, and pathological samples are represented by P.

The loss function for SCL, L_{con}, was defined based on different combinations of loss functions (L_{PR}, L_{RP}, L_{PP} and L_{RR}). Six SCLM loss-combination variants were constructed accordingly. Three models were (SCLM-a to SCLM-c) built with radiology features extracted from only the primary lesion: SCLM-a ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{RP}} + L_{\mathrm{RR}} + L_{\mathrm{PP}}$), SCLM-b ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{PR}}$), and SCLM-c ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{PR}} + L_{\mathrm{RR}}$). Three models (SCLM-d to SCLM-f) were built with radiology features extracted from both the primary lesion and LNs: SCLM-d ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{PR}} + L_{\mathrm{RR}} + L_{\mathrm{PP}}$), SCLM-e ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{PR}}$), and SCLM-f ($L_{\mathrm{con}} = L_{\mathrm{RP}} + L_{\mathrm{PR}} + L_{\mathrm{RR}}$). Experimental comparisons were conducted to evaluate the training performance of these variants.

CT image features were inputted into a multilayer perceptron with cross-entropy loss. The extracted features were input into a classifier employed for the purpose of predictive classification. The cross-entropy loss was computed as loss function L_{cel} as follows:

$$L_{cel} = -\frac{1}{N} \sum_i \left[y_i \cdot \ln(p_i) + (1 - y_i) \cdot \ln(1 - p_i) \right]$$

Here, y_i represents the label of sample i, with a value of 1 denoting the positive class and 0 indicating the negative class. Concurrently, p_i signifies the probability assigned to sample i being predicted as belonging to the positive class.

The supervised contrastive and cross-entropy losses were combined as the final loss of the proposed network. The loss function L was computed as follows:

$$L = L_{con} + L_{cel}$$

Incremental SCLM construction

To enhance training stability, incremental learning[25] was employed to build the iSCLM, following SCLM establishment, enhancing the training of feature extractors for radiology (ResNet-34) and pathology (GAT). This approach was aimed to achieve continuous learning and adapt to new data while preserving previously learned knowledge.

In the first training, the model was trained to initialize use of the CT and WSIs of the incremental cohort. This was used to develop a classification model for predicting chemotherapy response in patients with unresectable tumors. In the second training, datasets pertaining to patients undergoing neoadjuvant chemotherapy were fed into the previously trained extractor. The model parameters were updated, and the response probability of the neoadjuvant patients was the output.

The extractors were trained with the dual objective of improving their performance and robustness. The primary objective was to enhance the predictive performance of the radiological or pathological extractors for neoadjuvant chemotherapy response. Loss_a is the cross-entropy loss for new tasks of predicting the response to neoadjuvant chemotherapy, and the loss encourages predictions $\hat{y}_n$ to be consistent with the ground truth y_n. Loss_a was computed as follows:

$$Loss_a = -\frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{l} y_n \cdot \log \hat{y}_n$$

where l is the number of labels, y_n denotes the one-hot ground truth label, and $\hat{y}_n$ denotes the output probability of the network. The second goal of the model is to ensure that the output probabilities for each image are close to the output from the original network of the first training. This approach facilitates the retention of previously acquired knowledge within the model, thereby enhancing the robustness and performance of the model itself during subsequent training iterations. We refined the model's predictions to align the output closely with the preceding model and make the output probabilities of each neoadjuvant chemotherapy patient in the second training close to the output from the first training. Loss_b was computed as follows:

$$Loss_b = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{l} {y'_0}^{(i)} \log \hat{y}_0'^{(i)}$$

$$y_0'^{(i)} = \frac{\left(y_0^{(i)}\right)^{1/T}}{\sum_j \left(y_0^{(j)}\right)^{1/T}}, \hat{y}_0'^{(i)} = \frac{\left(\hat{y}_0^{(i)}\right)^{1/T}}{\sum_j \left(\hat{y}_0^{(j)}\right)^{1/T}}$$

where $y_0^{(i)}$ and $\hat{y}_0^{(i)}$ denote recorded and current probabilities, and ${y'_0}^{(i)}$ and $\hat{y}_0'^{(i)}$ are the modified versions of recorded and current probabilities $y_0^{(i)}$, $\hat{y}_0^{(i)}$ ($T = 2$).

In the training of incremental mechanism, we defined the set of shared parameters of the extractor θ_s, task-specific parameters of classifiers for previously learned tasks in the first training θ_o, and task-specific parameters of classifiers for new tasks in the second training θ_n. We first trained θ_s and θ_o with samples from the incremental cohort. Then, we jointly trained all weights θ_s, θ_o, and θ_n until convergence to build the model based on data from the patients in the development cohort. $L(\theta)$ was optimized to train the radiology and pathology extractors using incremental learning.

$$L(\theta) = Loss_{\mathrm{a}} + Loss_{\mathrm{b}}$$

The parameters θ_s, θ_o, and θ_n were determined as follows:

$$\theta_s, \theta_o, \theta_n = \mathrm{argmin}\,(L(\theta))$$

We evaluated models built using a single modality: a radiological model (RM) based on CT, a pathology model (PM) based on WSI, and a radiological and pathological model (RPM) constructed using a direct concatenation of radiological and pathological features during validation. For the CT feature extractor, we compared the performance of radiomics and deep learning approaches. For the pathology feature extractor, we compared several architectures, including GAT, graph isomorphism network (GIN),[53] and graph convolutional network (GCN).[54] The model demonstrating the best performance was selected as the feature extractor. Furthermore, to integrate features from different modalities, we assessed six different SCLM loss-combination variants to determine the optimal training strategy.

The development cohort was divided into training and validation groups at a 9:1 ratio. The model was trained over 200 epochs and output a binary prediction of the neoadjuvant chemotherapy response. The batch size was set to 72 during each iteration, owing to memory constraints. The Adam optimizer was used with a learning rate of 1e−4 to update the model parameters using gradient descent. Specifically, the epoch with the highest AUROC (computed using the validation dataset) was selected for each model. Finally, iSCLM based on incremental learning was constructed as described above. Similar to that used in the iSCLM, we employed incremental learning to effectively train the RM, PM, and RPM. Below are the specific network frameworks for iRM, iPM, and iRPM (Figure S11).

iRM: We employed a deep residual neural network (ResNet-34) to independently extract 1000 features from the primary lesion and lymph nodes. The extracted features were then concatenated and subjected to dimensionality reduction via a fully connected layer, resulting in a 256-dimensional feature vector. This vector was subsequently input into another fully connected layer, culminating in the application of the softmax function to estimate the probability of response to neoadjuvant chemotherapy. Additionally, we conducted training on data from patients who received chemotherapy with unresected tumors and then employed incremental learning to improve the model's performance.

iPM: We extracted 64 graph-wise pathological features using a GAT. The extracted features were then fed into fully connected layers, resulting in a 256-dimensional feature vector. Subsequently, this vector was passed through an additional fully connected layer, and the softmax function was used to map the probability of response to neoadjuvant chemotherapy. Additionally, we employed incremental learning to enhance model performance.

iRPM: We concatenated the 256-dimensional feature vectors derived from both iRM and iPM, creating a comprehensive feature representation. This concatenated vector was then efficiently processed through fully connected layers, culminating in the application of the softmax function to accurately estimate the probability of response to neoadjuvant chemotherapy.

Model evaluation

Then, the iSCLM was compared with each model (iRM, iPM, and iRPM) based on incremental learning in the external and prospective test cohorts. The ground truth was defined as the actual pathological tumor response (patients with TRG 0–2 were considered responders, whereas those with TRG 3 were defined as non-responders) reported by pathologists following surgical excision. The AUROC was used to evaluate model performance. The optimal threshold for prediction in different cohorts was determined using the Youden index in the development cohort. Additionally, we conducted a stratified analysis to investigate whether the conclusions drawn from the study were applicable to patients with different histological types, tumor locations, and chemotherapy regimens.

Visualization of CT attention analysis

To visualize alterations among models with and without supervised contrastive learning, we used SHAP[26] to illustrate the attention map of ResNet-34 for chemotherapy response prediction on CT. Based on the "SHAP" package in Python, we obtained the importance of the features with interpretations on their contribution to the prediction. We used GradientExplainer to illustrate the model output with respect to each block of the pretrained ResNet-34 network, using 300 evaluations to obtain explanations. To investigate whether supervised contrastive learning influences the spatial autocorrelation of SHAP value matrix, we calculated the global Moran's index,[55,56] which has been used to elucidate the level of clustering and dispersion within spatial data. We computed the global Moran's index for the last layer of the last block of the pretrained ResNet-34 network on the external test cohort in both iRM and iSCLM.

Biological interpretations of iSCLM

To interpret the performance of iSCLM in predicting the neoadjuvant chemotherapy response in a biological context, we ranked the significance of pathology tiles from WSI for model output using global sort pooling[42] before and after employing supervised contrastive learning. The ranked results were visualized to assess the contribution of the tiles and their positions within the WSI. The results for each tile in the GAT of iPM and iSCLM are presented. The labels (stroma or tumor) of the adjacent tiles according to manual delineation, situated at the top, bottom, left, and right, were examined individually by an iterative process to spatially locate each tumor tile. If the central tile exhibited neighboring tiles containing stromal elements, it was defined as a tile within the tumor-invasive border of the tumor. Subsequently, the Euclidean distance from each tile in the tumor and stromal regions to the corresponding tumor-invasive border was calculated. For each tile, the minimum distance among these Euclidean distances was considered the distance from the tile to the tumor-invasive border. The tiles in the tumor region are denoted as t_i, those in the stromal region as s_j, and those in the tumor-invasive border region as b_k. The distance from each tile in the tumor region to the tumor-invasive border can be expressed as follows:

$$\text{distance}\ (t_i, b_k) = \min_{ik}\{\|t_i - b_k\|\}$$

The distance from each tile in the stromal region to the tumor-invasive border can be expressed as follows:

$$\text{distance}\ (s_j, b_k) = \min_{jk}\{\|s_j - b_k\|\}$$

The geographical distribution of the top-ranked tiles (i.e., model output by global sort pooling) was identified to measure its distance to tumor-invasive border.

Furthermore, we conducted a cell component analysis on the top-ranked tiles using HoVer-Net[27] to explore biological characteristics. We utilized HoVer-Net to classify various cellular components of all the tiles, including neoplastic epithelial cells, non-neoplastic epithelial cells, inflammatory cells, connective tissue cells, and necrotic cells. The counts of individual cell types in each tile were calculated. We then summarized the cellular composition of each slide and standardized it based on the content of neoplastic epithelial cells. We compared the differences in cellular content before and after supervised contrastive learning.

To confirm the precision of HoVer-Net in quantifying diverse cellular elements within tissues, we evaluated the consistency of HoVer-Net cell counts compared with manual counting and other automated cell classification tools. We randomly sampled 100 tiles and submitted them to a certified pathologist with 20 years of experience in pathology for manual enumeration of neoplastic epithelial cells and inflammatory cells, yielding more than 6486 categorized cells. In addition, we used TSFD-Net, a well-established deep-learning cell classifier, to identify the five distinct cellular components within these 100 tiles. Subsequently, we performed Spearman correlation analysis and inter-class correlation coefficient analysis to evaluate the agreement between the cell counts derived from HoVer-Net and those derived from these two approaches.

Biological validation with RNA-seq

RNA purification, reverse transcription, library construction, and sequencing were performed at Sequanta Technologies Co., Ltd. (Shanghai, China) according to the manufacturer's instructions. The captured coding regions of the transcriptome from total RNA were prepared using a TruSeq RNA Exome Library Preparation Kit (Illumina). Approximately 100 ng of total RNA (DV200 > 30%) was used as total RNA input. The RNA was then fragmented into small pieces using divalent cations at elevated temperature. cDNA was generated from the cleaved RNA fragments using random priming during first- and second-strand synthesis, and sequencing adapters were ligated to the resulting double-stranded cDNA fragments. To create the final library, coding regions of the transcriptome were captured using sequence-specific probes.

After library construction, a Qubit 3.0 fluorometer dsDNA HS Assay (Thermo Fisher Scientific) was used to quantify the concentration of the resulting sequencing libraries, and the size distribution was analyzed using an Agilent BioAnalyzer (Agilent). Sequencing was performed using an Illumina NovaSeq 6000 following the Illumina-provided protocols for 2 × 150 paired-end sequencing at Sequanta Technologies Co., Ltd.

Raw RNA-seq data were processed to filter out low-quality reads. Clean reads were obtained from each sample and used for subsequent analyses. Read counts and transcripts per million values were calculated based on the pseudoalignment of RNA-seq reads to reference transcripts downloaded from the GENCODE (v38) database, as implemented in the feature counts. Gene expression levels were determined from the transcript levels.

The infiltration levels of immune cell types were quantified using single-sample gene set enrichment analysis (ssGSEA) in the R package "GSVA".[57] ssGSEA applies the gene signatures expressed by immune cell populations to individual cancer samples. The approach used in our study included immune cells that are T cells, CD8 T cells, T helper cells, T helper 1 cells, T helper 2 cells, T helper 17 cells, T follicular helper cells, central memory T cells, effector memory T cells, regulatory T cells, gamma delta T cells, B cells, cytotoxic cells, natural killer cells, natural killer CD56bright cells, natural killer CD56dim cells, dendritic cells, activated dendritic cells, immature dendritic cells, mast cells, neutrophils, eosinophils, and macrophages. The "GSVA" package was also used for gene set variation analysis (GSVA).

Biological validation with IHC

In addition, we validated the molecular features identified in the above analysis, which are associated with the response prediction of neoadjuvant chemotherapy, using IHC. Formalin-fixed paraffin-embedded samples were cut into 4-μm-thick sections for IHC staining. The sections were dewaxed in xylene and rehydrated in descending concentration gradients of ethanol. Prior to staining, the sections were subjected to endogenous peroxidase blocking in a 3% H_2O_2 solution diluted in methanol for 20 min and then heated in a microwave for 30 min with 10 mmol/L citrate buffer (pH 6.0). Serum blocking was performed using 5% normal bovine serum for 1 h. The slides were incubated overnight at 4°C with antibodies targeting human immune cell biomarkers, including CD4 (1:500, Abcam, ab133616), CD8 (1:400, Cell Signaling Technology, 85336), CD11c (1:500, Abcam, ab52632), CD20 (1:200, Cell Signaling Technology, 48750), CD56 (1:400, Cell Signaling Technology, 99746), CD68 (1:3000, Abcam, ab955), and CD163 (1:500, Abcam, ab182422). The sections were then washed with 1% phosphate-buffered saline (PBS) and incubated with secondary antibodies (Absin, abs20040, and abs20039) for 1 h at room temperature. The slides were then washed with PBS. Biotinylated secondary antibodies were incubated with an avidin–biotin complex staining kit (VECTOR, VECTASTAIN ELITE ABC Kits, PK6100, USA) for 30 min at room temperature. The slides were washed again with PBS and stained with 3,3N-diaminobenzidine tetrahydrochloride (DAB, VECTOR, SK-4105, USA) for approximately 1.5 min. The sections were counterstained with hematoxylin, dehydrated in ethanol, and sealed with gum. For all assays, negative control slides without primary antibodies were included. Each staining run included a slide containing a positive control. All slides were stained with DAB for the same duration for each antibody.

IHC slides were scanned at 20× magnification using a Zeiss Axioscan 7 digital slide scanner. The digital pathology images were uploaded to the software Qupath. Two pathologists who were blinded to the clinical data annotated the ROIs that covered the tumors on the WSI. We chose a threshold of 300,000 μm^2 for the ROI area and excluded data from two patients whose areas did not meet the standard. For the tumor-invasive border, we shrank the ROI by 100 μm and acquired a ring area close to the stroma. A third pathologist was consulted to reach a consensus when different opinions arose between the two pathologists. For CD4, CD8, CD11c, CD20, CD56, CD68, and CD163, all cells were calculated using default nuclear DAB optical density settings in tumor areas. Positive cells were then detected using a Random Trees-based classifier, and finally the positive cell rate results were exported. The same procedures were performed on the tumor-invasive border to explore the differences between the tumor area and the tumor-invasive border area.

Biological validation with multiplex IHC

The multiplex IHC assay was performed using multiplex IHC staining kits (Absin, abs50013) according to the manufacturer's instructions. First, the sections (4 μm) were dewaxed and rehydrated before antigen retrieval. The slides were then subjected to endogenous peroxidase blocking in a 3% H_2O_2 solution diluted in methanol for 20 min, followed by heating in a microwave for 30 min with 10 mmol/L citrate buffer (pH 6.0). Serum blocking was performed with 5% normal bovine serum for 1 h. The primary antibody incubation was carried out at 37°C for 1 h, followed by incubation with the secondary antibody (horseradish peroxidase) at 37°C for 10 min and visualization using tyramide signal amplification kits. This antibody incubation and visualization process was repeated. The nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) before sealing. Primary antibodies recognizing CD11c (1:800, Abcam, ab52632), CD163 (1:800, Abcam, ab182422), CD4 (1:800, Abcam, ab133616), and CD8 (1:800, Cell Signaling Technology, 85336) were used for the incubation. All sections were scanned using the Leica DMI8 THUNDER (Leica, Germany).

Survival analysis

We assessed the predictive capability of the proposed models for postoperative outcomes using overall survival, which was derived from the diagnosis date to the last contact date—either death or the most recent follow-up. Follow-up visits were scheduled every 3 months; patients with a follow-up period <36 months and those who died within 1 month for various reasons were excluded from the survival analysis.

To compare the overall survival difference between the groups predicted by the models with those of conventional pathological evaluation, we performed Kaplan–Meier analysis and the log rank test on data from patients stratified according to pathological evaluation or model prediction. Additionally, we conducted univariate and multivariate analyses using the Cox proportional hazards model for iSCLM, along with clinical features and cancer biomarkers.

We employed propensity score matching[58] to adjust for potential biases, utilizing a 1:1 matching protocol without replacement (K nearest neighbor matching algorithm), with a caliper width equal to 0.2 times the standard deviation of the logit of the propensity score. Matching variables included patient age, cT, and cN category.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was performed using R software (version 4.3.1). Patients were categorized as responders or non-responders according to the TRG. For baseline statistics, the two groups were compared using the chi-squared test for categorical variables and the Wilcoxon test for continuous variables. Quantitative data are presented as median (interquartile range).

For predictive validation, receiver operating characteristic curves were performed with 95% CIs, and the AUROC was used to evaluate model performance. Differences in AUROCs among the models were evaluated using the DeLong test. Additionally, to comprehensively quantify model performance, standard metrics of accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were computed. To determine reproducibility, the intraclass correlation coefficient was used to evaluate variations in iSCLM-based response predictions per the manual annotations provided by the aforementioned radiologists. A paired *t*-test was employed to assess the disparity of Moran's index for the same patients between the iRM and iSCLM.

For biological interpretation, the distance between each tile and the defined tumor-invasive border was quantified. The distance of top-ranked tiles was statistically compared with the average distance of all tiles, utilizing a paired Wilcoxon test. Differences in cell components between responders and non-responders were analyzed using *t*-tests, and cell count comparisons identified by HoVer-Net and those of manual counts were validated using Spearman's rank correlation. All statistical analyses were two-sided with a significance level of $p < 0.05$.

ADDITIONAL RESOURCES

Registration of the prospective trial on ChiCTR (ChiCTR2300068917): https://www.chictr.org.cn/showproj.html?proj=189882.