



# Deep Learning Model for Predicting the Pathological Complete Response to Neoadjuvant Chemoradiotherapy of Locally Advanced Rectal Cancer

OPEN ACCESS

**Edited by:**Qi Liu,  
Fudan University, China**Reviewed by:**Fevzi Cengiz,  
Tinaztepe University, Turkey  
Simon J. Furney,  
Royal College of Surgeons in Ireland,  
Ireland  
Jacopo Lenkovicz,  
Agostino Gemelli University Polyclinic  
(IRCCS), Italy**\*Correspondence:**Yan Huang  
huangy@mail.sysu.edu.cn  
Jianhua Yao  
jianhuayao@tencent.com  
Jingping Yun  
yunjp@sysucc.org.cn

†These authors have contributed  
equally to this work and share  
first authorship

**Specialty section:**

This article was submitted to  
Gastrointestinal Cancers:  
Colorectal Cancer,  
a section of the journal  
Frontiers in Oncology

**Received:** 01 November 2021**Accepted:** 02 May 2022**Published:** 08 June 2022**Citation:**

Lou X, Zhou N, Feng L, Li Z, Fang Y,  
Fan X, Ling Y, Liu H, Zou X, Wang J,  
Huang J, Yun J, Yao J and Huang Y  
(2022) Deep Learning Model for  
Predicting the Pathological Complete  
Response to Neoadjuvant  
Chemoradiotherapy of Locally  
Advanced Rectal Cancer.  
*Front. Oncol.* 12:807264.  
doi: 10.3389/fonc.2022.807264

Xiaoying Lou<sup>1,2†</sup>, Niyun Zhou<sup>3†</sup>, Lili Feng<sup>2,4†</sup>, Zhenhui Li<sup>5†</sup>, Yuqi Fang<sup>3,6†</sup>, Xinjuan Fan<sup>1,2</sup>,  
Yihong Ling<sup>7</sup>, Hailing Liu<sup>1,2</sup>, Xuan Zou<sup>1,2</sup>, Jing Wang<sup>1,2</sup>, Junzhou Huang<sup>3</sup>, Jingping Yun<sup>7\*</sup>,  
Jianhua Yao<sup>3\*</sup> and Yan Huang<sup>1,2\*</sup>

<sup>1</sup> Department of Pathology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, <sup>2</sup> Guangdong Institute of Gastroenterology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, <sup>3</sup> Tencent AI Lab, Shenzhen, China, <sup>4</sup> Department of Radiation Oncology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, <sup>5</sup> Department of Pathology, Yunnan Cancer Hospital, Kunming, China, <sup>6</sup> Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, Hong Kong SAR, China, <sup>7</sup> Department of Pathology, Cancer Center of Sun Yat-sen University, Guangzhou, China

**Objective:** This study aimed to develop an artificial intelligence model for predicting the pathological complete response (pCR) to neoadjuvant chemoradiotherapy (nCRT) of locally advanced rectal cancer (LARC) using digital pathological images.

**Background:** nCRT followed by total mesorectal excision (TME) is a standard treatment strategy for patients with LARC. Predicting the pCR to nCRT of LARC remains difficult.

**Methods:** 842 LARC patients treated with standard nCRT from three medical centers were retrospectively recruited and subgrouped into the training, testing and external validation sets. Treatment response was classified as pCR and non-pCR based on the pathological diagnosis after surgery as the ground truth. The hematoxylin & eosin (H&E)-stained biopsy slides were manually annotated and used to develop a deep pathological complete response (DeepPCR) prediction model by deep learning.

**Results:** The proposed DeepPCR model achieved an AUC-ROC of 0.710 (95% CI: 0.595, 0.808) in the testing cohort. Similarly, in the external validation cohort, the DeepPCR model achieved an AUC-ROC of 0.723 (95% CI: 0.591, 0.844). The sensitivity and specificity of the DeepPCR model were 72.6% and 46.9% in the testing set and 72.5% and 62.7% in the external validation cohort, respectively. Multivariate logistic regression analysis showed that the DeepPCR model was an independent predictive factor of nCRT ( $P=0.008$  and  $P=0.004$  for the testing set and external validation set, respectively).

**Conclusions:** The DeepPCR model showed high accuracy in predicting pCR and served as an independent predictive factor for pCR. The model can be used to assist in clinical treatment decision making before surgery.

**Keywords:** rectal cancer, deep learning, neoadjuvant chemoradiotherapy, pathological complete response, artificial intelligence

## INTRODUCTION

Colorectal cancer remains one of the leading causes of cancer death (1). For patients with locally advanced rectal cancer (LARC), neoadjuvant chemoradiotherapy (nCRT) followed by total mesorectal excision (TME) is recommended as a standard treatment strategy. nCRT can significantly reduce local recurrence and treatment-associated toxicity and more importantly, make tumors more amenable to resection. However, the treatment response to nCRT varies greatly among patients. Approximately 15–38% of patients could obtain a pathological complete response (pCR) and are recommended the watch and wait approach to avoid the side effects of surgery (2), while 20% of patients have little to no response to nCRT and might even suffer significant side effects and miss their best opportunity for surgery (3–5). More importantly, patients with pCR have better long-term outcomes, indicating a favorable prognosis (6). However, how to predict treatment response, especially to identify pCR candidates prior to nCRT, remains challenging for LARC.

Previous studies have shown that tumor stage, serum tumor markers before neoadjuvant therapy, and lymphocyte infiltration in the tumor microenvironment are associated with tumor regression to nCRT (7). Recently, with the development of artificial intelligence algorithms, radiological imaging has been used to evaluate the treatment response of LARC (8–14). The commonly adopted imaging techniques include diffusion-weighted magnetic resonance imaging (MRI) (11), diffusion kurtosis and T2-weighted MRI (8), and a multiparametric MRI protocol with dynamic-contrast-enhanced MRI (13). For instance, Zhang *et al.* (10) developed a pCR prediction model based on diffusion kurtosis and T2-weighted MRI, and the area under the curve (AUC) was 0.70 (95% confidence interval (CI): 0.59, 0.79). Histopathological images prevail as the gold standard for patient diagnosis and contain abundant biological information. Therefore, we anticipate that more accurate predictions can be achieved by analyzing pathological images than by analyzing radiological images.

Compared with conventional machine learning, deep learning can automatically extract features from an image without the necessity of feature predefinition and is suitable for mining the most relevant feature representations. Multi-instance learning (MIL), as a weakly supervised deep learning technique, has achieved promising results on the topic of patient prognosis and outcome prediction (15–18). MIL enables the network to learn more holistic information from whole-slide images (WSIs). To the best of our knowledge, there has been little investigation on the prediction of pCR based on histopathological images prior to nCRT with the MIL technique. The aim of this study was to develop a deep pathological complete response (DeepPCR) prediction model for the prediction of pCR directly from conventional hematoxylin & eosin (H&E)-stained histopathological images.

## MATERIALS AND METHODS

### Study Cohort and Availability

Two different cohorts, i.e., the primary cohort and external validation cohort, were adopted for training and internal and

external validation and included retrospectively identified LARC patients from January 1, 2010, to January 1, 2018, from three hospitals in China (the Sixth Affiliated Hospital of Sun Yat-sen University, Cancer Center of Sun Yat-sen University, and Yunnan Cancer Hospital). A total of 842 patients were recruited; among them, the primary cohort (783 patients from the Sixth Affiliated Hospital of Sun Yat-sen University and Cancer Center of Sun Yat-sen University) was randomly subgrouped into the training set (666 patients, 85%) and testing set (117 patients, 15%), and the external validation cohort (from Yunnan Cancer Hospital) contained 102 patients. The inclusion criteria were as follows: (1) patients had locally advanced disease determined by pretreatment TNM stage (T3/T4, and/or N+); (2) biopsy was performed, and the biopsy specimen was pathologically diagnosed as adenocarcinoma; and (3) patients underwent nCRT followed by rectal resection. The exclusion criteria were as follows: (1) patients with familial adenomatous polyposis, distant metastases, or Lynch syndrome; and (2) patients with no information on tumor regression grade (TRG) and no available H&E-stained slides.

All patients accepted a standard treatment strategy based on the National Comprehensive Cancer Network (NCCN) guidelines (version 3, 2017). The nCRT regimen was 50 Gy pelvic radiation therapy with concurrent 5-fluorouracil-based chemotherapy (FOLFIRI or FOLFOX regimens). TME was performed by either anterior resection or abdominoperineal resection after nCRT of 4–8 weeks. The TRG after nCRT was used to divide patients into two groups based on H&E-stained slides after surgery: pCR (with no remaining viable cancer cells) and non-pCR (with small clusters of cancer cells or no response with extensive residual cancer). The flow diagram of patient enrollment into the two cohorts is shown in **Figure 1**.

Clinicopathological variables, such as age, sex, TNM stage, histological grade, TRG after surgery, and blood testing parameters, including lymphocytes, neutrophils, carcinoembryonic antigen (CEA), carbohydrate antigen 19-9 (CA19-9), and lactate dehydrogenase (LDH) prior to nCRT treatment, were collected. This study was approved by the Institutional Review Board of the Sixth Affiliated Hospital of Sun Yat-sen University.

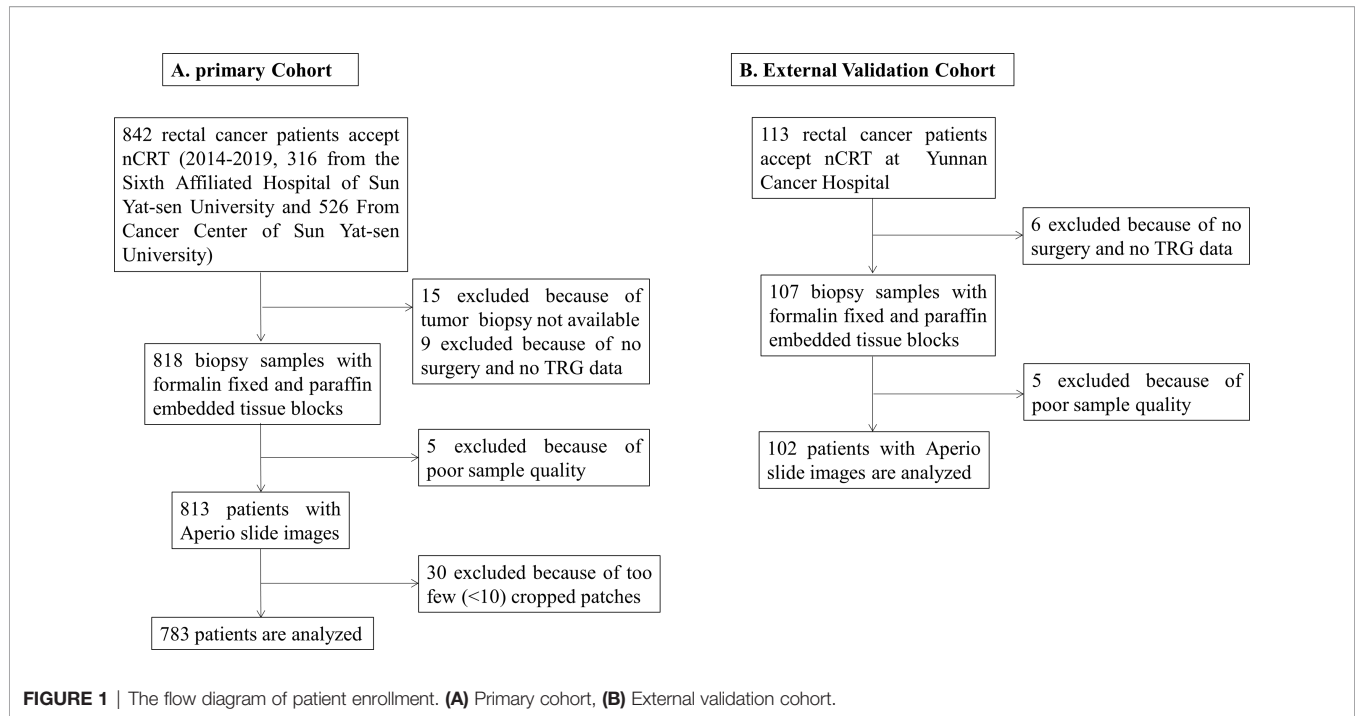
### Data Preparation

Formalin-fixed paraffin-embedded (FFPE) biopsy tissue blocks were cut into 4- $\mu$ m sections for H&E staining. All slides were checked by a pathologist who ascertained that they contained tumor areas. WSIs were acquired at a magnification of 20 $\times$  on an Aperio scanner.

Tumor tissue regions were hand-delineated by pathologists (Dr. XYL and Dr. HLL) using Aperio Image Scope software and subsequently cropped into patches with a size of 299 $\times$ 299 pixels at a magnification of 20 $\times$ . The distribution of the number of patches per slide followed a long-tail distribution, with the majority of slides containing approximately 100 patches. For slides with more than 1000 patches, we randomly chose 1000 cropped patches.

### pCR Candidate Classification

Four models were designed for classifying the input biopsy histological images, with patients' distinct TRG outcomes as



the ground truth. The first three models were trained on 102,728 patches and tested on 18475 patches in the primary cohort, namely, the DeepPCR model, *patch-based combined* model, and *patch-based individual model*. The DeepPCR model was built upon the MIL strategy (Figure 2). Specifically, a pretrained ResNet-18 model (19) was leveraged to extract the pathological feature representations of each cropped patch, i.e., the patch-wise phenotype representation (patchPR). Based on the patchPRs, the unsupervised K-means algorithm was used to categorize these features into six clusters (see Supplementary Figure 1). Each cluster occupied a subspace of the features and comprised a distinctive phenotype group. The patches in each cluster were further processed by a multi-instance fully convolutional model (MI-FCM) (20) to generate cluster-wise phenotype representation (clusPR). Herein, the MI-FCM was comprised of two pairs of *Conv-ReLU* layers, followed by a *pooling* layer. Afterwards, WSI-wise phenotype representation (wsiPR) was constructed by concatenating the clusPRs from the same WSI. The wsiPR sufficiently exploited the intercluster feature difference and intracluster feature dependence, constituting the most informative phenotype representation. Based on wsiPR, a two-layer fully connected network was leveraged to generate the final prediction. The DeepPCR model built a hierarchical feature structure from patch to WSI and explicitly modeled the mutual dependence between different phenotype groups for patient outcome prediction.

The *patch-based combined model* and *patch-based individual model* used patch-based approaches in which the cropped patches shared the same label with the original histopathological WSI and the prediction of patch-based methods was made for each patch rather than each WSI. Similar to DeepPCR, the pretrained ResNet-18 model was adopted to extract the phenotype representations of

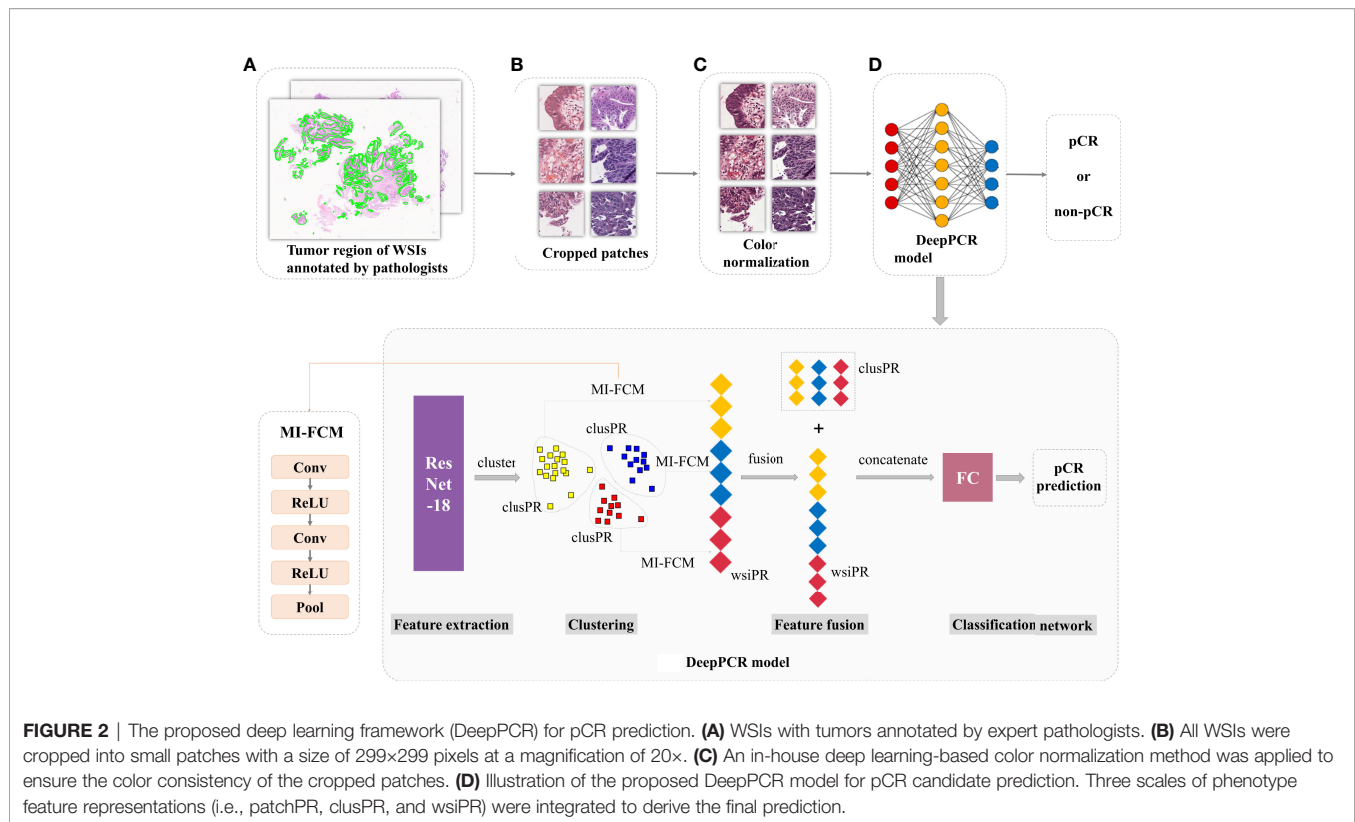
each cropped patch. According to the aggregation method of the patch-level prediction, we implemented these patch-based models in two ways. One was to predict each individual patch's label, and then combined them *via* majority voting, which was called the *patch-based individual model*. The other was to aggregate the patch-level predictions of each subject by removing the clustering step in DeepPCR, called the *patch-based combined model* (remaining modules are the same as DeepPCR). To validate the effectiveness of pathological imaging data in pCR outcome prediction compared with nonpathological data, the fourth model (*hematology model*), based on clinical hematology data, including CEA, CA19-9, LDH, lymphocytes, and neutrophils, was built. A two-layer multilayer perceptron (MLP) model was adopted in the hematology model.

## Phenotype Visualization

To visualize the representative phenotypes in each K-means cluster, t-distributed Stochastic Neighbor Embedding (t-SNE) (21) and the Raster Fairy method (22) were applied on the patchPRs. t-SNE is a technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional data. The Raster Fairy method aims to transform the two-dimensional clustering data derived from t-SNE into a regular grid without destroying the neighborhood relations emerging from the clustering. The GradCAM method (23) was used to calculate the patch importance for target prediction.

## Statistical Analysis

The predictive efficacy of the model was evaluated by the area under the receiver operating characteristic curve (AUC-ROC), area under the precision-recall curve (AUC-PR), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Univariate and multivariate logistic regression analyses



were performed to investigate the predictive value for all biomarkers. The statistical significance of the differences in the clinicopathological characteristics of pCR and non-pCR patients were calculated using the Mann-Whitney test (two-tailed) for continuous variables and Fisher's exact test (two-tailed) for dichotomous variables. Comparisons of clinicopathological factors in the primary and external validation cohorts were performed using Student's t test for continuous variables and Fisher's exact test (two-tailed) for dichotomous variables. A two-sided p value of less than 0.05 was considered statistically significant.

## RESULTS

### Patient Characteristics

The primary cohort included 783 patients: 295 patients from the Sixth Affiliated Hospital of Sun Yat-sen University and 488 patients from the Cancer Center of Sun Yat-sen University. A total of 201 and 582 patients were classified as pCR and non-pCR, respectively. The external validation cohort from Yunnan Cancer Hospital included 102 patients, of which 24 and 78 patients were classified as pCR and non-pCR, respectively. The clinicopathological characteristics of the patients in the primary and external validation cohorts are provided in **Table 1**. The clinicopathological characteristics, including clinical T stage and histological grade, were different between the primary and external validation cohorts ( $P < 0.001$  and  $P < 0.001$ , respectively) (**Supplementary Table 1**).

### pCR Candidate Prediction in the Primary Cohort

The DeepPCR model had a higher discriminative power, with an AUC-ROC of 0.710 (95% CI: 0.595, 0.808) and an AUC-PR of 0.875 (95% CI: 0.795, 0.935) in the primary cohort (**Figure 3A** and **Table 2A**). The sensitivity, specificity, PPV and NPV were 72.6%, 46.9%, 70.4%, and 54.0%, respectively (**Table 2A**). The other three models showed inferior performance. Specifically, the *hematology model* had an AUC-ROC of 0.403 (95% CI: 0.274, 0.534) and an AUC-PR of 0.698 (95% CI: 0.591, 0.805). The *patch-based individual model* and *patch-based combined model* achieved an AUC-ROC of 0.544 (95% CI: 0.432, 0.653) and an AUC-PR of 0.805 (95% CI: 0.717, 0.885) and an AUC-ROC of 0.627 (95% CI: 0.516, 0.733) and an AUC-PR of 0.842 (95% CI: 0.762, 0.909), respectively (**Figures 3A, C** and **Table 2A**). As shown in **Figure 3E**, the AUC-ROC of the DeepPCR model was significantly higher than that of the *hematology model* ( $P < 0.001$ ) and *patch-based individual model* ( $P < 0.05$ ).

### pCR Candidate Prediction in the External Validation Cohort

To investigate the effectiveness and generalizability of the DeepPCR model, it was validated in the external cohort. In the external validation cohort, the DeepPCR model achieved a similar AUC-ROC of 0.723 (95% CI: 0.591, 0.844) and an AUC-PR of 0.887 (95% CI: 0.805, 0.949) (**Figures 3B, D** and **Table 2B**). The sensitivity, specificity, PPV and NPV were 0.725 (95% CI: 0.637, 0.814), 0.627 (95% CI: 0.463, 0.773), 0.758 (95%

**TABLE 1 |** Clinicopathological characteristics of patients in the training, testing, and external validation cohorts.

	Training set (n=666)			Testing set (n=117)			ExternalValidation set (n=102)		
	PCR (%) (n=171)	Non-PCR (%) (n=495)	P value	PCR (%) (n=30)	Non-PCR (%) (n=87)	P value	PCR (%) (n=24)	Non-PCR (%) (n=78)	P value
Age, mean(SD), y	52.77 ± 12.02	54.71 ± 11.78	0.078	53.90 ± 11.71	55.38 ± 11.47	0.549	54.08 ± 11.01	57.17 ± 10.37	0.182
Sex, No. (%)			0.849			0.376			0.081
Female	55(32.2)	154(31.1)		8(26.7)	32(36.8)		12(50.0)	22(28.2)	
Male	116(67.8)	341(68.9)		22(73.3)	55(63.2)		12(50.0)	56(71.8)	
Clinical T stage									
cT2	10(5.9)	16(3.2)	0.124	2(6.7)	2(2.3)	0.271	1(4.2)	0(0.0)	0.235
cT3	113(66.1)	323(65.3)	0.926	20(66.7)	55(63.2)	0.827	9(37.5)	24(30.8)	0.612
cT4	48(28.0)	156(31.5)	0.442	8(26.6)	30(34.5)	0.503	14(58.3)	54(69.2)	0.333
Clinical N stage									
cN0	34(19.9)	76(15.4)	0.189	5(16.7)	13(14.9)	0.777	0(0.0)	18(23.1)	0.006
cN1	86(50.3)	249(50.3)	1	13(43.3)	40(46.0)	0.834	17(70.8)	43(55.1)	0.236
cN2	51(29.8)	170(34.3)	0.301	12(40.0)	34(39.1)	1	7(29.2)	17(21.8)	0.582
TNM stage									
Stage II	35(20.5)	76(15.3)	0.167	5(16.6)	13(14.9)	0.777	0(0.0)	18(23.1)	0.006
Stage III	136(79.5)	419(84.7)	0.124	25(83.4)	74(85.1)	0.777	24(100.0)	60(76.9)	0.006
Histological grade									
1	22(12.9)	55(11.1)	0.579	3(10.0)	16(18.4)	0.394	1(4.2)	0(0)	0.235
2	125(73.1)	382(77.2)	0.299	22(73.3)	65(74.7)	1	23(95.8)	71(91.0)	0.677
3	24(14.0)	58(11.7)	0.421	5(16.7)	6(6.9)	0.147	0(0)	7(9.0)	0.194
Patch No.	102,728			18475			46599		

CI: 0.671, 0.847), and 0.536 (95% CI: 0.368, 0.688), respectively (**Table 2B**). In external cohorts, the AUC-ROC of the DeepPCR model was significantly higher than that of the hematology model ( $P < 0.001$ ) and patch-based individual model ( $P < 0.05$ ) (**Figure 3F**).

## Univariate and Multivariate Analyses

In the primary cohort, the univariate logistic regression analysis showed that CEA and DeepPCR model were significantly correlated with pCR ( $P=0.033$  and  $0.0001$ , respectively) (**Table 3A**). Multivariate logistic regression analysis showed that only DeepPCR was an independent factor for predicting pCR (95% CI: 1.646, 28.743;  $P=0.008$ ) (**Table 3B**).

In the external validation cohort, age, CEA, neutrophil-to-lymphocyte ratio (NLR), patch-based combined model and DeepPCR model were significantly correlated with pCR ( $P=0.042$ ,  $0.029$ ,  $0.04$ ,  $0.023$ , and  $0.0001$ , respectively) (**Table 3A**). Multivariate logistic regression analysis showed that only DeepPCR was an independent factor for predicting pCR (95% CI: 2.138, 51.186;  $P=0.004$ ) (**Table 3B**).

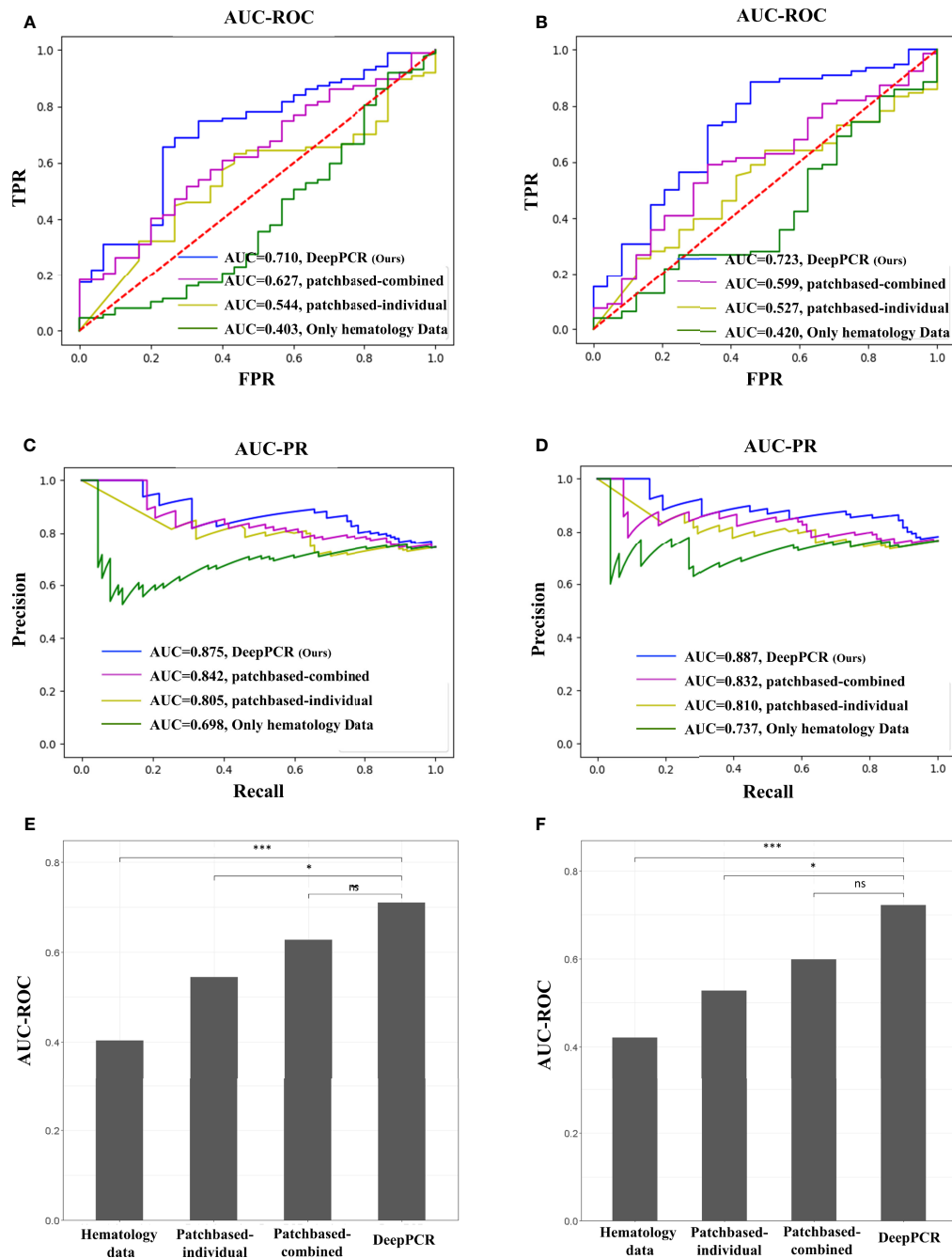
## Histological Patterns Associated With TRG

To find some important clinical insights based on the DeepPCR model, we determined which types of histological patterns were most relevant to patient TRG, and the pipeline of this process is displayed in **Figure 4**. In **Figure 4A**, each grid represented an individual patch, and the patchPRs obtained from all these patches were categorized into six phenotype clusters (**Figure 4B**), which were reduced into a two-dimensional feature space based on t-SNE and the Raster Fairy method. Here, the phenotypes could be color, edges, texture, curve and/or shape of cancer and normal tissues. To conduct an investigation into which types of phenotypes contribute

the most to pCR prediction, the GradCAM method (23) was adopted to calculate the importance of patches. The importance heatmap is shown in **Figure 4C**, and darker colors indicate that the patches played a more important role in pCR prediction. We also calculated the sum of the importance values of the patches in each cluster (**Figure 4D**). It can be seen that different clusters had different predictive powers for pCR prediction, and a larger value indicated that the corresponding cluster contributed more to DeepPCR. The size of bubbles represents the number of patches in the corresponding cluster. We found that patches in clusters 0 and 1 played more important roles in pCR candidate prediction. Specifically, the patch importance value of cluster 1 was significantly larger than that of clusters 2, 3, 4, and 5 ( $P<0.001$ ,  $P<0.05$ ,  $P<0.01$ , and  $P<0.001$ , respectively). There were no significant differences between cluster 0 and cluster 1 in terms of the patch importance value (**Figure 4D**). **Figure 4E** shows the representative patches of cluster 1 and their distribution in a WSI, which also represented a special histological pattern and spatial pattern highly associated with pCR. Similarly, **Figures 4F–J** demonstrates the same patchPR visualization process but for the non-pCR group. The patch importance value of cluster 2 was significantly larger than that of clusters 0, 1, 3, 4, and 5 ( $P<0.001$ ,  $P<0.001$ ,  $P<0.001$ ,  $P<0.001$ , and  $P<0.001$ , respectively) (**Figure 4I**) in non-pCR candidate prediction.

## DISCUSSION

In the present study, we developed a novel model to predict pCR in LARC using digital pathological images. We found that the DeepPCR model could achieve a relatively high AUC-ROC score of 0.710. Multivariate logistic analysis showed that the DeepPCR model was indeed an independent factor for predicting pCR,



**FIGURE 3 | (A, B)** AUC-ROC of the four comparative methods in the **(A)** primary and **(B)** external validation cohorts (top row). **(C, D)** AUC-PR of the four comparative methods in the **(C)** primary and **(D)** external validation cohorts (middle row). **(E, F)** DeLong test for the four comparative methods in the **(E)** primary and **(F)** external validation cohorts (bottom row). In this work, we used a probability threshold of 0.7 (that is, any patient with a pCR prediction probability greater than 0.7 was reported as a pCR candidate). No significant difference (ns);  $P > 0.05$ ,  $*P < 0.05$ ,  $***P < 0.001$ .

indicating that the model could assist in treatment decision making prior to surgery for LARC.

In recent years, there has been increasing interest in digital pathology image analysis based on machine learning algorithms to assist in pathological diagnosis (24, 25). With the development

of deep learning, an increasing number of studies have focused on clinical-grade detection and the prediction of outcomes. For example, Cao et al. (24) developed a pathomics-based model for microsatellite instability prediction from pathological images. Ole-Johan Skrede et al. (25) developed a deep learning-based

**TABLE 2 |** Results of DeepPCR and the comparative models in the (a) primary and (b) external validation cohorts.

(a) Model/Outcome	AUC-ROC	AUC-PR	Sen (%)	Spe (%)	PPV (%)	NPV (%)
Hematology model	0.403 (0.274, 0.534)	0.698 (0.591, 0.805)	72.6 (64.1, 80.3)	27.2 (18.8, 36.2)	61.7 (60.0, 71.1)	37.7 (30.8, 51.2)
Patch-based individual model	0.544 (0.432, 0.653)	0.805 (0.717, 0.885)	68.4 (59.8, 76.9)	25.8 (17.0, 34.7)	57.2 (45.2, 69.6)	27.0 (15.4, 46.8)
Patch-based combined model	0.627 (0.516, 0.733)	0.842 (0.762, 0.909)	69.2 (60.7, 77.8)	30.4 (20.7, 40.7)	61.6 (50.5, 73.1)	37.6 (18.0, 59.4)
DeepPCR model	0.710 (0.595, 0.808)	0.875 (0.795, 0.935)	72.6 (64.1, 80.3)	46.9 (32.6, 61.0)	70.4 (61, 79.9)	54.0 (35.8, 70.9)
(b) Model/Outcome	AUC-ROC	AUC-PR	Sen (%)	Spe (%)	PPV (%)	NPV (%)
Hematology model	0.420 (0.293, 0.548)	0.737 (0.623, 0.846)	70.6 (61.8, 79.4)	21.7 (14.2, 30.0)	57.4 (48.5, 67.4)	17.6 (14.3, 20.4)
Patch-based individual model	0.527 (0.402, 0.657)	0.810 (0.712, 0.895)	73.5 (64.7, 81.4)	22.6 (15.3, 31.4)	57.9 (62.6, 72.4)	17.8 (17.4, 18.1)
Patch-based combined model	0.599 (0.474, 0.726)	0.832 (0.732, 0.919)	69.6 (60.8, 78.4)	27.2 (16.3, 38)	62.3 (49.9, 74.5)	31.7 (14.8, 54.1)
DeepPCR model	0.723 (0.591, 0.844)	0.887 (0.805, 0.949)	72.5 (63.7, 81.4)	62.7 (46.3, 77.3)	75.8 (67.1, 84.7)	53.6 (36.8, 68.8)

The CI value is inside the parentheses. Sen, sensitivity; Spe, specificity; PPV, positive predictive value; NPV, negative predictive value. In this work, we used a probability threshold of 0.7 (that is, any patient with a pCR prediction probability greater than 0.7 was reported as a pCR candidate).

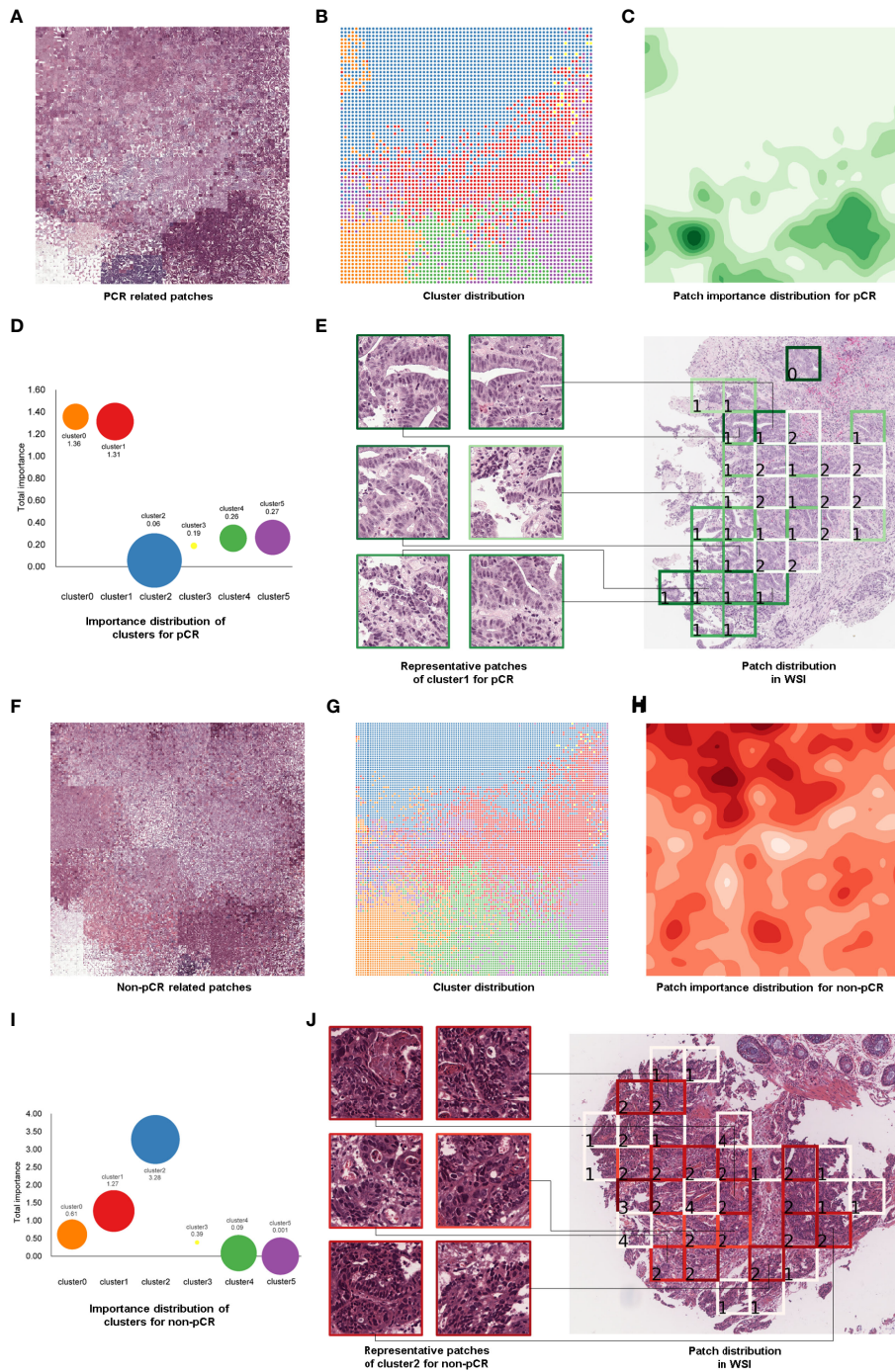
**TABLE 3 |** Univariate and multivariate logistic regression analyses.

(a) Univariate logistic regression	Testing Set		External Validation Set	
	P value	Exp (B) (95% CI)	P value	Exp (B) (95% CI)
Sex	0.316	1.6 (0.638, 4.011)	0.051	0.393 (0.153, 1.006)
Age	0.051	2.679 (0.995, 7.212)	0.042	2.768 (1.039, 7.376)
TNM stage	0.822	1.138 (0.369, 3.513)	0.998	0 (0, -)
CEA	0.033	2.796 (1.087, 7.197)	0.029	3.667 (1.145, 11.74)
CA-199	0.087	2.128 (0.896, 5.055)	0.054	2.505 (0.985, 6.37)
CRP	0.198	2.348 (0.639, 8.621)	-	-
LDH	0.999	5.80e8 (0, -)	0.207	2.4 (0.617, 9.339)
Lymphocytes	0.24	2.186 (0.593, 8.062)	0.133	2.2 (0.788, 6.146)
Neutrophils	0.414	1.524 (0.555, 4.186)	0.097	2.508 (0.846, 7.436)
NLR	0.142	3.155 (0.681, 14.623)	0.04	3.045 (1.054, 8.804)
Patch-indi	0.06	2.248 (0.967, 5.224)	0.219	1.786 (0.709, 4.5)
Patch-comb	0.053	2.548 (0.989, 6.564)	0.023	3.143 (1.171, 8.437)
DeepPCR	0.0001	6.125 (2.462, 15.239)	0.0001	7 (2.575, 19.028)
(b) Multivariate logistic regression	Test Cohort		External Validation Cohort	
	Sig.	Exp (B) (95% CI)	Sig.	Exp (B) (95% CI)
Sex	0.143	2.45 (0.739, 8.124)	0.011	0.122 (0.024, 0.621)
Age	0.489	1.576 (0.434, 5.72)	0.705	1.346 (0.289, 6.261)
TNM stage	0.965	1.034 (0.233, 4.582)	0.998	0 (0, -)
CEA	0.101	2.718 (0.823, 8.973)	0.189	3.211 (0.564, 18.284)
CA-199	0.124	2.413 (0.785, 7.415)	0.059	4.137 (0.945, 18.108)
CRP	0.104	4.607 (0.732, 29.003)	-	-
LDH	0.999	2.6e8 (0, -)	0.118	7.334 (0.604, 89.051)
Lymphocytes	0.128	3.412 (0.704, 16.539)	0.203	3.418 (0.514, 22.723)
Neutrophils	0.979	0.981 (0.239, 4.023)	0.874	0.846 (0.107, 6.699)
NLR	0.138	4.242 (0.628, 28.678)	0.05	8.854 (0.995, 78.749)
Patch-indi	0.346	1.657 (0.58, 4.732)	0.831	0.855 (0.204, 3.591)
Patch-comb	0.8	0.819 (0.175, 3.842)	0.642	1.453 (0.301, 7.023)
DeepPCR	0.008	6.879 (1.646, 28.743)	0.004	10.461 (2.138, 51.186)

(a) Univariate logistic regression analysis of the testing set and external validation set. (b) Multivariate logistic regression analysis of the testing set and external validation set. The covariates were sex, age, TNM stage, CEA, CA19-9, CRP, LDH, lymphocytes, neutrophils, neutrophil-to-lymphocyte ratio (NLR), patch-based individual (patch-indi) model, patch-based combined (patch-comb) model, and DeepPCR model.

biomarker model for colorectal cancer outcome by analyzing H&E-stained sections. The successful applications of artificial intelligence in digital pathology indicate that digital pathology images contain important information for the diagnosis and prognosis of cancer. Due to the complexity of pathological imaging, there have been few relevant studies on the prediction of neoadjuvant efficacy based on preoperative pathological biopsy with artificial intelligence. Some studies used MRI to predict neoadjuvant efficacy. For instance, Petresc et al. (8) utilized pretreatment T2-weighted radiomic features to predict

LARC responders, and least absolute shrinkage and selection operator (LASSO) regression analysis was applied to derive a predicted AUC of 0.80 (95% CI: 0.58, 0.94). Although their model's performance was better than ours, they used a small cohort of patients. In a retrospective study, Zhang et al. (10) developed a deep learning-based model for pCR prediction based on diffusion kurtosis and T2-weighted MRI, and the AUC was 0.70 (95% CI: 0.59, 0.79), which was similar to that of our proposed model. The limitation of their model was that they did not validate the model in independent external cohorts.



**FIGURE 4 |** Patch-level feature interpretation in the pCR group (A–E) and non-pCR group (F–J). Patches in the correctly predicted pCR group (A) and correctly predicted non-pCR group (F). PatchPRs were categorized into six phenotype clusters based on t-SNE and the Raster Fairy method, and each grid represented an individual patch (B, G). The importance distribution of the patches in the pCR group (C) and non-pCR group (H). Darker colors represent the patches that played a more important role in pCR or non-pCR prediction. Demonstration of patch importance and the number of patches in each cluster; the size of the bubble represents the number of patches in the corresponding cluster (D, I). Representative patches of cluster 1 (E) and cluster 2 (J) and the part of the WSI from which they were selected.



Radiological imaging has its own limitation in distinguishing inflammatory lesions from neoplastic lesions. As the gold standard of disease diagnosis, conventional preoperative pathological biopsy is of great significance for the diagnosis and prognosis of tumors.

The discriminative power of DeepPCR model was significantly higher than that of the hematology model ( $P < 0.001$ ) and the other two patch-based models ( $P < 0.001$  and  $P < 0.001$ , respectively). The number of patients in our study was larger than other reported works (8–14). Moreover, the DeepPCR model was evaluated in independent cohorts. The external validation cohort came from another center with a different sample handling procedure and using a different scanner. Although the external validation cohort was different from the primary cohort in terms of the clinicopathological characteristics, such as clinical T stage and histological grade ( $P < 0.001$  and  $P < 0.001$ , respectively), the proposed model achieved similar results as those in the primary cohort, indicating its generalizability and robustness.

The proposed model leveraged an MIL-based deep learning model and showed a superior performance compared to previous patch-based learning methods. Existing patch-based approaches can be categorized into two classes based on the level of the employed annotations. For the first class, patch-wise annotations are used to train deep learning models (26–30), and strong supervision is typically performed, benefiting from the precise labeling information. Nevertheless, these methods depend on pixel-level annotations by expert pathologists, and it would be labor intensive and hard to obtain sufficient high-quality annotation data. For the second class of methods, the ground-truth labels are provided for the whole images rather than the patches (31, 32). When performing the learning process, the global image-level label of each WSI is taken as the patch-level label directly, and the final prediction is generated by combining the patch-level outputs. Although this type of method is very straightforward, there are two crucial problems. First, the cropped patches of WSIs are processed independently, and the spatial constraints of these patches are neglected. The second problem is that the patches in the same image indiscriminately share the same label and thus introduce a substantial disturbance to model training. To address these problems, several MIL-based approaches that aim to leverage the feature representations of all image patches to collaboratively predict the patient outcome have been developed (15–18, 33). Building upon these methods, our proposed model can effectively mine the dependence of feature representations at three different scales, i.e., patch-level, cluster-level, and WSI-level phenotype representations. In this patient outcome prediction task, the MIL-based learning method outperformed the patch-based learning methods. Specifically, the MIL-based methods were able to jointly consider intrapatch dependence; thus, the spatial relationships between tumor tissues (including cancer cells and surrounding stromal cells) were exploited. These tissues form the tumor microenvironment (34), and the characterization of the microenvironment plays an important role in tumor progression and the response to treatment. However, the patch-based learning methods only

processed patches independently, and the spatial information among patches was neglected; thus, these methods showed poor performance. Our findings suggest that MIL-based learning models can handle the spatial information inherent in the tumor microenvironment.

Some deep learning-based studies visualized and interpreted the learned feature representations (31, 35–38), which may provide some important clinical insights. For instance, Courtiol et al. (35) identified regions that contributed to patient outcome prediction (mesothelioma classification) by visualizing various scenarios predicted by the deep learning model. They found that these regions are mostly located in the stroma and are associated with inflammation, cellular diversity and vacuolization. Campanella et al. (36) assessed the model by visualizing the features reduced in a 2D space and found that a set of top-ranked patches with probabilities close to 0.5 contained glands suspicious of being malignant. In our study, patchPRs were categorized into six phenotype clusters based on the DeepPCR model. We determined that different clusters had different predictive powers for pCR prediction. We calculated the sum of the importance values of the patches in each cluster and found that the patches in cluster 0 and cluster 1 played more important roles in pCR candidate prediction. Although we did not analyze each cluster in more detail, we proposed that some histological patterns may be associated with the predicted TRG. The novel histological pattern may be associated with the morphological features and microenvironment of the tumor.

Previous studies showed that pretreatment serum CEA levels were significantly correlated with pCR (39). In our study, the univariate logistic regression analysis showed that CEA levels significantly correlated with pCR in the primary cohort ( $P = 0.033$ ) and in the external validation cohort ( $P = 0.042$ ). However, in multivariate logistic regression analysis, this association did not persist, and only the DeepPCR model was an independent factor for predicting pCR (95% CI: 1.646, 28.743;  $P = 0.008$ ). We also conducted pCR prediction experiments based on clinical data, i.e., CEA, CA19-9, LDH, lymphocytes, and neutrophils. In the experimental studies, an AUC-ROC of 0.403 was achieved based on these nonpathological data, showing that they may not be sufficient for prognostic pCR prediction.

Although promising results and relevant clinical insights were found, there are some limitations in this study. First, this study was a retrospective study. A multicenter prospective study is needed to confirm the performance of the prediction model. Second, due to the prevalence of tumor heterogeneity, the representativeness of biopsy specimens was limited. Another limitation of this study was that deep learning has the disadvantage of its black-box nature. Although we determined some histological patterns relevant to patient TRG, the morphological features and microenvironment of each histological pattern should be further investigated.

In conclusion, our study was the first to investigate the nCRT outcome prediction problem in LARC patients using presurgical biopsy pathological images. A clinically useful prediction model was developed using deep learning. The DeepPCR model was evaluated in an independent cohort and achieved stable results. This model has the potential to guide clinicians in making nCRT choices.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

This study was approved by the Institutional Review Board of the Sixth Affiliated Hospital of Sun Yat-sen University.

## AUTHOR CONTRIBUTIONS

JH, JPY, JHY, and YH contribute to conception and design. LF, ZL, XZ, JW, and YL contribute to acquisition of data. NZ and YF contribute to analysis and interpretation of data. NZ, YF, XL, and XF participate in drafting the article. All authors give final approval of the version to be published.

## REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Roh MS, Colangelo LH, O'Connell MJ, Yothers G, Deutsch M, Allegra CJ, et al. Preoperative Multimodality Therapy Improves Disease-Free Survival in Patients With Carcinoma of the Rectum: NSABP R-03. *J Clin Oncol* (2009) 27(31):5124. doi: 10.1200/JCO.2009.22.0467
- Capirci C, Valentini V, Cionini L, Paoli AD, Rodel C, Jones RG, et al. Prognostic Value of Pathologic Complete Response After Neoadjuvant Therapy in Locally Advanced Rectal Cancer: Long-Term Analysis of 566 ypCR Patients. *Int J Radiat Oncol Biol Phys* (2008) 72(1):99–107. doi: 10.1016/j.ijrobp.2007.12.019
- Ciccocioppo A, Stephens JH, Hewett PJ, Rieger NA. Complete Pathologic Response After Preoperative Rectal Cancer Chemoradiotherapy. *ANZ J Surg* (2009) 79(6):481–4. doi: 10.1111/j.1445-2197.2009.04950.x
- Yeo SG, Kim DY, Kim TH, Chang HJ, Oh JH, Park W, et al. Pathologic Complete Response of Primary Tumor Following Preoperative Chemoradiotherapy for Locally Advanced Rectal Cancer: Long-Term Outcomes and Prognostic Significance of Pathologic Nodal Status (KROG 09-01). *Ann Surg* (2010) 252(6):998–1004. doi: 10.1097/SLA.0b013e3181f3f1b1
- Maas M, Nelemans PJ, Valentini V, Das P, Rödel C, Kuo LJ, et al. Long-Term Outcome in Patients With a Pathological Complete Response After Chemoradiation for Rectal Cancer: A Pooled Analysis of Individual Patient Data. *Lancet Oncol* (2010) 11(9):835–44. doi: 10.1016/S1470-2045(10)70172-8
- Huang Y, Lou X, Zhu Y, Wang YC, Zhang L, Liu HL, et al. Local Environment in Biopsy Better Predict the Pathological Response to Neoadjuvant Chemoradiotherapy in Rectal Cancer. *Biosci Rep* (2019) 39(3). doi: 10.1042/BSR20190003
- Petresc B, Lebovici A, Caraiani C, Feier DS, Graur F, Buruian MM. Pre-Treatment T2-WI Based Radiomics Features for Prediction of Locally Advanced Rectal Cancer Non-Response to Neoadjuvant Chemoradiotherapy: A Preliminary Study. *Cancers* (2020) 12(7):1894. doi: 10.3390/cancers12071894
- Shaish H, Aukerman A, Vanguri R, Spinelli A, Armenta P, Jambawalikar S, et al. Radiomics of MRI for Pretreatment Prediction of Pathologic Complete Response, Tumor Regression Grade, and Neoadjuvant Rectal Score in

## FUNDING

This work was supported by the Guangdong Science and Technology Project (No. 2019B030316003 to XF), Natural Science Foundation of Guangdong Province (No. 2019A1515010901 to XF), the National Science Fund for Excellent Young Scholars (No.82122057 to XF); Guangdong Natural Science Funds for Distinguished Young Scholars (No. 2021B1515020022 to XF).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.807264/full#supplementary-material>

**Supplementary Figure 1 |** Effects of the number of clusters on pCR prediction. The AUC-ROC in the test cohort was optimal when the number of clusters was set to 6.

**Supplementary Figure 2 |** Demonstration of pCR and non-pCR patient differentiation using (A) patch-level, (B) cluster-level, (C) WSI-level feature representations based on the t-SNE results. With the increase in level, pCR and non-pCR candidates were easier to differentiate.

- Patients With Locally Advanced Rectal Cancer Undergoing Neoadjuvant Chemoradiation: An International Multicenter Study. *Eur Radiol* (2020) 1–11. doi: 10.1007/s00330-020-06968-6
- Zhang XY, Wang L, Zhu HT, Li ZW, Ye M, Li XT, et al. Predicting Rectal Cancer Response to Neoadjuvant Chemoradiotherapy Using Deep Learning of Diffusion Kurtosis MRI. *Radiology* (2020) 190936. doi: 10.1148/radiol.2020190936
- Fu J, Zhong X, Li N, Lewis J, Sung K, Raldow AC, et al. Deep Learning-Based Radiomic Features for Improving Neoadjuvant Chemoradiation Response Prediction in Locally Advanced Rectal Cancer. *Phys Med Biol* (2020) 65(7):075001. doi: 10.1088/1361-6560/ab7970
- Shayesteh SP, Alikhass A, Farhan F, Ghalehtaki R, Soltanabadi M, Haddad P, et al. Prediction of Response to Neoadjuvant Chemoradiotherapy by MRI-Based Machine Learning Texture Analysis in Rectal Cancer Patients. *J Gastrointest Cancer* (2020) 51(2):601–9. doi: 10.1007/s12029-019-00291-0
- Shi L, Zhang Y, Nie K, Sun X, Niu T, Yue N, et al. Machine Learning for Prediction of Chemoradiation Therapy Response in Rectal Cancer Using Pre-Treatment and Mid-Radiation Multi-Parametric MRI. *Magn Reson Imaging* (2019) 61:33–40. doi: 10.1016/j.mri.2019.05.003
- Liao W, Pei H. MRI-Based Radiomics Predicts Tumor Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Front Oncol* (2019) 9:552. doi: 10.3389/fonc.2019.00552
- Mercan C, Aksoy S, Mercan E, Shapiro LG, Weaver DL, Elmore JG. Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images. *IEEE Trans Med Imaging* (2017) 37(1):316–25. doi: 10.1109/TMI.2017.2758580
- Sudharshan PJ, Petitjean C, Spanhol F, Oliveira LE, Heutte L, Honeine P, et al. Multiple Instance Learning for Histopathological Breast Cancer Image Classification. *Expert Syst Appl* (2019) 117:103–11. doi: 10.1016/j.eswa.2018.09.049
- Zhu W, Lou Q, Vang YS, Xie XH. Deep Multi-Instance Networks With Sparse Label Assignment for Whole Mammogram Classification. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer (2017). p. 603–11.
- Das K, Conjeti S, Roy AG, Chatterjee J, Sheet D. Multiple Instance Learning of Deep Convolutional Neural Networks for Breast Histopathology Whole Slide Classification. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE (2018). p. 578–81.
- He K, Zhang X, Ren S. Deep Residual Learning for Image Recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016). p. 770–8.

20. Yang H, Tianyi Zhou J, Cai J, Ong YS. Mimi-Fcn+: Multi-Instance Multi-Label Learning via Fully Convolutional Networks With Privileged Information. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017). p. 1577–85.
21. Maaten L, Hinton G. Visualizing Data Using T-SNE. *J Mach Learn Res* (2008) 9(Nov):2579–605.
22. Available at: <https://github.com/Quasimondo/RasterFairy>.
23. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, et al. Grad-Cam: Visual Explanations From Deep Networks via Gradient-Based Localization. In: *Proceedings of the IEEE International Conference on Computer Vision* (2017). p. 618–26.
24. Cao R, Yang F, Ma SC, Liu L, Zhao Y, Li Y, et al. Development and Interpretation of a Pathomics-Based Model for the Prediction of Microsatellite Instability in Colorectal Cancer. *Theranostics* (2020) 10(24):11080–91. doi: 10.7150/thno.49864
25. Skrede OJ, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, et al. Deep Learning for Prediction of Colorectal Cancer Outcome: A Discovery and Validation Study. *Lancet* (2020) 395(10221):350–60. doi: 10.1016/S0140-6736(19)32998-8
26. Roy K, Banik D, Bhattacharjee D, Nasipuri M. Patch-Based System for Classification of Breast Histology Images Using Deep Learning. *Comput Med Imaging Graphics* (2019) 71:90–103. doi: 10.1016/j.compmedimag.2018.11.003
27. Bandi P, Geessink O, Manson Q, Van Dijk M, Balkenhol M, Hermsen M, et al. From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The Camelyon17 Challenge[J]. *IEEE Trans Med Imaging* (2018) 38(2):550–60. doi: 10.1109/TMI.2018.2867350
28. Bejnordi BE, Veta M, Van Diest PJ, Ginneken BV, Karssemeijer N, Litjens G. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* (2017) 318(22):2199–210. doi: 10.1001/jama.2017.14580
29. Kong B, Wang X, Li Z, Qi S, Zhang S. Cancer Metastasis Detection via Spatially Structured Deep Network. In: *International Conference on Information Processing in Medical Imaging*. Cham: Springer (2017). p. 236–48.
30. Zanjani FG, Zinger S, With P. Cancer Detection in Histopathology Whole-Slide Images Using Conditional Random Fields on Deep Embedded Spaces. In: *Medical Imaging 2018: Digital Pathology* (2018) 1058101. doi: 10.1117/12.2293107
31. Kather JN, Pearson AT, Halama N. Deep Learning can Predict Microsatellite Instability Directly From Histology in Gastrointestinal Cancer. *Nat Med* (2019) 25(7):1054–6. doi: 10.1038/s41591-019-0462-y
32. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification. In: *Proceedings of the Ieee Conference on Computer Vision and Pattern Recognition* (2016). p. 2424–33.
33. Yao J, Zhu X, Huang J. Deep Multi-Instance Learning for Survival Prediction From Whole Slide Images. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer (2019). p. 496–504.
34. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
35. Courtiol P, Maussion C, Moarii M, Pronier E, Pilcer S, Sefta M, et al. Deep Learning-Based Classification of Mesothelioma Improves Prediction of Patient Outcome. *Nat Med* (2019) 25(10):1519–25. doi: 10.1038/s41591-019-0583-3
36. Campanella G, Hanna MG, Geneslaw L, Miralflor A, Silva VWK, Busam KJ, et al. Clinical-Grade Computational Pathology Using Weakly Supervised Deep Learning on Whole Slide Images. *Nat Med* (2019) 25(8):1301–9. doi: 10.1038/s41591-019-0508-1
37. Lee H, Yune S, Mansouri M, Kim M, Tajmir SH, Guerrier CE, et al. An Explainable Deep-Learning Algorithm for the Detection of Acute Intracranial Haemorrhage From Small Datasets. *Nat Biomed Eng* (2019) 3(3):173. doi: 10.1038/s41551-018-0324-9
38. Mitani A, Huang A, Venugopalan S, Corrado GS, Peng L, Webster DR, et al. Detection of Anaemia From Retinal Fundus Images via Deep Learning. *Nat Biomed Eng* (2020) 4(1):18–27. doi: 10.1038/s41551-019-0487-z
39. Garland ML, Vather R, Bunkley N, Pearse M, Bissett IP. Clinical Tumour Size and Nodal Status Predict Pathologic Complete Response Following Neoadjuvant Chemoradiotherapy for Rectal Cancer. *Int J Colorectal Dis* (2014) 29(3):301–7. doi: 10.1007/s00384-013-1821-7

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lou, Zhou, Feng, Li, Fang, Fan, Ling, Liu, Zou, Wang, Huang, Yun, Yao and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.