REVIEW

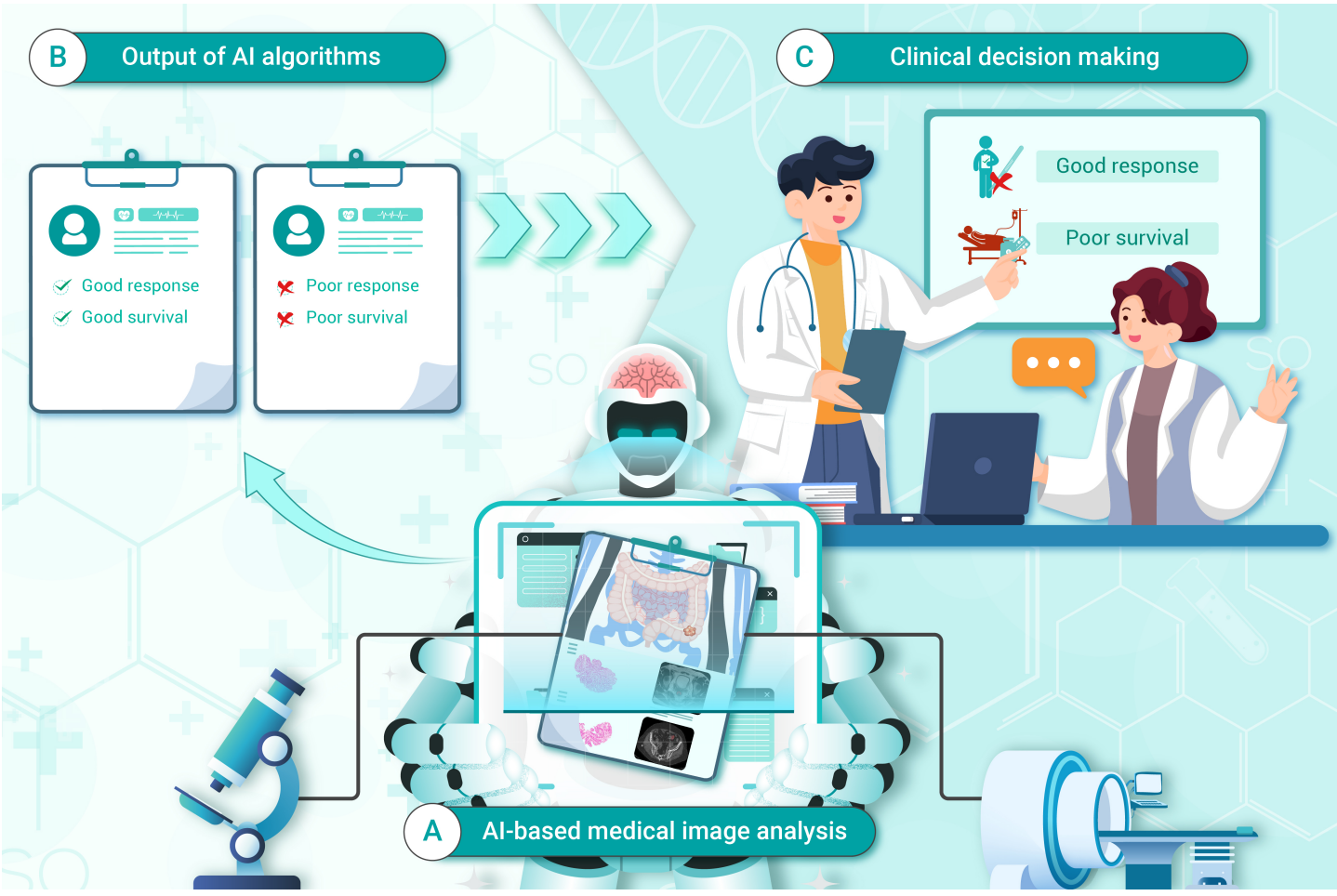
Improving prediction of treatment response and prognosis in colorectal cancer with AI-based medical image analysis

Xiangyu Liu,[1,2] Song Zhang,[2,6] Lizhi Shao,[2] Caixia Sun,[2,3,4] Bao Li,[2,5] Wei Wei,[2] Zuobin Ying,[7,*] Zhenyu Liu,[2,6,*] and Jie Tian[1,2,3,4,*]

*Correspondence: zbying@cityu.mo (Z.Y.); zhenyu.liu@ia.ac.cn (Z.L.); jie.tian@ia.ac.cn (J.T.)

Received: August 29, 2023; Accepted: April 3, 2024; Published Online: April 11, 2024; https://doi.org/10.59717/j.xinn-med.2024.100069

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Patients with colorectal cancer respond heterogeneously to the same treatment.
- Accurately predict patient response and prognosis to treatment helps to select the optimal treatment plan.
- Predictions of treatment response and prognosis can be used to tailor treatment plans at different stages.
- Artificial intelligence-based medical image analysis can be used for prediction tasks.

REVIEW

www.the-innovation.org/medicine

Improving prediction of treatment response and prognosis in colorectal cancer with AI-based medical image analysis

Xiangyu Liu,[1,2] Song Zhang,[2,6] Lizhi Shao,[2] Caixia Sun,[2,3,4] Bao Li,[2,5] Wei Wei,[2] Zuobin Ying,[7,*] Zhenyu Liu,[2,6,*] and Jie Tian[1,2,3,4,*]

[1]Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an 710126, China
[2]CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[3]Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Engineering Medicine, Beihang University, Beijing 100191, China
[4]Key Laboratory of Big Data-based Precision Medicine, Beihang University, Ministry of Industry and Information Technology, Beijing 100191, China
[5]Center for Biomedical Imaging, University of Science and Technology of China, Hefei 230026, China
[6]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100080, China
[7]Faculty of Data Science, City University of Macau, Macau 999078, China
*Correspondence: zbying@cityu.mo (Z.Y.); zhenyu.liu@ia.ac.cn (Z.L.); jie.tian@ia.ac.cn (J.T.)

Received: August 29, 2023; Accepted: April 3, 2024; Published Online: April 11, 2024; https://doi.org/10.59717/j.xinn-med.2024.100069

Citation: Liu X., Zhang S., Shao L., et al., (2024). Improving prediction of treatment response and prognosis in colorectal cancer with AI-based medical image analysis. The Innovation Medicine **2(2)**: 100069.

The heterogeneous response and prognosis of patients with colorectal cancer (CRC) to standard treatment regimens remains a challenge for clinical management. Individually weak prognostic markers, defined by gene mutations and protein expression, are difficult to apply in routine clinical practice because of their high acquisition cost and mediocre prediction accuracy. Visual evaluation of medical images, including radiology and digital pathology images, is an important part of CRC management. With the rapid development of artificial intelligence (AI), high-dimensional imaging features other than visual information are increasingly being used to develop imaging markers. At different stages of treatment, accurate predictions of treatment response and prognosis may help in selecting patients and tailoring their treatment. Here, we review the current state of AI applied to the medical imaging of CRC and describe its recent progress in short-term response and long-term survival prediction. In addition, we illustrate how these AI-based approaches may affect clinical decision-making. Although few approaches have been applied in routine clinical practice, their results are promising. Finally, we discuss the challenges in applying AI in clinical practice and possible future solutions from three perspectives: model interpretability, model generalizability, and patient privacy protection. This comprehensive assessment underscores the transformative potential of AI in CRC management and emphasizes the need for further exploration and integration into routine clinical workflows.

INTRODUCTION

Colorectal cancer (CRC) is the third most common malignant tumor, accounting for approximately 10% of new cancer cases and cancer-related deaths worldwide each year.[1] Despite the different primary treatments recommended by the National Comprehensive Cancer Network (NCCN) guidelines for patients at various clinical stages,[2] there is still heterogeneity in the responses to the same treatment at the same clinical stage. For example, preoperative neoadjuvant chemoradiotherapy is recommended in patients with locally advanced disease to reduce the risk of postoperative local recurrence (LR). Adjuvant chemotherapy is recommended for patients with stage III colon cancer to increase the probability of disease-free survival (DFS) and 5-year overall survival (OS). However, even after receiving the same treatment, the prognosis of patients in the above clinical stages varies greatly.[3] This necessitates the development of biomarkers that can predict treatment responses to facilitate precision therapy for the benefit of patients. During the treatment process, the preoperative, surgical, and postoperative treatment plans of patients can be adjusted by predicting short-term response and long-term survival (Figure 1).

In this study, short-term response specifically refers to patient response to preoperative neoadjuvant chemoradiotherapy. Short-term response prediction can be used to adjust the patient's preoperative and surgical treatment plans. First, approximately 7% of patients do not respond to neoadjuvant therapy. These patients experience less benefit from neoadjuvant therapy, but are still exposed to its associated toxicity.[4–6] Therefore, predicting patient response to neoadjuvant therapy before initiation can avoid overtreatment and aid in the early selection of alternative interventions. Second, approximately 20% of patients achieve pathologic complete response (pCR) after neoadjuvant therapy.[7] Through further assessment, these patients may be able to avoid surgery and obtain excellent rectal preservation and pelvic tumor control through the 'Watch-and-Wait' strategy.[8] However, pCR must be confirmed by histopathological examination of surgically resected specimens. Preoperative prediction of pCR can help clinically adjust the surgical plan.

Long-term survival prediction can be used to adjust postoperative and patient-specific treatment plans such as immunotherapy and targeted therapy. The guidelines[2,9] recommend that patients undergo adjuvant chemotherapy after neoadjuvant treatment and surgery regardless of the surgical pathology results. However, identifying the subgroup of patients that could benefit from adjuvant chemotherapy remains controversial, even among those who have achieved pCR or down-staging.[10,11] The development of prognostic biomarkers independent of treatment response to achieve an accurate stratification of patient prognosis can facilitate the selection of postoperative adjuvant treatment options. Additionally, the status of molecular pathways and key mutations is related to patient prognosis and the potential benefits of specific treatment modalities.[12,13] However, traditional molecular biomarker measurements are labor-intensive and difficult to apply in routine clinical practices.[14,15] Therefore, the prediction of molecular markers related to patient prognosis will further assist in the formulation of patient-specific treatment decisions.

Medical imaging plays an important role in various CRC diagnoses and treatments (Table 1). Radiology imaging is often used to confirm suspected malignancies, diagnose clinical staging, and detect potential prognostic factors.[16] Histopathological slides are the gold standard for diagnosing and staging malignant tumors. With the advancement of digital technology, multi-gigapixel whole-slide images (WSI) are increasingly used to generate information beyond simple diagnosis.[17] Therefore, radiology and digital pathology images are always available for the routine diagnosis and treatment of patients with CRC. However, traditional "semantic" features obtained from images, such as tumor density, tumor size, intratumoral cellular and acellular composition, do not adequately reflect the heterogeneity of the tumor and its dynamic evolution. Interpreting such large amounts of data poses a significant challenge to clinical management.[18]

To that end, artificial intelligence (AI) paves the way for medical image analysis. For clinical requirements, such as lesion detection and segmentation, the performance of AI approaches and surpasses that of human experts.[19,20] More importantly, AI has produced satisfactory results in clinical tasks such as treatment response assessment[21] and prognosis prediction.[22] AI analysis techniques for medical images can be divided into those based on predefined features and those based on deep learning. Radiomics is the method of analyzing radiology images based on predefined features.[23] It aims to build downstream diagnostic or predictive models by encoding images as high-dimensional quantitative features that capture their shapes, wavelets, and textures. Deep-learning methods learn task-related feature representations directly from raw image data using a backpropagation algorithm.[24] Although deep learning is less interpretable and requires a large amount of

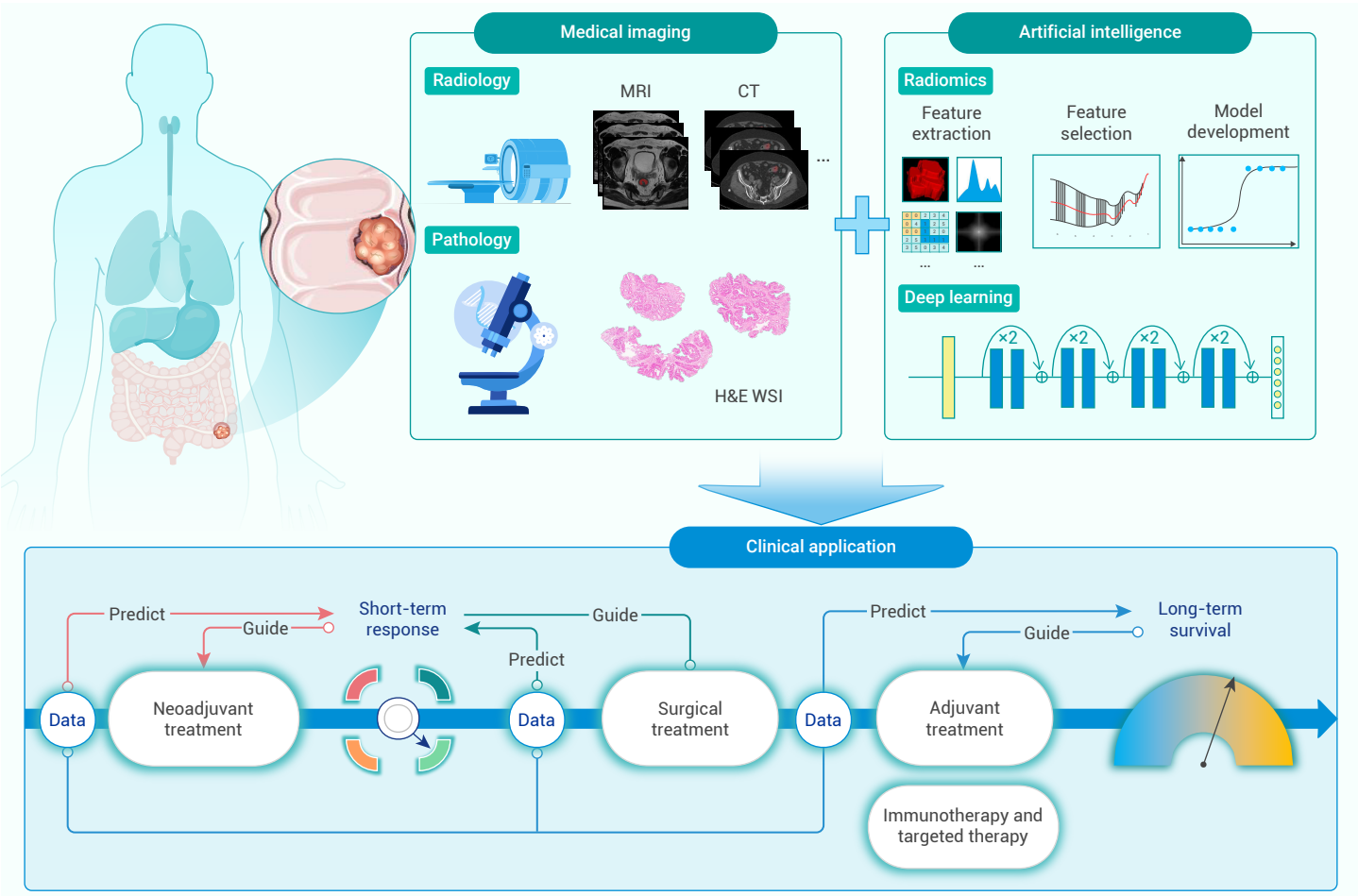


Figure 1. Schematic of artificial intelligence application on CRC medical imaging for clinical support CRC, colorectal cancer; MRI, Magnetic resonance imaging; CT, computed tomography; H&E, hematoxylin and eosin; WSI, whole slide image.

data for training, its robustness to noise[25] and generalization to different imaging domains[26] make it one of the most practical methods for medical image analysis.

AI-based medical image analysis provides a new paradigm for deciding treatment options at different treatment stages of CRC. We review the current AI analysis methods applied to the prediction of short-term response and long-term survival in CRC. These methods can be used to develop novel image-based biomarkers directly or indirectly by analyzing radiology or pathology images. These image-based biomarkers are expected to guide treatment decisions in clinical practice (Table 2). In addition, given that approximately 25% of patients with CRC present with metastasis at initial diagnosis,[27] we further review the current AI-based medical image analysis techniques aimed at assisting the diagnosis and treatment of metastatic CRC (mCRC). Finally, we discuss the challenges and prospects for the practical clinical applications of AI in CRC.

RADIOMICS AND DEEP LEARNING-BASED AI ANALYSIS OF CRC MEDICAL IMAGES

Radiomics analysis of CRC medical images

The concept of radiomics was first proposed in 2012.[23] It defines a medical image analysis workflow that includes tumor segmentation, feature extraction, feature selection, and model construction.

Segmentation. Progress has been made in the automatic segmentation of tumors in radiology images,[28] including methods derived from intensity-based analysis,[29] nondeep learning,[30] and deep-learning-based segmentation.[31] However, owing to the high heterogeneity of colorectal tumors and blurred boundaries between lesions and normal colorectal tissue, automatic segmentation remains a challenging task. In a study on colorectal tumor segmentation based on a fully convolutional network (FCN) and T2-weighted imaging (T2WI), the average dice similarity coefficient (DSC) was 83.56.[32] However, diffusion-weighted imaging (DWI) suffers from noise and artifacts, which makes automatic segmentation more difficult. A deep learning method based on the U-Net architecture had an accuracy of 0.675 DSC for rectal tumor segmentation on volumetric DWI[33] In summary, considerable challenges remain in applying automatic segmentation methods to the radiomics analysis of CRC. Therefore, most recent radiomics studies on CRC still adopt manual or semiautomatic segmentation methods.

Feature extraction. The construction of an accurate predictive model relies on the extraction of robust high-throughput features. In some early studies, high-dimensional feature extraction of colorectal tumor images was usually performed using in-house software.[34,35] However, this approach hampers the reproducibility and comparability of results. Van Griethuysen et al. open-sourced a Python-based radiomics feature extraction platform called Pyradiomics.[36] This platform supports the pre-processing of a wide variety of medical image formats. Additionally, it can support the extraction and standardization of different types of imaging features. This establishes a reference standard for subsequent radiomics research that mostly conforms to the feature definitions set by the Image Biomarker Standardization Initiative (IBSI).[37] Pyradiomics has been widely used in imaging feature extraction of colorectal tumors for efficacy and prognosis prediction tasks.[38–41]

Model development. After the extraction and selection of high-throughput features, a classical radiomic model used for efficacy or survival prediction is constructed through supervised learning. First, the patient's responses to treatment, such as pCR/non-pCR or downstaging/non-downstaging, are one-hot encoded as classification labels. Patient survival outcomes, such as OS and DFS, are defined as regression labels based on paired follow-up time and censoring status. Second, a suitable classification/regression model and loss function are selected to learn the mapping between paired classification/

Table 1. Medical images generated during the diagnosis and treatment of CRC

Image type		Characteristics	
		Advantages	Disadvantages
Radiology image	ERUS	Non-radiative; real time imaging; identification available for earliest tumors	Depth-limited, limited by the patient's body size
	CT	Widely used; rapid imaging; identification available for metastases and tumor stage	Radiation exposure; limited soft tissue resolution
	PET/CT	Malignant tumor detection and localization; identification available for distant metastasis and lymph node metastasis	Radiation exposure; relatively low spatial resolution
	MRI	Non-radiative; multiparametric imaging; high resolution for soft tissue	Relatively slow imaging; higher cost
Pathology image	H&E WSI	Widely used; nuclear characteristics available	Low specificity; relatively limited molecular level information
	IHC WSI	High specificity; characterization of specific biomolecules	Explanatory subjectivity; staining complexity

CRC, colorectal cancer; ERUS, endoscopic rectal ultrasound; CT, computed tomography; PET, positron emission tomography; MRI, magnetic resonance imaging; H&E, hematoxylin and eosin; IHC, immunohistochemistry; WSI, whole-slide imaging

regression labels and high-throughput features. After model development, its performance is validated using various metrics. Different classification algorithms, such as the support vector machine (SVM),[42] LASSO logistic regression,[43,44] and eXtreme Gradient Boosting,[45] have shown promising results in the treatment response prediction of CRC. The widely used Cox regression model and its variant, LASSO-Cox, have achieved good performance in predicting OS, DFS,[46] relapse-free survival,[47] and distant metastasis-free survival (DMFS).[48]

Deep learning analysis of CRC medical images

In some large-scale multi-center or prospective studies, deep learning-based radiology or pathology image analysis has achieved promising results in predicting the treatment response[49] and prognosis[22] of CRC. The convolutional neural network (CNN) and its variants are the most common deep learning methods used for medical image analysis. Using supervised, semi-supervised, or unsupervised methods, a CNN-based framework converts medical images into high-level representations using nonlinear convolution modules. Unlike predefined features, these high-level representations, also called deep learning features, are relevant to downstream tasks such as CRC histopathological classification.[50] Deep learning features have been shown to specifically improve model prediction performance.[51] In addition, with the great success of transformers,[52] especially vision transformers (ViT)[53] in computer vision, more researchers have turned their attention to transformer-based medical image analysis. Compared to CNN, transformers can capture the global context and encode the long-range dependencies of medical images[54] A recent study performed magnetic resonance imaging (MRI) analysis through ViT and successfully predicted the outcomes of rectal cancer.[55] The predictive model achieved good OS and DFS prediction performances in the validation and independent test cohorts. Despite progress, the application of deep learning to medical image analysis faces several challenges. In addition to the general application paradigms described above, we review the methods and strategies for applying deep learning to the specific characteristics of radiology or pathology images.

Radiology image-based deep learning analysis. Deep learning is a data-driven approach whose performance depends on the availability of large-scale datasets. However, owing to strict legal and ethical requirements to protect patient privacy,[56] the availability of radiology images for training deep-learning models is often limited. This leads to the overfitting of deep learning models on small-sample data. Data augmentation is widely used in the preprocessing of radiology images to adaptively increase the size of the training samples. This operation includes random rotation, mirroring, and noise injection. Cropping is generally rejected because it may distort the region of interest (ROI).[57] On the other hand, the principle of transfer learning pertains to pre-training the deep learning model on a large dataset, such as ImageNet,[58] and then fine-tuning the pre-trained model on a smaller medical dataset.[59] This strategy enables the sharing of common low-level features, thereby reducing the risk of overfitting. In addition, multi-center data validation is a key step in measuring the robustness and generalization of deep learning methods. However, data discrepancies between centers caused by different scanning instruments and parameters, also known as domain shifts, hinder the generalization of deep learning models. Numerous studies have focused on image- or feature-level alignments for domain adaptation using supervised, semi-supervised, or unsupervised methods.[60] However, there are few reports on domain adaptation methods for radiology images of CRC, which may be an important future direction to further improve the generalization performance of deep learning models. Finally, unlike images from natural scenes, radiology imaging sequences usually consist of multiple grayscale images. These continuous images contain tumor spatial information, which serves as important prior knowledge for building predictive models. Some deep-learning-based studies that encoded the spatial information of tumors through strategies such as 3D CNN,[61] attention mechanisms,[62] and reinforcement learning[63] achieved performance improvements in downstream tasks.

Pathology image-based deep learning analysis. Gigapixel WSI cannot be handled directly using a CNN. The usual processing flow is to patch the WSI through the ROI and aggregate the patch-level features or predicted values using the multi-instance learning (MIL) method to obtain the WSI-level analysis results. However, manual delineation of the WSI ROI is labor intensive. In some large-scale international multi-center studies analyzing CRC WSIs, ROIs were defined using pre-trained segmentation models[22] or tissue-type classification models.[50,64] In addition, different staining materials and methods may lead to color differences between the WSIs from different centers, which is not conducive to the generalization of the predictive model. Staining normalization methods such as SPCN[65] and Macenko[66] are therefore widely used in WSI pre-processing. Multiscale information on WSI at different magnifications (40X, 20X, 10X) is often utilized during model development. Consistent with the pathologist's diagnostic process, this combination of macroscopic and microscopic features enables a more comprehensive understanding of tumor data. Finally, the effective aggregation of patch-level information into WSI-level decisions is an important issue in WSI research. In some large-scale studies that analyzed the WSIs of other cancer types, different MIL methods showed good patch-level information aggregation performance and were well suited for downstream prediction tasks. These MIL approaches include tile probability ranking,[67] attention pooling,[68] and graph deep learning,[69] and have achieved good results in cancer grading[70] and gene mutation prediction.[71] Such results provide a promising outlook for efficacy and prognosis prediction in CRC and prompt further validation of the aforementioned methods using CRC datasets. The AI-based WSI processing flow (excluding staining normalization) is shown in Figure 2.

AI-BASED MEDICAL IMAGE ANALYSIS FOR SHORT-TERM RESPONSE PREDICTION OF CRC

Short-term response prediction in CRC is usually defined as a binary classification task that is used to assist in the personalized formulation of preoperative treatment plans and surgical protocols. According to specific applica-

Table 2. Overview of radiology and pathology image application for short-term response and long-term survival prediction

Target	Research type		References	Label	Data		Technology	Evaluation		
		Summary			Type	Size (N)		Main metric	Performance	External validation
SHORT-TERM RESPONSE PREDICTION	Single-time point data-based	Pre-treatment prediction	[35, 39–41, 43–45, 72–75, 77, 78]	pCR/TRG/ LNM	MRI/CT/ PET/ H&E WSI	48–1033	Radiomics/ Multimodal fusion	AUC	0.77–0.97	[39, 45, 73]
		Post-treatment prediction	[38, 81]	pCR/TRG	MRI	114–898	Radiomics	AUC	0.82–0.93	
	Multi-time points data-based	Pre- and post-treatment prediction: Direct combination	[42, 49, 82]	pCR/TRG/ ypT0-1N0	MRI	125–383	Radiomics/ CNN	AUC	0.70–0.99	[82]
		Pre- and post-treatment prediction: Dynamic information	[83–87]	pCR/TRG	MRI	39–622	Delta-radiomics/ CNN	AUC	0.91–0.97	[83, 87]
		Pre- and mid-treatment prediction	[88]	pCR and TRG	MRI	51	Radiomics and CNN	AUC	0.86 and 0.93	
LONG-TERM SURVIVAL PREDICTION	Surrogate approaches	MSI	[64, 95–99]	MSI	MRI/CT/ H&E WSI	276–762	Radiomics/ CNN	AUC	0.77–0.90	[64, 99]
		KRAS	[101–105]	KRAS	CT/MRI	83–390	Radiomics/ CNN/ Fourier analysis	AUC	0.68–0.94	[102]
		Multiple molecular biomarkers	[71, 106, 107]	MSI/BRAF/ TP53/KRAS, etc.	H&E WSI	499–1888	GCN/MIL/ WSL/ViT	AUC	0.52–0.91	[71, 106, 107]
	End-to-end approaches	Radiology image-based	[46, 48, 55, 74, 109, 110]	OS/DFS/ DMFS	CT/PET/ MRI	86–1028	Radiomics/ CNN/RNN/ ViT	C-index	0.67–0.83	[48, 55, 109]
		Pathology image-based	[22, 50, 108, 111]	OS/DFS/ RFS/Specific survival	H&E WSI	210–4515	CNN/GAN	HR; C-index	1.63–10.71; 0.65–0.79	[22, 50, 111]

pCR, pathologic complete response; TRG, tumor regression grade; LNM, lymph node metastasis; MSI, microsatellite instability; KRAS, Kirsten rat sarcoma; OS, overall survival; DFS, disease-free survival; DMFS, distant metastasis-free survival; RFS, relapse-free survival; AUC, area under the ROC curve.

tion scenarios, short-term response is usually defined as pCR (the absence of viable tumor cells) versus non-pCR, and good response (GR) versus non-GR according to the tumor regression grade (TRG). As previously mentioned, patients who are predicted to be responders can avoid unnecessary surgery through a 'Watch-and-Wait' strategy. For patients predicted to be non-responders, unnecessary treatment toxicities can be avoided and preoperative treatment strategies can be changed.

Binary classification models based on high-dimensional image features have achieved promising results in predicting treatment responses. Preoperative imaging examinations for patients with CRC usually include computed tomography (CT), positron emission tomography (PET)/CT, and MRI. These imaging examinations are performed at different time points during tumor diagnosis and treatment, including before, during, and after treatment. Most studies aim to predict patient responses through AI analysis of images at a single time point. As image data at a single point in time are relatively easy to obtain, large-scale datasets of this type can be constructed for AI model training and validation. However, information from these images cannot fully reflect the dynamic evolution of a tumor as well as multi-time point images, although problems such as incomplete data and missing modals in the application persist. In the following section, we review the current AI approaches for short-term response prediction based on preoperative single-time-point and multi-time-point CRC images.

Treatment response prediction based on single-time point images

Predicting patient responses using pre-treatment images is useful for developing personalized neoadjuvant treatment plans. MRI, a tool recommended by the guidelines, is the most commonly used diagnostic imaging method for colorectal tumors.[2,9] To predict pCR/non-pCR and GR/non-GR after neoadjuvant treatment, Nie et al.[35] for the first time developed classification models by extracting the radiomics features of pre-treatment multiparameter MRIs and feeding them into a three-layer artificial neural network. In the four-fold cross-validation, the area under the ROC curve (AUC) was 0.84 for pCR prediction and 0.89 for GR prediction. However, the study was limited by its small sample size (48 patients), which may have caused some bias in the results. In some follow-up studies, different strategies have been used to overcome this bias. These strategies include expanding the sample size of the dataset,[43,44,72] including data on patients from multiple countries[41] and using leave-one-out nested cross-validation.[40] These studies achieved good prediction results for neoadjuvant treatment response by extracting the radiomics features of tumors or tumor margins. The imaging modality used was either multiparametric MRIs or only T2WIs. AUC values in predicting pCR, GR, non-response or lymph node status were mostly/all greater than 0.8 in these studies. In addition to MRI, pre-treatment enhanced abdominal CT[73] and PET/CT[74,75] have been used to predict patient responses to neoadjuvant treatment. The validation results on an independent external validation cohort of 72 patients showed that pre-treatment CT had an AUC of 0.9 in predicting lateral lymph node metastases after neoadjuvant treatment. A meta-analysis of fluorodeoxyglucose (FDG) PET/CT-based prediction of response to neoadjuvant treatment showed that a global cohort of 1526 patients had good pooled accuracy, with an AUC of 0.83.[76]

The aforementioned pre-treatment image-based research focuses on the application of single-modal images. The combination of multimodal data further improve the prediction performance of the model. Multimodal data usually consist of a combination of different radiology images, such as MRI and PET/CT,[77,78] or a combination of radiology and pathology images.[39,45] In a prospective cohort validation study, the AUC for predicting pCR using a radiopathomics model combining radiology and pathology information was 0.812.[39] Compared with the MRI-based, pathomics nucleus-based, and pathomics microenvironment-based models, the accuracy of the radiopathomics model increased by 10%, 8%, and 18%, respectively. To date, the main method for AI-based multimodal fusion of CRC is late fusion. In other words, quantitative features are extracted from images of different modalities and concatenated, then used to develop a classification model. Although this strategy has achieved preliminary results, it does not fully consider the interactions and related information between the different modalities of data. Hence, the early fusion approaches for multimodal data require further investigation.

Compared to pre-treatment images, post-treatment images are closer to the gold standard for the pathological assessment of patient response to treatment. In earlier studies, MRI was used to assess clinical complete

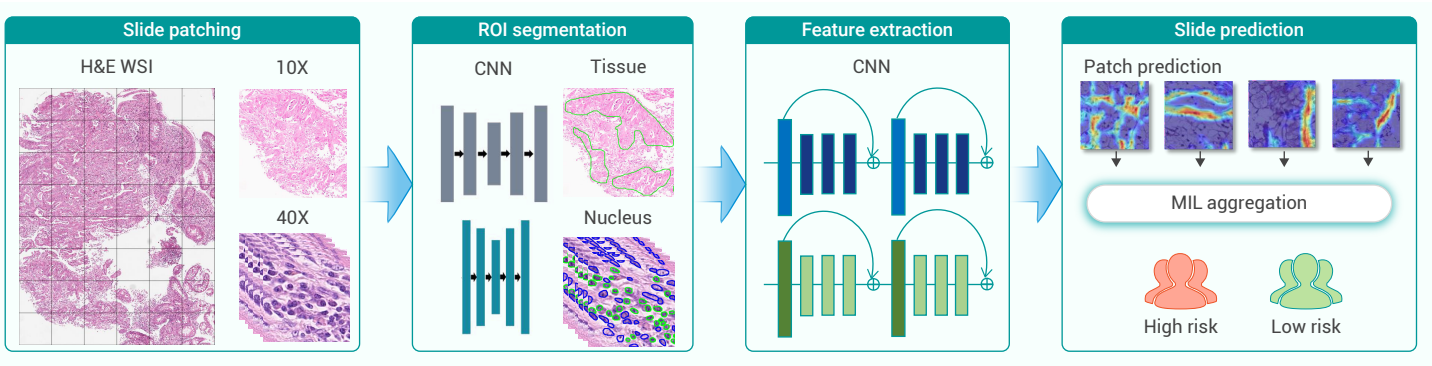


Figure 2. Schematic of WSIs processing through AI methods WSI, whole slide image; AI, artificial intelligence; ROI, region of interest; H&E, hematoxylin and eosin; CNN, convolutional neural network; MIL, multi-instance learning.

response (cCR),[79,80] a surrogate imaging marker of pCR. However, cCR is insufficient to support nonsurgical strategies.[2] The purpose of using post-treatment images to predict (or assess) patient response to treatment is mainly to guide surgical strategies, such as organ preservation and "Watch-and-Wait." Through cross-validation, Horvat et al.[81] proposed that radiomics analysis using only post-treatment T2WI had a better pCR classification performance than qualitative assessment using T2WI and DWI. In addition, a recent study with a large-scale validation dataset (n = 306) showed that the post-treatment T2WI-based radiomics model had better classification performance than expert-assessed TRG on MRI (AUC, 0.82 vs. 0.74, P = 0.009).[38] This study also showed that the addition of apparent diffusion coefficient (ADC) radiomics analysis did not improve the model classification performance, which may be because the advantage of radiomics lies in its ability to extract difficult-to-perceive pattern changes in images.

Treatment response prediction based on multi-time points images

Combining image information at multiple time points to develop a predictive model may improve the accuracy of the response prediction. Liu et al.[42] for the first time used pre- and post-treatment T2WI and DWI images to establish a radiomics model to predict pCR for patients with LARC. The predictive model achieved an AUC of 0.9756 for the internal validation cohort. Bulens et al.[82] further validated the performance of a radiomics model based on pre- and posttreatment MRI in a prospective cohort. Although not superior to the semantic model, the radiomics model achieved an AUC of 0.83, demonstrating its potential as an additional tool for identifying patients eligible for organ-preserving treatment. In addition, a deep learning-based prospective study showed that pCR can be effectively predicted by CNN analysis of pre- and post-treatment diffusion kurtosis MRI.[49] The model outputs can be used to assist radiologists in clinical decision-making.

A limitation of the aforementioned studies is that they did not fully consider the dynamic evolution of colorectal tumors on images over time. Delta radiomics is defined as changes in radiomics features on a set of longitudinal images that often reflect more individualized information. Several studies have performed pre- and post-treatment MRI analysis based on delta radiomics to predict pCR or tumor downstaging in CRC after neoadjuvant treatment.[83-86] Although good response prediction results were obtained, these studies have drawn controversial conclusions. Shayesteh et al.[83] pointed out that the delta radiomics model outperformed models based on both pre- and post-treatment features. However, according to the results for Chen et al.,[84] the pre-CRT, post-CRT, and delta radiomic-based models showed no significant difference in predicting three different responses. These contradictory conclusions may be due to the small sample size and retrospective nature of the data collection. This necessitates the development of prospective multi-center studies to demonstrate the reliability and robustness of delta radiomics for response prediction. In addition to reflecting the dynamic evolution of tumors through changes in predefined high-dimensional features, deep-learning-based strategies have gained popularity in recent years. Jin et al.[87] designed a multitasking deep-learning framework based on the Siamese network architecture. Feature representation was achieved through an in-depth comparison of pre- and post-treatment images; rectal tumor segmentation and neoadjuvant response prediction were performed simultaneously. In different validation cohorts, the AUCs of the model containing multi-time point dynamic information were improved by 14–19% compared to the model trained using only single-time point images. This promising result demonstrates the potential of deep learning for extracting information regarding the dynamic evolution of tumors. The above discussion mainly focused on the application of pre- and post-treatment images at two time points. Shi et al.[88] predicted patient response by analyzing multiparametric MRI scans before and during treatment. The advantage of this strategy is that it not only improves prediction accuracy by using information from multi-time point images, but also proactively changes the treatment plan according to the prediction results mid-treatment. This method has achieved a certain predictive effect, and the AUCs for predicting pCR and GR in the cross-validation were 0.86 and 0.93, respectively. However, the sample size of the dataset included in this study was small (n=51), which may have been limited by the strict inclusion and exclusion criteria.

AI-BASED MEDICAL IMAGE ANALYSIS FOR LONG-TERM SURVIVAL PREDICTION OF CRC

Long-term survival of patients was defined by different follow-up endpoints, including LR, distant metastasis (DM), and OS. By accurately predicting the prognosis and realizing prognostic risk stratification of different patients, clinicians can formulate personalized treatment plans for immune/targeted therapy and postoperative adjuvant therapy. For example, postoperative adjuvant chemotherapy is recommended for patients considered to be at high risk of DM. Biomarkers such as genetic mutations and copy number alterations are associated with the prognosis of patients with CRC.[89] Therefore, two ways to predict the prognosis of CRC through AI-based image analysis are: "surrogate" and "end-to-end".[90] The pipeline of "surrogate" methods is to predict the status of prognostic biomarkers or mine information related to these biomarkers from images, thereby guiding treatment strategies. In "end-to-end" methods, patient prognosis is directly predicted by extracting predefined features or training deep learning models through images. The prognostic risk in different patients can then be stratified using prognosis-related imaging markers. Next, we review the progress of "surrogate" or "end-to-end" approaches for predicting CRC prognosis, respectively.

Prognosis prediction based on surrogate approaches

The NCCN guidelines[2] recommend that all patients newly diagnosed with CRC undergo microsatellite instability (MSI) testing. MSI is defined as a change in the length or base composition of microsatellite sequences caused by an abnormal DNA mismatch repair (MMR).[91] This change can be divided into MSI-high (MSI-H), MSI-low (MSI-L), and microsatellite stability (MSS) according to the degree. MSI status correlates with the prognosis of patients with CRC[92] and can be used to evaluate their response to fluoropyrimidine-based chemotherapy[93] and immunotherapy.[94] Some studies have performed MSI prediction using radiology image analysis, in which the prediction task was defined as a binary classification between MSS and MSI. The imaging modalities used for the analysis included CT[95,96] and T2WI.[97,98] In general, the radiomics- and CNN-based models performed well in predicting MSI from

radiology images, with AUCs ranging 0.77–0.90 on validation cohorts. However, the validation cohorts in the above studies were randomly divided from multiple centers, and the results require further validation through prospective studies. Furthermore, the interpretability of high-dimensional imaging features in predicting MSI remains unexplored. On the other hand, Hematoxylin and eosin (H&E)-stained WSI contain human-recognizable MSI-related morphological information.[64] Prediction of MSI based on WSI has achieved promising results in international multi-center studies with large-scale datasets. Kather et al.,[99] for the first time, performed MSI prediction of gastrointestinal cancer using H&E-stained WSI. Automatic tumor detection and MSI status classification were performed using deep residual networks. The model achieved a prediction AUC of 0.84 on the CRC dataset. Based on the above findings, Yamashita et al.[64] further evaluated the added value of deep learning in predicting MSI in the clinical detection paradigm, where the deep learning model outperformed experienced experts in gastrointestinal pathology in predicting MSI.

Kirsten rat sarcoma (KRAS) mutation is a negative biomarker for targeted therapy with epidermal growth factor receptor and is associated with drug resistance.[100] The analysis results of CT[101] and T2WI[102–105] showed the potential of using radiology images to predict mutant and wild-type KRAS. Because of the common visual similarity between images of mutant and wild-type KRAS, Ma et al.[104] combined information from T2WI in the spatial and frequency domains to predict KRAS based on fully preserved features. This strategy achieved an AUC of 0.94 in ten-fold cross-validation and outperformed other deep learning-based methods.

In addition to predicting a single molecular biomarker, several studies have aimed to predict the status of various molecular pathways and key mutations using deep learning methods to analyze specific paradigms of CRC pathology images. Bilal et al.[106] designed a weakly supervised learning framework comprising three CNNs. The framework implements WSI-based prediction through tumor tile detection, molecular status-related tile ranking, and nuclear segmentation (classification). For biomarker prediction, including hypermutation, MSI, chromosomal instability, BRAF/TP53/KRAS, and CpG island methylator phenotype, the framework outperformed previously published state-of-the-art results. In this study, specific WSI phenotypes were associated with MSI and hypermutations. Ding et al.[71] developed a feature that encodes tumor spatial heterogeneity in WSI using a graph-neural-network-based method. Using this strategy, a colon cancer dataset was used to build predictive models for genetic mutations, copy number alterations, functional protein expression, and MSI. These models can be further generalized to the RC to achieve relatively robust prediction results. In a recent study,[107] CNN and ViT were used to jointly analyze WSI in CRC. In the identification of genomic and proteomic statuses, and the prediction of OS and DFS, this framework achieved good prediction performance without detailed region-level annotations. More importantly, a method for model interpretation was proposed in this study. Model interpretations based on expert pathological knowledge were obtained by correlating the WSI phenotypes with predicted probabilities to generate concept scores.

Prognosis prediction based on end-to-end approaches

The "end-to-end" prediction of prognosis for CRC patient has been formulated as regression tasks[108] or binary classification tasks[22,109,110] in previous studies. The labels of the regression tasks were jointly defined by follow-up time and event status, and the predictive models assigned patients a continuous risk score by estimating the risk of endpoint events. However, the existence of censored data makes it difficult to apply deep learning approaches to regression tasks on medical datasets with small sample sizes. For example, censored data make it more difficult for the model training to converge. Therefore, classification-based studies define a cut-off for follow-up time, such as the median follow-up time, and directly define endpoint events such as the presence or absence of DM, RF, etc., for predictive model development.

Radiomics and deep learning approaches have been applied for the direct prediction of patient prognosis from radiology images. In an MRI-based study,[48] Liu et al. developed radiomic signatures for DMFS prediction by extracting radiomic features from T2WI and ADC images of patients with LARC. In addition to being independent prognostic factors for DMFS, radiomic signatures can be used to assess the benefit of adjuvant chemotherapy in specific patient subgroups. On this basis, a deep learning-based study was conducted to predict DMFS in patients with LARC receiving neoadjuvant chemoradiotherapy[109] The nomogram developed in this study can effectively stratify the risk of DM in patients, regardless of whether they achieve pCR or downstaging. In addition, some MRI-based studies have predicted multiple clinical outcomes, including OS, DFS, and DMFS, using radiomics or deep learning approaches.[46,55] The main added value of these studies is the determination of treatment benefits for patients. In addition to the application of multiparametric MRIs, FDG PET/CT[74] and serial CT images[110] have been used to predict patient prognoses, such as OS and DFS. Lovinfosse et al.[74] pointed out that texture analysis of FDG PET/CT provided a better prognostic predictive ability than intensity- and volume-based parameters. The model proposed by Lu et al.[110] used CNN and RNN structures to represent tumor morphological changes in serial CT images, thereby achieving an accurate prediction of OS in mCRC.

To predict CRC prognosis based on pathology images, different deep learning models have been developed to fully extract prognosis-related information from WSI. In a multi-center study,[50] more than 100,000 patches were generated from hand-delineated ROIs on CRC tissue slides. Each patch belonged to one of nine tissue categories. Pretrained CNN models with transfer learning were used for patch classification. From the output of the classification model, a biomarker significantly associated with OS was calculated, called the "deep stroma score". Another international multi-center study proposed a pipeline for predicting patient cancer-specific survival using WSI,[22] with one CNN for the automatic segmentation of cancer foci and ten CNNs for classification. The DoMore-v1-CRC biomarker developed using this pipeline can guide the selection of adjuvant treatment for stage II and III patients, such as avoiding therapy and undergoing more intensive treatment regimens. Based on this, Kleppe et al.[111] created a clinical decision support system. By integrating DoMore-v1-CRC and pathological staging markers, patients were divided into three risk groups. Low-risk patients with stage II or III CRC could avoid additional adjuvant chemotherapy, which was confirmed using an external validation dataset. Furthermore, a recent study encoded the WSI using an unsupervised deep learning network to help identify patients with stage III CRC who required prolonged chemotherapy to improve their prognosis.[108]

AI-BASED MEDICAL IMAGE ANALYSIS FOR METASTATIC CRC

For CRC diagnosis, preoperative knowledge of lymph node metastasis can provide valuable information for determining the need for adjuvant therapy and the appropriateness of surgical resection. In 2016, Huang et al.[34] introduced the concept of a nomogram into radiomics, visually demonstrating the contribution of the radiomics signature and independent clinical factors to the final prediction of lymph node metastasis in CRC. Since then, various studies have predicted lymph node metastasis using radiomics[44,73] or deep-learning methods.[112] The results of these studies suggest that AI-based models predict lymph node metastasis in patients with CRC using MRI or CT more accurately than radiologist assessments.

In the treatment of patients with mCRC, oxaliplatin-based first-line chemotherapy and targeted therapy have made significant progress in recent years, thereby improving prognosis. However, given the heterogeneous response of patients to this treatment, several studies have developed medical imaging markers to predict the treatment response, thereby assisting in personalized early treatment decision-making. CT is the preferred imaging modality for such analyses. To assess the response of mCRC to chemotherapy, Nakanishi et al.[113] classified patients into responders and non-responders based on the Response Evaluation Criteria for Solid Tumors (RECIST). The radiomic signature constructed based on pre-treatment CT showed a good predictive value in the validation cohort (AUC 0.779). Furthermore, Giannini et al.[114] comprehensively analyzed CT images at different treatment stages using delta radiomics. This strategy had a higher prediction accuracy for the RECIST assessment-based chemotherapy response, with an AUC higher than 0.9 in both the training and validation datasets. For targeted therapy response assessment, Dohan et al.[115] and Lu et al.[110] performed a CT analysis based on radiomics and deep learning to predict the OS of patients treated with anti-vascular endothelial growth factor (anti-VEGF) agents. The results showed that these methods were better at identifying patients' treat-

ment responses than RECIST. The 2-month radiomic nomogram proposed by Dohan et al. had the same prognostic value as the 6-month RECIST1·1 but had better predictive performance than the 2-month RECIST1·1 (HR: 2.07 *vs.* 1.40, $P < 0.0008$ *vs.* $P = 0.238$). The DL-Nomo model proposed by Lu et al. incorporated the DL prediction score into the Size-Nomo model and achieved the best performance among all models (C-index 0.694 (95% CI: 0.661, 0.720), all $P < 0.001$, z-test). Although showing acceptable discrimination performance (0.657 (95% CI: 0.546, 0.769)), the RECIST 1.1 model added little clinical value to the early prediction of OS.

CHALLENGES AND PROSPECTS

With the rapid development of AI in recent years, studies on the prediction of short-term response and long-term survival of patients with CRC from medical images have shown promising results. In addition to improving the accuracy of prediction, researchers are increasingly interested in the development of imaging biomarkers that can influence and guide specific treatment plans for patients. However, the application of AI-based imaging analyses in real-world clinical practice remains challenging. These challenges include technical as well as moral and ethical considerations. In the following section, we summarize the challenges and potential solutions to model interpretability, generalizability, and data privacy protection.

Model interpretability

As the size of medical datasets increases, data-driven deep learning approaches have achieved better performance than handcrafted feature-based models in various downstream tasks. However, the "black box" nature of deep learning makes it difficult for clinicians to fully trust the decision-making process. Therefore, in-depth research on model interpretability is beneficial for deep learning visualization applications. Some strategies, such as class activation maps,[116] aim to reveal regions in the original input image that the model pays the most "attention" towards. For vision tasks, such as cancer segmentation and lesion detection, such visualization strategies allow for an effective comparison of the model output to expert knowledge and subsequent evaluation of the model performance. However, for response and prognosis prediction tasks, these strategies result in poor interpretability. This is because the response and prognosis prediction tasks are usually non-What You See is What You Recognize. This makes visualization strategies less attractive and useful only for excluding erroneous outputs, causing them to seek explicability rather than interpretability.[117] Nevertheless, some progress has been made in radiology and pathology imaging to seek the biological interpretation of features and models.

For radiology image analysis, habitat imaging is a way to describe tumor heterogeneity. It associates high-dimensional features with specific tissue phenotypes, such as necrosis and edema, by classifying different tumor regions.[118–120] Since tumor heterogeneity is related to the treatment response and prognosis of patients, model development using features obtained via habitat imaging is interpretable to a certain extent. However, there is currently no unified measurement template for habitat imaging, and more standardized and prospective research is required. For pathology image analysis, the phenotypes of WSIs can be obtained by automatic clustering of patches[121] or by building a patch classifier.[50] After phenotype recognition, the contextual information of WSIs can be further obtained using a graph convolutional network[122] or attention-based MIL methods,[68] and then applied to downstream tasks. Finally, molecular biomarkers such as MSI, KRAS, and tumor mutational burden are related to patient response or prognosis. Predictive model development by mining information related to these molecular biomarkers in radiology[123] or pathology images[124] may provide more modeling details than end-to-end prediction.

Model generalizability

Formulation of standard protocols such as CLAIM,[125] IBSI,[37] and TRIPOD[126] has enhanced the repeatability of AI model results in medical image applications and improved the quality of research. However, despite the excellent performance of AI models on specific datasets, their performance degradation on external datasets has significantly hindered their application in clinical practice. This phenomenon may have been caused by data heterogeneity and inappropriate model training strategies. From the perspective of data, non-independent and identically distributed (non-IID) data can result from inconsistent collection protocols and label shifts among different centers. The IID of data is an idealized assumption for many AI model applications. From the perspective of model development, overfitting caused by small training sets is a common problem when AI is applied to medical image datasets. Building AI models for domain adaptation is currently a research hotspot. Ganin et al.[127] suggested achieving domain adaptation to some extent by training the model to be more separable from the learning task in the source domain and more invariant to different domains.

First, the performance of the predictive model in the source domain should be improved as much as possible while preventing overfitting. For the segmentation and classification of medical image datasets, some general few-shot learning methods, such as transfer learning,[59] meta learning,[128] and metric learning,[129] have made progress. Multimodal fusion is a promising research direction because conclusions regarding clinical diagnosis and treatment can be drawn by integrating different modal data. Combining multimodal information with AI approaches may lead to predictions with higher accuracy and improved generalization. Combining easily available clinicopathological factors and high-throughput imaging features is an effective way to improve the model's prediction performance. The feasibility of this strategy has been confirmed in prediction tasks including pCR,[42] DM,[109] and lymph node metastasis.[34] The fusion of high-dimensional representations of radiology and pathology images and genomic data has attracted the attention of an increasing number of researchers. Recent studies have adopted the visual question answering mechanism-based method[130] or correlation-guided method[131] to obtain interpretable multimodal-fused features.

Second, model variability across domain transfers should be minimized. Typically, the pre-processing of data is necessary to train AI-based predictive models. Grayscale or color normalization, resampling, and ROI selection are widely used to preprocess radiology and pathology images. Domain adaptation through frequency domain pre-processing is a simple and effective method. Yang et al.[132] achieved unsupervised domain adaptation by exchanging low-frequency spectra of the source and target domain images. This method performs style transfer while preserving the high-level semantics perception of the target domain images. In terms of model building, adversarial learning approaches[133,134] minimized domain shifts by maximizing the confusion between the source and target representations through trained discriminators.

Finally, while early research relied on densely annotated medical image data for training robust neural networks, obtaining such detailed manual annotations is often expensive, laborious, and prone to intra- and inter-observer variability. This flaw is particularly evident when gigapixel WSIs are annotated. With the great success of self-supervised pre-training in natural languages, many studies have focused on exploring the potential of this method in medical images. The core concept of this method is to effectively learn image representations through pretext tasks and unlabeled data. The learned image representations are then transferred to specific downstream tasks, thereby accelerating the learning process of the model using limited labeled data. For CRC recognition,[135] detection,[136] tissue type classification,[137] survival prediction,[138] and MSI prediction,[139] the results of different studies show promising prospects for self-supervised pretraining. For example, according to the method proposed by Azizi et al.,[138] in out-of-distribution settings, only 1–33% of the data were required for retraining to match the performance of supervised models retrained using all available data. The advantage of self-supervised pre-training is that it is beneficial for model generalization. However, at present, there are few large-scale public datasets of CRC radiology images, which pose certain difficulties for the application of self-supervised pre-training methods. Whether radiology image analysis of CRC can benefit from this strategy remains to be studied.

Patient privacy protection

The limited availability of datasets for AI algorithm training and validation hinders their widespread use in clinical practice. The main reasons for this are the lack of standard electronic medical records and strict patient privacy protection laws.[56,140] Sharing data between different centers must be approved by the institutional review board or ethical committee. To address this challenge, federated learning (FL)[141] has been proposed to effectively

utilize multi-center data through distributed training while protecting patient privacy and data security. The basic principle of FL involves setting up a central server, learning the distribution of different central data under the premise of integrating model parameters instead of data, and applying them to downstream tasks. Federated averaging (FedAvg) is a simple method of aggregating model parameters. Based on the proportion of data in different centers, the FedAvg algorithm weighs and averages the model parameters on the central server and feeds them back to each center. FedAvg achieved preliminary results in the clinical tasks of COVID-19 lung abnormalities detection,[142] rare cancer boundary detection,[143] response prediction to neoadjuvant treatment,[144] and prognosis prediction.[145] However, the FedAvg strategy may exacerbate the negative impact of non-IID data on model training and affect model convergence. Some studies have realized the development of domain adaptation and domain generalization algorithms based on the FL framework and applied them to multi-center datasets under patient privacy protection.[128,146–148] The results of these studies indicate that designing novel federated domain adaptation or generalization methods is a promising research direction for medical image analysis. They can maintain the accuracy of the predictive models while protecting patient privacy.

DISCUSSION

AI-based medical image analysis provides opportunities for precision medicine development. This article comprehensively reviews the primary AI methods for the medical image analysis of CRC and the current research status of such methods in predicting the neoadjuvant chemoradiotherapy response and prognosis of patients with CRC. Current research results show that fully utilizing medical image information can assist doctors in making clinical diagnoses and treatment decisions at different stages, resulting in potential clinical benefits. First, accurate prediction of a patient's short-term response from pre- and post-treatment images can help tailor neoadjuvant treatment plans and surgical options, respectively. The comprehensive utilization of image information at different time points is expected to further improve prediction performance. Second, the prediction of a patient's long-term prognosis using end-to-end or surrogate methods can help clinicians make decisions regarding adjuvant treatment options. For example, accurate prediction of DM can help formulate postoperative adjuvant chemotherapy regimens. Additionally, in surrogate approaches, mining images for information related to molecular markers can increase the interpretability of prognostic prediction models.

Despite these promising results, the methods proposed in the aforementioned studies still face challenges in their application in clinical practice. A limited sample size is one of the main factors hindering the success of these methods. AI models built using small-sample datasets may not be sufficient for handling complex clinical scenarios. Furthermore, the proportion of studies with external validations remains low. Among the 48 studies on short-term response and long-term survival predictions introduced in this review, only 18 conducted an external validation of the proposed method (Table 2). Expanding the sample size and conducting multi-center prospective studies are important for the clinical promotion of AI methods. Moreover, medical image analysis methods such as multi-modal fusion and self-supervised pre-training should be considered in future research. Model performance and interpretability should be reported in sufficient detail in the corresponding studies. Finally, constructing AI models while protecting patient privacy is an important aspect of applying these analytics to a wide range of clinical practices. Thus, addressing model performance degradation under a federated learning framework has broad research significance.

REFERENCES

1. Sung, H., Ferlay, J., Siegel, R.L., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **71**: 209–249. DOI: 10.3322/caac.21660.
2. Benson, A.B., Venook, A.P., Al-Hawary, M.M., et al. (2022). Rectal cancer, version 2.2022, NCCN clinical practice guidelines in oncology. J. Natl. Compr. Cancer Netw. Network **20**: 1139–1167. DOI: 10.6004/jnccn.2022.0051.
3. Keller, D.S., Berho, M., Perez, R.O., et al. (2020). The multidisciplinary management of rectal cancer. Nat. Rev. Gastroenterol. Hepatol. **17**: 414–429. DOI: 10.1038/s41575-020-0275-y.
4. Bosset, J.-F., Collette, L., Calais, G., et al. (2006). Chemotherapy with preoperative radiotherapy in rectal cancer. N. Engl. J. Med. **355**: 1114–1123. DOI: 10.1056/NEJMoa060829.
5. Roh, M.S., Colangelo, L.H., O'Connell, M.J., et al. (2009). Preoperative multimodality therapy improves disease-free survival in patients with carcinoma of the rectum: NSABP R-03. J. Clin. Oncol. **27**: 5124. DOI: 10.1200/JCO.2009.22.0467.
6. Rödel, C., Liersch, T., Becker, H., et al. (2012). Preoperative chemoradiotherapy and postoperative chemotherapy with fluorouracil and oxaliplatin versus fluorouracil alone in locally advanced rectal cancer: initial results of the German CAO/ARO/AIO-04 randomised phase 3 trial. Lancet Oncol. **13**: 679–687. DOI: 10.1016/S1470-2045(12)70187-0.
7. Park, I.J., You, Y.N., Agarwal, A., et al. (2012). Neoadjuvant treatment response as an early response indicator for patients with rectal cancer. J. Clin. Oncol. **30**: 1770–1776. DOI: 10.1200/JCO.2011.39.7901.
8. Smith, J.J., Strombom, P., Chow, O.S., et al. (2019). Assessment of a watch-and-wait strategy for rectal cancer in patients with a complete response after neoadjuvant therapy. JAMA Oncol. **5**: e185896. DOI: 10.1001/jamaoncol.2018.5896.
9. Glynne-Jones, R., Wyrwicz, L., Tiret, E., et al. (2017). Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol. **28**: iv22–iv40. DOI: 10.1093/annonc/mdx224.
10. Maas, M., Nelemans, P.J., Valentini, V., et al. (2015). Adjuvant chemotherapy in rectal cancer: defining subgroups who may benefit after neoadjuvant chemoradiation and resection: A pooled analysis of 3,313 patients. Int. J. Cancer **137**: 212–220. DOI: 10.1002/ijc.29355.
11. Dossa, F., Acuna, S.A., Rickles, A.S., et al. (2018). Association between adjuvant chemotherapy and overall survival in patients with rectal cancer and pathological complete response after neoadjuvant chemotherapy and resection. JAMA Oncol. **4**: 930–937. DOI: 10.1001/jamaoncol.2017.5597.
12. Singh, M.P., Rai, S., Pandey, A., et al. (2021). Molecular subtypes of colorectal cancer: An emerging therapeutic opportunity for personalized medicine. Genes Dis. **8**: 133–145. DOI: 10.1016/j.gendis.2019.10.013.
13. Shia, J., Schultz, N., Kuk, D., et al. (2017). Morphological characterization of colorectal cancers in The Cancer Genome Atlas reveals distinct morphology–molecular associations: Clinical and biological implications. Mod. Pathol. **30**: 599–609. DOI: 10.1038/modpathol.2016.198.
14. Kawakami, H., Zaanan, A., and Sinicrope, F.A. (2015). Microsatellite instability testing and its role in the management of colorectal cancer. Curr. Treat Options. Oncol. **16**: 1–15. DOI: 10.1007/s11864-015-0348-2.
15. Boland, C.R., and Goel, A. (2010). Microsatellite instability in colorectal cancer. Gastroenterology **138**: 2073–2087. DOI: 10.1053/j.gastro.2009.12.064.
16. Lord, A.C., D'Souza, N., Shaw, A., et al. (2022). MRI-diagnosed tumor deposits and EMVI status have superior prognostic accuracy to current clinical TNM staging in rectal cancer. Ann. Surg. **276**: 334–344. DOI: 10.1097/SLA.0000000000004499.
17. Snead, D.R.J., Tsang, Y., Meskiri, A., et al. (2016). Validation of digital pathology imaging for primary histopathological diagnosis. Histopathology **68**: 1063–1072. DOI: 10.1111/his.12879.
18. Bi, W.L., Hosny, A., Schabath, M.B., et al. (2019). Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J. Clin. **69**: 127–157. DOI: 10.3322/caac.21552.
19. Esteva, A., Kuprel, B., Novoa, R.A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature **542**: 115–118. DOI: 10.1038/nature21056.
20. Gulshan, V., Peng, L., Coram, M., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**: 2402–2410. DOI: 10.1001/jama.2016.17216.
21. Sun, R., Limkin, E.J., Vakalopoulou, M., et al. (2018). A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: An imaging biomarker, retrospective multicohort study. Lancet Oncol. **19**: 1180–1191. DOI: 10.1016/S1470-2045(18)30413-3.
22. Skrede, O.-J., De Raedt, S., Kleppe, A., et al. (2020). Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. Lancet **395**: 350–360. DOI: 10.1016/S0140-6736(19)32998-8.
23. Lambin, P., Rios-Velazquez, E., Leijenaar, R., et al. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. Eur. J. Cancer **48**: 441–446. DOI: 10.1016/j.ejca.2011.11.036.
24. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature **521**: 436–444. DOI: 10.1038/nature14539.
25. Zhu, B., Liu, J.Z., Cauley, S.F., et al. (2018). Image reconstruction by domain-transform manifold learning. Nature **555**: 487–492. DOI: 10.1038/nature25988.
26. Moeskops, P., Wolterink, J.M., Van Der Velden, B.H.M., et al. (2016). Deep learning for multi-task medical image segmentation in multiple modalities. Proc. Med. Image Comput. Comput. Interv. **9901**: 478–486. DOI: 10.1007/978-3-319-46723-8_55.
27. Biller, L.H., and Schrag, D. (2021). Diagnosis and treatment of metastatic colorectal cancer: a review. JAMA **325**: 669–685. DOI: 10.1001/jama.2021.0106.
28. Harrison, K., Pullen, H., Welsh, C., et al. (2022). Machine learning for auto-segmentation in radiotherapy planning. Clin. Oncol. **34**: 74–88. DOI: 10.1016/j.clon.2021.12.003.
29. Grau, V., Mewes, A.U.J., Alcaniz, M., et al. (2004). Improved watershed transform for medical image segmentation using prior information. IEEE Trans. Med. Imaging **23**:

447–458. DOI: 10.1109/TMI.2004.824224.
30. Seo, H., Badiei Khuzani, M., Vasudevan, V., et al. (2020). Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. Med. Phys. **47**: e148–e167. DOI: 10.1002/mp.13649.
31. Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. Proc. Med. Image Comput. Comput. Interv. **9351**: 234–241. DOI: 10.1007/978-3-319-24574-4_28.
32. Jian, J., Xiong, F., Xia, W., et al. (2018). Fully convolutional networks (FCNs)-based segmentation method for colorectal tumors on T2-weighted magnetic resonance images. Phys. Eng. Sci. Med. **41**: 393–401. DOI: 10.1007/s13246-018-0636-9.
33. Zhu, H., Zhang, X., Shi, Y., et al. (2021). Automatic segmentation of rectal tumor on diffusion - weighted images by deep learning with U - Net. J. Appl. Clin. Med. Phys. **22**: 324–331. DOI: 10.1002/acm2.13381.
34. Huang, Y., Liang, C., He, L., et al. (2016). Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J. Clin. Oncol. **34**: 2157–2164. DOI: 10.1200/JCO.2015.65.9128.
35. Nie, K., Shi, L., Chen, Q., et al. (2016). Rectal cancer: Assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. Clin. Cancer Res. **22**: 5256–5264. DOI: 10.1158/1078-0432.CCR-15-2997.
36. Van Griethuysen, J.J.M., Fedorov, A., Parmar, C., et al. (2017). Computational radiomics system to decode the radiographic phenotype. Cancer Res. **77**: e104–e107. DOI: 10.1158/0008-5472.CAN-17-0339.
37. Zwanenburg, A., Vallières, M., Abdalah, M.A., et al. (2020). The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology **295**: 328–338. DOI: 10.1148/radiol.2020191145.
38. Shin, J., Seo, N., Baek, S.-E., et al. (2022). MRI radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy. Radiology **303**: 351–358. DOI: 10.1148/radiol.211986.
39. Feng, L., Liu, Z., Li, C., et al. (2022). Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: A multicentre observational study. Lancet Digit. Health **4**: e8–e17. DOI: 10.1016/S2589-7500(21)00215-6.
40. Delli Pizzi, A., Chiarelli, A.M., Chiacchiaretta, P., et al. (2021). MRI-based clinical-radiomics model predicts tumor response before treatment in locally advanced rectal cancer. Sci. Rep. **11**: 5379. DOI: 10.1038/s41598-021-84816-3.
41. Shaish, H., Aukerman, A., Vanguri, R., et al. (2020). Radiomics of MRI for pretreatment prediction of pathologic complete response, tumor regression grade, and neoadjuvant rectal score in patients with locally advanced rectal cancer undergoing neoadjuvant chemoradiation: An international multicenter study. Eur. Radiol. **30**: 6263–6273. DOI: 10.1007/s00330-020-06968-6.
42. Liu, Z., Zhang, X.-Y., Shi, Y.-J., et al. (2017). Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Clin. Cancer Res. **23**: 7253–7262. DOI: 10.1158/1078-0432.CCR-17-1038.
43. Zhou, X., Yi, Y., Liu, Z., et al. (2019). Radiomics-based pretherapeutic prediction of non-response to neoadjuvant therapy in locally advanced rectal cancer. Ann. Surg. Oncol. **26**: 1676–1684. DOI: 10.1245/s10434-019-07300-3.
44. Zhou, X., Yi, Y., Liu, Z., et al. (2020). Radiomics-based preoperative prediction of lymph node status following neoadjuvant therapy in locally advanced rectal cancer. Front. Oncol. **10**: 604. DOI: 10.3389/fonc.2020.00604.
45. Shao, L., Liu, Z., Feng, L., et al. (2020). Multiparametric MRI and whole slide image-based pretreatment prediction of pathological response to neoadjuvant chemoradiotherapy in rectal cancer: A multicenter radiopathomic study. Ann. Surg. Oncol. **27**: 4296–4306. DOI: 10.1245/s10434-020-08659-4.
46. Cui, Y., Yang, W., Ren, J., et al. (2021). Prognostic value of multiparametric MRI-based radiomics model: Potential role for chemotherapeutic benefits in locally advanced rectal cancer. Radiother. Oncol. **154**: 161–169. DOI: 10.1016/j.radonc.2020.09.039.
47. Dai, W., Mo, S., Han, L., et al. (2020). Prognostic and predictive value of radiomics signatures in stage I - III colon cancer. Clin. Transl. Med. **10**: 288–293. DOI: 10.1002/ctm2.31.
48. Liu, Z., Meng, X., Zhang, H., et al. (2020). Predicting distant metastasis and chemotherapy benefit in locally advanced rectal cancer. Nat. Commun. **11**: 4308. DOI: 10.1038/s41467-020-18162-9.
49. Zhang, X., Wang, L., Zhu, H., et al. (2020). Predicting rectal cancer response to neoadjuvant chemoradiotherapy using deep learning of diffusion kurtosis MRI. Radiology **296**: 56–64. DOI: 10.1148/radiol.2020190936.
50. Kather, J.N., Krisam, J., Charoentong, P., et al. (2019). Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. PLoS Med. **16**: e1002730. DOI: 10.1371/journal.pmed.1002730.
51. Liu, Z., Wang, S., Dong, D., et al. (2019). The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. Theranostics **9**: 1303–1322. DOI: 10.7150/thno.30309.
52. Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. Proc. Adv. Neural Inf. Process. Syst. **30**: 5998–6008.
53. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv.cs.AI. DOI: arxiv-2010.11929.
54. Shamshad, F., Khan, S., Zamir, S.W., et al. (2023). Transformers in medical imaging: A survey. Med. Image Anal. **88**: 102802. DOI: 10.1016/j.media.2023.102802.
55. Jiang, X., Zhao, H., Saldanha, O.L., et al. (2023). An MRI deep learning model predicts outcome in rectal cancer. Radiology **307**: e222223. DOI: 10.1148/radiol.222223.
56. Kaissis, G.A., Makowski, M.R., Rückert, D., et al. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. Nat. Mach. Intell. **2**: 305–311. DOI: 10.1038/s42256-020-0186-1.
57. Candemir, S., Nguyen, X. V, Folio, L.R., et al. (2021). Training strategies for radiology deep learning models in data-limited scenarios. Radiol. Artif. Intell. **3**: e210014. DOI: 10.1148/ryai.2021210014.
58. Russakovsky, O., Deng, J., Su, H., et al. (2015). Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. **115**: 211–252. DOI: 10.1007/s11263-015-0816-y.
59. Shin, H.-C., Roth, H.R., Gao, M., et al. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging **35**: 1285–1298. DOI: 10.1109/TMI.2016.2528162.
60. Guan, H., and Liu, M. (2021). Domain adaptation for medical image analysis: A survey. IEEE Trans. Biomed. Eng. **69**: 1173–1185. DOI: 10.1109/TBME.2021.3117407.
61. Li, X., Gao, H., Zhu, J., et al. (2021). 3D deep learning model for the pretreatment evaluation of treatment response in esophageal carcinoma: A prospective study (ChiCTR2000039279). Int. J. Radiat. Oncol. Biol. Phys. **111**: 926–935. DOI: 10.1016/j.ijrobp.2021.06.033.
62. Shao, L., Liu, Z., Liu, J., et al. (2022). Patient-level grading prediction of prostate cancer from mp-MRI via GMINet. Comput. Biol. Med. **150**: 106168. DOI: 10.1016/j.compbiomed.2022.106168.
63. Shao, L., Liu, Z., Yan, Y., et al. (2021). Patient-level prediction of multi-classification task at prostate MRI based on end-to-end framework learning from diagnostic logic of radiologists. IEEE Trans. Biomed. Eng. **68**: 3690–3700. DOI: 10.1109/TBME.2021.3082176.
64. Yamashita, R., Long, J., Longacre, T., et al. (2021). Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. Lancet Oncol. **22**: 132–141. DOI: 10.1016/S1470-2045(20)30535-0.
65. Vahadane, A., Peng, T., Sethi, A., et al. (2016). Structure-preserving color normalization and sparse stain separation for histological images. IEEE Trans. Med. Imaging **35**: 1962–1971. DOI: 10.1109/TMI.2016.2529665.
66. Macenko, M., Niethammer, M., Marron, J.S., et al. (2009). A method for normalizing histology slides for quantitative analysis. IEEE Int. Symp. Biomed. imaging **2009**: 1107–1110. DOI: 10.1109/ISBI.2009.5193250.
67. Campanella, G., Hanna, M.G., Geneslaw, L., et al. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nat. Med. **25**: 1301–1309. DOI: 10.1038/s41591-019-0508-1.
68. Lu, M.Y., Williamson, D.F.K., Chen, T.Y., et al. (2021). Data-efficient and weakly supervised computational pathology on whole-slide images. Nat. Biomed. Eng. **5**: 555–570. DOI: 10.1038/s41551-020-00682-w.
69. Lee, Y., Park, J.H., Oh, S., et al. (2022). Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning. Nat. Biomed. Eng. Advance online publication DOI: 10.1038/s41551-022-00923-0.
70. Dabass, M., Vashisth, S., and Vig, R. (2022). A convolution neural network with multi-level convolutional and attention learning for classification of cancer grades and tissue structures in colon histopathological images. Comput. Biol. Med. **147**: 105680. DOI: 10.1016/j.compbiomed.2022.105680.
71. Ding, K., Zhou, M., Wang, H., et al. (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. DOI: 10.1016/S2589-7500(22)00168-6.
72. Cui, Y., Yang, X., Shi, Z., et al. (2019). Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Eur. Radiol. **29**: 1211–1220. DOI: 10.1007/s00330-018-5683-9.
73. Nakanishi, R., Akiyoshi, T., Toda, S., et al. (2020). Radiomics approach outperforms diameter criteria for predicting pathological lateral lymph node metastasis after neoadjuvant (chemo) radiotherapy in advanced low rectal cancer. Ann. Surg. Oncol. **27**: 4273–4283. DOI: 10.1245/s10434-020-08974-w.
74. Lovinfosse, P., Polus, M., Van Daele, D., et al. (2018). FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. Eur. J. Nucl. Med. Mol. Imaging **45**: 365–375. DOI: 10.1007/s00259-017-3855-5.
75. Bang, J.-I., Ha, S., Kang, S.-B., et al. (2016). Prediction of neoadjuvant radiation chemotherapy response and survival using pretreatment [18 F] FDG PET/CT scans in locally advanced rectal cancer. Eur. J. Nucl. Med. Mol. Imaging **43**: 422–431. DOI: 10.1007/s00259-015-3180-9.
76. Maffione, A.M., Marzola, M.C., Capirci, C., et al. (2015). Value of (18)F-FDG PET for predicting response to neoadjuvant therapy in rectal cancer: Systematic review and meta-analysis. Am. J. Roentgenol. **204**: 1261–1268. DOI: 10.2214/AJR.14.13210.

77. Giannini, V., Mazzetti, S., Bertotto, I., et al. (2019). Predicting locally advanced rectal cancer response to neoadjuvant therapy with 18 F-FDG PET and MRI radiomics features. Eur. J. Nucl. Med. Mol. Imaging **46**: 878–888. DOI: 10.1007/s00259-018-4250-6.
78. Schurink, N.W., van Kranen, S.R., Berbee, M., et al. (2021). Studying local tumour heterogeneity on MRI and FDG-PET/CT to predict response to neoadjuvant chemoradiotherapy in rectal cancer. Eur. Radiol. **31**: 7031–7038. DOI: 10.1007/s00330-021-07724-0.
79. Patel, U.B., Taylor, F., Blomqvist, L., et al. (2011). Magnetic resonance imaging–detected tumor response for locally advanced rectal cancer predicts survival outcomes: MERCURY experience. J. Clin. Oncol. **29**: 3753–3760. DOI: 10.1200/JCO.2011.34.9068.
80. Patel, U.B., Brown, G., Rutten, H., et al. (2012). Comparison of magnetic resonance imaging and histopathological response to chemoradiotherapy in locally advanced rectal cancer. Ann. Surg. Oncol. **19**: 2842–2852. DOI: 10.1245/s10434-012-2309-3.
81. Horvat, N., Veeraraghavan, H., Khan, M., et al. (2018). MR imaging of rectal cancer: Radiomics analysis to assess treatment response after neoadjuvant therapy. Radiology **287**: 833–843. DOI: 10.1148/radiol.2018172300.
82. Bulens, P., Couwenberg, A., Intven, M., et al. (2020). Predicting the tumor response to chemoradiotherapy for rectal cancer: Model development and external validation using MRI radiomics. Radiother. Oncol. **142**: 246–252. DOI: 10.1016/j.radonc.2019.07.033.
83. Shayesteh, S., Nazari, M., Salahshour, A., et al. (2021). Treatment response prediction using MRI-based pre-, post-, and delta-radiomic features and machine learning algorithms in colorectal cancer. Med. Phys. **48**: 3691–3701. DOI: 10.1002/mp.14896.
84. Chen, H., Shi, L., Nguyen, K.N.B., et al. (2020). MRI radiomics for prediction of tumor response and downstaging in rectal cancer patients after preoperative chemoradiation. Adv. Radiat. Oncol. **5**: 1286–1295. DOI: 10.1016/j.adro.2020.04.016.
85. Wan, L., Peng, W., Zou, S., et al. (2021). MRI-based delta-radiomics are predictive of pathological complete response after neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Acad. Radiol. **28**: S95–S104. DOI: 10.1016/j.acra.2020.10.026.
86. Peng, J., Wang, W., Jin, H., et al. (2023). Develop and validate a radiomics space-time model to predict the pathological complete response in patients undergoing neoadjuvant treatment of rectal cancer: An artificial intelligence model study based on machine learning. BMC Cancer **23**: 365. DOI: 10.1186/s12885-023-10855-w.
87. Jin, C., Yu, H., Ke, J., et al. (2021). Predicting treatment response from longitudinal images using multi-task deep learning. Nat. Commun. **12**: 1–11. DOI: 10.1038/s41467-021-22188-y.
88. Shi, L., Zhang, Y., Nie, K.E., et al. (2019). Machine learning for prediction of chemoradiation therapy response in rectal cancer using pre-treatment and mid-radiation multi-parametric MRI. Magn. Reson. Imaging **61**: 33–40. DOI: 10.1016/j.mri.2019.05.003.
89. Armaghany, T., Wilson, J.D., Chu, Q., et al. (2012). Genetic alterations in colorectal cancer. Gastrointest Cancer Res. **5**: 19–27.
90. Ghaffari Laleh, N., Ligero, M., Perez-Lopez, R., et al. (2023). Facts and hopes on the use of artificial intelligence for predictive immunotherapy biomarkers in cancer. Clin. Cancer Res. **29**: 316–323. DOI: 10.1158/1078-0432.CCR-22-0390.
91. Cohen, R., Hain, E., Buhard, O., et al. (2019). Association of primary resistance to immune checkpoint inhibitors in metastatic colorectal cancer with misdiagnosis of microsatellite instability or mismatch repair deficiency status. JAMA Oncol. **5**: 551–555. DOI: 10.1001/jamaoncol.2018.4942.
92. Merok, M.A., Ahlquist, T., Røyrvik, E.C., et al. (2013). Microsatellite instability has a positive prognostic impact on stage II colorectal cancer after complete resection: Results from a large, consecutive Norwegian series. Ann. Oncol. **24**: 1274–1282. DOI: 10.1093/annonc/mds614.
93. Li, L.S., Morales, J.C., Veigl, M., et al. (2009). DNA mismatch repair (MMR)-dependent 5-fluorouracil cytotoxicity and the potential for new therapeutic targets. Br. J. Pharmacol. **158**: 679–692. DOI: 10.1111/j.1476-5381.2009.00423.x.
94. Mandal, R., Samstein, R.M., Lee, K.-W., et al. (2019). Genetic diversity of tumors with mismatch repair deficiency influences anti-PD-1 immunotherapy response. Science **364**: 485–491. DOI: 10.1126/science.aau0447.
95. Pei, Q., Yi, X., Chen, C., et al. (2022). Pre-treatment CT-based radiomics nomogram for predicting microsatellite instability status in colorectal cancer. Eur. Radiol. **32**: 714–724. DOI: 10.1007/s00330-021-08167-3.
96. Ying, M., Pan, J., Lu, G., et al. (2022). Development and validation of a radiomics-based nomogram for the preoperative prediction of microsatellite instability in colorectal cancer. BMC Cancer **22**: 1–13. DOI: 10.1186/s12885-022-09584-3.
97. Zhang, W., Huang, Z., Zhao, J., et al. (2021). Development and validation of magnetic resonance imaging-based radiomics models for preoperative prediction of microsatellite instability in rectal cancer. Ann. Transl. Med. **9**: 134. DOI: 10.21037/atm-20-7673.
98. Zhang, W., Yin, H., Huang, Z., et al. (2021). Development and validation of MR-based deep learning models for prediction of microsatellite instability in rectal cancer. Cancer Med. **10**(12): 4164–4173. DOI: 10.1002/cam4.3957.
99. Kather, J.N., Pearson, A.T., Halama, N., et al. (2019). Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat. Med. **25**: 1054–1056. DOI: 10.1038/s41591-019-0462-y.
100. Sorich, M.J., Wiese, M.D., Rowland, A., et al. (2015). Extended RAS mutations and anti-EGFR monoclonal antibody survival benefit in metastatic colorectal cancer: A meta-analysis of randomized, controlled trials. Ann. Oncol. **26**: 13–21. DOI: 10.1093/annonc/mdu378.
101. Xue, T., Peng, H., Chen, Q., et al. (2022). Preoperative prediction of KRAS mutation status in colorectal cancer using a CT-based radiomics nomogram. Br. J. Radiol. **95**: 20211014. DOI: 10.1259/bjr.20211014.
102. Cui, Y., Liu, H., Ren, J., et al. (2020). Development and validation of a MRI-based radiomics signature for prediction of KRAS mutation in rectal cancer. Eur. Radiol. **30**: 1948–1958. DOI: 10.1007/s00330-019-06572-3.
103. Zhang, Z., Shen, L., Wang, Y., et al. (2021). MRI Radiomics signature as a potential biomarker for predicting KRAS status in locally advanced rectal cancer patients. Front. Oncol. **11**: 614052. DOI: 10.3389/fonc.2021.614052.
104. Ma, Y., Wang, J., Song, K., et al. (2021). Spatial-frequency dual-branch attention model for determining KRAS mutation status in colorectal cancer with T2-weighted MRI. Comput. Methods Programs Biomed. **209**: 106311. DOI: 10.1016/j.cmpb.2021.106311.
105. Liu, H., Yin, H., Li, J., et al. (2022). A deep learning model based on MRI and clinical factors facilitates noninvasive evaluation of KRAS mutation in rectal cancer. J. Magn. Reson. Imaging **56**: 1659–1668. DOI: 10.1002/jmri.28237.
106. Bilal, M., Raza, S.E.A., Azam, A., et al. (2021). Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: A retrospective study. Lancet Digit. Health **3**: e763–e772. DOI: 10.1016/S2589-7500(21)00180-1.
107. Tsai, P.-C., Lee, T.-H., Kuo, K.-C., et al. (2023). Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. Nat. Commun. **14**: 2102. DOI: 10.1038/s41467-023-37179-4.
108. Sun, C., Li, B., Wei, G., et al. (2022). Deep learning with whole slide images can improve the prognostic risk stratification with stage III colorectal cancer. Comput. Methods Programs Biomed. **221**: 106914. DOI: 10.1016/j.cmpb.2022.106914.
109. Liu, X., Zhang, D., Liu, Z., et al. (2021). Deep learning radiomics-based prediction of distant metastasis in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy: A multicentre study. EBioMedicine **69**: 103442. DOI: 10.1016/j.ebiom.2021.103442.
110. Lu, L., Dercle, L., Zhao, B., et al. (2021). Deep learning for the prediction of early on-treatment response in metastatic colorectal cancer from serial medical imaging. Nat. Commun. **12**: 6654. DOI: 10.1038/s41467-021-26990-6.
111. Kleppe, A., Skrede, O.-J., De Raedt, S., et al. (2022). A clinical decision support system optimising adjuvant chemotherapy for colorectal cancers by integrating deep learning and pathological staging markers: A development and validation study. Lancet Oncol. **23**: 1221–1232. DOI: 10.1016/S1470-2045(22)00391-6.
112. Lu, Y., Yu, Q., Gao, Y., et al. (2018). Identification of metastatic lymph nodes in MR imaging with faster region-based convolutional neural networks. Cancer Res. **78**: 5135–5143. DOI: 10.1158/0008-5472.CAN-18-0494.
113. Nakanishi, R., Oki, E., Hasuda, H., et al. (2021). ASO author reflection: Radiomics-based prediction for the responder to first-line oxaliplatin-based chemotherapy in patients with colorectal liver metastasis. Ann. Surg. Oncol. **28**: 2986–2987. DOI: 10.1245/s10434-020-09584-2.
114. Giannini, V., Pusceddu, L., Defeudis, A., et al. (2022). Delta-radiomics predicts response to first-line oxaliplatin-based chemotherapy in colorectal cancer patients with liver metastases. Cancers **14**: 241. DOI: 10.3390/cancers14010241.
115. Dohan, A., Gallix, B., Guiu, B., et al. (2020). Early evaluation using a radiomic signature of unresectable hepatic metastases to predict outcome in patients with colorectal cancer treated with FOLFIRI and bevacizumab. Gut **69**: 531–539. DOI: 10.1136/gutjnl-2018-316407.
116. Zhou, B., Khosla, A., Lapedriza, A., et al. (2016). Learning deep features for discriminative localization. Proc IEEE Conf Comput Vis Pattern Recognit pp: 2921–2929. DOI:10.1109/CVPR.2016.319.
117. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. **1**: 206–215. DOI: 10.1038/s42256-019-0048-x.
118. Napel, S., Mu, W., Jardim - Perassi, B. V, et al. (2018). Quantitative imaging of cancer in the postgenomic era: Radio (geno) mics, deep learning, and habitats. Cancer **124**: 4633–4649. DOI: 10.1002/cncr.31630.
119. Dextraze, K., Saha, A., Kim, D., et al. (2017). Spatial habitats from multiparametric MR imaging are associated with signaling pathway activities and survival in glioblastoma. Oncotarget **8**: 112992. DOI: 10.18632/oncotarget.22947.
120. Jardim-Perassi, B. V, Huang, S., Dominguez-Viqueira, W., et al. (2019). Multiparametric MRI and coregistered histology identify tumor habitats in breast cancer mouse models. Cancer Res. **79**: 3952–3964. DOI: 10.1158/0008-5472.CAN-19-0213.
121. Yao, J., Zhu, X., Jonnagaddala, J., et al. (2020). Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. Med. Image Anal. **65**: 101789. DOI: 10.1016/j.media.2020.101789.
122. Chen, R.J., Lu, M.Y., Shaban, M., et al. (2021). Whole slide images are 2d point clouds: Context-aware survival prediction using patch-based graph convolutional

networks. Proc Med Image Comput Comput Assist Interv **12908**: 339–349. DOI: 10.1007/978-3-030-87237-3_33.
123. He, B., Dong, D., She, Y., et al. (2020). Predicting response to immunotherapy in advanced non-small-cell lung cancer using tumor mutational burden radiomic biomarker. J. Immunother. Cancer **8**: e000550. DOI: 10.1136/jitc-2020-000550.
124. Liu, X., Liu, Z., Yan, Y., et al. (2023). Development of prognostic biomarkers by TMB-guided WSI analysis: A two-step approach. IEEE J. Biomed. Health Inform. **27**: 1780–1789. DOI: 10.1109/JBHI.2023.3249354.
125. Mongan, J., Moy, L., and Kahn Jr, C.E. (2020). Checklist for artificial intelligence in medical imaging (CLAIM): A guide for authors and reviewers. Radiol. Artif. Intell. **2**: e200029. DOI: 10.1148/ryai.2020200029.
126. Collins, G.S., Reitsma, J.B., Altman, D.G., et al. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Ann. Intern. Med. **162**: 55–63. DOI: 10.7326/M14-0697.
127. Ganin, Y., and Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. Proc Int Conf Mach Learn **37**: 1180–1189. DOI: 10.5555/3045118.3045244.
128. Liu, Q., Chen, C., Qin, J., et al. (2021). Feddg: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space. Proc IEEE Int Conf Comput Vis pp: 1013–1023. DOI: 10.1109/CVPR46437.2021.00107.
129. Kaya, M., and Bilge, H.Ş. (2019). Deep metric learning: A survey. Symmetry **11**: 1066. DOI: 10.3390/sym11091066.
130. Chen, R.J., Lu, M.Y., Weng, W.-H., et al. (2021). Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. Proc IEEE Int Conf Comput Vis pp: 3995-4005. DOI: 10.1109/ICCV48922.2021.00398.
131. Bhattacharya, I., Seetharaman, A., Kunder, C., et al. (2022). Selective identification and localization of indolent and aggressive prostate cancers via CorrSigNIA: An MRI-pathology correlation and deep learning framework. Med. Image Anal. **75**: 102288. DOI: 10.1016/j.media.2021.102288.
132. Yang, Y., and Soatto, S. (2020). FDA: Fourier domain adaptation for semantic segmentation. Proc IEEE Conf Comput Vis Pattern Recognit pp: 4084-4094. DOI: 10.1109/CVPR42600.2020.00414.
133. Tzeng, E., Hoffman, J., Saenko, K., et al. (2017). Adversarial discriminative domain adaptation. Proc IEEE Conf Comput Vis Pattern Recognit pp: 2962-2971. DOI: 10.1109/CVPR.2017.316.
134. Kumar, A., Sattigeri, P., Wadhawan, K., et al. (2018). Co-regularized alignment for unsupervised domain adaptation. Proc Adv neural Inf Process Syst **31**: 9367–9378. DOI: 10.5555/3327546.3327607.
135. Yu, G., Sun, K., Xu, C., et al. (2021). Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images. Nat. Commun. **12**: 6311. DOI: 10.1038/s41467-021-26643-8.
136. Abbet, C., Studer, L., Fischer, A., et al. (2022). Self-rule to multi-adapt: Generalized multi-source feature learning using unsupervised domain adaptation for colorectal cancer tissue detection. Med. Image Anal. **79**: 102473. DOI: 10.1016/j.media.2022.102473.
137. Srinidhi, C.L., Kim, S.W., Chen, F.-D., et al. (2022). Self-supervised driven consistency training for annotation efficient histopathology image analysis. Med. Image Anal. **75**: 102256. DOI: 10.1016/j.media.2021.102256.
138. Azizi, S., Culp, L., Freyberg, J., et al. (2023). Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging. Nat. Biomed. Eng. **7**: 756–779. DOI: 10.1038/s41551-023-01049-7.
139. Schirris, Y., Gavves, E., Nederlof, I., et al. (2022). DeepSMILE: Contrastive self-supervised pre-training benefits MSI and HRD classification directly from H&E whole-slide images in colorectal and breast cancer. Med. Image Anal. **79**: 102464. DOI: 10.1016/j.media.2022.102464.
140. Price, W.N., and Cohen, I.G. (2019). Privacy in the age of medical big data. Nat. Med. **25**: 37–43. DOI: 10.1038/s41591-018-0272-7.
141. McMahan, B., Moore, E., Ramage, D., et al. (2017). Communication-efficient learning of deep networks from decentralized data. Proc Artif Intell Stat **54**: 1273–1282. DOI: 10.48550/arXiv.1602.05629.
142. Dou, Q., So, T.Y., Jiang, M., et al. (2021). Federated deep learning for detecting COVID-19 lung abnormalities in CT: A privacy-preserving multinational validation study. NPJ Digit. Med. **4**: 60. DOI: 10.1038/s41746-021-00431-6.
143. Pati, S., Baid, U., Edwards, B., et al. (2022). Federated learning enables big data for rare cancer boundary detection. Nat. Commun. **13**: 7346. DOI: 10.1038/s41467-022-33407-5.
144. Ogier du Terrail, J., Leopold, A., Joly, C., et al. (2023). Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer. Nat. Med. **29**: 135–146. DOI: 10.1038/s41591-022-02155-w.
145. Lu, M.Y., Chen, R.J., Kong, D., et al. (2022). Federated learning for computational pathology on gigapixel whole slide images. Med. Image Anal. **76**: 102298. DOI: 10.1016/j.media.2021.102298.
146. Jiang, M., Wang, Z., and Dou, Q. (2022). Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images. Proc AAAI Conf Artif Intell **36**: 1087–1095. DOI: 10.1609/aaai.v36i1.19993.
147. Feng, C.-M., Yan, Y., Wang, S., et al. (2022). Specificity-preserving federated learning for MR image reconstruction. IEEE Trans Med Imaging **42**: 2010–2021. DOI: 10.1109/TMI.2022.3202106.
148. Yan, R., Qu, L., Wei, Q., et al. (2023). Label-efficient self-supervised federated learning for tackling data heterogeneity in medical imaging. IEEE Trans Med Imaging **42**: 1932–1943. DOI: 10.1109/TMI.2022.3233574.

FUNDING AND ACKNOWLEDGMENTS

This review was supported by the National Natural Science Foundation of China under Grants 62333022, the Beijing Natural Science Foundation under Grant JQ23034, and the Natural Science Basic Research Program of Shaanxi under Grant 2023-JC-YB682. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

X.L. and Z.L. was responsible for conceptualization and writing the original draft. S.Z., L.S., C.S., B.L. and W.W. wrote and edited the manuscript. Z.Y., Z.L. and J.T. supervised this work. All authors contributed to the article and approved the submitted version.

DECLARATION OF INTERESTS

Jie Tian is an Editorial Board member of The Innovation Medicine. He was blinded from reviewing or making final decisions on the manuscript. The article was subject to the journal's standard procedures, with peer review handled independently of this Editorial Board member and their research groups. The authors declare no competing interests.

DATA AND CODE AVAILABILITY

Not applicalbe.

LEAD CONTACT WEBSITE

Jie Tian: https://web.xidian.edu.cn/tian/

Zhenyu Liu: http://www.ia.cas.cn/sourcedb_ia_cas/cn/iaexpert/202006/t20200608_5603067.html

Zuobin Ying: https://fds.cityu.edu.mo/members/188