

SHORT COMMUNICATION

Predicting treatment response to neoadjuvant chemoradiotherapy in local advanced rectal cancer by biopsy digital pathology image features

Fang Zhang^{1,2} | Su Yao³ | Zhi Li³ | Changhong Liang² | Ke Zhao^{2,4} | Yanqi Huang^{2,5} | Ying Gao¹ | Jinrong Qu⁶ | Zhenhui Li⁷  | Zaiyi Liu² 

¹ School of Computer Science and Engineering, South China University of Technology, Guangzhou, Guangdong, China

² Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, China

³ Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, China

⁴ School of Medicine, South China University of Technology, Guangzhou, Guangdong, China

⁵ Southern Medical University, Guangzhou, Guangdong, China

⁶ Department of Radiology, the Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, Henan, China

⁷ Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, Yunnan, China

Correspondence

Zhenhui Li, Department of Radiology, the Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, 519 Kunzhou Road, Xi Shan District, Kunming, 650118, China

Email: lizhenhui621@qq.com

Zaiyi Liu, MD, Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, 106 Zhongshan Er Rd, Yue Xiu District, Guangzhou 510080, China

Email: zylu@163.com

Fang Zhang, Su Yao, and Zhi Li contributed equally to this study.

Funding information

National Key Research and Development Program of China, Grant/Award Number: 2017YFC1309100; National Science Fund for Distinguished Young Scholars, China, Grant/Award Number: 81925023; National Natural Science Foundation of China, Grant/Award Numbers: 81771912, 81671854, 81702322; National Science Foundation for Young Scientists of China, Grant/Award Number: 81701782; Key Research and Development Program of Guangdong Province, China, Grant/Award Number: 2018B030339001

Colorectal cancer has third-highest morbidity and second-highest mortality worldwide. Around one-third of the tumors are localized in the rectum, and about 70% of rectal cancer is locally advanced rectal cancer (LARC).^{1,2}

List of abbreviations: AUC, area under the curve; CEA, carcinoembryonic antigen; CI, confidence interval; H&E, hematoxylin and eosin; H-L, Hosmer-Lemeshow; LARC, locally advanced rectal cancer; LASSO, least absolute shrinkage and selection operator; MRI, magnetic resonance imaging; nCRT, neoadjuvant chemoradiotherapy; PR, poor response; ROC, receiver operating characteristic; ROI, region of interest; SVM, support vector machine; TRG, tumor regression grade; WSI, whole slide images

The standard treatment for LARC involves neoadjuvant chemoradiotherapy (nCRT) followed by surgery.³ By downstaging and downsizing the tumor, nCRT increases the chance of subsequent successful R0 resection and sphincter-preserving surgery, and decrease the chance of local relapse.⁴ However, nCRT may weaken the immune system and cause delayed surgery for patients who cannot benefit from it.^{5,6} Therefore, it is necessary to identify biomarkers for the treatment response to nCRT for LARC, and to pinpoint the patients who will not benefit from it to improve treatment strategy and reduce unnecessary pain and cost.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

To predict and monitor the treatment response to nCRT in LARC, several tumor-related biomarkers have been proposed, including pathological, radiological, clinical, and molecular ones. Certain radiological and molecular markers have shown promise in the response prediction, the reported sensitivity and specificity were limited.⁷ Besides, although some preoperative clinicopathological features like enlarged size and tumor stage have been proposed to predict response to nCRT, but their prediction performance was unstable.^{8,9} Therefore, robust biomarkers with high accuracy still need to be identified and validated.

Biopsy samples are indispensable for the diagnosis of gastrointestinal tumors in current clinical practice. Advances in computerized image processing technology have generated automated histopathological analysis based on the digital whole slide images (WSIs) of biopsy specimens. As a useful approach for tumor diagnosis and prognosis, it has been increasingly investigated in oncology in recent years, with works reported in counting mitoses,¹⁰ quantifying tumor-infiltrating immunocyte,¹¹ and predicting the grade of tumor differentiation.¹² Yu et al. selected areas of dense tumor cells in hematoxylin and eosin (H&E) stained WSIs and quantifies features to predict the non-small cell lung cancer prognosis.¹³ Another study predicted microsatellite instability from the tumor areas of the H&E histology slides in gastrointestinal cancer.¹⁴ These studies have suggested that WSIs and machine learning approaches can be used to identify and quantify image features beyond simple densities in traditional pathologic interpretation and to explore the potential correlation with the features and treatment response.

For our knowledge, there is no published study on digital-pathology-based biomarkers that uses biopsy H&E histology images to predict the treatment response to nCRT in LARC. Therefore, we aim to investigate whether the quantitative features of H&E stained histology slides can predict treatment response.

This study was retrospective and single-centered. We obtained the approval from the institutional review board of our hospital and observed the Helsinki Declaration and relevant guidelines throughout the work.

A total of 151 LARC patients with adenocarcinomas who received nCRT treatment between January 2013 and June 2018 were recruited by the criteria in Appendix S1. Their baseline clinicopathologic data, including age, gender, pretreatment clinical T and N stage, pretreatment carcinoembryonic antigen (CEA), tumor location, and size, were derived from medical records (Table 1). The tumor location was the distance from the lower edge of colonoscopy to anus, and the tumor size was measured by the length and thickness of tumor from computed tomography imaging. The patients were split randomly into primary and validation datasets according to the ratio of 80%:20%. No sig-

nificant difference existed in the clinicopathological data between the two datasets (Appendix S2). The complete workflow of data analysis is shown in Figure 1.

The nCRT regimen is provided in Appendix S3. The treatment response to nCRT was evaluated according to tumor regression grade (TRG).⁶ TRG consists of 0, 1, 2, and 3, whose details are described in Appendix S4. We divided the treatment response into two categories, with TRG 2/3 being poor response (PR), where patients did not benefit from nCRT, and TRG 0/1 being non-poor response (non-PR). There were no significant differences between non-PR and PR groups in all included clinical characteristics concerning age, gender, pretreatment clinical T stage, and CEA-level except clinical N stage ($P = .015$ and $P = .521$ in primary and validation datasets, respectively) and age ($P = .023$ in validation dataset) (Table 1).

All the H&E WSIs used in this study were obtained from Yunnan Cancer hospital. The sample collection process is described in Appendix S5. The regions of interest (ROIs) encompassing tumor regions at $\times 400$ magnification were roughly annotated by an experienced pathologist (with 10 years of clinical experience) by using Aperio ImageScope and were confirmed by another pathologist (with 25 years of clinical experience). To make computation feasible, the ROIs were cropped into minimum bounding rectangle images. Those images were downsampled to lower resolution of $\times 200$ magnification and tiled into 512×512 pixels without overlapping. From those tiles, the tiles in the regions of the tumor cells were selected by the first pathologist and were confirmed by second. In the end, a total of 667 tiles were selected (525 from the primary cohort and 142 from the validation one).

A total of 104 texture features were extracted from each of the selected tiles,^{15,16} and the most helpful predictive features were selected by using the least absolute shrinkage and selection operator (LASSO) method with 10-fold cross-validation from the primary cohort (Figure 2A,B). LASSO has sparse solutions to reduce the dimensionality of the data.¹⁷ At last, 17 potential predictors were distilled from the 104 texture features in the primary cohort (Appendix S6).

The pathology signature was built as follow steps. The first step was to build a support vector machine (SVM) model with a radial basis function kernel, namely tile-level classifier, based on selected features to determine whether the selected tiles were classified. The second step was to compute the mean of all selected tiles' probabilities that the tile-level classifier output in each patient, and then to regard the mean as each patient's characteristic for further statistical analyses. The third step was to calculate a pathology score via the logistic regression method to discriminate pathology-reported PR status based on the characteristic for each patient in this study (Appendix S7 and S8).

TABLE 1 Clinical characteristic in the primary and validation datasets

Characteristic	Primary dataset		P-value	Validation dataset		P-value
	Non-PR	PR		Non-PR	PR	
Age, mean \pm SD	56.0 \pm 11.4	55.4 \pm 10.9	.465	51.7 \pm 11.8	60.4 \pm 9.18	.012*
Gender, No. (%)			.401			.800
Male	38 (62.3%)	42 (71.2%)		8 (72.7%)	14 (70.0%)	
Female	23 (37.7%)	17 (28.8%)		3 (27.3%)	6 (30.0%)	
T staging, No. (%)			.698			.378
T0	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T1	0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T2	2 (3.3%)	1 (1.7%)		1 (9.1%)	0 (0%)	
T3	23 (37.7%)	26 (44.1%)		4 (36.4%)	7 (35.0%)	
T4	36 (59.0%)	32 (54.2%)		6 (54.5%)	13 (65.0%)	
N staging, No. (%)			.015*			.521
N0	6 (9.8%)	17 (28.8%)		2 (18.2%)	4 (20.0%)	
N1	46 (75.4%)	31 (52.5%)		8 (72.7%)	11 (55.0%)	
N2	9 (14.8%)	11 (18.6%)		1 (9.1%)	5 (25.0%)	
CEA level, No. (%)			.134			.724
Normal	30 (49.2%)	38 (64.4%)		5 (45.5%)	9 (45.0%)	
Abnormal	31 (50.8%)	21 (35.6%)		6 (54.5%)	11 (55.0%)	
Tumor location (cm)			.313			.298
< 5	36 (59.1%)	40 (67.8%)		8 (72.7%)	9 (45.0%)	
5-10	25 (40.9%)	18 (30.5%)		3 (27.3%)	10 (50.0%)	
\geq 10	0 (0%)	1 (1.7%)		0 (0.0%)	1 (5.0%)	
Length of tumor (cm), mean \pm SD	4.87 \pm 1.88	4.92 \pm 1.52	.483	4.56 \pm 1.19	4.89 \pm 1.70	.917
Thickness of tumor (cm), mean \pm SD	1.77 \pm 0.69	1.76 \pm 0.65	.945	1.27 \pm 0.38	1.76 \pm 0.53	.03*
Pathology score, median (interquartile range)	0.128 (0.089 to 0.321)	0.891 (0.626 to 0.929)	<.001*	0.207 (0.130 to 0.483)	0.814 (0.481 to 0.874)	<.001*

Note. P-value is derived from the univariable association analyses between each of the clinicopathological variables and treatment response. The clinical characters were the data from the initial diagnosis. The threshold value for CEA level was \leq 5 ng/mL and $>$ 5 ng/mL according to the universally normal range used.

Abbreviations: CEA, pre-treatment carcinoembryonic antigen; SD, standard deviation.

* $P < .05$.

The multivariable logistic regression analysis, including age, gender, pretreatment clinical T and N stage, CEA level as well as the pathology score, showed that the pathology score was the only independent predictor (Appendix S7).

To evaluate the discriminative power of the pathology signature, the receiver operating characteristics (ROC) curve was applied. The tile-level classifier for distinguishing PR from non-PR produced an area under the ROC curve (AUC) of 0.887 (95% confidence interval (CI), 0.858-0.916), and 0.797 (95% CI, 0.718-0.866) in the primary and validation datasets, respectively (Figure 2C). The pathology signature provided an AUC of 0.930 (95% CI, 0.883-0.966), and 0.877 (95% CI, 0.719-0.97) in the two datasets, respectively (Figure 2D). Meanwhile, we use accuracy, sensitivity, specificity, and F1 score to assess the signature performance (Table 2). The calibration

of the signature was evaluated by using the calibration curves and Hosmer-Lemeshow (H-L) test (Figure 2E). The clinical usefulness of the signature was determined by using the decision curve analysis, which could quantify the net benefits at different threshold probabilities in the primary and validation datasets (Figure 2F). The H-L test generated a non-significant statistic ($P = .332$ in primary dataset, $P = .213$ in validation dataset), suggesting no departure from a good fit.

The tools and methods of the statistical analysis are shown in Appendix S9. The reported statistical significance levels were all two-sided, with statistical significance set at 0.05.

In this study, we presented a digital-pathology-based signature to predict treatment response to nCRT in LARC. We showed that quantitative features extracted from H&E

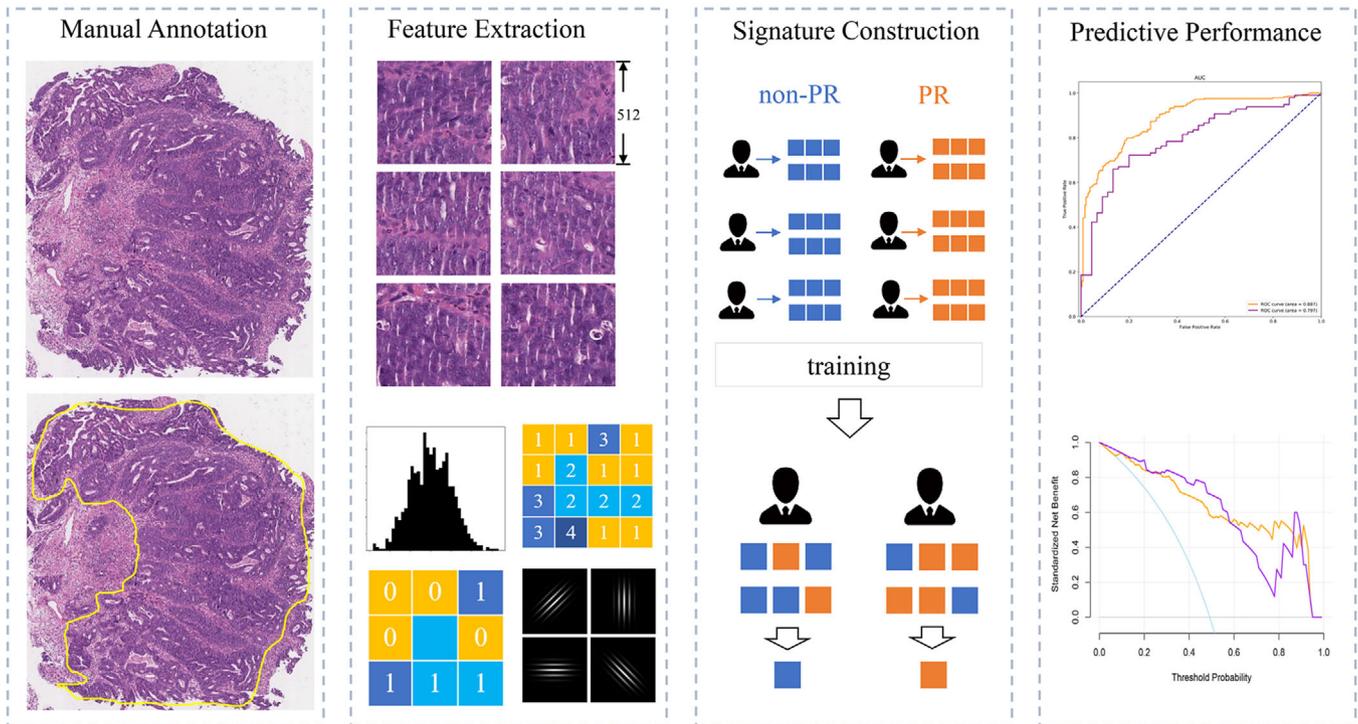


FIGURE 1 Pathology signature construction in hematoxylin and eosin stained whole slide images (WSIs). With manually annotated tumor areas, we cropped WSIs into tile images; quantitative features were extracted and reduced from the selected patches of tumor cell dense area. Next, we built a tile-level classifier via a support vector machine (SVM) model, and then the pathology signature was constructed with a logistic regression model. Finally, the predictive power of the signature was evaluated

histology slides could provide predictive information for treatment response.

The digital-pathology-based signature we proposed has advantages over the previous methods or complements them. Magnetic resonance imaging (MRI) is the major method to predict the treatment response to nCRT in LARC.^{18,19} Zhou et al. predicted non-response to neoadjuvant therapy in LARC by using pretreatment MRI images and achieved an AUC of 0.773.¹⁹ We yielded an AUC of 0.877 by using pathological images, demonstrating a better prediction performance than that achieved by Zhou et al. Besides, Nie et al. used pretreatment multiparametric MRI images for nCRT efficacy evaluation and obtained an AUC of 0.89,⁵ but the approach is limited partly because some patients cannot receive MRI examinations. In such cases, our method can serve as an alternative.

Our work provides a new and easy-to-apply method based on the biopsy-acquired H&E slides for the treatment response prediction in LARC, which could be well suited for routine clinical practice. The previous studies showed the quantitative image features could be mined through approaches of digital pathology and machine learning, but their clinical applications faced obstacles, like the heavy annotation workload.²⁰ However, the pathology signature in our study required a small amount of annotating work,

namely circling parts of tumor areas and selecting a few tiles. Besides, the tiles in our study were cropped from the $\times 200$ magnification WSIs that is different from the previous study using the $\times 400$ magnification WSIs.¹³ The tiles from the $\times 400$ magnification WSIs may provide more details but have a narrow field of view, while the tiles from the $\times 200$ magnification WSIs could reflect a wider range of tumor tissue patterns and make more pathological information comprehensible.²¹

A key part of our image processing approach is to select tiles of dense tumor cell areas. Because the normal rectum area is composed predominantly of regular epithelium and normal cell with typical morphological characteristics, the areas of dense tumor cell show exhibit pathological changes, such as different cell nucleus and gland morphology, and image feature extraction from the tiles of the tumor area is potential to be biologically informative.²² The strategy of using representative patch samples was adopted when selecting tiles to avoid the time-consuming way of utilizing all image patches of each WSIs.²³

Some limitations exist in our study. First, the size of patient samples was relatively small considering the large number of predictors. Second, the quality of H&E biopsy slides may not be consistent in some cases, where the slide quality may have been affected by where the material was

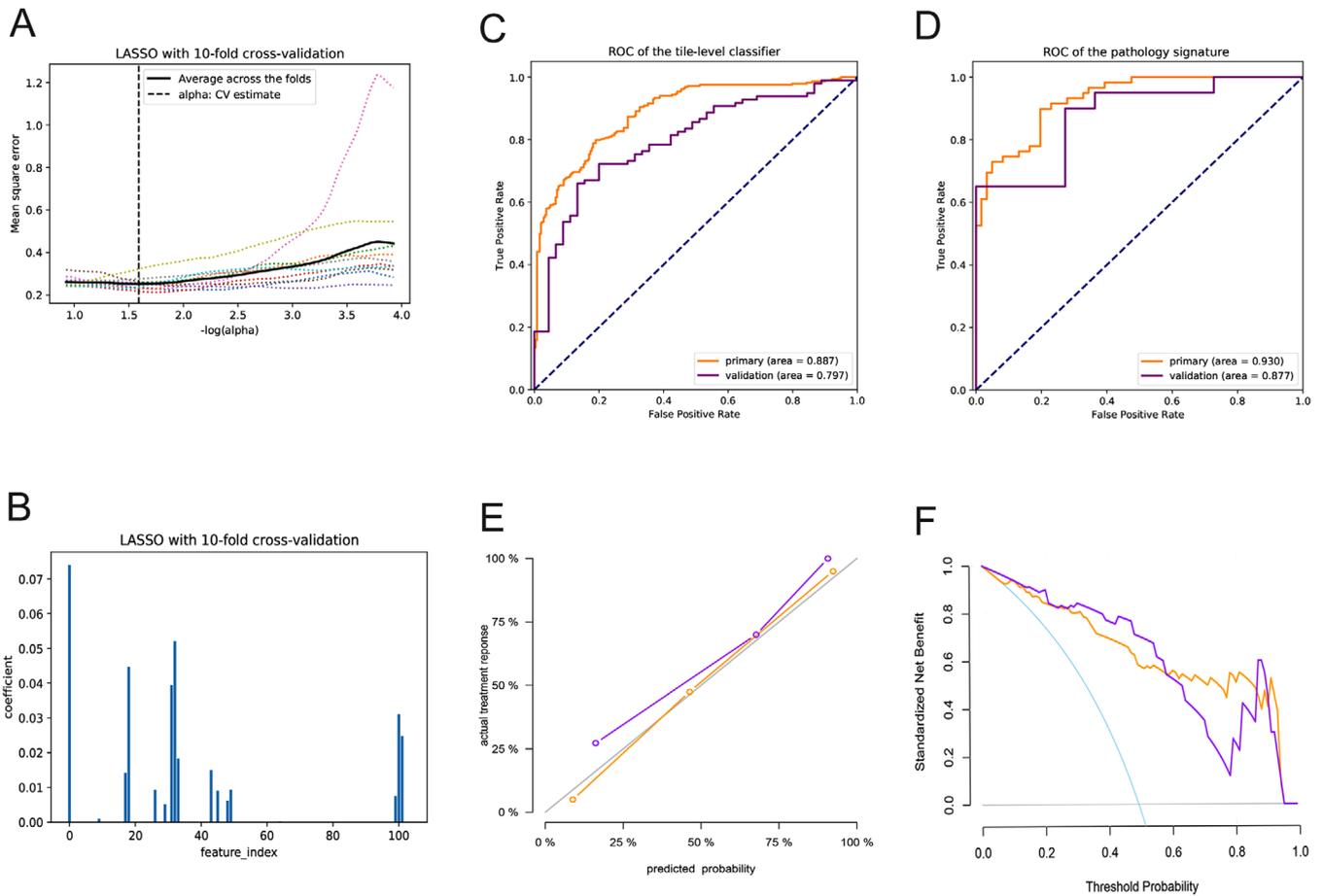


FIGURE 2 Texture feature selection and pathology signature's performance. A, The parameter alpha selection in the least absolute shrinkage and selection operator (LASSO) model used the 10-fold cross-validation. B, LASSO coefficient profiles of the 104 texture features using 10-fold cross-validation. C, the receiver operating characteristics (ROC) curve of the tile-level classifier in the primary and validation dataset. D, ROC curve of the pathology signature in the primary and validation datasets. E, calibration curves of the pathology signature in the primary and validation datasets. F, Decision curve analysis for the pathology signature in the primary and validation datasets

TABLE 2 Performance of pathology signature

	Accuracy(95%CI)	Sensitivity(95% CI)	Specificity(95% CI)	F1-score(95% CI)	AUC(95%CI)
TL-p	0.790(0.756-0.82)	0.760(0.714-0.811)	0.826(0.778-0.872)	0.796(0.760-0.833)	0.887(0.858-0.916)
TL-v	0.732(0.655-0.803)	0.753(0.663-0.833)	0.688(0.553-0.821)	0.793(0.720-0.853)	0.797(0.718-0.866)
P-p	0.792(0.708-0.858)	0.780(0.661-0.873)	0.803(0.692-0.897)	0.786(0.685-0.855)	0.930(0.883-0.966)
P-v	0.710(0.548-0.871)	0.700(0.476-0.889)	0.727(0.444-1.0)	0.757(0.564-0.878)	0.877(0.719-0.97)

Abbreviations: CI, confidence interval; P-p, the performance of primary dataset in pathology signature; P-v, the performance of validation dataset in pathology signature; TL-p, the performance of primary dataset in tile-level classifier; TL-v, the performance of validation dataset in tile-level classifier

taken from and which microscopical section was chosen. Hence measures are needed to control slide quality, and a large, independent, and prospective multicenter validation queue is needed to examine whether the proposed model can be extended and the clinical potential be transformed.

In conclusion, we developed a pathology signature as a new method to predict the treatment response to nCRT in LARC, which can help make personalized treatment plans and improve outcomes for patients.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (No. 2017YFC1309100), the National Science Fund for Distinguished Young Scholars, China (No. 81925023), the National Natural Science Foundation of China (No. 81771912, No. 81671854, and No. 81702322), the National Science Foundation for Young Scientists of China (No. 81701782), and the Key Research and Develop-

ment Program of Guangdong Province, China (No. 2018B030339001).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study follows the Declaration of Helsinki and good clinical practice guidelines, and was approved by Medical Ethics Committee of Guangdong Provincial People's Hospital.

CONFLICT OF INTEREST

The authors declare that they have no potential conflicts of interest.

ORCID

Zhenhui Li  <https://orcid.org/0000-0001-7221-2588>

Zaiyi Liu  <https://orcid.org/0000-0003-3296-9759>

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: gLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
- Benson AB, Venook AP, Al-Hawary MM, et al. NCCN guidelines insights: colon cancer, version 2.2018. *J Natl Compr Canc Netw*. 2018;16(4):359-369.
- Liu Z, Zhang XY, Shi YJ, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res*. 2017;23(23):7253-7262.
- Lovinfosse P, Polus M, Van Daele D, et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. *Eur J Nucl Med Mol Imaging*. 2018;45(3):365-375.
- Nie K, Shi L, Chen Q, et al. Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. *Clin Cancer Res*. 2016;22(21):5256-5264.
- Song C, Chung JH, Kang SB, et al. Impact of tumor regression grade as a major prognostic factor in locally advanced rectal cancer after neoadjuvant chemoradiotherapy: a proposal for a modified staging system. *Cancers (Basel)*. 2018;10(9):319.
- Ryan JE, Warrier SK, Lynch AC, Ramsay RG, Phillips WA, Heriot AG. Predicting pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a systematic review. *Colorectal Dis*. 2016;18(3):234-246.
- Huh JW, Kim HR, Kim YJ. Clinical prediction of pathological complete response after preoperative chemoradiotherapy for rectal cancer. *Dis Colon Rectum*. 2013;56(6):698-703.
- Kalady MF, de Campos-Lobato LF, Stocchi L, et al. Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer. *Ann Surg*. 2009;250(4):582-589.
- Balkenhol MCA, Tellez D, Vreuls W, et al. Deep learning assisted mitotic counting for breast cancer. *Lab Invest*. 2019;99(11):1596-1606.
- Turkki R, Linder N, Kovanen PE, Pellinen T, Lundin J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *J Pathol Inform*. 2016;7:38.
- Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301-1309.
- Yu KH, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun*. 2016;7:12474.
- Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*. 2019;25(7):1054-1056.
- Vallieres M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol*. 2015;60(14):5471-5496.
- Kather JN, Weis CA, Bianconi F, et al. Multi-class texture analysis in colorectal cancer histology. *Sci Rep*. 2016;6:27988.
- Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol*. 2016;34(18):2157-2164.
- Cui Y, Yang X, Shi Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol*. 2019;29(3):1211-1220.
- Zhou X, Yi Y, Liu Z, et al. Radiomics-based pretherapeutic prediction of non-response to neoadjuvant therapy in locally advanced rectal cancer. *Ann Surg Oncol*. 2019;26(6):1676-1684.
- Komura D, Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J*. 2018;16:34-42.
- Tokunaga H, Teramoto Y, Yoshizawa A, Bise R. Adaptive weighting multi-field-of-view CNN for semantic segmentation in pathology. Paper presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019.
- Bychkov Dmitrii, Linder Nina, Turkki Riku, Nordling Stig, Kovanen Panu E., Verrill Clare, Walliander Margarita, Lundin Mikael, Haglund Caj, Lundin Johan. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports*. 2018;8(1):1-11.
- Luo X, Zang X, Yang L, et al. Comprehensive computational pathological image analysis predicts lung cancer prognosis. *J Thorac Oncol*. 2017;12(3):501-509.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zhang F, Yao S, Li Z, et al. Predicting treatment response to neoadjuvant chemoradiotherapy in local advanced rectal cancer by biopsy digital pathology image features. *Clin Transl Med*. 2020;10:e110.

<https://doi.org/10.1002/ctm2.110>