

RESEARCH

Open Access



Association of the tumor microenvironment collagen score and immunoscore with colon cancer lymph node metastasis

Chenyang Long^{1,2,3,4}, Jiaxin Cheng^{1,2}, Mingyuan Feng^{1,2}, Botao Yan^{1,2}, Yiran Li^{1,2}, Wei Jiang^{1,2}, Dexin Chen^{1,2,3} and Jun Yan^{1,2,3*}

Abstract

Background In clinical practice, lymph node status has an important impact on colon cancer (CC) management and treatment. The role of the tumor microenvironment collagen score and immunoscore in colon cancer lymph node metastasis remains unknown.

Methods A total of 249 CC patients who underwent laparoscopic-assisted D3 lymphadenectomy from June 2016 to May 2019 were included. The patients' clinicopathological data were collected retrospectively. A total of 142 collagen features were extracted by multiphoton imaging and collagen quantification. A collagen score was constructed using a LASSO logistic regression model. Antibodies against CD3 and CD8 were used for immunostaining. The immunoscore was constructed based on the mean densities of CD3+ and CD8+T cells both in the tumor center and invasion margin on imaging.

Results The lymph node metastasis rate among colon cancer patients was 42.2% (105/249). The multivariate analysis indicated that lymphatic invasion (OR: 3.892, 95% CI: 1.784–8.491, $p=0.001$), vascular invasion (OR, 3.234, 95% CI: 1.544–6.776); $p=0.002$), mucus adenocarcinoma and signet-ring cell carcinoma (OR: 2.990, 95% CI: 1.413–6.328, $p=0.004$), the collagen score (OR: 6.304, 95% CI: 2.145–18.527, $p=0.001$) and the immunoscore [intermediate group (OR, 2.473; 95% CI, 1.192–5.130; $p=0.015$); low group (OR, 5.877; 95% CI, 2.423–14.257; $p<0.01$)] were independent risk factors for colon cancer lymph node metastasis. The newly developed model comprising these five independent predictors showed good discrimination with an AUROC of 0.809 (95% CI: 0.755–0.862). The new model performed significantly better than the traditional clinicopathological model [AUROC: 0.715 (95% CI: 0.649–0.780), $p<0.001$].

Conclusions The tumor microenvironment collagen score and immunoscore are associated with colon cancer lymph node metastasis.

Keywords Colon cancer, Collagen score, Immunoscore, Lymph node metastasis

*Correspondence:

Jun Yan

yanjunfudan@163.com

Full list of author information is available at the end of the article



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

Introduction

Colon cancer (CC) is one of the most common malignancies of the gastrointestinal tract. [1] The presence of lymph node metastasis (LNM) is a feature of poor CC patient prognosis. Patients with LNM have a lower 5-year survival rate than those without LNM. [2] Researchers have recently focused on constructing signatures using the characteristics of the tumor itself to analyze LNM. [3–5] Bae JH et al. reported that vascular invasion was significantly associated with LNM. [6] Mou A et al. found that tumor size was related to LNM. [7] However, few studies have focused on the impact of the tumor micro-environment (TME) on LNM.

Collagen fibers are the main component of the TME extracellular matrix (ECM), and changes in these fibers are associated with cancer invasion and prognosis. [8] Chen D et al. found that collagen signatures in the TME were associated with LNM in early gastric cancer. [9] Dong X et al. reported that the collagen score could predict the prognosis of rectal cancer patients after neoadjuvant chemoradiotherapy. [10] Multiphoton imaging technology is an effective tool for analyzing the quantity, quality, and conformation of collagen changes. [11, 12] However, whether collagen changes in the microenvironment are related to LNM in CC remains unknown.

Immunocytes in the TME are also generally believed to play a significant role in the development of tumors. [13] There is a consensus that the immune scoring system (immunoscore) based on the location and density of CD3+ and CD8+ T cells can predict the prognosis of early and advanced colorectal cancer. [14, 15] Nevertheless, the association between the immunoscore and LNM in CC has not been investigated in detail.

We hypothesized that the TME collagen and immunoscores are associated with CC LNM. The primary objective of this study was to examine the evidence that the collagen and immunoscores are associated with LNM in CC. This is the first study to analyze the relationship between LNM CC and changes in TME structure and immune infiltration.

Methods

Patients and specimens

A total of 249 CC patients who underwent laparoscopic-assisted D3 lymphadenectomy in the Department of General Surgery at Nanfang Hospital of Southern Medical University from June 2016 to May 2019 were included in this study.

The inclusion criteria were as follows: age ≥ 18 years; ASA score of 1–3; no neoadjuvant chemoradiotherapy performed before the operation; D3 lymphadenectomy (R0 resection); and postoperative pathological stage I–III. We excluded patients who underwent preoperative

neoadjuvant therapy or palliative resection and those who had other tumors, stage IV disease, incomplete clinicopathological data or missing formalin-fixed paraffin-embedded (FFPE) samples.

Basic clinicopathologic data were collected for each patient from medical records, including demographics such as age (≥ 60 years, < 60 years), sex (male, female), and BMI (≥ 24 kg/m², < 24 kg/m²) according to the Chinese BMI classification guideline [16] s. Additional clinical data included ASA scores (1, 2, 3) and preoperative serum CEA levels (≥ 5 ng/ml, < 5 ng/ml). Pathological characteristics of surgical specimens were also recorded, including tumor location (left-sided colon, right-sided colon), size (maximum diameter ≥ 5 cm, < 5 cm) [17], histologic type (adenocarcinoma, mucinous adenocarcinoma with mucinous components, and signet-ring cell carcinoma), differentiation (well, moderate, poor), and invasion status—lymphatic, vascular, and perineural (present or absent for each). Tumor T category (T1–2, T3–4) was also documented.

Tumor budding was assessed according to the method recommended by the International Tumor Budding Consensus Conference (ITBCC) [18] and classified as mild (0–4 buds), moderate (5–9 buds), or marked (≥ 10 buds). Patients were categorized into two groups—LNM+ and LNM—based on the presence or absence of postoperative LNM in the pathology report. The FFPE specimens of all patients were used.

Collagen score construction

First, samples from all patients were sectioned consecutively. One slide of continuously sectioned samples from each patient was randomly selected for hematoxylin–eosin (H&E) staining. The histological evaluation was performed by a pathologist. Five random corresponding unstained regions of interest (ROIs, 500 \times 500 μ m) in invasive margin were selected for multiphoton imaging and magnified 20 times. A multiphoton imaging system based on two-photon fluorescence excitation/second harmonic generation (SHG/TPEF) was conducted as previously described. [19] The excitation wavelength used in this study was 810 nm. MATLAB 2018b (MathWorks) was used to extract collagen features. [20] A total of 142 features comprising 8 morphologic features, 6 intensity features, 80 Gy-level co-occurrence matrix-based features and 48 Gabor wavelet transform features, were extracted. Least-absolute shrinkage and selection operator (LASSO) logistic regression was used to select the features most associated with LNM. [21] The method used an L1 penalty to shrink some regression coefficients to exactly zero. The penalty parameter λ , known as the tuning constant, controls the intensity of the penalty. The λ value can be increased to select fewer predictors

to enter the model. In our study, tenfold cross-validation was used to determine the optimal value of λ , and the final λ value was selected based on the minimum standard. LASSO logistic regression was conducted using R software (version 4.0.3) with the “glmnet” package. The collagen score of each patient was calculated from the selected collagen characteristics and the LASSO regression coefficients to form a standard linear equation.

Immunoscore construction

FFPE tumor sections were deparaffinized and stained with antibodies against CD3 and CD8 (Maixin Biotech. Co., Ltd., Fuzhou, China). Immunohistochemical staining was performed according to the manufacturer's instructions. Then, the slices were stained with DAB and counterstained with hematoxylin. Following immunohistochemistry, all the stained slices were digitized by Aperio ImageScope (Leica Biosystems, CA, USA) at 20 \times magnification. The immunoscore was based the mean densities of CD3+ and CD8+T cells both in the tumor center (TC) and invasion margin (IM) on imaging and included the following steps. First, five ROIs in the TC and IM were manually annotated. Second, QuPath software (version 0.2.3) was used to calculate the number and density of positively stained cells. [22] The maximum Youden index of mean density was used as the cutoff value to distinguish ‘high’ and ‘low’ immune responses. A high immune response score was set as 1, and a low immune response score was set as 0. The CD3_{TC}, CD3_{IM}, CD8_{TC}, and CD8_{IM} scores were added and converted into an immunoscore (I0–I4). Patients were divided into three groups based on their immunoscores. Immunoscores I0–I1 were classified as “low”, I2 was classified as “intermediate”, and I3–I4 were classified as “high”.

Statistical analysis and nomogram development

Continuous variables were compared by a 2-tailed *t* test (or the Mann–Whitney U test when appropriate). Categorical variables were compared by a χ^2 test. Statistical analysis was conducted using R version 4.0.2 (R Foundation for Statistical Computing, Vienna, Austria) or SPSS 26.0 (SPSS, Chicago, IL, USA). The independent risk factors for LNM were explored by binary logistic regression analysis, which was applied to calculate the odds ratio (OR) and its corresponding 95% confidence interval (CI). [23] Differences with a 2-sided $p < 0.05$ were considered statistically significant. The tolerance and variance inflation factor (VIF) were used to evaluate the multicollinearity of the model. [24]

A nomogram was constructed based on the independent risk factors. This nomogram provided a visual presentation of the factors associated with LNM and the corresponding points, thus visualizing the probability

of LNM in each CC patient. The area under the receiver operating characteristic curve (AUROC) was measured to assess the accuracy of the nomogram. The calibration curve was plotted to assess the goodness of fit of the nomogram, accompanied by the Hosmer–Lemeshow test. Decision curve analysis (DCA) was performed to assess the net benefits of the model at different threshold probabilities to estimate its clinical value. [25] The maximum Youden index was selected as the cutoff value. The sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of the new model and traditional model were calculated.

Result

Participants

The clinicopathological characteristics of the 249 patients are summarized in Table 1. There were 105 patients in the LNM+ group, the LNM rate was 42.2% (105/249), the median age [interquartile range (IQR)] was 61 (51–68.5) years, and 65 (61.9%) of the patients were male. There were 144 patients in the LNM- group, accounting for 57.8% (144/249) of the patients. The median age (IQR) in the LNM- group was 58 (46.3–65) years, and 86 (59.7%) of the patients were male.

Lymph node data

The number of lymph nodes (LNs) harvested per patient was as follows: in the LNM- group, the median (IQR) was 37.00 (27.25, 50.75) LNs, while in the LNM+ group, the median (IQR) was 35.00 (27.00, 48.00) LNs. For the LNM+ group, the positive lymph node ratio [median (IQR)] was calculated as 0.05 (0.03, 0.14). Statistical analysis showed no significant correlation between the positive lymph node ratio and the total number of harvested lymph nodes ($p > 0.05$).

Collagen score

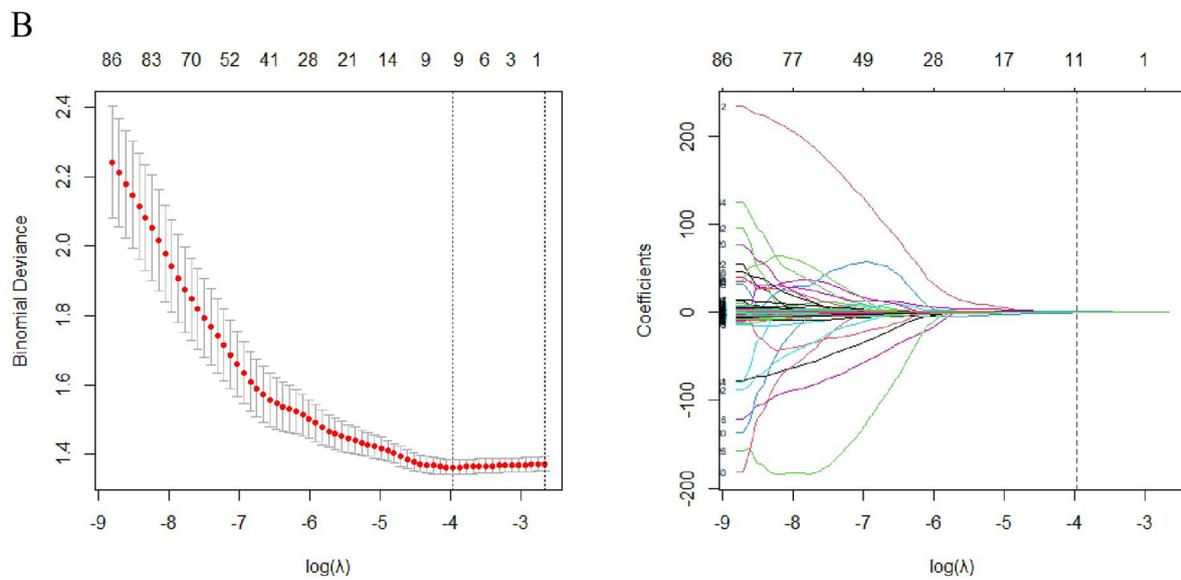
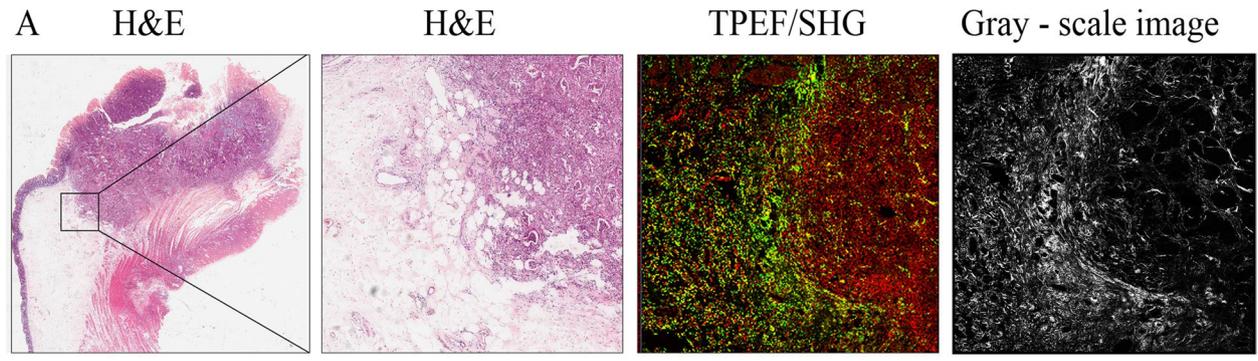
A flowchart demonstrating the construction of the collagen score is shown in Fig. 1. The coefficient profiles of the collagen features in CC were extracted from the LASSO logistic regression. Furthermore, the collagen score of each patient was calculated as follows: Collagen score = $-2.97289133 \cdot \text{Morphology_Fiber width} \cdot 0.13822854 + \text{Morphology_CrosslinkSpace} \cdot 0.02059122 - \text{Intensity_Kurtosis} \cdot 0.0148381 + \text{GLCM_90}^\circ_2\text{displacement_correlation} \cdot 1.79545484 + \text{GLCM_90}^\circ_5\text{displacement_homogeneity} \cdot 0.93172095 + \text{GLCM_135}^\circ_1\text{displacement_correlation} \cdot 0.24059861 + \text{GLCM_135}^\circ_2\text{displacement_homogeneity} \cdot 0.41317486 + \text{GLCM_135}^\circ_2\text{displacement_homogeneity} \cdot 0.01091541 - \text{Gabor_3scale_4orientation_mean} \cdot 0.1395055 + \text{Gabor_4scale_6orientation_variance} \cdot 0.30749449$.

Table 1 The clinicopathological characteristics of the 249 patients

Clinicopathological feature	LNM+	LNM-	p value
Age, No. (%), year			0.063
< 60	48(45.7)	83(57.6)	
≥ 60	57(54.3)	61(42.4)	
Sex, No. (%)			0.728
Male	65(61.9)	86(59.7)	
Female	40(38.1)	58(40.3)	
BMI, No. (%), kg/m ²			0.524
≥ 24	24(22.9)	38(26.4)	
< 24	81(77.1)	106(73.6)	
ASA, No. (%)			0.316
1	17(16.2)	34(23.6)	
2	79(75.2)	101(70.1)	
3	9(8.6)	9(6.3)	
Preoperative serum CEA, No. (%), ng/ml			0.260
< 5	84(80)	123(85.4)	
≥ 5	21(20)	21(14.6)	
Size, No. (%), cm			0.326
< 5	43(41)	68(47.2)	
≥ 5	62(59)	76(52.8)	
Location, No. (%),			0.087
Left-sided colon	71(67.6)	82(56.9)	
Right-sided colon	34(32.4)	62(43.1)	
T category, No. (%)			0.012
T1-2	4(3.8)	19(13.2)	
T3-4	101(96.2)	125(86.8)	
Differentiation, No. (%),			0.198
Well	5(4.8)	12(8.4)	
Moderate	69(65.7)	102(70.8)	
Poor	31(29.5)	30(20.8)	
Histologic type, No. (%),			0.003
Adenocarcinoma	75(71.4)	125(86.8)	
Mucus adenocarcinoma and signet-ring cell carcinoma	30(28.6)	19(13.2)	
Lymphatic invasion, No. (%),			< 0.001
Negative	63(60)	123(85.4)	
Positive	42(40)	21(14.6)	
Vascular invasion, No. (%),			< 0.001
Negative	63(60)	121(84)	
Positive	42(40)	23(16)	
Perineural invasion, No. (%),			< 0.001
Negative	68(64.8)	113(78.5)	
Positive	37(35.2)	31(21.5)	
Budding, No. (%),			0.098
Mild	15(14.3)	32(22.2)	
Moderate	17(16.2)	31(21.5)	
Marked	73(69.5)	81(56.3)	
Collagen score, [median (IQR)]	-0.285(-0.399, -0.130)	-0.332 (-0.538, -0.181)	0.006
Immunoscore			< 0.01
High	27(25.7%)	73(50.7%)	
Intermediate	46(43.8%)	53(36.8%)	

Table 1 (continued)

Clinicopathological feature	LNM+	LNM-	p value
Low	32(30.5%)	18(12.5%)	



C

$$\text{Collagen score} = \sum \text{coefficient} \times \text{Collagen feature}$$

Fig. 1 Collagen score construction flowcharts. **A** Five random regions of interest (ROIs, 500 × 500 μm) were selected from H&E staining corresponding to the unstained IM region for multiphoton imaging with a 20 × objective. Then, the TPEF/SHG images were transferred to gray-scale images for collagen feature extraction. **B** LASSO logistic regression was performed to select the prognostic factors. **C** The collagen score formula comprised the extracted collagen features multiplied by the corresponding coefficients, and then added

After calculation, the collagen score [median (IQR)] in the LNM+ group was $-0.236(-0.399, -0.070)$ and that in the LNM- group was $-0.366(-0.570, -0.193)$. There was a significant difference between the two groups ($p < 0.01$).

Immunoscore

The density of CD3+ and CD8+ tumor-infiltrating lymphocytes (TILs) was higher in IM than in TC. Percentiles for evaluating the immune response were based on the density (cells/mm²) of CD3+ and CD8+ cells (eTable 1 in supplement). The cutoff value was based on the maximum Youden index of positive cell density.

In the LNM- group, 4 (2.8%), 14 (9.7%), 53 (36.8%), 26 (18.1%), and 47 (32.6%) patients had immunoscores of I0, I1, I2, I3, and I4, respectively. In the LNM+ group, 19 (18.1%), 13 (12.4%), 46 (43.8%), 18 (17.1%), and 9 (8.6%) patients had immunoscores of I0, I1, I2, I3, and I4, respectively. The immunoscore was divided into three categories: low, intermediate and high. In the LNM+ group, 32 (30.5%), 46 (43.8%) and 27 (25.7%)

patients had low, intermediate and high immunoscores, respectively; in the LNM- group, these numbers were 18 (12.5%), 53 (36.8%) and 73 (50.7%). There was a significant difference between the two groups ($p < 0.01$). This result indicates that the immunoscore is associated with LNM. A flow chart showing the construction of the immunoscore is shown in Fig. 2.

Univariate and multivariate analyses of LNM and nomogram construction

The univariate analysis revealed that T3-4 (OR, 3.838; 95% CI,1.265–11.642, $p=0.018$), lymphatic invasion (OR, 3.905; 95% CI,2.131–7.154, $p < 0.01$), vascular invasion (OR,3.507; 95% CI,1.939–6.344 $p < 0.01$), perineural invasion (OR, 1.983; 95% CI,1.128–3.487, $p=0.017$), mucus adenocarcinoma and signet-ring cell carcinoma (OR, 2.632; 95% CI, 1.385–5.001, $p=0.003$), the collagen score (OR, 5.596; 95% CI,2.189–14.306, $p < 0.01$) and the immunoscore [intermediate group (OR, 2.347; 95% CI, 1.298–4.243, $p=0.005$); low group (OR, 4.807; 95% CI,

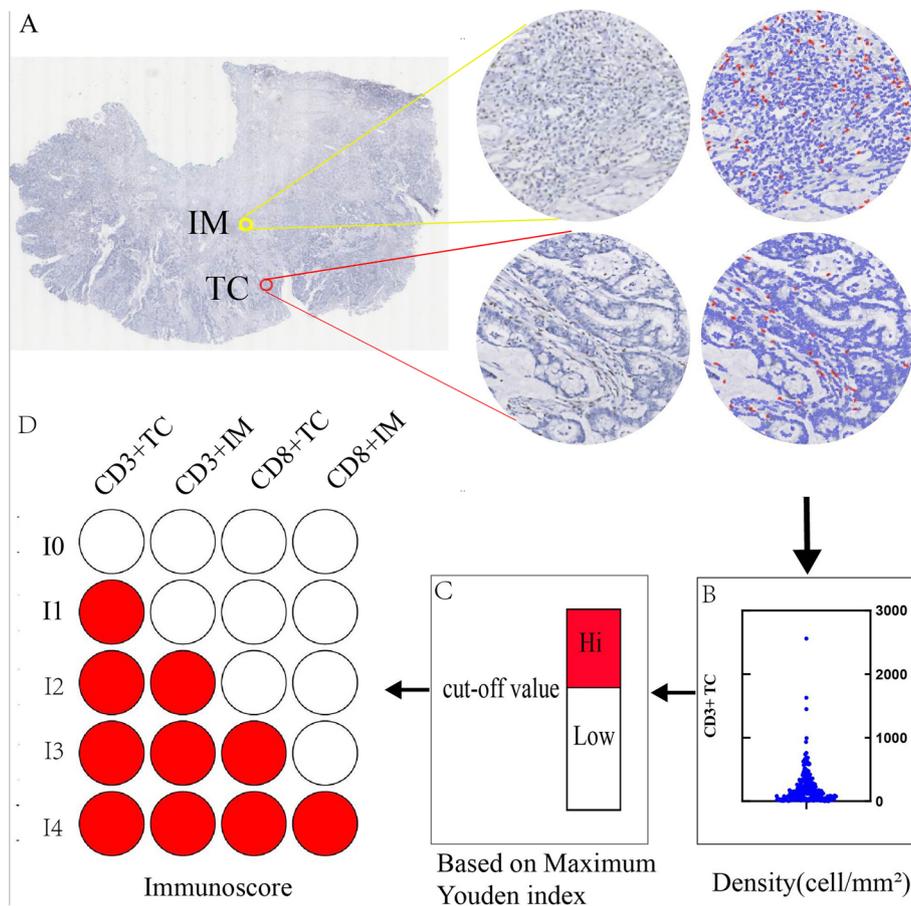


Fig. 2 Immunoscore construction flowcharts. **A** ROIs in TC and IM were manually annotated. **B** The number and density of positively stained cells were calculated. **C** The maximum Youden index of mean density was used as the cutoff value to distinguish 'high' and 'low' immune responses. **D** the CD3_{TC}, CD3_{IM}, CD8_{TC}, and CD8_{IM} scores were added and converted into an immunoscore (I0–I4)

324.943, $p < 0.01$) were statistically significant with LNM in CC. Moreover, the multivariate analysis indicated that lymphatic invasion (OR: 3.892, 95% CI: 1.784–8.491, $p = 0.001$), vascular invasion (OR, 3.234, 95% CI: 1.544–6.776); $p = 0.002$), mucus adenocarcinoma and signet-ring cell carcinoma (OR: 2.990, 95% CI: 1.413–6.328, $p = 0.004$), the collagen score (OR: 6.304, 95% CI: 2.145–18.527, $p = 0.001$) and the immunoscore [intermediate group (OR, 2.473; 95% CI, 1.192–5.130; $p = 0.015$); low group (OR, 5.877; 95% CI, 2.423–14.257; $p < 0.01$)] were independent risk factors for CC LNM. We summarized the univariate and multivariate logistic regression analyses of LNM in Table 2. The VIF of each predictor was < 10 , and the tolerance was > 0.1 , indicating no multicollinearity in the traditional model. [24] No multicollinearity among these factors was observed (eTable 2 in supplement). A nomogram was constructed based on these five independent factors (Fig. 3a). The newly developed model showed good discrimination with an AUROC of 0.809 (95% CI: 0.755–0.862), and the calibration curve showed good agreement between the nomogram-estimated probability of LNM and the actual LNM rate (Fig. 3c). The Hosmer–Lemeshow test demonstrated a $p = 0.058$, indicating no departure from a good fit.

Comparison between the traditional clinicopathological nomogram and the new nomogram integrating clinicopathological features, collagen and immunoscores.

To illustrate the superiority of the new model, we removed the immunoscore and collagen score. The traditional clinicopathological model was constructed based on clinicopathological characteristics (Fig. 3b). After univariate analysis, T category, histologic type, lymphatic invasion, vascular invasion, and perineural invasion were related to LNM ($p < 0.05$). A further multivariate logistic regression analysis showed that mucus adenocarcinoma and signet-ring cell carcinoma ($p < 0.01$, OR: 3.703, 95% CI: 1.860–7.375), lymphatic invasion ($p < 0.01$, OR: 3.226, 95% CI: 1.635–6.361), and vascular invasion ($p = 0.006$, OR: 2.651, 95% CI: 1.358–5.176) were positively correlated with LNM (eTable 3 in the Supplement). There was no multicollinearity among the independent risk factors in the model (eTable 4 in the Supplement). Furthermore, we developed an ROC curve to estimate the accuracy of the model, which had an AUC of 0.715 (95% CI: 0.649–0.780), implying moderate concordance. The calibration curve showed that the nomogram was in good agreement, with a Hosmer–Lemeshow test p value of 0.982 (Fig. 3d).

Compared with the traditional clinicopathological model, the new model integrating clinicopathological features, the collagen score and immunoscore, showed a more robust ability to estimate the risk for CC LNM ($p < 0.001$) (Fig. 3e). The area under the ROC curve of the

new model was 0.809 (95% CI: 0.755–0.862) and that of the traditional model was 0.715 (95% CI: 0.649–0.780), and this difference was significant ($p < 0.001$). The sensitivity, specificity, accuracy, PPV, and NPV of the new model were 71.43%, 79.17%, 75.90%, 71.43% and 79.17%, respectively. The sensitivity, specificity, accuracy, PPV, and NPV of the old model were 71.42%, 63.89%, 67.07%, 59.06% and 75.41%, respectively (Supplementary eTable 5). A comparison of the DCA between the new model and traditional model is shown in Fig. 3f. DCA demonstrated that the nomogram might indicate a better net benefit for determining LNM risk in CC than nontreatment or all-treatment strategies. The new model was superior to the traditional model.

Discussion

LNM status determines whether patients with CC need lymphadenectomy after local excision. Local lymph nodes should be removed from patients with advanced CC to reduce the recurrence rate. The 5-year survival rate of patients with stage III is over 20% lower than that of those with stage II CC (59.5% vs. 82.5%). [26] Thus, LNM status has an important impact on CC treatment and management. [27]

In our study, the LNM rate among 249 CC patients was 42.2% (105/249). We integrated the collagen score, immunoscore and clinicopathological factors to construct a model for analyzing LNM risk. After univariate and multivariate analyses, lymphatic invasion, vascular invasion, mucus adenocarcinoma and signet-ring cell carcinoma, the collagen score and immunoscore were independent risk factors for CC LNM. The AUROC of the new model was 0.809, 95% CI: 0.755–0.862, which was significantly higher than that of the traditional model (0.715, 95% CI: 0.649–0.780).

Tumors are not just a collection of single malignant cells. Tumors also interact with their surrounding TME. [28] Collagen is the main component of the TME ECM. In the 1970s, Wolfe et al. found a relationship between collagen fiber density and breast cancer. [29] Collagen around normal epithelial structures in breast tissue is usually coiled and smooth. [29] However, with the development of tumors, collagen gradually thickens, linearizes and hardens, promoting tumor migration and metastasis. In 2007, Wyckoff et al. reported that breast cancer cells and white blood cells migrate rapidly along collagen fibers in vivo. [30] Collagen fibers are considered the “highway” of tumor escape, and the number and arrangement of these “highways” are directly involved in the migration process of tumor cells. [30] Fang et al. also observed a linear invasion of these ‘highways’ in hepatocellular carcinoma. [31] Adur et al. mentioned that under physiological conditions, collagen fibers were arranged at an angle

Table 2 Univariate and multivariate logistic regression of LNM

Variable	Univariate Logistic Regression		Multivariate Logistic regression	
	OR (95% CI)	p value	OR (95% CI)	p value
Age (year)				
< 60	1 (Reference)			
≥ 60	1.616(0.974–2.682)	0.063		
Sex				
Male (%)	1 (Reference)			
Female (%)	0.633 (0.325–1.233)	0.179		
BMI				
< 24	1 (Reference)			
≥ 24	0.827 (0.459–1.487)	0.525		
ASA				
1	1 (Reference)			
2	1.564 (0.815–3.004)	0.179		
3	2.000 (0.671–5.961)	0.213		
Preoperative serum CEA (ng/ml)				
< 5	1 (Reference)			
≥ 5	1.464 (0.753–2.849)	0.261		
Lymphadenectomy on CT				
< 10 mm	1 (Reference)			
≥ 10 mm	0.557 (0.261–1.189)	0.130		
Size (cm)				
< 5	1 (Reference)			
≥ 5	0.775 (0.466–1.289)	0.326		
Location				
Left-sided colon	1 (Reference)			
Right-sided colon	1.579 (0.934–2.670)	0.087		
T category				
T1-2	1 (Reference)			
T3-4	3.838 (1.265–11.642)	0.018		
Differentiation				
Well	1 (Reference)			
Moderate	1.624 (0.547–4.815)	0.382		
Poor	2.480 (0.779–7.893)	0.124		
Histologic type				
Adenocarcinoma	1 (Reference)			
Mucus adenocarcinoma and signet-ring cell carcinoma	2.632(1.385–5.001)	0.003	2.990(1.413–6.328)	0.004
Lymphatic invasion				
Negative	1 (Reference)			
Positive	3.905 (2.131–7.154)	<0.01	3.892(1.784–8.491)	0.001
Vascular invasion				
Negative	1 (Reference)			
Positive	3.507 (1.939–6.344)	<0.01	3.234(1.544–6.776)	0.002
Perineural invasion				
Negative	1 (Reference)			
Positive	1.983 (1.128–3.487)	0.017		
Budding				
Mild	1 (Reference)			
Moderate	1.170(0.499–2.743)	0.718		

Table 2 (continued)

Variable	Univariate Logistic Regression		Multivariate Logistic regression	
	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value
Marked	1.923 (0.964–3.833)	0.063		
Collagen score	5.596 (2.189–14.306)	< 0.01	6.304 (2.145–18.527)	0.001
Immunoscore				< 0.01
High	1 (Reference)		1 (Reference)	1
Intermediate	2.347(1.298–4.243)	0.005	2.473(1.192–5.130)	0.015
Low	4.807(2.324–9.943)	< 0.01	5.877(2.423–14.257)	< 0.01

of 10° in the epithelial matrix of the intestinal mucosa, while in colorectal tumors, collagen fibers were thicker and arranged at an angle of 50°. [32] However, it was still unclear whether the change in collagen fibers would promote LNM in CC. In our research, we used multiphoton imaging technology to quantitatively analyze collagen, one of the main components of the ECM, and constructed a collagen score. [33] The collagen score was closely related to LNM (OR: 6.304, 95% CI: 2.145–18.527, $p < 0.01$). The collagen score included two morphological parameters (fiber width and crosslink space), one intensity parameter (kurtosis) and seven texture parameters.

Immune cell infiltration, especially antitumor type I lymphocyte infiltration, predicts prognosis in many different tumor types, including CC, ovarian cancer, lung cancer and breast cancer. [34–37] The immunoscore was based on the level of CD3 and CD8 T lymphocyte infiltration of the TME. The analysis of solid tumor tissue IHC is the gold standard for evaluating tumor immune infiltration as it allows accurate quantification of the type, density and location of immune cells. [38, 39] To date, more attention has been paid to the prognostic role of CD3+ and CD8+ lymphocyte density in TC and IM represented by immunohistochemistry (IHC) staining intensity. [13, 15, 40–42] Although there was a consensus that the immunoscore was associated

with the CC patient prognosis, [15] it had not previously been used to analyze CC LNM risk. In our study, we found that the immunoscore was correlated with LNM in CC. The higher the immunoscore was, the less prone to the CC was to LNM. For early-stage colorectal cancer (stages I–II), a higher Immunoscore was observed, whereas for late-stage colorectal cancer (stage III), the Immunoscore was lower. This is consistent with the findings of Pagès F et al.'s study. [15]

This effect might be related to the activation of a highly invasive antitumor immune response to cancer cells in the high immunoscore group.

In this research, vascular invasion, lymphatic invasion, histologic type and the immunoscore were classified variables, and the collagen score was a continuous variable. The new model was constructed based on these five factors. For example, in a patient without lymphatic and vascular invasion, adenocarcinoma, a high immunoscore and a collagen score of –1.6, the risk for LNM was less than 0.1. If a patient had lymphatic invasion, vascular invasion, signet-ring cell carcinoma and mucinous adenocarcinoma, a low immunoscore and a collagen score of 0.6, they had a probability of LNM over 0.9. Therefore, the nomogram could intuitively and conveniently evaluate the risk for LNM in CC patients. The decision curve

(See figure on next page.)

Fig. 3 Comparison of the new model and traditional model. **A** Nomogram for estimating LNM in CC based on the collagen score and immunoscore. The probability of LNM involvement in CC was weighed. A line was drawn to the point on the axis for either of the following parameters: vascular invasion, lymphatic invasion, histologic type, immunoscore and collagen score. The scores of either variable were summed and located on the total point line. Next, a vertical line was projected from the total point line to the predicted probability bottom scale to obtain the individual probability of LNM involvement. **B** Nomograms constructed from clinicopathological features. **C** Calibration curve of the new model. The diagonal dotted line represents a perfect prediction by an ideal model. The other dotted line represents the performance of the nomogram. The solid line represents the bias-corrected performance of the nomogram. The calibration curve of the nomogram had a mean absolute error of 0.01. **D** Calibration curve of the traditional model. The calibration curve of the nomogram had a mean absolute error of 0.028. **E** Comparison of ROC curves between the traditional and new predictive models. The area under the ROC curve of the new model and the traditional model was 0.809 (95% CI: 0.755–0.862) and 0.715 (95% CI: 0.649–0.780), respectively, and there was a significant difference between them ($p < 0.01$). **F** Comparison of decision curve analysis between the traditional and new models. The x axis indicates the threshold probability, and the y axis indicates the net benefit. The black line represents the assumption that no patient had LNM, and the gray line represents the assumption that all patients had LNM. The red line represented the new model. The blue line represented the traditional model

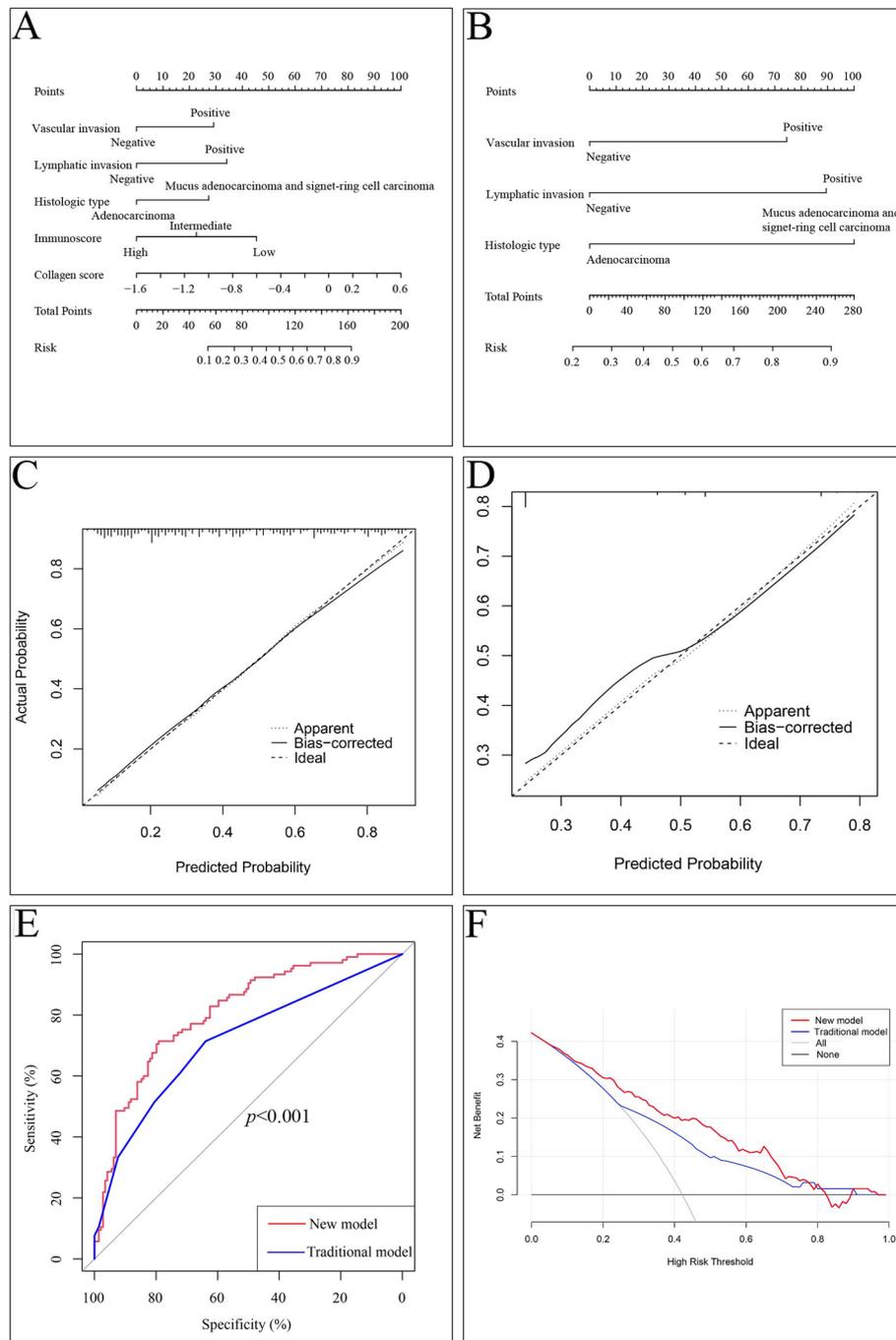


Fig. 3 (See legend on previous page.)

showed that the model could provide more net benefits than stratifying by all treatments or no treatment.

Compared with the traditional clinicopathological model based on vascular invasion, lymphatic invasion and histologic type, the new model integrating clinicopathological features, the collagen score and immunoscore performed better (AUROC: 0.809, 95% CI:

0.755–0.862 vs. 0.715, 95% CI: 0.649–0.780, $p < 0.001$). The overall accuracy of the new model was improved from 67.07% to 75.90%, the sensitivity was improved from 71.42% to 71.43%, and the specificity was improved from 63.89% to 79.17%. This indicates that the collagen score and immunoscore play an important role in CC LNM. In short, the TME collagen score and immunoscore are

associated with CC LNM. Integrating the collagen score, immunoscore, and clinicopathological features enables individualized prediction of the probability of lymph node metastasis (LNM) in colon cancer, assisting clinicians in precisely performing surgical interventions and planning postoperative adjuvant therapies.

Limitations

This study has limitations. First, it was retrospective in nature, and all specimens were acquired from 1 medical center in China; thus, potential bias was inevitable. A prospective, multicenter trial is needed to validate the performance of collagen score and immunoscore. Second, the underlying mechanism of the collagen score and immunoscore for the association with CC LNM remains unclear; therefore, further investigations are needed to better understand the role of collagen score and immunoscore in CC LNM. Third, the purpose of this study was to investigate the association of the tumor microenvironment collagen score and immunoscore with colon cancer lymph node metastasis. Therefore, the primary outcome and end point was lymph node metastasis collected from post-operative pathological report. We did not collect follow-up survival. This is also a limitation of this study. Fourth, recent studies reported detection rates of vascular invasion ranging from 19 to 34% [6, 43–45]. The rate of vascular invasion in our study was 26.1% (65/249). This discrepancy may be due to differences in pathological assessment criteria or interpretation by pathologists at different institutions need for external validation.

Conclusions

In summary, the TME collagen score and immunoscore are associated with LNM in CC.

Abbreviations

CC	Colon cancer
LNM	Lymph node metastasis
TME	Tumor microenvironment
ECM	Extracellular matrix
Immunoscore	Immune scoring system
FFPE	Formalin-fixed paraffin-embedded
ITBCC	International Tumor Budding Consensus Conference
H&E	Hematoxylin–eosin
SHG/TPEF	Two-photon fluorescence excitation/second harmonic generation
LASSO	Least-absolute shrinkage and selection operator
TC	Tumor center
IM	Invasion margin
OR	Odds ratio
CI	Confidence interval
VIF	Variance inflation factor
AUROC	Area under the receiver operating characteristic curve
DCA	Decision curve analysis
PPV	Positive predictive value
NPV	Negative predictive value

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-025-13842-5>.

- Supplementary Material 1.
- Supplementary Material 2.
- Supplementary Material 3.
- Supplementary Material 4.
- Supplementary Material 5.

Authors' contributions

Drs. Chenyan Long, Jiabin Cheng and Mingyuan. Feng contributed equally to this study. Dr Yan had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Chenyan. Long, Mingyuan Feng, Dexin Chen, Jun Yan. Acquisition, analysis, or interpretation of data: Chenyan Long, Jiabin Cheng, Mingyuan Feng, Botao Yan, Yiran Li, Wei Jiang, Dexin Chen, Jun Yan. Drafting of the manuscript: Chenyan Long, Jiabin Cheng, Mingyuan Feng, Jun Yan. Critical revision of the manuscript for important intellectual content: Chenyan Long, Jiabin Cheng, Mingyuan Feng, Botao Yan, Yiran Li, Wei Jiang, Dexin Chen, Jun Yan. Statistical analysis: Chenyan Long, Botao Yan, Mingyuan Feng. Obtained funding: Jun Yan. Administrative, technical, or material support: Jun Yan. Supervision: Dexin Chen, Jun Yan.

Funding

This work was supported by the following: grant 81773117 from the National Natural Science Foundation of China; grant 2020B121201004 from the Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Cancer; grant 320.2710.1851 from the Special Fund from the Clinical Research of Wu Jieping Medical Foundation; grant LC2016PY010 from the Clinical Research Project of Southern Medical University; grant 2019Z023 from the President Fund of Nanfang Hospital; and grant 2018CR034 from the Clinical Research Project of Nanfang Hospital.

Data availability

The data supports the findings of this study are available in tables and supplementary materials of this article. The relevant data of colon cancer patients treated in our institution involved during the current research period can be obtained from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This retrospective research was approved by the Institutional Review Board of Nanfang Hospital. Patient informed consent was waived, and patient information was protected. This study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Consent to publish

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Colorectal Surgery, Fujian Medical University Union Hospital, Fuzhou, Fujian 350001, People's Republic of China. ²Department of General Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong 510515, People's Republic of China. ³Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Cancer, Guangzhou, Guangdong 510515, People's Republic of China. ⁴Department of Colorectal & Anal Surgery, Cancer Hospital affiliated to, Guangxi Medical University, Nanning, Guangxi 530021, People's Republic of China.

Received: 5 May 2024 Accepted: 28 February 2025

Published online: 19 March 2025

References

- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49. <https://doi.org/10.3322/caac.21660>.
- Kuo YT, Tsai WS, Hung HY, et al. Prognostic value of regional lymph node involvement in patients with metastatic colorectal cancer: palliative versus curative resection. *World J Surg Oncol*. 2021;19(1):150. <https://doi.org/10.1186/s12957-021-02260-z>.
- Xu W, He Y, Wang Y, et al. Risk factors and risk prediction models for colorectal cancer metastasis and recurrence: an umbrella review of systematic reviews and meta-analyses of observational studies. *BMC Med*. 2020;18(1):172. <https://doi.org/10.1186/s12916-020-01618-6>.
- Macias-Garcia F, Celeiro-Munoz C, Lesquereux-Martinez L, et al. A clinical model for predicting lymph node metastasis in submucosal invasive (T1) colorectal cancer. *Int J Colorectal Dis*. 2015;30(6):761–8. <https://doi.org/10.1007/s00384-015-2164-3>.
- 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):2157–64. <https://doi.org/10.1200/JCO.2015.65.9128>.
- Bae JH, Kim JH, Kye BH, et al. Comparison of vascular invasion with lymph node metastasis as a prognostic factor in stage I–III colon cancer: an observational cohort study. *Front Surg*. 2021;8:773019. <https://doi.org/10.3389/fsurg.2021.773019>.
- Mou A, Li H, Chen XL, et al. Tumor size measured by multidetector CT in resectable colon cancer: correlation with regional lymph node metastasis and N stage. *World J Surg Oncol*. 2021;19(1):179. <https://doi.org/10.1186/s12957-021-02292-5>.
- Chen D, Liu Z, Liu W, et al. Predicting postoperative peritoneal metastasis in gastric cancer with serosal invasion using a collagen nomogram. *Nat Commun*. 2021;12(1):179. <https://doi.org/10.1038/s41467-020-20429-0>.
- Chen D, Chen G, Jiang W, et al. Association of the collagen signature in the tumor microenvironment with lymph node metastasis in early gastric cancer. *JAMA Surg*. 2019;154(3):e185249. <https://doi.org/10.1001/jamasurg.2018.5249>.
- Dong X, Huang Y, Yu X, et al. Collagen score in the tumor microenvironment predicts the prognosis of rectal cancer patients after neoadjuvant chemoradiotherapy. *Radiother Oncol*. 2021. <https://doi.org/10.1016/j.radonc.2021.12.023>.
- Wen BL, Brewer MA, Nadiarynykh O, et al. Texture analysis applied to second harmonic generation image data for ovarian cancer classification. *J Biomed Opt*. 2014;19(9):096007. <https://doi.org/10.1117/1.JBO.19.9.096007>.
- Kakkad SM, Solaiyappan M, Argani P, et al. Collagen I fiber density increases in lymph node positive breast cancers: pilot study. *J Biomed Opt*. 2012;17(11):116017. <https://doi.org/10.1117/1.JBO.17.11.116017>.
- Fridman WH, Pages F, Sautes-Fridman C, et al. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer*. 2012;12(4):298–306. <https://doi.org/10.1038/nrc3245>.
- Mlecnik B, Van den Eynde M, Bindea G, et al. Comprehensive intrametastatic immune quantification and major impact of immunoscore on survival. *J Natl Cancer Inst*. 2018;110(1). <https://doi.org/10.1093/jnci/djx123>.
- Pages F, Mlecnik B, Marliot F, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. 2018;391(10135):2128–39. [https://doi.org/10.1016/S0140-6736\(18\)30789-X](https://doi.org/10.1016/S0140-6736(18)30789-X).
- Chen C, Lu FC, Department of Disease Control Ministry of Health PRC. The guidelines for prevention and control of overweight and obesity in Chinese adults. *Biomed Environ Sci*. 2004;17 Suppl:1–36.
- Yu N, Lin S, Wang X, et al. Endoscopic obstruction predominantly occurs in right-side colon cancer and endoscopic obstruction with tumor size. *Front Oncol*. 2024;14:1415345. <https://doi.org/10.3389/fonc.2024.1415345>.
- Lugli A, Kirsch R, Ajioka Y, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod Pathol*. 2017;30(9):1299–311. <https://doi.org/10.1038/modpathol.2017.46>.
- Zhuo S, Chen J, Luo T, et al. Multimode nonlinear optical imaging of the dermis in ex vivo human skin based on the combination of multichannel mode and Lambda mode. *Opt Express*. 2006;14(17):7810–20. <https://doi.org/10.1364/oe.14.007810>.
- Xu S, Kang CH, Gou X, et al. Quantification of liver fibrosis via second harmonic imaging of the Glisson's capsule from liver surface. *J Biophotonics*. 2016;9(4):351–63. <https://doi.org/10.1002/jbio.201500001>.
- Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med*. 2007;26(30):5512–28. <https://doi.org/10.1002/sim.3148>.
- Kather JN, Suarez-Carmona M, Charoentong P, et al. Topography of cancer-associated immune cells in human solid tumors. *Elife*. 2018;7. <https://doi.org/10.7554/eLife.36967>.
- Luomaranta A, Leminen A, Loukovaara M. Prediction of lymph node and distant metastasis in patients with endometrial carcinoma: a new model based on demographics, biochemical factors, and tumor histology. *Gynecol Oncol*. 2013;129(1):28–32. <https://doi.org/10.1016/j.ygyno.2013.01.008>.
- Dormann CF, Elith J, Bacher S, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013;36(1):27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>.
- Kerr KF, Brown MD, Zhu K, et al. Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. *J Clin Oncol*. 2016;34(21):2534–40. <https://doi.org/10.1200/JCO.2015.65.5654>.
- O'Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. *J Natl Cancer Inst*. 2004;96(19):1420–5. <https://doi.org/10.1093/jnci/djh275>.
- Weitz J, Koch M, Debus J, et al. Colorectal cancer. *Lancet*. 2005;365(9454):153–65. [https://doi.org/10.1016/S0140-6736\(05\)17706-X](https://doi.org/10.1016/S0140-6736(05)17706-X).
- Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci*. 2012;125(Pt 23):5591–6. <https://doi.org/10.1242/jcs.116392>.
- Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*. 1976;37(5):2486–92. [https://doi.org/10.1002/1097-0142\(197605\)37:5%3c2486::aid-cnrc2820370542%3e3.0.co;2-8](https://doi.org/10.1002/1097-0142(197605)37:5%3c2486::aid-cnrc2820370542%3e3.0.co;2-8).
- Wyckoff JB, Wang Y, Lin EY, et al. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res*. 2007;67(6):2649–56. <https://doi.org/10.1158/0008-5472.CAN-06-1823>.
- Fang M, Yuan J, Peng C, et al. Collagen as a double-edged sword in tumor progression. *Tumour Biol*. 2014;35(4):2871–82. <https://doi.org/10.1007/s13277-013-1511-7>.
- Adur J, Bianchi M, Pelegati VB, et al. Colon adenocarcinoma diagnosis in human samples by multicontrast nonlinear optical microscopy of hematoxylin and eosin stained histological sections. *J Cancer Ther*. 2014;05(13):1259–69. <https://doi.org/10.4236/jct.2014.513127>.
- Yan J, Zhuo S, Chen G, et al. Real-time optical diagnosis for surgical margin in low rectal cancer using multiphoton microscopy. *Surg Endosc*. 2014;28(1):36–41. <https://doi.org/10.1007/s00464-013-3153-7>.
- Pages F, Kirilovsky A, Mlecnik B, et al. In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *J Clin Oncol*. 2009;27(35):5944–51. <https://doi.org/10.1200/JCO.2008.19.6147>.
- Hwang WT, Adams SF, Tahirovic E, et al. Prognostic significance of tumor-infiltrating T cells in ovarian cancer: a meta-analysis. *Gynecol Oncol*. 2012;124(2):192–8. <https://doi.org/10.1016/j.ygyno.2011.09.039>.
- Dieu-Nosjean MC, Antoine M, Danel C, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J Clin Oncol*. 2008;26(27):4410–7. <https://doi.org/10.1200/JCO.2007.15.0284>.
- Denkert C, Loibl S, Noske A, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol*. 2010;28(1):105–13. <https://doi.org/10.1200/JCO.2009.23.7370>.
- Fridman WH, Zitvogel L, Sautes-Fridman C, et al. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol*. 2017;14(12):717–34. <https://doi.org/10.1038/nrclinonc.2017.101>.
- Becht E, Giraldo NA, Germain C, et al. Immune contexture, immunoscore, and malignant cell molecular subgroups for prognostic and theranostic

- classifications of cancers. *Adv Immunol.* 2016;130:95–190. <https://doi.org/10.1016/bs.ai.2015.12.002>.
40. Galon J, Mlecnik B, Bindea G, et al. Towards the introduction of the “immunoscore” in the classification of malignant tumours. *J Pathol.* 2014;232(2):199–209. <https://doi.org/10.1002/path.4287>.
 41. Angell H, Galon J. From the immune contexture to the immunoscore: the role of prognostic and predictive immune markers in cancer. *Curr Opin Immunol.* 2013;25(2):261–7. <https://doi.org/10.1016/j.coi.2013.03.004>.
 42. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science.* 2006;313(5795):1960–4. <https://doi.org/10.1126/science.1129139>.
 43. Gibson KM, Chan C, Chapuis PH, et al. Mural and extramural venous invasion and prognosis in colorectal cancer. *Dis Colon Rectum.* 2014;57(8):916–26. <https://doi.org/10.1097/DCR.000000000000162>.
 44. Parnaby CN, Scott NW, Ramsay G, et al. Prognostic value of lymph node ratio and extramural vascular invasion on survival for patients undergoing curative colon cancer resection. *Br J Cancer.* 2015;113(2):212–9. <https://doi.org/10.1038/bjc.2015.211>.
 45. Amri R, England J, Bordeianou LG, et al. Risk stratification in patients with stage II colon cancer. *Ann Surg Oncol.* 2016;23(12):3907–14. <https://doi.org/10.1245/s10434-016-5387-9>.

Publisher's Note

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