

RESEARCH

Open Access



Artificial intelligence for quantifying immune infiltrates interacting with stroma in colorectal cancer

Jing Yang^{1,2†}, Huifen Ye^{1,3,4†}, Xinjuan Fan^{5†}, Yajun Li^{1†}, Xiaomei Wu⁶, Minning Zhao⁴, Qingru Hu⁴, Yunrui Ye⁴, Lin Wu⁷, Zhenhui Li^{1,8}, Xueli Zhang⁹, Changhong Liang^{1,3}, Yingyi Wang¹⁰, Yao Xu^{1,11}, Qian Li¹, Su Yao^{12*}, Dingyun You^{13*}, Ke Zhao^{1,3,14*} and Zaiyi Liu^{1,3*}

Abstract

Background: We proposed an artificial intelligence-based immune index, Deep-immune score, quantifying the infiltration of immune cells interacting with the tumor stroma in hematoxylin and eosin-stained whole-slide images of colorectal cancer.

Methods: A total of 1010 colorectal cancer patients from three centers were enrolled in this retrospective study, divided into a primary (N = 544) and a validation cohort (N = 466). We proposed the Deep-immune score, which reflected both tumor stroma proportion and the infiltration of immune cells in the stroma region. We further analyzed the correlation between the score and CD3⁺T cells density in the stroma region using immunohistochemistry-stained whole-slide images. Survival analysis was performed using the Cox proportional hazard model, and the endpoint of the event was the overall survival.

Result: Patients were classified into 4-level score groups (score 1–4). A high Deep-immune score was associated with a high level of CD3⁺T cells infiltration in the stroma region. In the primary cohort, survival analysis showed a significant difference in 5-year survival rates between score 4 and score 1 groups: 87.4% vs. 58.2% (Hazard ratio for score 4 vs. score 1 0.27, 95% confidence interval 0.15–0.48, P < 0.001). Similar trends were observed in the validation cohort (89.8% vs. 67.0%; 0.31, 0.15–0.62, < 0.001). Stratified analysis showed that the Deep-immune score could distinguish high-risk and low-risk patients in stage II colorectal cancer (P = 0.018).

[†]Jing Yang, Huifen Ye, Xinjuan Fan and Yajun Li contributed equally to this work

*Correspondence: ysjulien@126.com; youdingyun@qq.com; zhaoke@gdph.org.cn; liuzaiyi@gdph.org.cn

¹ Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangdong Provincial People's Hospital, Guangzhou, China

¹² Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, 106 Zhongshan Er Road, Guangzhou 510080, China

¹³ School of Public Health, Kunming Medical University, 191 West Renmin Road, Kunming 650500, China

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



Conclusion: The proposed Deep-immune score quantified by artificial intelligence can reflect the immune status of patients with colorectal cancer and is associated with favorable survival. This digital pathology-based finding might advocate change in risk stratification and consequent precision medicine.

Keywords: Deep learning, Whole-slide images, Deep-immune score, Colorectal cancer, Digital pathology

Background

The immune system was an essential component of the tumor microenvironment (TME) in colorectal cancer (CRC) [1]. It played a central role in tumorigenesis and progression, affecting the treatment and prognosis of CRC [2]. The need to go beyond the tumor-node-metastasis (TNM) staging system has been addressed by detecting the tumor immune microenvironment (TIME), which was affected by the type, density, and location of tumor-infiltrating lymphocytes (TILs) [3–5]. The Immunoscore[®], calculated based on the density of CD3⁺ and CD8⁺ T cells in the tumor core and invasive margin, has been shown to provide superior prognosis to TNM stage in CRC [4, 6]. However, most of the current immune scores needed to be stained with immunohistochemistry (IHC) or other special staining (such as multiplexed immunofluorescence), which is not commonly used in clinical practice, hindering their widespread clinical application.

In most cases, high-quality hematoxylin and eosin (HE)-stained slide was sufficient to confirm the diagnosis [7]. Immune-related features within HE-stained whole-slide images (WSIs), such as TILs and Crohn-like lymphoid reaction (CLR), can be quantified using deep learning. HE-stained slides also have immunological information, with a larger sample size and lower cost. Deep learning has recently entered the field of computational pathology and shows excellent promise for task automation [8]. Deep learning with digital pathology has been successfully applied to breast, prostate, lung, and CRC [9–15]. The tissue-level components of CRC, such as tumor-stroma ratio (TSR) and CLR [16, 17], can be quantified using deep learning. Therefore, the artificial intelligence (AI)-based method has the potential to quantify the tissue composition and immune status using HE-stained WSIs.

The spatial distribution of TILs was important for CRC prognosis. TILs can interact with the tumor through direct contact or cytokine signaling to produce tumor-killing immune cells for protection of the organism, mainly in the interstitial region [18]. The immune cells in the stroma are produced by the surrounding lymphoid follicles or migrate from the blood to the tumor area. As the stroma increases, the number of immune cells in the stroma decreases and the anti-tumor effect decreases [16]. In addition, the distribution of immune cells in the

stroma impacted the prognosis of CRC [19]. We hypothesized, therefore, that a comprehensive consideration of the stroma proportion and the immune cell infiltration in stroma would further refine the prognostic stratification of CRC patients.

The aim of this study was two-fold. First, we proposed a deep learning-based immune index, the Deep-immune score, quantifying immune infiltration interaction with the stroma in HE-stained WSIs. A further investigation of its prognostic value is performed in CRC patients from three centers.

Methods

Patients

Our study recruited patients with histologically confirmed stage I–III CRC who had undergone surgery with the intent of curing their cancer, and had paraffin-embedded tumor samples available (Additional file 1: Fig. S1). The primary cohort consists of patients from Guangdong Provincial People's Hospital (from Apr 2008 to Jun 2016), and the validation cohort includes patients from Yunnan Cancer Hospital (from Dec 2012 to Apr 2015) and The Sixth Affiliated Hospital of Sun Yat-sen University (from Jan 2013 to Oct 2016). This study was approved by the Research Ethics Committees of the respective hospitals, with the need for informed consent waived for this retrospective study. The exclusion criteria are presented in Additional file 1: Methods.

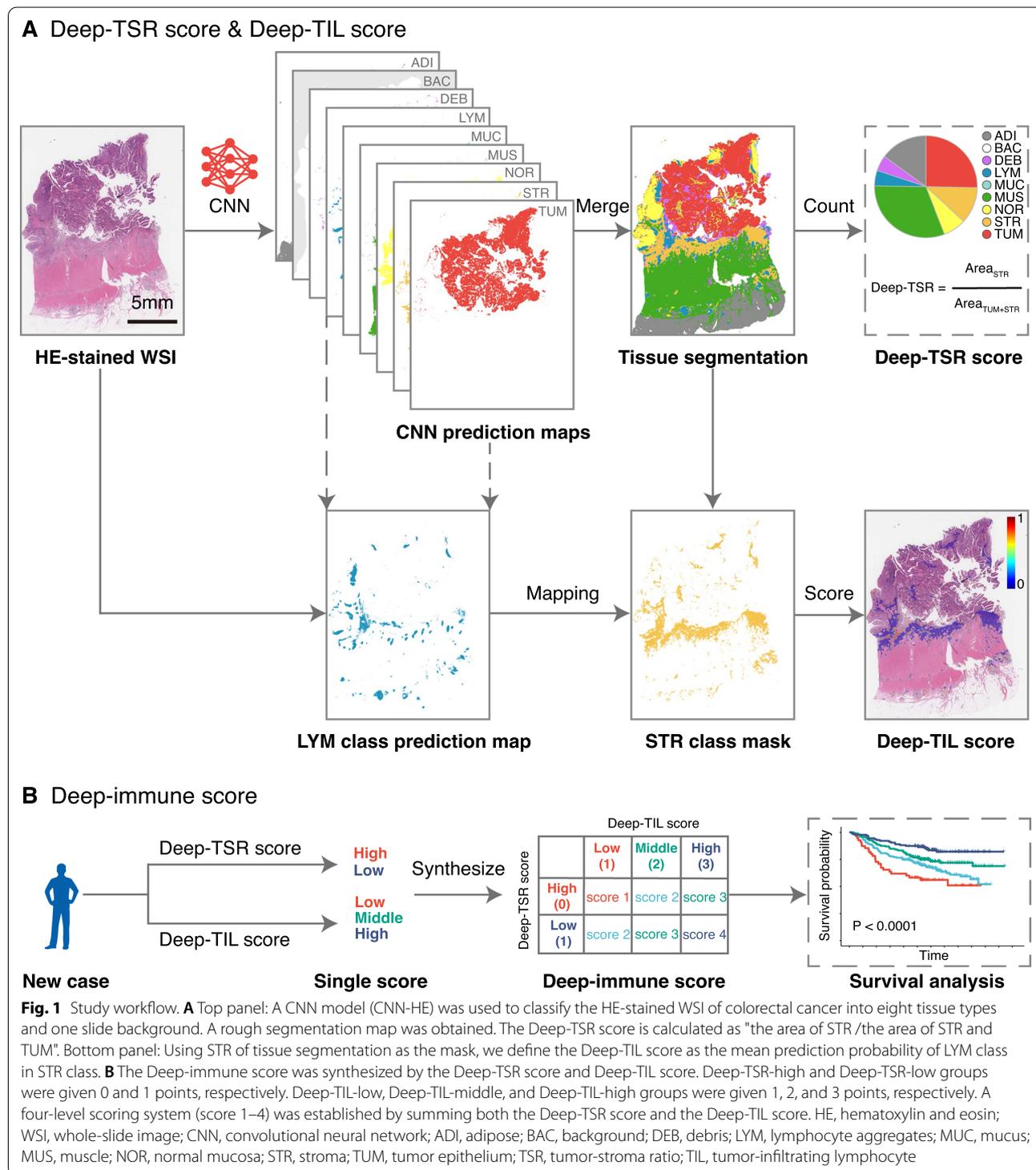
Clinicopathological characteristics information was collected, including age, sex, stage, tumor location, grade, and carcinoembryonic antigen (CEA) level. Stage was performed according to the Union for International Cancer Control guideline [20]. Preoperative CEA level was binarized as normal and abnormal groups, with the cut-off of 5 ng/mL. We focused on overall survival (OS), defined as the time from surgery to death for any reason.

Whole-slide images acquisition and segmentation

The HE-stained slide selection and digitization process are shown in Additional file 1: Methods. The stain type, stain location, and digital slide scanner for three different hospitals are presented in Additional file 1: Table S1. A convolutional neural network (CNN) model, VGG-19, was used to classify WSIs into eight tissue types and background. The detailed training process of the deep learning model (CNN-HE) has been reported in our

previous work [16]. Then, HE-stained WSIs were tiled into overlapped patches (224×224 pixel² at $20\times$). The model's prediction maps were reserved, and the tissue segmentation was obtained by setting each image patch as the tissue class with the max probability (Fig. 1A). The

model training and segmentation were done in MATLAB environment (R2019a, MathWorks, USA). Additional file 1: Fig. S2 shows example of HE-stained WSIs and corresponding tissue segmentation results from three hospitals.



Deep-TSR score and Deep-TIL score

With deep learning, CRC tissues within WSIs can be segmented into nine categories. We defined Deep-TSR score as the area of stroma divided by the area of stroma plus tumor epithelium. The calculation process of the Deep-TSR in HE-stained WSIs can be seen in Fig. 1A (top panel). Patients were divided into Deep-TSR-low and Deep-TSR-high groups using a cutoff of 50%.

We also noticed that the infiltration level of lymphocytes in stroma could be quantified by using the stroma region of tissue segmentation as the mask, combined with the prediction probability map of LYM class (Fig. 1A, bottom panel). We define the Deep-TIL score as the mean prediction probability of LYM class in the stroma. For each patient, a percentile value was calculated according to the distribution of the Deep-TIL score in the primary cohort. With 33% and 66% percentiles as thresholds, patients were classified into a 3-level group: low, middle, and high score.

Deep-immune score

We further developed the Deep-immune score, synthesized by the Deep-TSR and Deep-TIL scores. The calculation details are as follows: Deep-TSR-high and Deep-TSR-low groups were given 0 and 1 points, respectively. Deep-TIL-low, Deep-TIL-middle, and Deep-TIL-high groups were given 1, 2, and 3 points, respectively. A four-level scoring system was established by summing both the Deep-TSR and Deep-TIL scores (Fig. 1B).

Immunohistochemistry validation

We used a subgroup of patients for immunohistochemical validation to explore whether the Deep-immune score extracted from HE-stained WSIs could reflect the patient's immune status. We used the CD3⁺T cells to measure immune infiltration in CRC. A consecutive section was processed for immunohistochemistry (Additional file 1: Fig. S1), and the details of the IHC (CD3⁺) staining procedure were presented in our previous work [19].

We used a second deep learning model (CNN-IHC), for tissue-level segmentation, to automatically obtain stroma regions in IHC-stained WSIs. The segmented stroma was used as the region of interest (ROI) for WSI, and all CD3⁺T cells were segmented and counted using our previously developed software [21]. Then, the stroma-CD3 density was calculated by using the number of CD3⁺T cells divided by the stroma area. The calculation process of the stroma-CD3 density is shown in Fig. 2A.

Statistical analysis

Clinicopathological characteristics were compared by Student t-test for a continuous variable or Chi-square

test for a category variable. Kaplan–Meier curves were plotted to determine difference in survival rates among groups, and log-rank tests were used to calculate P values. For multivariate analysis, univariate variables with $P < 0.05$ were selected. Hazard ratio (HR) with 95% confidence interval (CI) was calculated by using the Cox model. The discrimination performance of factor or model was assessed using Harrell's C-statistics (C-index) with 95% CI and the integrated area under the ROC curve (iAUC) of 0–5 years. All statistical analyses were conducted in R environment (version 4.0.3), with statistical significance set at 0.05.

Results

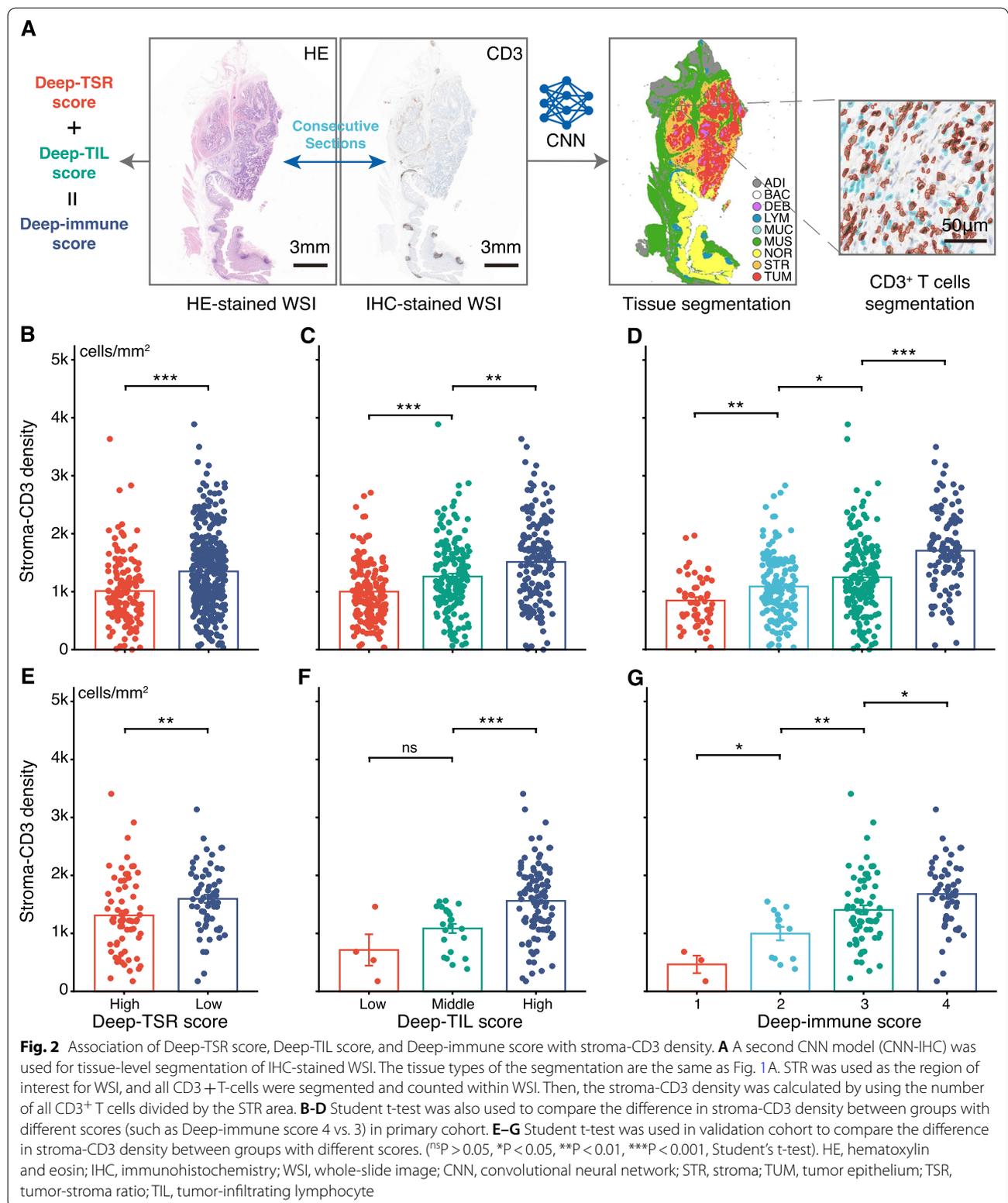
Patients

A total of 1010 CRC patients from three hospitals were included. Among them, 544 patients formed the primary cohort, with a median follow-up time for censored of 6.46 (interquartile range [IQR], 5.46–8.50) years; 466 patients formed the validation cohort, with a median follow-up time of 5.09 (IQR, 4.39–5.82) years. A comparison of clinicopathologic characteristics between the two cohorts is shown in Additional file 1: Table S2. For each patient, one HE-stained WSI was used to construct Deep-TSR, Deep-TIL, and Deep-immune scores. Among them, 477 from the primary cohort and 129 patients from the validation cohort had the paired IHC-stained WSIs for CD3⁺T cells density evaluation (Additional file 1: Fig. S1).

Association of Deep-TSR, Deep-TIL, and Deep-immune scores with stroma-CD3 density

For patients with IHC-stained WSIs in the primary cohort, 338 (71%) patients were grouped as Deep-TSR-low, and 139 (29%) patients were grouped as Deep-TSR-high. The distribution of stroma-CD3 density versus Deep-TSR is shown in Fig. 2B. Deep-TSR-low was associated with a high level of stroma-CD3 density ($P < 0.001$). For the Deep-TIL score, 164 (34%) cases were classified as a low score, 162 (34%) as middle score, and 151 (32%) as a high score. The mean stroma-CD3 density of the high-score group was 1.5 times higher than that of the low-score group (1513 vs. 1001 cells/mm²; $P < 0.001$; Fig. 2C). We also observed that with the increase in Deep-immune score, the mean density of the stroma-CD3 also increased (density in score 1–4: 844, 1088, 1246, and 1708 cells/mm²; Fig. 2D).

In the validation cohort, the Deep-TSR score, Deep-TIL score, and Deep-immune score had a similar trend to the primary cohort (Fig. 2E–G; Additional file 1: Table S3). Based on this analysis, high Deep-immune scores were associated with increasing levels of CD3⁺T cells infiltration in the stroma region.



Prognostic value of Deep-TIL score and Deep-TSR score

In the primary cohort, overall survival was significantly longer for patients with higher Deep-TIL scores. The 5-year survival rates were 67.4% in Deep-TIL-low group, 76.6% in middle group, and 82.9% in high group ($P=0.0003$; Fig. 3A). Patients with high and low Deep-TIL scores experienced a significant difference in survival (unadjusted HR 0.45, 95% CI 0.30–0.67; $P<0.001$; Table 1). In the validation cohort, these findings were confirmed (0.49, 0.31–0.77; 0.002; Fig. 3B; Table 1).

In the primary cohort, the 5-year overall survival for CRC patients with low and high Deep-TSR scores were 78.9% and 67.3% respectively (unadjusted HR for low vs. high 0.62, 95% CI 0.45–0.86; $P=0.004$; Fig. 3C). Validation cohort results confirm the initial findings: the 5-year survival rates for low-stroma and high-stroma patients were 82.7% and 72.2%, respectively, with unadjusted HR (low vs. high) of 0.57 (0.38–0.85; 0.005; Fig. 3D).

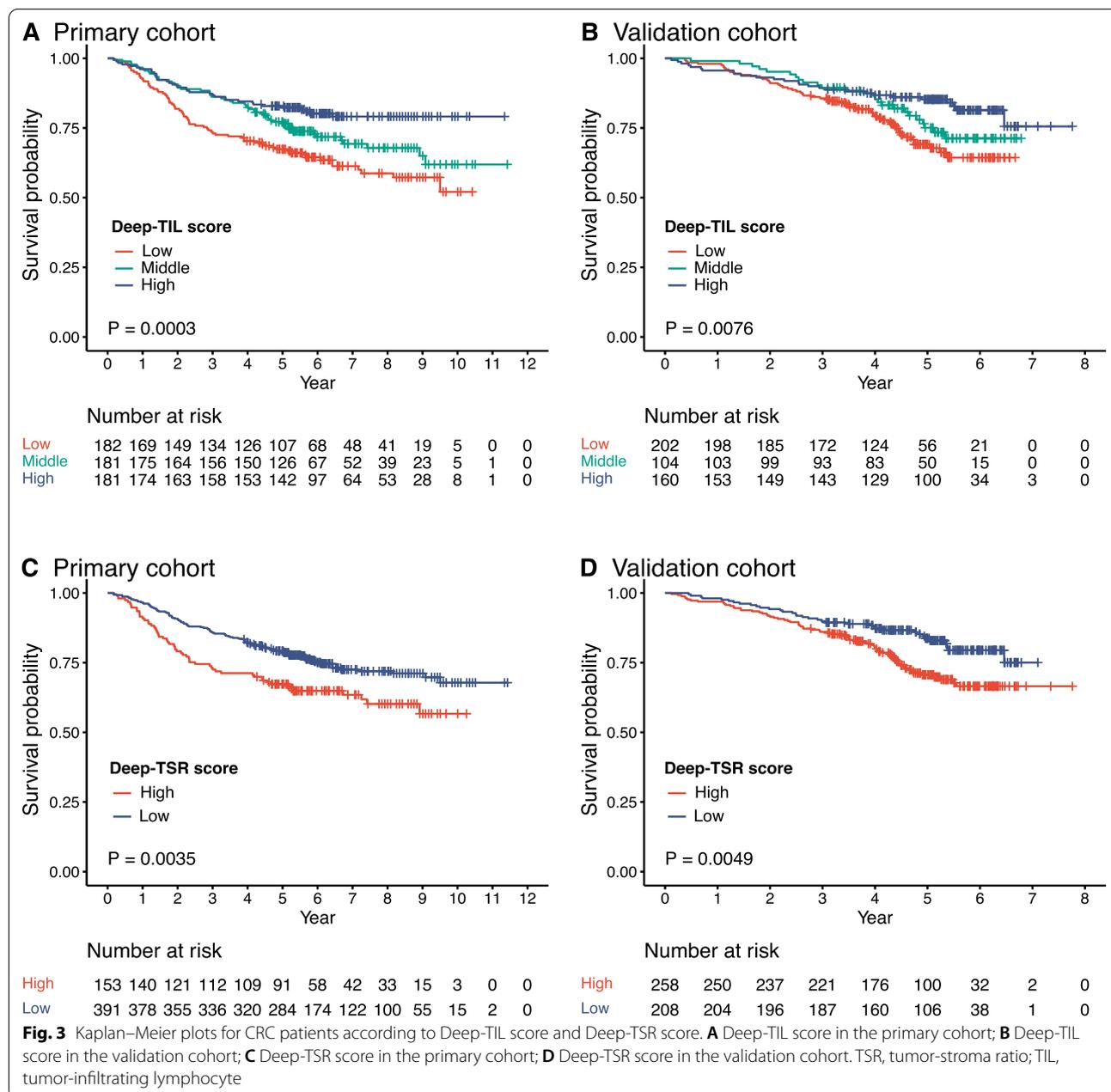


Table 1 Uni- and multivariate analyses including TNM, sex, age, location, CEA, grade, Deep-TSR score, Deep-TIL score, and Deep-immune score for OS in the two cohorts

	Univariate analysis				Multivariate analysis			
	Primary cohort		Validation cohort		Primary cohort		Validation cohort	
	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
<i>TNM</i>								
I	1		1		1		1	
II	2.48 (1.12–5.50)	0.025	1.53 (0.54–4.34)	0.400	2.56 (1.01–6.51)	0.048	1.14 (0.39–3.36)	0.800
III	6.17 (2.87–13.3)	<0.001	3.94 (1.44–10.8)	0.008	5.78 (2.33–14.4)	<0.001	2.70 (0.94–7.79)	0.066
<i>Sex</i>								
Male	1		1					
Female	1.05 (0.71–1.56)	0.800	0.78 (0.33–1.84)	0.600				
Age	1.03 (1.01–1.04)	<0.001	1.03 (1.01–1.04)	0.001	1.03 (1.01–1.04)	<0.001	1.03 (1.01–1.04)	<0.001
<i>Location</i>								
Colon	1		1					
Rectum	1.00 (0.73–1.37)	0.999	1.48 (1.00–2.21)	0.053				
<i>CEA</i>								
Normal	1		1		1		1	
Abnormal	2.58 (1.87–3.56)	<0.001	1.98 (1.35–2.91)	<0.001	1.94 (1.40–2.70)	<0.001	1.46 (0.46–1.18)	0.200
<i>Grade</i>								
Low	1		1					
High	1.44 (0.91–2.28)	0.120	1.80 (1.22–2.67)	0.003				
<i>Deep-TSR score</i>								
High	1		1					
Low	0.62 (0.45–0.86)	0.004	0.57 (0.38–0.85)	0.005				
<i>Deep-TIL score</i>								
Low	1		1					
Middle	0.69 (0.48–0.99)	0.044	0.73 (0.46–1.17)	0.200				
High	0.45 (0.30–0.67)	<0.001	0.49 (0.31–0.77)	0.002				
<i>Deep-immune score</i>								
1	1		1		1		1	
2	0.68 (0.43–1.08)	0.100	0.72 (0.45–1.14)	0.200	0.67 (0.42–1.10)	0.110	0.73 (0.46–1.18)	0.200
3	0.48 (0.30–0.77)	0.002	0.48 (0.29–0.79)	0.004	0.54 (0.33–0.90)	0.019	0.64 (0.38–1.07)	0.091
4	0.27 (0.15–0.48)	<0.001	0.31 (0.15–0.62)	<0.001	0.36 (0.20–0.66)	0.001	0.41 (0.20–0.84)	0.016

TNM, tumor-node-metastasis; CEA, carcinoembryonic antigen; TSR, tumor-stroma ratio; TIL, tumor-infiltrating lymphocytes; OS, overall survival; HR, Hazard ratio; CI, confidence interval

Prognostic value of Deep-immune score

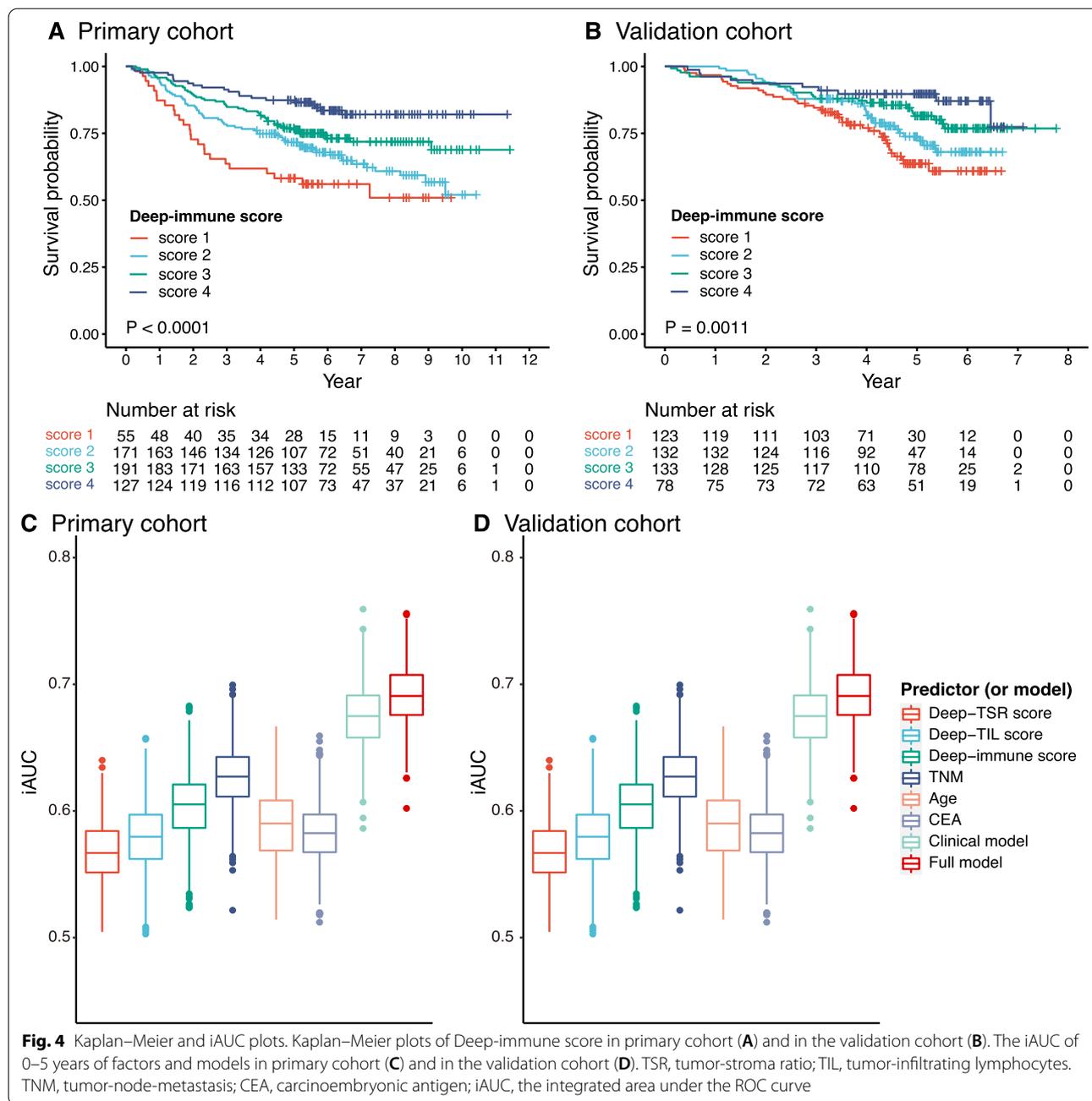
The Deep-immune score was synthesized by the Deep-TSR score and the Deep-TIL score. In the primary cohort, a significant difference was observed in 5-year survival rates between score 4 and 1 groups: 87.4% vs. 58.2% (P < 0.001; Fig. 4A). Patients with the highest score had the most favorable OS (unadjusted HR for score 4 vs. score 1 0.27, 95% CI 0.15–0.48; P < 0.001; Table 1).

Similar trends were observed in the validation cohort. Compared with score 1, patients with score 4 of Deep-immune score had favorable outcomes (unadjusted HR for score 4 vs. score 1 0.31, 95% CI 0.15–0.62, P < 0.001; Table 1). The 5-year survival rates of score 4 and score 1

groups were 89.8% and 67.0%, respectively (Fig. 4B). The results showed that a combined analysis of two scores proved to be more accurate at predicting outcomes.

Survival analysis of scores stratified with stage

In addition, when we examined whether Deep-TSR score could be applied to subgroups of patients with stage, based on stratification by stage I and II, the score had no statistically significant impact on the prognosis. (All P > 0.05; Additional file 1: Fig. S3A, C, E). As the Deep-TIL score, only a marginally statistically significance was found among stage II individuals (P = 0.051; Additional file 1: Fig. S3F). However, the Deep-immune score



showed good prognostic ability in all stage I–III subgroups (All $P < 0.05$; Additional file 1: Fig. S4), especially in the stage II group ($P = 0.018$).

The added value of the Deep-immune score

In Table 1, the univariate association between clinicopathological characteristics and OS is presented. We identified age, stage, CEA level, Deep-TSR score, Deep-TIL score, and Deep-immune score as prognostic factors for OS ($P < 0.05$). In multivariate analysis, Deep-immune

score was still associated with OS, independent of age, CEA level, and stage. There was an association between a lower Deep-immune score and a shorter OS in the primary (adjusted HR for score 4 vs. score 1 0.36, 95% CI 0.20–0.66, $P = 0.001$) and validation cohorts (0.41, 0.20–0.84, 0.016; Table 1).

Two Cox models were developed in order to evaluate the added prognostic value of the Deep immune score: the full model incorporated the independent predictors (stage, age, CEA level, and Deep-immune score) in

multivariate analysis of the primary cohort, and a clinical model incorporated clinical factors. The model with the Deep-immune score had a better discrimination performance than the clinical model (C-index: 0.732 vs. 0.720; iAUC: 0.726 vs. 0.713; Fig. 4C). The results of the validation cohort were similar as well (0.701 vs. 0.689; 0.691 vs. 0.676; Additional file 1: Table S4, Fig. 4D).

Discussion

The AI-based method could quantify the tissue composition in TME with HE-stained WSIs of CRC. To understand the basis of tumor heterogeneous clinical behavior, many scholars have focused on TME [22, 23]. Studies have shown that TME characterization provides additional insight into the prognosis of patients with solid tumors [24, 25]. The stroma of TME was the focus of studying the prognosis of CRC. Our previous analysis and studies by other scholars have shown that in CRC patients, abundant stroma in tumor tissue was associated with poor prognosis [16, 26]. Results of our present work also suggested that stroma proportion quantified in HE-stained WSIs can help stratify risk of CRC patients. Patients who have Deep-TSR-high scores have a much lower 5-year survival rate than those who have low scores (67.3% vs. 78.9%). Remodeling of stroma can serve as a physical barrier to prevent tumor cells from coming into contact with immune cells [27–29]. The mean stroma-CD3 density in the Deep-TSR-low group was 1350 cells/mm², higher than that in the Deep-TSR-high group (1011 cells/mm²). Additionally, the stroma contains special connective tissues such as fibroblasts, mesenchymal stromal cells, osteoblasts and chondrocytes, along with extracellular matrix. The endothelial cells within it provide nourishment for tumor growth, constitute a pathway for metastatic spread through angiogenesis, and lead to resistance to chemotherapy and radiation therapy [23, 30, 31]. Therefore, the more stromal components, the lower the OS of patients.

Furthermore, the stroma was incredibly intricate. For example, desmoplastic reaction was classified as immature, intermediate, or mature according to the different connective tissue-promoting reactions in the stroma [32]. Moreover, according to cancer-immune phenotypes, anticancer immunity in humans can be categorized into three main types: the immune-desert, the immune-excluded, and the inflamed phenotypes [33]. Studies have shown that the content and density of TILs in the stroma were also attached to OS [15, 34]. A patch-level segmentation was performed in our work without dissecting each lymphocyte with precision, which cannot accurately quantify the density and spatial location of lymphocytes. However, we noted that the class of lymphocytes, one of the tissue categories in our model, were structured

made up of clusters of TILs. Therefore, we tried to take the result of stroma segmentation as ROI and defined the mean predictive probability of the ROI for this category of LYM as the Deep-TIL score. Besides, we found that the mean stroma-CD3 density of the high-score group was 1.5 times higher than the low group. The automatic quantification of the Deep-TIL score could reflect the immune cells infiltration in the stroma region. Survival analysis showed that the Deep-TIL score could stratify the prognosis of CRC. The higher the Deep-TIL score, the longer the survival time. The 5-year survival rate was recorded for 70.2% of patients with a low score, 75.7% of patients with a middle score, and 85.4% of patients with a high score. This scoring method was kind to use, and this method only needed the label of the patch, which was less computationally intensive. More worth mentioning was that it also took into account the spatial distribution of TILs in the stroma.

Tumor growth pattern, aggressiveness, metastasis, and patient prognosis are the result of a combination of multiple factors. These include the interaction between components of TME from the cellular level to the tissue level [35, 36]. Based on this, after completing the above two scores, the Deep-TSR score, and Deep-TIL score, we raised the conjecture whether the combination of the two scores could reflect more prognostic information. Patients with the highest score had the most favorable OS (unadjusted HR for score 4 vs. score 1: 0.27). Similar results were also found in the validation cohort, which revealed that our score was robust. Furthermore, we found that the full model, including Deep-immune score and clinicopathological factors, had a higher prognostic value than a clinicopathological model (iAUC, 0.726 vs. 0.713). Combined with clinicopathological factors, the prognosis of patients with CRC could be evaluated in a more comprehensive and integrated manner. We also observed that with the increase in the Deep-immune score, the stroma-CD3 density also increased. CD3⁺T cells are membrane markers of mature T lymphocytes that can be used to quantify the total number of T lymphocytes [19]. When CD3⁺ cells was increased, it represented a higher tumor lymphatic infiltration and a higher amount of tumor-killing immune cells, which has a protective effect on the organism [37, 38]. This result supports the idea that our proposed Deep-immune score may be sufficient to predict prognosis of CRC. In addition, the Deep-immune score, which was fully automated and with HE-stained as a routine staining method and IHC-stained as a special staining, has certain economic benefits. There have been many studies suggesting that TSR and TILs could predict prognosis of other solid tumors. Take breast cancer as an example. Studies on TSR found a significant association between high tumor

stroma content and poor prognosis [39, 40]. The results of related studies on TILs showed that increased TILs concentrations were associated with increased frequency of adjuvant chemotherapy responses in all breast cancer subtypes, and they were also associated with longer survival in patients with triple-negative breast cancer and HER2-positive breast cancer [41, 42]. The results were similar to those of our study in CRC. This suggested that if we had segmented breast cancer tissues and defined the corresponding types of tissues, using our method to calculate tumor-stroma ratio, Deep-TIL score, and Deep-immune score, could also predict prognosis for breast cancer patients. Therefore, our method has the potential ability to be applied to other solid tumors.

In stage II patients, neither Deep-TSR nor Deep-TIL score can distinguish between high-risk and low-risk CRC individuals (all $P > 0.05$). However, the composite score can stratify patients' prognostic risk ($P = 0.018$). This indicated that Deep-immune score had the potential to guide clinical risk stratification of patients with stage II CRC, which in turn could influence clinical decision-making.

Conclusions

In conclusion, we proposed a Deep-immune score that can fully quantify the HE-stained WSI of CRC by artificial intelligence. Evidence suggests that a Deep-immune score may reflect the immune state of CRC patients and be associated with better survival. Findings based on digital pathology could be particularly useful for adjusting risk stratification of CRC and impacting subsequent precision medicine.

Abbreviations

TME: Tumor microenvironment; CRC: Colorectal cancer; TNM: Tumor-node-metastasis; TIME: Tumor immune microenvironment; TILs: Tumor-infiltrating lymphocytes; IHC: Immunohistochemistry; HE: Hematoxylin and eosin; WSIs: Whole-slide images; CLR: Crohn-like lymphoid reaction; TSR: Tumor-stroma ratio; AI: Artificial intelligence; CEA: Carcinoembryonic antigen; OS: Overall survival; CNN: Convolutional neural network; ROI: Region of interest; HR: Hazard ratio; CI: Confidence intervals; iAUC: Integrated area under the ROC curve; IQR: Interquartile range.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-022-03666-3>.

Additional file 1.

Acknowledgements

This work was supported by Key-Area Research and Development Program of Guangdong Province (2021B0101420006), the National Science Fund for Distinguished Young Scholars (81925023), the National Key R&D Program of China (2021YFF1201003), Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application (2022B1212010011),

NSFC Incubation Project of Guangdong Provincial People's Hospital (KY0120220037), and High-level Hospital Construction Project (DFJH201914 and DFJHBF202105).

Author contributions

Study design: JY and ZYL. Performed the research and collected data: XJF, LW, ZHL, and SY. Analyzed the data: KZ and YJL. Manuscript drafting: HFY and KZ. Provided discussion, critical feedback and manuscript editing: JY, XMW, MNZ, QRH, YRY, XLZ, CHL, YYW, YX, QL, DYY, KZ, and ZYL. All authors read and approved the final manuscript.

Funding

Funding sources can be found in the acknowledgements section.

Data sharing

The data sets for training the model were available online (<https://doi.org/10.5281/zenodo.4024676>, <https://doi.org/10.5281/zenodo.4023999>).

Declarations

Ethics approval and consent to participate

This study was approved by the Research Ethics Committees of Guangdong Provincial People's Hospital (GDREC2020011H), the Institutional Review Board of Yunnan Cancer Hospital (KY201824), and the Institutional Review Board of the Sixth Affiliated Hospital of Sun Yat-sen University (2019ZSLYEC-169), with the need for informed consent waived for this retrospective study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Guangdong Provincial Key Laboratory of Artificial Intelligence in Medical Image Analysis and Application, Guangdong Provincial People's Hospital, Guangzhou, China. ²Department of Cardiology, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, China. ³Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, 106 Zhongshan Er Road, Guangzhou 510080, China. ⁴The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China. ⁵Department of Pathology, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China. ⁶Department of Radiology, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China. ⁷Department of Pathology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China. ⁸Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China. ⁹Department of Ophthalmology, Guangdong Eye Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China. ¹⁰Department of Radiology, Zhuhai People's Hospital, Zhuhai Hospital Affiliated With Jinan University, Zhuhai, China. ¹¹School of Medicine, South China University of Technology, Guangzhou, China. ¹²Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, 106 Zhongshan Er Road, Guangzhou 510080, China. ¹³School of Public Health, Kunming Medical University, 191 West Renmin Road, Kunming 650500, China. ¹⁴Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China.

Received: 3 June 2022 Accepted: 25 September 2022

Published online: 04 October 2022

References

- Schürch CM, Bhate SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell*. 2020;182(5):1341-59.e19.

2. Gunter MJ, Alhounou S, Arnold M, Brenner H, Burn J, Casey G, et al. Meeting report from the joint IARC-NCI international cancer seminar series: a focus on colorectal cancer. *Ann Oncol*. 2019;30(4):510–9.
3. Suzuki K, Kadota K, Sima CS, Nitadori J, Rusch VW, Travis WD, et al. Clinical impact of immune microenvironment in stage I lung adenocarcinoma: tumor interleukin-12 receptor $\beta 2$ (IL-12R $\beta 2$), IL-7R, and stromal FoxP3/CD3 ratio are independent predictors of recurrence. *J Clin Oncol*. 2013;31(4):490–8.
4. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet (London, England)*. 2018;391(10135):2128–39.
5. Corró C, Dutoit V, Koessler T. Emerging Trends for Radio-Immunotherapy in Rectal Cancer. *Cancers*. 2021;13:6.
6. Trabelsi M, Farah F, Zouari B, Jaafoura MH, Kharrat M. An Immunoscore System Based On CD3(+) And CD8(+) infiltrating lymphocytes densities to predict the outcome of patients with colorectal adenocarcinoma. *Onco Targets Ther*. 2019;12:8663–73.
7. Dorantes-Heredia R, Ruiz-Morales JM, Cano-García F. Histopathological transformation to small-cell lung carcinoma in non-small cell lung carcinoma tumors. *Translational lung cancer research*. 2016;5(4):401–12.
8. Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. *J Intern Med*. 2020;288(1):62–81.
9. Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. Automated Tubule Nuclei Quantification and Correlation with Oncotype DX risk categories in ER+ Breast Cancer Whole Slide Images. *Sci Rep*. 2016;6:32706.
10. Mercan E, Mehta S, Bartlett J, Shapiro LG, Weaver DL, Elmore JG. Assessment of machine learning of breast pathology structures for automated differentiation of breast cancer and high-risk proliferative lesions. *JAMA Netw Open*. 2019;2(8):e198777.
11. Nagpal K, Foote D, Liu Y, Chen PC, Wulczyn E, Tan F, et al. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *NPJ digital medicine*. 2019;2:48.
12. Bulten W, Balkenhol M, Belinga JA, Brilhante A, Çakır A, Egevad L, et al. Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Modern Pathol*. 2021;34(3):660–71.
13. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559–67.
14. Geessink OGF, Baidoshvili A, Klaase JM, Ehteshami Bejnordi B, Litjens GJS, van Pelt GW, et al. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. *Cell Oncol (Dordr)*. 2019;42(3):331–41.
15. Yoo SY, Park HE, Kim JH, Wen X, Jeong S, Cho NY, et al. Whole-Slide Image Analysis Reveals Quantitative Landscape of Tumor-Immune Microenvironment in Colorectal Cancers. *Clin Cancer Res*. 2020;26(4):870–81.
16. Zhao K, Li Z, Yao S, Wang Y, Wu X, Xu Z, et al. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer. *EBioMedicine*. 2020;61: 103054.
17. Zhao M, Yao S, Li Z, Wu L, Xu Z, Pan X, et al. The Crohn's-like lymphoid reaction density: a new artificial intelligence quantified prognostic immune index in colon cancer. *Cancer Immunol Immunother*. 2022;71(5):1221–31.
18. Xie Y, Xie F, Zhang L, Zhou X, Huang J, Wang F, et al. Targeted Anti-Tumor Immunotherapy Using Tumor Infiltrating Cells. *Advan Sci*. 2021;8(22):e2101672.
19. Xu Z, Li Y, Wang Y, Zhang S, Huang Y, Yao S, et al. A deep learning quantified stroma-immune score to predict survival of patients with stage II-III colorectal cancer. *Cancer Cell Int*. 2021;21(1):585.
20. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17(6):1471–4.
21. Zhao K, Wu X, Li Z, Wang Y, Xu Z, Li Y, et al. Prognostic value of a modified Immunoscore in patients with stage I-III resectable colon cancer. *Chinese journal of cancer research = Chung-kuo yen cheng yen chiu*. 2021;33(3):379–90.
22. Maman S, Witz IP. A history of exploring cancer in context. *Nat Rev Cancer*. 2018;18(6):359–76.
23. Khalaf K, Hana D, Chou JT, Singh C, Mackiewicz A, Kaczmarek M. Aspects of the Tumor Microenvironment Involved in Immune Resistance and Drug Resistance. *Front Immunol*. 2021;12: 656364.
24. Giraldo NA, Sanchez-Salas R, Peske JD, Vano Y, Becht E, Petitprez F, et al. The clinical role of the TME in solid cancer. *Br J Cancer*. 2019;120(1):45–53.
25. Belli C, Trapani D, Viale G, D'Amico P, Duso BA, Della Vigna P, et al. Targeting the microenvironment in solid tumors. *Cancer Treat Rev*. 2018;65:22–32.
26. Wu J, Liang C, Chen M, Su W. Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis. *Oncotarget*. 2016;7(42):68954–65.
27. Sangaletti S, Chiodoni C, Tripodo C, Colombo MP. The good and bad of targeting cancer-associated extracellular matrix. *Curr Opin Pharmacol*. 2017;35:75–82.
28. Kaushik N, Kim S, Suh Y, Lee SJ. Proinvasive extracellular matrix remodeling for tumor progression. *Arch Pharmacol Res*. 2019;42(1):40–7.
29. Eble JA, Niland S. The extracellular matrix in tumor progression and metastasis. *Clin Exp Metas*. 2019;36(3):171–98.
30. Valkenburg KC, de Groot AE, Pienta KJ. Targeting the tumour stroma to improve cancer therapy. *Nat Rev Clin Oncol*. 2018;15(6):366–81.
31. Hida K, Akiyama K, Ohga N, Maishi N, Hida Y. Tumour endothelial cells acquire drug resistance in a tumour microenvironment. *J Biochem*. 2013;153(3):243–9.
32. Ueno H, Kanemitsu Y, Sekine S, Ishiguro M, Ito E, Hashiguchi Y, et al. Desmoplastic Pattern at the Tumor Front Defines Poor-prognosis Subtypes of Colorectal Cancer. *Am J Surg Pathol*. 2017;41(11):1506–12.
33. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017;541(7637):321–30.
34. Idos GE, Kwok J, Bonthala N, Kysch L, Gruber SB, Qu C. The Prognostic Implications of Tumor Infiltrating Lymphocytes in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Sci Rep*. 2020;10(1):3360.
35. Li C, Li S, Du K, Li P, Qiu B, Ding W. On-Chip Replication of Extremely Early-Stage Tumor Behavior. *ACS Appl Mater Interfaces*. 2021;13(17):19768–77.
36. Qi D, Wu E. Cancer prognosis: Considering tumor and its microenvironment as a whole. *EBioMedicine*. 2019;43:28–9.
37. Mukherji B, Chakraborty NG, Sivanandham M. T-cell clones that react against autologous human tumors. *Immunol Rev*. 1990;116:33–62.
38. Zhou S, Meng F, Du S, Qian H, Ding N, Sha H, et al. Bifunctional iRGD-anti-CD3 enhances antitumor potency of T cells by facilitating tumor infiltration and T-cell activation. *Journal for immunotherapy of cancer*. 2021;9(5).
39. Kramer CJH, Vangangelt KMH, van Pelt GW, Dekker TJA, Tollenaar R, Mesker WE. The prognostic value of tumour-stroma ratio in primary breast cancer with special attention to triple-negative tumours: a review. *Breast Cancer Res Treat*. 2019;173(1):55–64.
40. Dekker TJ, van de Velde CJ, van Pelt GW, Kroep JR, Julien JP, Smit VT, et al. Prognostic significance of the tumor-stroma ratio: validation study in node-negative premenopausal breast cancer patients from the EORTC perioperative chemotherapy (POP) trial (10854). *Breast Cancer Res Treat*. 2013;139(2):371–9.
41. Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol*. 2018;19(1):40–50.
42. Afghahi A, Purington N, Han SS, Desai M, Pierson E, Mathur MB, et al. Higher Absolute Lymphocyte Counts Predict Lower Mortality from Early-Stage Triple-Negative Breast Cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2018;24(12):2851–8.

Publisher's Note

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