Cancer Immunology, Immunotherapy (2026) 75:163
https://doi.org/10.1007/s00262-026-04409-9

RESEARCH

Artificial intelligence-powered H&E-based quantification of spatial tumor-infiltrating lymphocyte distribution identifies prognostic immune niches in colorectal cancer

Hyun-Hee Koh[1] · Seungeun Lee[2] · Chiyoon Oum[2] · Sanghoon Song[2] · Soo Ick Cho[2] · Sérgio Pereira[2] · Chang Ho Ahn[2] · Jun Yong Kim[1] · Milim Kim[1] · Minsun Jung[1]

Received: 9 March 2026 / Accepted: 28 April 2026
© The Author(s) 2026

Abstract

Purpose The prognostic significance of tumor-infiltrating lymphocytes (TILs) in colorectal cancer (CRC) is well established; however, existing approaches inadequately capture their spatial distribution. We investigated the prognostic implications of TIL spatial distribution in CRC using an artificial intelligence (AI)-based method.

Methods A total of 202 patients with stage II–III CRC were included. TIL densities in intratumoral (iTIL) and stromal (sTIL) regions were quantified using AI-based analysis of hematoxylin and eosin (H&E)–stained images. Based on proximity to the tumor–stromal border (TSB), TILs were subclassified into core iTIL, bounding iTIL, bounding sTIL, and outermost sTIL. Immunoscore was calculated from $CD3^+$ and $CD8^+$ T-cell densities in the tumor center and invasive margin.

Results Correlations between AI-based and pathologist assessments (iTIL: $r = 0.57$; sTIL: $r = 0.70$) were comparable to inter-pathologist correlations (iTIL: $r = 0.47$; sTIL: $r = 0.70$). In univariate Cox regression analysis, bounding iTIL, bounding sTIL, and outermost sTIL were significantly associated with recurrence-free survival (RFS), whereas core iTIL was not. Composite TIL and TSB scores were developed by incorporating the prognostically significant regions. In multivariable analysis, the TIL score ($p = 0.001$), TSB score ($p < 0.001$), and Immunoscore ($p < 0.001$) independently predicted RFS. In microsatellite instability–high tumors, only the TSB score remained prognostically significant.

Conclusion AI-powered spatial analysis of TILs, particularly the TSB score, demonstrated prognostic performance comparable to conventional Immunoscore, thereby supporting the value of spatial immune profiling and AI-driven analysis of H&E-stained slides for improved risk stratification in CRC.

Keywords Artificial intelligence · Colorectal cancer · Spatial analysis · Tumor-infiltrating lymphocyte · Tumor-stromal border

Abbreviations

AI	Artificial intelligence
CA	Cancer area
CI	Confidence interval
CRC	Colorectal cancer
CS	Cancer stroma
H&E	Hematoxylin and eosin
HR	Hazard ratio
IHC	Immunohistochemistry
IM	Invasive margin
iTIL	Intratumoral tumor-infiltrating lymphocyte
MSI	Microsatellite instability
MSS	Microsatellite stable
PCC	Pearson correlation coefficient
RFS	Recurrence-free survival
ROI	Region of interest
sTIL	Stromal tumor-infiltrating lymphocyte
TC	Tumor center
TIL	Tumor-infiltrating lymphocyte
TME	Tumor immune microenvironment
TSB	Tumor-stromal border
WSI	Whole-slide image

Hyun-Hee Koh and Seungeun Lee have contributed equally to this work.

✉ Minsun Jung
jjunglammy@yuhs.ac

[1] Department of Pathology, Yonsei University College of Medicine, 50-1 Yonsei-Ro, Seodaemun-Gu, Seoul 03722, Republic of Korea

[2] Lunit Inc, Seoul, Republic of Korea

Published online: 04 May 2026

Introduction

In colorectal cancer (CRC), which ranks as the third most commonly diagnosed and second leading cause of cancer-related mortality worldwide, the tumor immune microenvironment (TME) plays a pivotal role in shaping antitumor immune responses, determining responsiveness to immunotherapy, and influencing patient prognosis [1–4]. Among immune-related features, the density of tumor-infiltrating lymphocytes (TILs) is a well-established biomarker of immune activity and clinical outcome in CRC. Beyond TIL density, the spatial distribution of T cells within the TME has emerged as an important biological and prognostic determinant [5–7]. As shown in our previous study [7], the spatial distribution of TILs adds important layers of information to the interpretation of the tumor–immune microenvironment in immuno-oncology. A common approach for spatial decomposition in CRC combines cancer cell islands and intervening stroma into a single tumor region, which is then separated from the peritumoral area. The Immunoscore is a well-established example in which immunohistochemistry (IHC) targeting CD3 and CD8 is used to separately quantify T cells in the invasive margin (IM), the region surrounding the forefront of the tumor, and the tumor center (TC). In CRC, the IM and TC do not distinguish between tumor epithelial and stromal cells, which may limit spatial resolution. An alternative TIL topology that separates intraepithelial TILs (iTILs) and stromal TILs (sTILs) has been proposed and shown to be prognostically significant in CRC [4, 5, 8, 9]. In this framework, the tumor–stromal border (TSB), which is a distinct feature of the TME associated with immunotherapy response in CRC, serves as a critical reference for evaluating spatial immune heterogeneity [10].

Recent advances in digital pathology and deep learning have transformed histological analysis, addressing limitations of conventional TIL assessments by pathologists, which are time-consuming, require substantial expertise, and are subject to substantial inter- and intra-observer variability [4, 9, 11]. While AI-powered methods provide faster analysis and higher spatial resolution, their clinical adoption requires rigorous validation and comparative studies with expert pathologists.

In this study, we quantified TILs using an artificial intelligence (AI)–powered analyzer and evaluated their prognostic significance in CRC by examining their density and high-resolution spatial distribution, with particular focus on the TSB. We developed composite TSB scores by integrating clinically meaningful spatial zones and demonstrated that these hematoxylin and eosin (H&E)–based scores provide prognostic stratification comparable to that of the Immunoscore.

Materials and methods

Patients

This study included 202 treatment-naïve patients with stage II or III CRC who underwent curative-intent resection at Severance Hospital (Seoul, South Korea) between January 2012 and June 2019. Microsatellite instability (MSI)–high tumors were intentionally enriched to better characterize this immunologically relevant subtype. Detailed sample information has been described previously [12]. While the previous study focused on cancer-associated fibroblasts and the tumor–stroma ratio, our study specifically investigated TILs. Clinical data, including recurrence-free survival (RFS), were retrospectively obtained from electronic medical records. RFS was defined as the interval from surgery to the first documented disease recurrence or the last follow-up. MSI status was determined using five quasi-mononucleotide repeat microsatellite loci (NR-21, NR-24, NR-27, BAT-25, and BAT-26) via peptide nucleic acid–based real-time polymerase chain reaction (U-TOP™ MSI Detection Kit, SeaSun Biomaterials, Daejeon, Republic of Korea) or by mismatch repair IHC markers, including MLH1 (760-5091, Ventana, AZ, US), MSH2 (760-5093, Ventana), PMS2 (288 M-16, Roche, Basel, Switzerland), and MSH6 (790-4455, Ventana) [13]. All procedures were conducted in accordance with the Declaration of Helsinki, and the Institutional Review Board of the Severance Hospital approved the study protocols (accession no. 4-2022-0900; approval date: August 29, 2022). Informed consent was waived due to the retrospective design, use of anonymized data, and minimal risk to participants.

AI-based analysis of TILs along the TSB in H&E-stained whole-slide images

Representative H&E-stained slides showing the deepest tumor invasion were selected for AI-powered analysis of whole-slide images (WSIs), which were scanned using an Aperio GT450 (Leica Biosystems, IL, US). Spatial quantification of TILs was performed using Lunit SCOPE IO, an enhanced version of a previously published algorithm [14]. Lunit SCOPE IO comprises two complementary deep learning models: (1) a tissue segmentation model that classifies tissue regions into cancer area (CA) and cancer stroma (CS), and (2) a cell detection model that identifies and classifies individual cells into tumor cells or lymphocytes, both utilizing ResNet-34 as the backbone architecture, with the cell detection model employing a DeepLabV3 + -based architecture. The tissue segmentation model was developed using 76,110 image patches from 18,935 WSIs (15,936 for training and 2999 for optimization), while the cell detection

model was constructed using 20,617 image patches from 5609 WSIs (3798 for training and 1811 for optimization). Model development involved annotation of 2,828,448 cells and 1.70×10^{10} μm^2 of tissue across more than 26 tumor types by board-certified pathologists [15]. By integrating these models, the system measured the lymphocyte densities within the CA (iTILs) and CS (sTILs) regions (Supplementary Fig. 1A–B).

The TSB was defined as the interface between the CA and CS regions. To capture TIL spatial distribution relative to the TSB, iTILs and sTILs were further subclassified based on their distance to the TSB: core iTILs are iTILs located more than 30 μm from the TSB; bounding iTILs are iTILs located within 0–30 μm from the TSB; bounding sTILs are sTILs located within 0–30 μm from the TSB; and outermost sTILs are sTILs located more than 30 μm from the TSB (Supplementary Fig. 1C–D). Within a 30 μm distance, biologically relevant short-range paracrine and mechanical cell–cell interactions can occur in the immune microenvironment, as is widely accepted [16, 17]. Supplementary analyses at boundary distances of 30, 50, 100, and 150 μm were conducted to evaluate the robustness of this threshold selection (Supplementary Table 2).

Development of AI-based TIL and TSB scores

To develop clinically applicable AI-based scoring systems, we created TIL and TSB composite scores analogous to the Immunoscore methodology [18]. Briefly, specific compartments that demonstrated prognostic significance in univariate Cox regression analysis were selected for inclusion in the composite scores. TIL density (cells/mm^2) was calculated independently for each included compartment, and each value was converted to a cohort-wide percentile rank (0–100%). The TIL score was computed as the mean of two percentile ranks (iTIL and sTIL), and the TSB score as the mean of three percentile ranks (bounding iTIL, bounding sTIL, and outermost sTIL). This mean percentile was then re-ranked cohort-wide to generate the final continuous score used in prognostic analyses.

Immunoscore calculation

Whole-slide IHC staining for CD3 (polyclonal, DAKO, Glostrup, Denmark) and CD8 (C8/144B; Thermo Fisher Scientific, MA, US) was performed on serial sections corresponding to the H&E slides analyzed by AI, using a Ventana XT automated stainer (Ventana). The stained slides were scanned using an Aperio GT450. QuPath (ver. 0.1.2) was used to quantify $CD3^+$ or $CD8^+$ TILs as previously described [7, 19]. Regions of interest (ROIs) for tumor and peritumoral stroma were manually delineated by experienced gastrointestinal pathologists (HHK and MJ), while excluding areas of necrosis, abscess, extracellular mucin pools, and mucosa-associated lymphoid tissue. The ROIs were subdivided into square tiles measuring 1 mm^2, and tiles at the invasive front were annotated as IM. Mean densities of $CD3^+$ and $CD8^+$ TILs were quantified separately within the IM and TC. The Immunoscore was measured based on the average of the four percentile values [18].

TIL scoring reader study

A total of 173 high-quality image grids were randomly selected for the reader study. Among these, 54 regions with CA ≥ 80% were assigned for iTIL evaluation, 60 regions with CS ≥ 80% for sTIL evaluation, and 59 regions with both CA and CS ≥ 40% were designated for both iTIL and sTIL. Each grid was independently reviewed by three board-certified pathologists (HHK, MK, and MJ) according to a predefined guideline [8]. TIL scores were assigned using a semi-quantitative scale, with 1-point increments up to a score of 5 and 5-point increments thereafter.

Statistical analysis

Pearson correlation coefficient (PCC) was used to assess associations between continuous variables, whereas Spearman's rank correlation coefficient was applied to evaluate correlations involving percentile-based scores. Paired comparisons of continuous variables were performed using the Wilcoxon signed-rank test. Univariate and multivariable Cox proportional hazards regression models were employed to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) associated with recurrence. Variables that were statistically significant in univariate analyses were included in the multivariable models. TIL densities were standardized using z-score transformation before inclusion in the regression model. Kaplan–Meier survival curves were generated and compared using the log-rank test. The predictive accuracy of each scoring system for survival was assessed using Harrell's C-index, and comparisons between C-indices were performed using a U-statistics–based method. All statistical analyses were conducted using R statistical software (version 4.2.2 or higher; R Foundation for Statistical Computing, Vienna, Austria), and statistical significance was defined as $p < 0.05$.

Results

Clinicopathological characteristics

The cohort included 108 males (53.5%) and 94 females (46.5%), with the median age of 63 years (range, 33–87 years). Tumors were classified as stage II in 106 cases

(52.5%) and stage III in 96 cases (47.5%). High-grade histology was observed in 20 tumors (9.9%), while 182 (90.1%) were low grade. Lymphovascular invasion, perineural invasion, and high-grade tumor budding were identified in 85 (42.1%), 42 (20.8%), and 77 (38.1%) tumors, respectively. Microsatellite status was microsatellite stable (MSS) in 132 tumors (65.3%) and MSI-high in 70 (34.7%). Table 1 summarizes the clinicopathological characteristics.

AI-powered evaluation of TILs showed close correlation to Immunoscore

AI-analyzed iTIL and sTIL densities using H&E-stained slides were $83.55 \pm 79.6/mm^2$ (mean ± standard deviation) and $884.71 \pm 439.87/mm^2$, respectively. Mean $CD3^+$ and $CD8^+$ TIL densities were $1009.20 \pm 576.41/mm^2$ and $650.46 \pm 626.75/mm^2$, respectively. Spatial subclassification revealed mean densities of $64.89 \pm 72.09/mm^2$ for core iTILs, $105.93 \pm 102.25/mm^2$ for bounding iTILs, $714.67 \pm 431.54/mm^2$ for bounding sTILs, and $1015.42 \pm 504.01/mm^2$ for outermost sTILs (Supplementary Fig. 2). Within the intratumoral region, core and bounding iTILs accounted for 40.43% and 59.57% of the total iTIL count on average, respectively, while occupying 53.07% and 46.93% of the total iTIL area. Similarly, within the stromal region, bounding and outermost sTILs comprised 32.80% and 67.20% of the total sTIL count on average, respectively, while occupying 41.28% and 58.72% of the total sTIL area.

To validate AI-based TIL quantification, PCCs were calculated against the conventional Immunoscore (Fig. 1). AI-evaluated iTIL density correlated significantly with $CD3^+$ or $CD8^+$ TIL densities in the TC (PCC = 0.62 and 0.67, respectively; both $p < 0.001$). Similarly, AI-evaluated sTIL density correlated with $CD3^+$ or $CD8^+$ TIL densities at the IM (PCC = 0.59 and 0.54, respectively; both $p < 0.001$). Overall, AI-evaluated iTILs (PCC = 0.52, $p < 0.001$) and sTILs (PCC = 0.59, $p < 0.001$) showed strong correlations with the Immunoscore.

Table 1 Clinicopathological characteristics of the study cohort (n = 202)

Characteristics	Value*
Age	63 (33–87)
Sex	
Male	108 (53.5%)
Female	94 (46.5%)
Size	4.5 (1.2–15.0)
Histological type	
Adenocarcinoma	186 (92.1%)
Mucinous carcinoma	15 (7.4%)
Undifferentiated carcinoma	1 (0.5%)
Stage	
II	106 (52.5%)
III	96 (47.5%)
Grade	
High (poorly differentiated or undifferentiated)	20 (9.9%)
Low (well or moderately differentiated)	182 (90.1%)
Microsatellite status	
MSI-high	70 (34.7%)
MSS	132 (65.3%)
Lymphovascular invasion	
Present	85 (42.1%)
Absent	117 (57.9%)
Perineural invasion	
Present	42 (20.8%)
Absent	160 (79.2%)
High-grade tumor budding	
Present	77 (38.1%)
Absent	125 (61.9%)

MSI, Microsatellite instability; MSS, Microsatellite stable

*Continuous variables are presented as median (range), and categorical variables are presented as n (%)

AI-powered evaluation of TILs showed close correlation to pathologists

Supplementary Table 1 summarizes the PCC between the AI model and three pathologists. For iTIL density, PCCs between the AI model and individual pathologists ranged from 0.23 to 0.80 (mean: 0.57), whereas inter-pathologist correlations ranged from 0.34 to 0.66 (mean: 0.47). For sTIL density, correlations between the AI model and pathologists ranged from 0.64 to 0.75 (mean: 0.70), while inter-pathologist correlations ranged from 0.62 to 0.79 (mean: 0.70). Collectively, correlations between the AI model and pathologists were comparable to the inter-pathologist correlations.

Prognostic value of spatially distributed TIL density evaluated by an AI model

In univariate Cox regression analysis, higher TIL density was significantly associated with a lower risk of recurrence (iTIL: HR = 0.47, 95% CI: 0.27–0.82, $p = 0.008$; sTIL: HR = 0.37, 95% CI: 0.25–0.56, $p < 0.001$). Among TSB-stratified TIL densities, bounding iTIL (HR = 0.23, 95% CI: 0.11–0.49, $p < 0.001$), bounding sTIL (HR = 0.38, 95% CI: 0.24–0.60, $p < 0.001$), and outermost sTIL (HR = 0.38, 95% CI: 0.26–0.57, $p < 0.001$) were significantly associated with reduced recurrence risk, while core iTIL showed no significant association (HR = 0.74, 95% CI: 0.49–1.11, $p = 0.145$) (Table 2). Among clinicopathological variables, the following were significantly associated with recurrence: tumor stage (stage III vs. II: HR = 2.43, 95% CI: 1.40–4.21, $p = 0.002$), MSI status (MSI-high vs. MSS: HR = 0.15, 95%

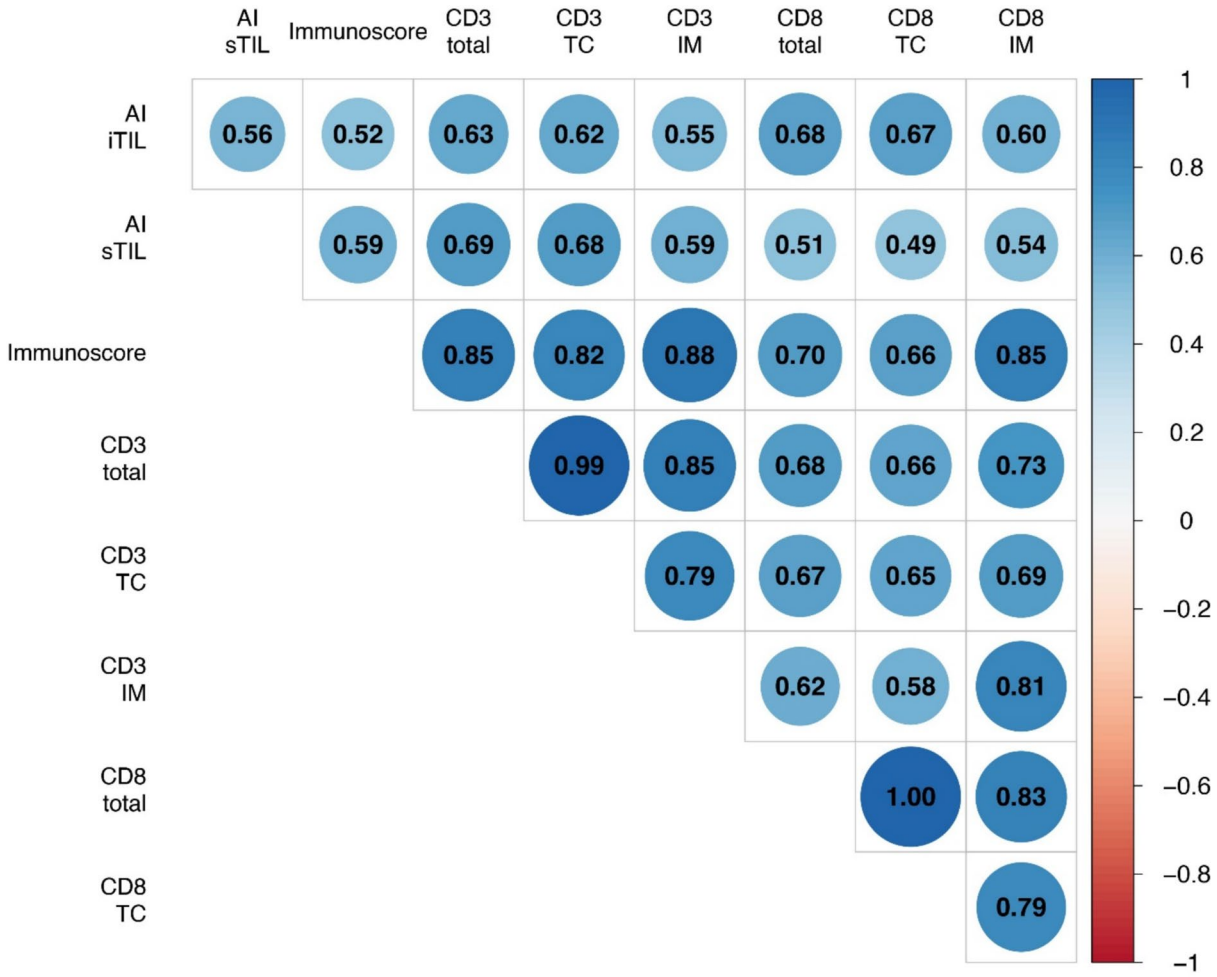

Fig. 1 Correlation matrix of AI-based TIL density, $CD3^+$ and $CD8^+$ TIL density, and Immunoscore. Color intensity and circle size denote the strength of Pearson correlation coefficients. All correlations were statistically significant ($p < 0.001$). AI, artificial intelligence; IM, invasive margin; iTIL, intratumoral tumor-infiltrating lymphocyte; sTIL, stromal tumor-infiltrating lymphocyte; TC, tumor center

CI: 0.06–0.36, $p < 0.001$), lymphovascular invasion (present vs. absent: HR = 2.33, 95% CI: 1.37–3.96, $p = 0.002$), perineural invasion (present vs. absent: HR = 2.51, 95% CI: 1.45–4.37, $p = 0.001$), and high-grade tumor budding (present vs. absent: HR = 3.65, 95% CI: 2.13–6.26, $p < 0.001$).

Correlation between AI-evaluated TIL scores and conventional immunoscore

Using the spatial compartments identified as significant in univariate analysis, we developed two composite scoring systems. The TIL score integrated iTIL and sTIL densities, while the TSB score incorporated prognostically significant TSB compartments, specifically bounding iTIL, bounding sTIL, and outermost sTIL. Core iTIL was excluded from the TSB score due to its lack of significant association with recurrence. Both the TIL score (Spearman's $\rho = 0.62$, $p < 0.001$) and TSB score ($\rho = 0.60$, $p < 0.001$) demonstrated moderate positive correlations with the conventional Immunoscore (Fig. 2).

Multivariable cox regression analysis of prognostic scoring systems

To assess the independent prognostic value of each scoring system, multivariable Cox regression analyses were performed, adjusting for tumor stage, MSI status, lymphovascular invasion, perineural invasion, and high-grade tumor budding (Table 3). The TIL score demonstrated independent prognostic significance (HR = 0.17, 95% CI: 0.06–0.50, $p = 0.001$), as did the TSB score (HR = 0.12, 95% CI: 0.04–0.36, $p < 0.001$). The conventional Immunoscore also remained an independent predictor of recurrence (HR = 0.07, 95% CI: 0.02–0.24, $p < 0.001$).

Among clinicopathological variables, when adjusted for each TIL score, high-grade tumor budding consistently

Table 2 Univariate cox regression analysis of the AI-based spatial TIL density, TIL score, TSB score, Immunoscore, and clinicopathological variables for recurrence-free survival

Variable	HR (95% CI)	*p-value*
iTIL density (continuous)	0.47 (0.27–0.82)	0.008
sTIL density (continuous)	0.37 (0.25–0.56)	< 0.001
TSB-Core iTIL density (continuous)	0.74 (0.49–1.11)	0.145
TSB-Bounding iTIL density (continuous)	0.23 (0.11–0.49)	< 0.001
TSB-Bounding sTIL density (continuous)	0.38 (0.24–0.60)	< 0.001
TSB-Outermost sTIL density (continuous)	0.38 (0.26–0.57)	< 0.001
TIL score (continuous)	0.10 (0.04–0.27)	< 0.001
TSB score (continuous)	0.07 (0.03–0.20)	< 0.001
Immunoscore (continuous)	0.04 (0.01–0.12)	< 0.001
Age (continuous)	1.01 (0.99–1.04)	0.195
Sex (male vs. female)	1.23 (0.73–2.08)	0.441
Size (continuous)	0.97 (0.86–1.09)	0.578
Stage (III vs. II)	2.43 (1.40–4.21)	0.002
Grade (high vs. low)	0.8 (0.32–2.01)	0.639
MSI-high (vs. MSS)	0.15 (0.06–0.36)	< 0.001
Lymphovascular invasion (present vs. absent)	2.33 (1.37–3.96)	0.002
Perineural invasion (present vs. absent)	2.51 (1.45–4.37)	0.001
High-grade tumor budding (present vs. absent)	3.65 (2.13–6.26)	< 0.001

AI, artificial intelligence; CI, confidence interval; HR, hazard ratio; iTIL, intratumoral tumor-infiltrating lymphocyte; MSI, microsatellite instability; MSS, microsatellite stable; sTIL, stromal tumor-infiltrating lymphocyte; TIL, tumor-infiltrating lymphocyte; TSB, tumor-stromal border

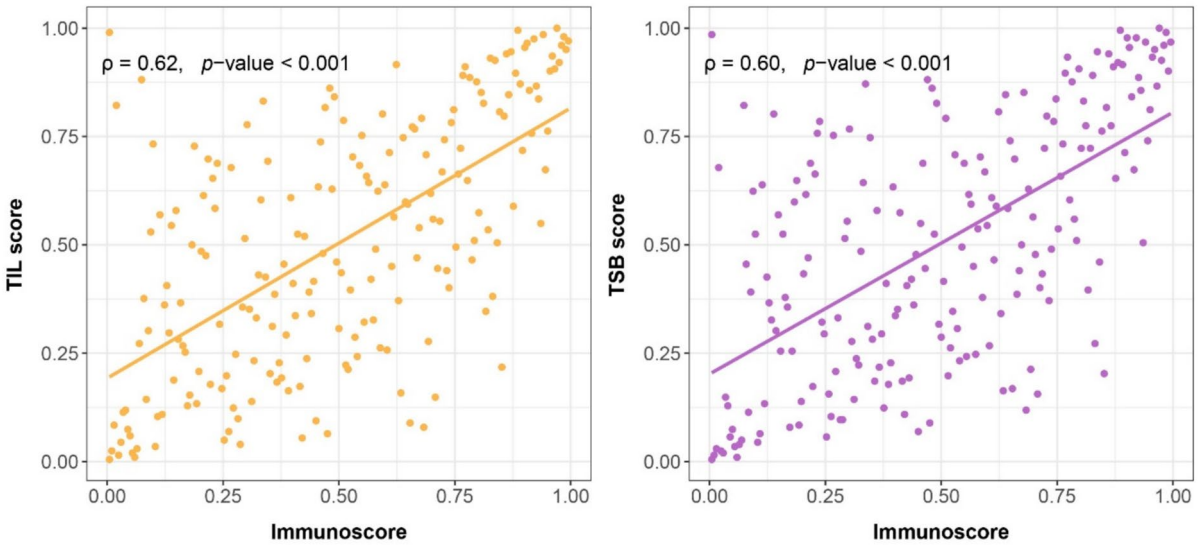

Fig. 2 Correlation between the Immunoscore and AI-based TIL or TSB scores. AI, artificial intelligence; TIL, tumor-infiltrating lymphocyte; TSB, tumor-stromal border

remained an independent adverse prognostic factor across all three TIL scores (HR range: 2.21–2.30, all $p < 0.01$). MSI-high status (vs. MSS) showed independent associations with reduced progression risk when adjusted for the TIL score (HR = 0.32, $p = 0.022$) and TSB score (HR = 0.32, $p = 0.019$), but not when adjusted for Immunoscore (HR = 0.44, $p = 0.104$). Tumor stage, lymphovascular invasion, and perineural invasion lost their independent prognostic significance in multivariable analyses across all TIL scores (all $p > 0.05$).

Table 3 Adjusted* multivariable Cox regression analysis for recurrence-free survival with TIL score, TSB score, and Immunoscore

Variable	TIL score + clinicopathological variables		TSB score + clinicopathological variables		Immunoscore + clinicopathological variables	
	HR (95% CI)	*p–value*	HR (95% CI)	*p–value*	HR (95% CI)	*p–value*
TIL score (continuous)	0.17 (0.06–0.50)	0.001	–	–	–	–
TSB score (continuous)	–	–	0.12 (0.04–0.36)	< 0.001		
Immunoscore (continuous)	–	–	–	–	0.07 (0.02–0.24)	< 0.001
Stage (III vs. II)	1.47 (0.82–2.62)	0.193	1.44 (0.81–2.58)	0.214	1.50 (0.84–2.67)	0.170
MSI–high (vs. MSS)	0.32 (0.12–0.85)	0.022	0.32 (0.12–0.83)	0.019	0.44 (0.16–1.18)	0.104
Lymphovascular invasion (present vs. absent)	1.68 (0.95–2.96)	0.075	1.72 (0.98–3.03)	0.058	1.56 (0.89–2.75)	0.123
Perineural invasion (present vs. absent)	1.46 (0.82–2.61)	0.197	1.35 (0.76–2.41)	0.310	1.33 (0.75–2.38)	0.328
High–grade tumor budding (present vs. absent)	2.26 (1.29–3.95)	0.004	2.21 (1.27–3.87)	0.005	2.30 (1.31–4.02)	0.004

CI, Confidence interval; HR, Hazard ratio; TIL, Tumor-infiltrating lymphocyte; TSB, Tumor-stromal border

*Adjusted for tumor stage, MSI status, lymphovascular invasion, perineural invasion, and high-grade tumor budding

Prognostic performance comparison between the TSB score and immunoscore

To further evaluate and compare the performance of the scoring systems, Harrell's C-indices were calculated across the three scoring methods. The TSB score showed a C-index of 0.70, which was significantly higher than that of the TIL score (0.68; $p = 0.027$) and numerically comparable to the Immunoscore (0.73; $p = 0.418$).

Using median values to define high- and low-score groups, patients with high scores had significantly longer RFS than those with low scores across all three scoring systems: TIL score (HR = 0.31, 95% CI: 0.17–0.57, $p < 0.001$), TSB score (HR = 0.20, 95% CI: 0.10–0.38, $p < 0.001$), and Immunoscore (HR = 0.20, 95% CI: 0.10–0.38, $p < 0.001$) (Fig. 3A–C). MSI status is a key immuno-oncologic biomarker, as MSI-high CRCs exhibit higher TIL densities and enhanced responsiveness to immunotherapy [20–22]. Notably, subgroup analysis by MSI status revealed differential prognostic performance among the three scoring systems. In the MSI-high subgroup, the TSB score retained significant prognostic stratification (HR = 0.10, 95% CI: 0.01–0.93, $p = 0.013$), whereas both the TIL score (HR = 0.31, 95% CI: 0.05–1.84, $p = 0.171$) and Immunoscore (HR = 0.40, 95% CI: 0.07–2.42, $p = 0.305$) did not reach statistical significance. In the MSS subgroup, all three scoring systems showed significant prognostic stratification: TIL score (HR = 0.44, 95% CI: 0.23–0.82, $p = 0.007$), TSB score (HR = 0.30, 95% CI: 0.15–0.60, $p < 0.001$), and Immunoscore (HR = 0.29, 95% CI: 0.14–0.62, $p = 0.001$) (Fig. 3D–F).

Discussion

This study aimed to improve prognostic immune profiling in CRC by leveraging AI-powered spatial analysis of iTILs and sTILs, with particular focus on their spatial distribution along the TSB. By incorporating this spatial context, we demonstrated that high-resolution spatial quantification of TILs using H&E-stained slides has comparable prognostic value to conventional Immunoscore and may offer superior prognostic stratification in MSI-high CRCs.

To validate the AI-derived TIL measurements from H&E-stained slides, we compared them with both IHC-based Immunoscore and expert pathologist assessments. AI-evaluated TIL densities showed moderate correlations with $CD3^+$ and $CD8^+$ TIL densities and the Immunoscore, indicating that the AI captures biologically relevant immune signals. Furthermore, correlations between AI and pathologist assessments were similar to, or slightly higher than, inter-pathologist correlations, especially for iTILs. Assessing iTILs is particularly challenging for pathologists due to their frequent admixture with tumor cells and subcellular structures. These findings indicate that the AI model operates within the range of expert-level interpretation and provides a reproducible, scalable alternative for immune profiling using standard H&E-stained slides.

To further examine spatial heterogeneity, TIL density was divided into four zones based on distance from the TSB: (1) core iTILs (distant from TSB), (2) bounding iTILs (near TSB), (3) bounding sTILs (near TSB), and (4) outermost sTILs (distant from TSB). All zones except core iTILs demonstrated significant prognostic value for RFS. Given its limited prognostic contribution, core iTILs were excluded from subsequent composite scoring models. This finding aligns with the well-established concept that the IM

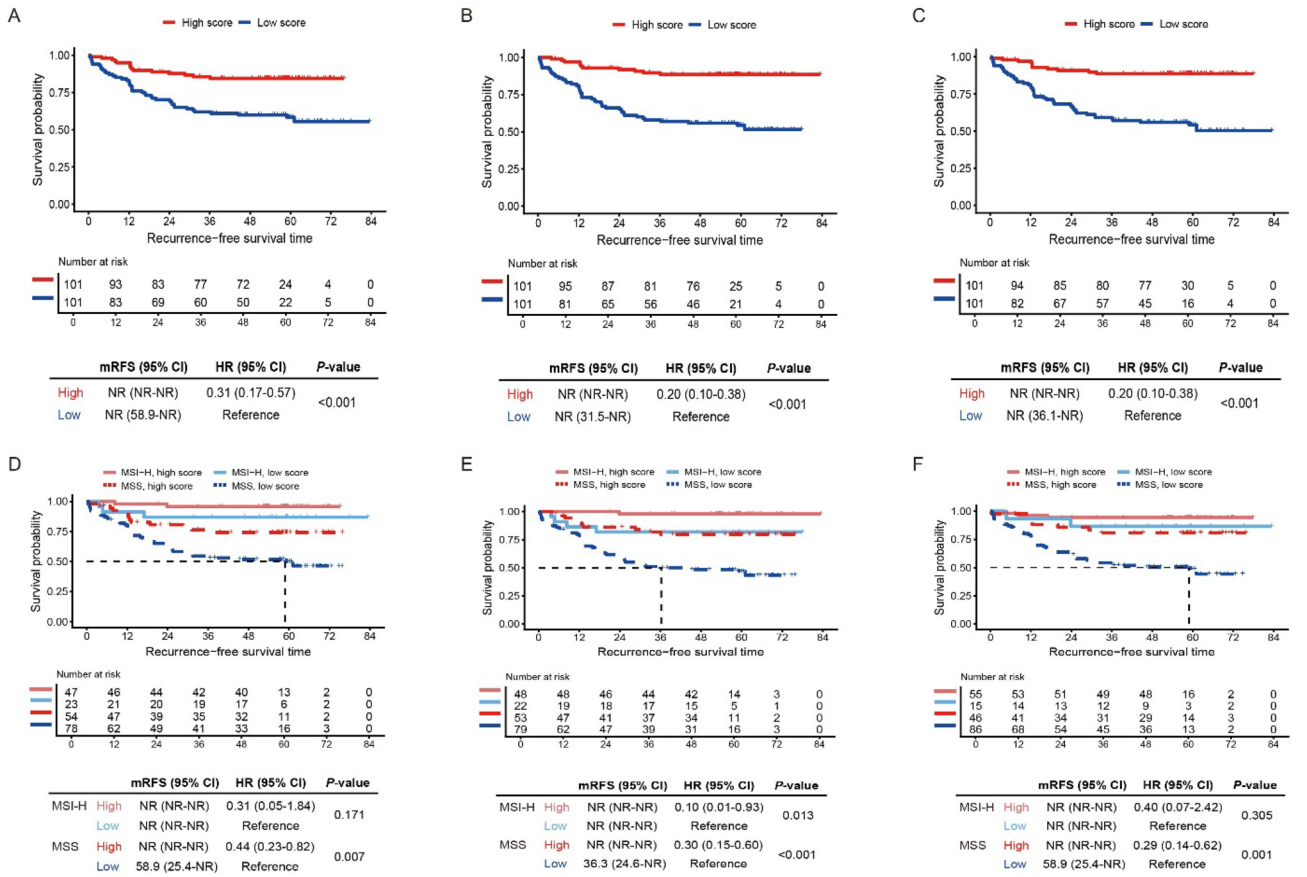

Fig. 3 Recurrence-free survival stratified by the median values of each scoring system. **A** TIL score in all patients, **B** TSB score in all patients, **C** Immunoscore in all patients, **D** TIL score by MSI status, **E** TSB score by MSI status, **F** Immunoscore by MSI status. Abbreviations: CI, confidence interval; HR, hazard ratio; MSI-H, microsatellite instability–high; MSS, microsatellite stable; NR, not reached; RFS, recurrence-free survival; TSB, tumor-stromal border

represents the principal interface where tumor cells interact with host immunity and initiate metastatic dissemination [4, 23]. Consequently, antitumor immunity at the TSB more directly reflects ongoing tumor–immune dynamics that determine clinical outcomes. In contrast, the tumor core typically exhibits an immunosuppressive microenvironment enriched with regulatory or exhausted immune populations. Additionally, iTIL density in the core is generally low and less dispersed than sTILs, reducing its statistical power to demonstrate meaningful prognostic associations. In our cohort, the core iTIL compartment constituted approximately one-half of the total iTIL area but displayed significantly lower TIL density compared with the bounding iTIL zone. This disproportion between area and density suggests that inclusion of the tumor core may dilute the prognostic signal of iTIL quantification. Since the core was defined as the region beyond 30 μm from the TSB, its substantial spatial contribution supports the methodological appropriateness of this operational definition. Therefore, excluding the core compartment when assessing iTILs may yield a more informative and clinically relevant metric, providing practical guidance for optimizing iTIL evaluation in future research and potential clinical applications.

Previous studies have emphasized the prognostic significance of T cell distribution in CRC. The pioneering work by Galon et al., which introduced the Immunoscore and was subsequently validated in numerous studies, demonstrated that TILs in TC and IM carry distinct prognostic value [24, 25]. However, these approaches generally treated the tumor as a single topological region, thereby limiting their ability to capture the heterogeneous spatial patterns of immune infiltration. Additionally, they largely rely on IHC, which constrains their scalability in routine clinical settings. Berthel et al. demonstrated that in CRC liver metastases, the prognostic impact of sTIL depends on their spatial distribution relative to the tumor–liver border [26]. However, their analysis did not capture iTILs and failed to incorporate distinct spatial zones.

Our study demonstrates the value in leveraging AI for high-resolution spatial analysis of CRC directly from

H&E-stained slides. This approach captures the complex spatial distribution of both iTILs and sTILs along the TSB while providing a practical and scalable method that does not require IHC. We developed two AI-based composite scores analogous to the Immunoscore methodology: (1) TIL score, a binary classification model that incorporates iTIL and sTIL densities, and (2) TSB score, a refined spatial model integrating the three prognostically significant TSB-related zones—bounding iTIL, bounding sTIL, and outermost sTIL. In multivariate Cox regression analysis, the TSB score demonstrated slightly superior prognostic performance over the TIL score (Table 3). Notably, in the MSI-high subgroup, a high TSB score was significantly associated with RFS, whereas Immunoscore was not (Fig. 3). These findings suggest that high-resolution spatial analysis of TILs may provide particular prognostic value in MSI-high CRC. In addition to its prognostic performance, the AI-based approach offers practical advantages in terms of cost-effectiveness and scalability. Unlike the conventional Immunoscore, which requires either centralized evaluation or additional CD3/CD8 IHC staining and scoring performed according to standardized protocols, entailing considerable labor and infrastructure demands, our approach uses routinely available H&E-stained slides without any additional tissue processing. This substantially reduces economic and logistical burdens and enables seamless application to existing digital pathology archives. Given that Immunoscore is not yet part of standard-of-care workflows in many centers, this H&E-based AI approach may provide a more accessible and readily implementable alternative for immune profiling.

MSI-high tumors typically exhibit an inflamed phenotype with abundant T-cell infiltration, creating a ceiling effect that limits the ability of Immunoscore to further stratify risk [27]. Consistently, in our cohort, 55 of 70 MSI-high tumors were classified as Immunoscore-high, leaving minimal variability for meaningful prognostic discrimination. Moreover, prolonged antigen stimulation inherent to the MSI-high microenvironment may drive T-cell exhaustion, evidenced by upregulation of co-inhibitory receptors including PD-1, LAG-3, and TIM-3, which may further attenuate the prognostic discriminatory capacity of CD3/CD8 density-based metrics even among highly infiltrated tumors [28]. Recent spatial analyses of MSI-high CRCs have shown that tumors with similar overall TIL densities can be stratified into prognostically distinct subgroups based on the spatial heterogeneity of TIL infiltration, underscoring that the prognostic relevance of the immune microenvironment in MSI-high tumors is determined not merely by the absolute quantity of TILs but also by their spatial organization [7, 29]. Furthermore, unlike Immunoscore, which exclusively quantifies $CD3^+$/$CD8^+$ T cells, H&E-based TIL assessment captures both T and B cell populations. Previous studies have shown that combining B and T cell densities at the IM enhances prognostic stratification [30]. Subsequent research has further revealed that B-cell infiltration toward the tumor provides independent prognostic value, particularly in MSI-high tumors, potentially explaining the retained discriminatory power of the TSB score in the MSI-high subgroups [31]. Collectively, these observations underscore the critical importance of spatially resolved immune metrics in MSI-high CRCs.

This study has some limitations. First, its retrospective, single-institution design may limit the generalizability of the findings. In addition, the proportion of MSI-high tumors in our cohort was higher than typically reported due to intentional enrichment to ensure adequate statistical power for subgroup analyses. While this design enabled a more robust evaluation in MSI-high disease, it may further limit generalizability and warrants external validation in population-based cohorts. Second, although MSI-high CRC was identified as a molecular subtype well suited for TSB analysis, we were unable to assess the ability of our approach to predict responses to immunotherapy. Finally, the AI-based analysis did not differentiate among TIL subsets (e.g., cytotoxic, helper, or regulatory T cells), which may have provided more refined immune profiling and potentially enhanced prognostic stratification. Future studies integrating spatial transcriptomics or multiplexed immunophenotyping would be valuable to elucidate the biological determinants of immune prognostic heterogeneity in MSI-high colorectal cancer, including the potential contribution of T-cell exhaustion to the limited discriminatory capacity of density-based metrics in this subgroup.

We demonstrated that CRC exhibits significant spatial heterogeneity of TILs, with distinct prognostic implications across TSB-defined regions. Notably, the H&E-based TSB score performed comparably to the IHC-based Immunoscore in the overall cohort and showed superior prognostic performance in MSI-high patients. These results underscore the clinical value of AI-powered, high-resolution spatial analysis of TILs for refined risk stratification in CRC.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00262-026-04409-9.

Acknowledgments We would like to thank Editage (www.editage.co.kr) for the English language editing.

Author contributions Conceptualization: HHK, SIC, MJ Data curation: HHK, MJ Formal analysis: HHK, SL, CO, SS, SIC, CHA, MK, MJ Funding acquisition: HHK, MJ Investigation: HHK, JYK, MJ Methodology: SL, CO, SS, SIC, SP, CHA Project administration: MJ Resources: HHK, JYK, MJ Software: SL, CO, SS, SIC, SP, CHA Supervision: MJ Validation: HHK, SL, MJ Visualization: HHK, SL, CO Writing—original draft: HHK, SL Writing—review & editing: HHK, MJ.

Funding This study was supported by the Korea–US Collaborative Cancer R&D Program funded by the Ministry of Health & Welfare, Republic of Korea (RS-2025—02223506); the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024—00341570); faculty research grants from Yonsei University College of Medicine (6—2023—0125 and 6—2023—0194); and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI20C2125).

Data availability All data generated and analyzed in this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest Financial interests: Seungeun Lee, Chiyoon Oum, Sanghoon Song, Soo Ick Cho, Sérgio Pereira, and Chang Ho Ahn are employed by Lunit Inc. (Seoul, Republic of Korea), which provided the AI model used in this study. The remaining authors declare no relevant financial or non-financial interests.

Ethical approval All procedures were conducted in accordance with the Declaration of Helsinki, and the Institutional Review Board of the Severance Hospital approved the study protocols (accession no. 4—2022—0900; approval date: August 29, 2022).

Consent to participate Informed consent was waived due to the retrospective design, use of anonymized data, and minimal risk to participants.

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/.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71:209–249. https://doi.org/10.3322/caac.21660
2. Guo L, Wang C, Qiu X, Pu X, Chang P (2020) Colorectal cancer immune infiltrates: significance in patient prognosis and immunotherapeutic efficacy. Front Immunol 11:1052. https://doi.org/10.3389/fimmu.2020.01052
3. Wozniakova M, Skarda J, Raska M (2022) The role of tumor microenvironment and immune response in colorectal cancer development and prognosis. Pathol Oncol Res 28:1610502. https://doi.org/10.3389/pore.2022.1610502
4. Lim Y, Choi S, Oh HJ et al (2023) Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes for prediction of prognosis in resected colon cancer. npj Precis Oncol 7:124. https://doi.org/10.1038/s41698-023-00470-0
5. Hendry S, Salgado R, Gevaert T et al (2017) Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International immuno-oncology biomarkers working group: part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. Adv Anat Pathol 24:311–335. https://doi.org/10.1097/pap.0000000000000161
6. Page DB, Broeckx G, Jahangir CA et al (2023) Spatial analyses of immune cell infiltration in cancer: current methods and future directions: a report of the International immuno-oncology biomarker working group on breast cancer. J Pathol 260:514–532. https://doi.org/10.1002/path.6165
7. Jung M, Lee JA, Yoo SY, Bae JM, Kang GH, Kim JH (2022) Intratumoral spatial heterogeneity of tumor-infiltrating lymphocytes is a significant factor for precisely stratifying prognostic immune subgroups of microsatellite instability-high colorectal carcinomas. Mod Pathol 35:2011–2022. https://doi.org/10.1038/s41379-022-01137-0
8. Hendry S, Salgado R, Gevaert T et al (2017) Assessing tumor-infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International immunooncology biomarkers working group: part 1: assessing the host immune response, TILs in invasive breast carcinoma and ductal carcinoma in situ, metastatic tumor deposits and areas for further research. Adv Anat Pathol 24:235–251. https://doi.org/10.1097/pap.0000000000000162
9. Fuchs TL, Sioson L, Sheen A, Jafari-Nejad K, Renaud CJ, Andrici J, Ahadi M, Chou A, Gill AJ (2020) Assessment of tumor-infiltrating lymphocytes using International TILs Working Group (ITWG) system is a strong predictor of overall survival in colorectal carcinoma: a study of 1034 patients. Am J Surg Pathol 44:536–544. https://doi.org/10.1097/pas.0000000000001409
10. Feng Y, Ma W, Zang Y et al (2024) Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients. Nat Commun 15:10259. https://doi.org/10.1038/s41467-024-54710-3
11. Chen C, Mejbel HA, Pathak T, Krasinskas A, Reid M, Corredor G, Fu P, Willis JE, Madabhushi A (2025) Artificial intelligence defines spatial patterns of tumor-infiltrating lymphocytes highly associated with outcome - a pan-GI cancer study. ESMO Open 10:105757. https://doi.org/10.1016/j.esmoop.2025.105757
12. Jung M, Kim JY, Jeong H, Valero Puche A, Song S, Cho SI, Jung M (2025) Artificial intelligence-driven quantification of tumor-stroma ratio and fibroblasts enables precise classification of stroma quality and quantity in predicting colorectal cancer recurrence. Pathobiology 92:276–287. https://doi.org/10.1159/000546021
13. Buhard O, Suraweera N, Lectard A, Duval A, Hamelin R (2004) Quasimonomorphic mononucleotide repeats for high-level microsatellite instability analysis. Dis Markers 20:251–257. https://doi.org/10.1155/2004/159347
14. Bang YH, Lee CK, Bang K et al (2024) Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes as a potential biomarker for immune checkpoint inhibitors in patients with biliary tract cancer. Clin Cancer Res 30:4635–4643. https://doi.org/10.1158/1078-0432.Ccr-24-1265
15. Bossowski JP, Pillai R, Kilian J et al (2026) The integrated stress response promotes immune evasion through lipocalin 2. Nature. https://doi.org/10.1038/s41586-026-10143-0

16. Mattes B, Scholpp S (2018) Emerging role of contact-mediated cell communication in tissue development and diseases. Histochem Cell Biol 150:431–442. https://doi.org/10.1007/s00418-018-1732-3
17. Su J, Song Y, Zhu Z, Huang X, Fan J, Qiao J, Mao F (2024) Cell-cell communication: new insights and clinical implications. Signal Transduct Target Ther 9:196. https://doi.org/10.1038/s41392-024-01888-z
18. Pagès F, Mlecnik B, Marliot F et al (2018) International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 391:2128–2139. https://doi.org/10.1016/s0140-6736(18)30789-x
19. Bankhead P, Loughrey MB, Fernández JA et al (2017) QuPath: open source software for digital pathology image analysis. Sci Rep 7:16878. https://doi.org/10.1038/s41598-017-17204-5
20. André T, Shiu KK, Kim TW et al (2020) Pembrolizumab in microsatellite-instability-high advanced colorectal cancer. N Engl J Med 383:2207–2218. https://doi.org/10.1056/NEJMoa2017699
21. Mulet-Margalef N, Linares J, Badia-Ramentol J, Jimeno M, Sanz Monte C, Manzano Mozo JL, Calon A (2023) Challenges and therapeutic opportunities in the dMMR/MSI-H colorectal cancer landscape. Cancers (Basel). https://doi.org/10.3390/cancers15041022
22. González-Montero J, Rojas CI, Burotto M (2024) Predictors of response to immunotherapy in colorectal cancer. Oncologist 29:824–832. https://doi.org/10.1093/oncolo/oyae152
23. Xu H, Cha YJ, Clemenceau JR, Choi J, Lee SH, Kang J, Hwang TH (2022) Spatial analysis of tumor-infiltrating lymphocytes in histological sections using deep learning techniques predicts survival in colorectal carcinoma. J Pathol Clin Res 8:327–339. https://doi.org/10.1002/cjp2.273
24. Galon J, Costes A, Sanchez-Cabo F et al (2006) Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 313:1960–1964. https://doi.org/10.1126/science.1129139
25. Bruni D, Angell HK, Galon J (2020) The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. Nat Rev Cancer 20:662–680. https://doi.org/10.1038/s41568-020-0285-7
26. Berthel A, Zoernig I, Valous NA, Kahlert C, Klupp F, Ulrich A, Weitz J, Jaeger D, Halama N (2017) Detailed resolution analysis reveals spatial T cell heterogeneity in the invasive margin of colorectal cancer liver metastases associated with improved survival. Oncoimmunology 6:e1286436. https://doi.org/10.1080/2162402x.2017.1286436
27. Mlecnik B, Bindea G, Angell HK et al (2016) Integrative analyses of colorectal cancer show Immunoscore is a stronger predictor of patient survival than microsatellite instability. Immunity 44:698–711. https://doi.org/10.1016/j.immuni.2016.02.025
28. Sorrentino C, D'Antonio L, Fieni C, Ciummo SL, Di Carlo E (2021) Colorectal cancer-associated immune exhaustion involves T and B lymphocytes and conventional NK cells and correlates with a shorter overall survival. Front Immunol 12:778329. https://doi.org/10.3389/fimmu.2021.778329
29. Kong H, Yang Q, Wu C et al (2024) Spatial context of immune checkpoints as predictors of overall survival in patients with resectable colorectal cancer independent of standard tumor-node-metastasis stages. Cancer Res Commun 4:3025–3035. https://doi.org/10.1158/2767-9764.Crc-24-0270
30. Bindea G, Mlecnik B, Tosolini M et al (2013) Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 39:782–795. https://doi.org/10.1016/j.immuni.2013.10.003
31. Nestarenkaite A, Fadhil W, Rasmusson A, Susanti S, Hadjimichael E, Laurinaviciene A, Ilyas M, Laurinavicius A (2020) Immuno-interface score to predict outcome in colorectal cancer independent of microsatellite instability status. Cancers (Basel). https://doi.org/10.3390/cancers12102902

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