Superpixel Image Segmentation of the Tumoral Microenvironment in Colorectal Cancer: Frontiers Beyond the Microscope

Sean Hacking (✉ shacking@northwell.edu)

Donald and Barbara Zucker School of Medicine at Northwell <https://orcid.org/0000-0001-6061-0781>

Dongling Wu

Donald and Barbara Zucker School of Medicine at Northwell

Claudine Alexis

Donald and Barbara Zucker School of Medicine at Northwell

Mansoor Nasim

Donald and Barbara Zucker School of Medicine at Northwell

Article

Keywords: Superpixel Segmentation, Tumoral Microenvironment, Colorectal Cancer, Frontiers Beyond the Microscope, Stromal Differentiation

Posted Date: April 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-379024/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Colorectal cancer (CRC) is the most common malignancy of the gastrointestinal (GI) tract and accounts for 9% of all cancers. The stroma and the tumoral microenvironment represent brave new frontiers for patients with colorectal cancer. Here we demonstrate novel superpixel image segmentation (SIS) techniques for whole slide images (WSI) to unravel this biology. Findings of significance include the association of low proportionated stromal area (PSA), high immature stromal percentage (ISP) and high myxoid stroma ratio (MSR) with worse prognostic outcomes in our CRC patients. Overall, stromal markers outperformed all others at predicting clinical outcomes. In particular, MSR may be able to prognosticate patients independent of tumor stage and may be the most optimal way to effectively prognosticate CRC patients which circumvents the need for more extensive deep learning (DL) based computational profiling. Approaches demonstrated here can be performed by a trained pathologist and very easily recorded during synoptic cancer reporting with appropriate quality assurance. Future well-designed, robust clinical trials will have the ultimate say in determining whether digital image analysis and superpixel image segmentation can better tailor the need for adjuvant therapy in patients with colorectal cancer.

Introduction

Colorectal cancer (CRC) is a diverse disease entity notable for being genetically inherited, although it also occurs sporadically[1]. Today, the field of histopathology is transitioning from isolated, traditional tumoral differentiation, to a more holistic approaches[2] which both encompass the tumoral microenvironment and incorporate molecular and computational findings to treat and prognosticate patients with CRC. The tumoral microenvironment is an ecosystem-like network of interacting cells, molecules, which regulate the extracellular matrix, as well as immune cell functioning[3].

Previous work has been undertaken to discover the essence of tumor budding (TB) [4] and stromal differentiation (SD)[2] in the CRC tumoral microenvironment. Histologically, SD has been demonstrated to be clinically significant in cancers of the breast[5], cervix[6] and stomach[7]; however, TB and SD has been most intensively study in colorectal cancer. Overall, immature SD and high TB are seen as bad players, being associated with reduced survival and higher pathologic stage[8].

The tumor-stroma-ratio (TSR) has also been examined by many groups in colon cancer[9]. It represents the ratio of the stroma to the tumor and measured by conventional light microscopy (CLM) manually by a pathologist and is restricted to the invasive tumoral front. Questions regarding manual assessment and reproducibility are concerning and traditional TSR is performed only in the invasive front of the tumor. It does not measure the tumor in its entirety. Despite this, the TSR is seen as a promising prognostic tool and tumours which harbor significant amounts of stroma have worse prognostic outcomes, suggesting a role for tumour-node-metastasis classification[9].

Current trends are pushing for more computation based approaches as we enter the digital era of pathology. Most recently, Guedj et al.[10] evaluated the prognostic potential of digital image analysis (DIA) for quantifying the stromal compartments of intrahepatic cholangiocarcinoma into order to calculate a proportionated stromal area (PSA) and found low values to be associated with worse outcomes.

In breast cancer, DIA has found triple negative biomarker status and high stroma was associated with poor prognosis, while in luminal tumors, high stroma was associated with favorable prognostic outcomes[11]. This suggests that the role of stroma may be tumoral subtype dependent in cancer, justifying the importance of validating DIA in CRC.

A pitfall of previous publications is that they did not differentiated stroma based on its myxoid differentiation. It is possible that it is not only important to look at stroma quantitatively (PSA), but to also look at the stroma qualitatively by assessing the degree of myxoid stroma differentiation.

Building on recent attempts to classify the proportionated stromal area (PSA)[10], here we will introduce two new concepts to DIA in the setting of stroma: immature stromal percentage (ISP) and myxoid stroma ratio (MSR). It is possible that tumors with immature (myxoid) differentiation may have worse survival outcomes, and if so, then both the quality and quantity of the stroma in CRC is important.

For pathologists, looking at the invasive front in isolation may be performed, but cautiously, digital image analysis may allow for a more holistic assessment of the tumoral microenvironment, which in addition to showing tumor-to-stroma ratios, also quantifies the degree of immature, myxoid stromal differentiation (degeneration).

Current paradigms are shifting away from manual qualitative assessment and towards quantitative scoring [12]. With so many open access digital pathology solutions, more opportunities are becoming available for transitioning to digital based methodologies. The tumoral microenvironment harbors significant biological diversity, which now must be unraveled from whole slide images (WSI).

Materials And Methods

Institutional Review Board:

This study received approval for all experimental protocols by the Institutional Review Board (IRB) of the Human Research Protection Program licensing committee at Northwell Health. All methods were carried out in accordance with all guidelines and regulations. Patient consent was not required due to the retrospective nature of the study (Northwell Health IRB number: 18–0128).

Study Design:

A total of 60 cases of colorectal carcinoma diagnosed in our health system were retrospectively analyzed and cases were collected from 2012–2017. Cases were only selected if they were primary resection

specimens, had nil adjuvant therapy and availability of sufficient non-frozen, formalin- fixed, paraffin- embedded tumor material for research purposes. One representative block was selected per case a single hematoxylin & eosin (H&E) slide was stained on the largest portion of tumor. Quality assurance steps were taken to further review slides for any folded, blurred, or obstructed morphologies. Slides were also evaluated for stromal differentiation, tumor budding and tumor-infiltrating lymphocytes. Further clinicopathological data was collected from the electronic medical records for each case and cancer free survival (CFS) data was provided by the cancer registry of Northwell Health. In following variables were included for exploratory analyses: cancer-free survival, age, gender, tumor budding score, pathologic stage, tumor-infiltrating lymphocytes (TIL), nodal status, tumor grade, stromal differentiation, Mis-match repair (MMR) status, Ki- *ras2* Kirsten rat sarcoma viral oncogene homolog (KRAS) and B-Raf Proto- Oncogene (BRAF).

Slide Digitalization:

Histological features in this study were assessed using virtual slides. Slides were scanned on Leica Aperio AT2 (Leica Biosystems, Buffalo Grove, Illinois, USA) whole slide scanner at 20×. Aperio vendor agnostic whole slide image viewer was used for pathologists to assess histology features described in our study. Digital slides were stored in the TIFF image format with JPEG image compression.

Superpixel Image Segmentation:

We incorporated QuPath[13], an open source for whole slides image analysis to evaluate the tumoral microenvironment. Qupath comes with built in trainable machine learning algorithms. All WSI files and were imported and orientated appropriately. This was performed on QuPath version 0.2.1 and superpixel image segmentation (SIS) was used to quantify stromal (myxoid and non-myxoid) and tumoral components. Superpixel image segmentation automatically grouped pixel similarities between different labeled cellular populations. In our study this was demonstrated as red (tumor), yellow (non-myxoid stroma), and blue (myxoid stroma). The entire area of the tumor was selected as shown in Fig. 1. In this publication we classified the proportionated straoma area (PSA) as performed by Guedj et al.[10] and defined by stromal to tumor area ratio. We developed the theoretical concepts of immature stromal percentage (ISP) and myxoid stromal ratio (MSR) for the purpose of this study. ISP is defined as the percentage of immature (myxoid) stroma to the entire tumoral area, while MSP is defined as the ratio of myxoid to non-myxoid stroma in the entire tumoral microenvironment.

Tumor Budding:

For tumor budding assessment, a detailed search was done for the area having the highest grade of tumour budding. After that, the counting of the buds took place in the hotspot region (20× objective lens). According to ITBCC protocol[14], the tumor budding is graded using 3-tier grading system as follows: Bd1: 0–4 buds, Bd2: 5–9 buds and Bd3: 10 or more buds.

Tumor-Infiltrating Lymphocytes:

TILs were defined as small mononuclear inflammatory cells which infiltrated between tumor cells. Tumors were assessed based on the 4-tier classification system previously validated for the quantification of inflammatory cells in colorectal cancer by Klintrup et al.[15]. Scoring was assessed at the deepest point of the invasive tumor. A score of 0 meant no inflammatory cells, 1 denoted a mild patchy increase in mononuclear cells, while a score of 2 and denoted a moderate (band-like) and florid (cuplike) infiltrate respectively, often accompanied by the destruction of cancer cell islands. Scoring was classified as low grade (0–1) or high grade (2–3).

Stromal Differentiation:

For stromal differentiation, scoring was primarily based on the grading system proposed by Ueno et al[16]. We analyzed the extramural desmoplastic front at low magnification. Myxoid stroma was defined as an amorphous stromal substance made of amphophilic material with a basophilic to grey extracellular matrix. This was usually intermingled with randomly oriented hyalinized collagen. Stroma was regarded as immature when at least a 40x field of myxoid change was observed. We categorized stroma as mature when the fibrotic stroma did not contain myxoid and most commonly comprised of fine mature collagen fibers stratified into multiple layers.

Mis Match Repair Status:

Immunohistochemistry was performed in order to determine MMR status on formalin-fixed and paraffin-embedded tumor sections cut at 4-micron thickness and stained on a Ventana Bench Mark Autostainer (Ventana Medical System, Tucson, Arizona). The following rabbit monoclonal pre-diluted primary antibodies were used: MLH-1(M1, Ventana), PMS2 (EPR 3947, Ventana), MSH2(G219-1129, Ventana), and MSH6 (44, Ventana). Technical methodologies and quality assurance were performed at the Immunopathology Laboratory of Long Island Jewish Medical Center (Northwell Health System, New Hyde Park, NY). Cases showing loss of immunohistochemical staining (< 1%) for any of the following stains: MLH1, PMS2, MSH2 and MSH6 was considered MMR deficient.

Next Generation Genomic Sequencing:

Molecular testing was performed on a subset of patients: 9 cases underwent BRAF and 12 cases underwent KRAS molecular testing. Genomic alterations of BRAF and KRAS were tested by next generation genomic sequencing on formalin-fixed, paraffin embedded tissue and mutational analysis was performed at Genpath laboratories (Elwood Park, NJ).

Statistical Analysis:

Comparative analysis was performed using the non-paired T test to examine differences in the clinicopathologic profile. The Kaplan-Meier method was used to evaluate cancer-free survival rate as a function of time. The log-rank method was used to compare differences between groups. These statistical Analysis was performed using Prism Graphpad version 8.4.2 Univariate and multivariate analyses of cancer free survival using Cox proportional-hazard regression were performed on SPSS statistics version 23.0.0.3. A p-value < 0.05 was considered statistically significant.

Results

Clinicopathologic and Patient Characteristics:

This study comprised data from 60 patients with colorectal adenocarcinoma who underwent surgical resection at our health system. Surgeries including block resection (1), right hemicolectomy (20), left hemicolectomy (27), transverse colectomy (5), rectosigmoidectomy (7). The mean age for our patient cohort was 67.4. There was a slight male predominance: 60% male (36), 40% female (24) to 51%. Tumor budding grades were as follows: Absent or low (24). Intermediate or high (36) Tumoral stage was as follows: pT1-2 (13), pT3-pT4 (47). For TIL, scoring was as follows: Low (40) and High (20). Lymph node status was as follows: N0 (32), N1-2 (28). Tumor differentiation was as follows: Well differentiated (44), moderately to Poorly differentiated (16). Stromal differentiation was as follows: 27 (45%) mature stroma, and 33 (55%) patients with immature. Out of 60 patients, 46 (77%) had intact mismatch repair, while 14 (23%) patients were mismatch repair deficient. Molecular testing was performed on a subset of patients: 9 (15%) had BRAF testing and 12 (20%) had KRAS mutation status.

Superpixel Image Segmentation:

We found a spectrum of tumoral, non-myxoid and myxoid stromal components by superpixel image segmentation. Heatmaps for percentages can be viewed in Fig. 2a, while non-linear regression curves can be seen in Fig. 2b for PSA, ISP and MSR respectively. Based off non-linear regression, cut-off values were designated as 0.9 (PSA), 10% (ISP) and 0.19 (MSR). Examples of superpixel hotspots with scores is demonstrated in Fig. 3.

Means were calculated for PSA, ISP and MSR in order to perform T-test analysis for the clinicopathologic profile. Firstly, tumor budding was found to be associated with a lower PSA (0.0116), high ISP (0.05) and high MSR (0.044). Pathologic stage was found to be associated with a lower PSA (0.0103) and a high ISP (0.013). Lymph node stage was found to be associated with high ISP (0.030) and high MSR (0.04). Tumor grade was found to be associated with high MSR (< 0.001). Stromal Differentiation was found to be associated with high ISP (< 0.0001) and high MSR (< 0.0001). MMR deficiency was found to correlate with PSA (0.0132) and ISP (0.04). KRAS mutation status (0.015). The remaining variables were not significant. The results for T-test analyses can be viewed in Table 1.

Table 1

T-test Analysis for Superpixel Image Segmentation. PSA, Proportionated Stromal Area; ISP, Immature Stromal Percentage; Myxoid Stromal Ratio; TIL, Tumor Infiltrating Lymphocytes; MMR, Mis-match repair; KRAS, Ki-*ras2* Kirsten rat sarcoma viral oncogene homolog; BRAF, B-Raf Proto-Oncogene. Significant features ($P \leq 0.05$) are shown in bold.

Ratio/ Percentage	PSA	P-Value	ISP	P-Value	MSR	P-Value
Variable	Mean		Mean		Mean	
Age		0.08		0.07		0.024
<=65	1.009		11.99		0.1944	
> 65	0.831		7.103		0.1921	
Gender		0.887		0.3771		0.648
Male	0.905		7.947		0.2011	
Female	0.890		10.32		0.1974	
Tumor Budding		0.0116		0.05		0.044
TBD1	1.052		12.05		0.1322	
TBD2 or TBD3	0.803		17.82		0.2840	
Pathologic Stage		0.0103		0.013		0.726
≤ PT2	1.121		8.144		0.1917	
PT3-PT4	0.831		18.30		0.1942	
TIL		0.8287		0.262		0.227
Low	0.887		7.668		0.1890	
High	0.912		10.83		0.2007	
Lymph node Stage		0.4941		0.030		0.04
N0	0.9355		10.01		0.1461	
N1-2	0.8674		18.07		0.2339	
Tumor grade		0.8296		0.3572		< 0.001
G1	0.9106		7.977		0.1407	
G2-G3	0.8860		10.77		0.3366	
Stromal Differentiation		0.5756		< 0.0001		< 0.0001
Immature	0.8714		0.322		0.2037	

Ratio/ Percentage	PSA	P-Value	ISP	P-Value	MSR	P-Value
Mature	0.9271		15.91		0.01	
MMR		0.0132		0.04		0.229
Intact	0.9624		7.485		0.1683	
Deficient	0.6798		13.53		0.204	
KRAS		0.7122		0.9512		0.015
Wild type	0.8578		7.738		0.2035	
Mutation	0.7799		7.350		0.0468	
BRAF		0.1353		0.2250		
Wild type	0.9582		9.075		0	
Mutation	0.375		2.5		0	

Cancer-free survival:

Cancer free survival (CFS) data was collected for all of the 60 patients with a mean follow up time of 1415 days. After setting the positivity cutoff to 0.9 for PSA, 10% for ISP and 0.19 for MSR we found that high PSA ($P = < 0.0001$), ISP ($P = 0.006$) and MSR ($P = 0.001$), SD ($P = 0.02$) and tumoral stage ($P = 0.04$) to be associated with poor CFS on Kaplan-Meier survival analysis, tumor grade was not significant ($P = 0.09$) (Fig. 3).

Based upon cox-regression of cancer-free survival (CFS), pathologic stage was found to be a poor prognostic factor on univariate analysis ($P = 0.03$) but not on univariate ($P = 0.062$). High TILs was found to have prognostic significant only on multivariate analysis ($P = 0.029$). Low PSA, was associated with worse prognostic outcomes on univariate (0.003) and multivariate analyses ($P = 0.0018$). High ISP was found to be associated with poor prognostic outcomes on univariate ($P = < 0.001$) and multivariate analysis ($P < 0.003$). High MSR was found to be associated with poor outcomes on univariate ($P = 0.028$) and multivariate ($P = 0.03$) analysis. Stromal differentiation was found to have prognostic significance on univariate (0.048) and multivariate analysis (0.019). The remaining clinicopathological variables were not significant ($P > 0.05$) on cox proportional hazard regression analysis (Table 2).

Table 2

Univariate and Multivariate analyses of Cancer Free Survival Using the Cox Proportional-Hazard Regression. CI, Confidence Interval; HR, Hazard Ratio; PSA, Proportionated Stromal Area; ISP, Immature Stromal Percentage; Myxoid Stromal Ratio; TIL, Tumor Infiltrating Lymphocytes; MMR, Mis-match repair; KRAS, Ki- *ras2* Kirsten rat sarcoma viral oncogene homolog; BRAF, B-Raf Proto-Oncogene. Significant features ($P \leq 0.05$) are shown in bold.

Variable	Frequency	Univariate HR	CI	P Value	Multivariate HR	CI	P Value
Age		0.987	0.943– 1.032	0.560	1.051	0.982– 1.125	0.151
<=65	22 (36.67%)						
> 65	38 (63.33%)						
Gender		1.247	0.429– 3.624	0.685	1.489	0.476– 4.652	0.494
Male	36 (60%)						
Female	24 (40%)						
Tumor Budding		2.815	0.784– 10.10	0.112	1.011	0.01– 7.83	0.173
TBD1	24 (40%)						
TBD2 or TBD3	36 (60%)						
Pathologic Stage		3.705	1.015– 13.53	0.047	0.405	0.140– 1.170	0.095
≤ PT2	13 (21.67%)						
PT3-PT4	47 (78.33%)						
TIL		0.404	0.111– 1.467	0.168	0.120	0.018– 0.809	0.029
Low	40 (66.67%)						
High	20 (33.33%)						
Lymph node stage		0.577	0.180– 1.848	0.355	0.909	0.205– 4.033	0.900
N0	32 (53.33%)						

Variable	Frequency	Univariate HR	CI	P Value	Multivariate HR	CI	P Value
N1-N2	28 (46.67%)						
PSA		0.046	0.06– 0.354	0.003	0.13	0.01– 0.451	0.018
< 0.9	35 (58.33%)						
>=0.9	25 (41.67%)						
ISP		1.189	1.106– 1.278	< 0.001	1.253	1.078– 1.458	0.003
< 10	37 (61.67%)						
>=10	23 (38.33%)						
MSR		5.613	1.947– 16.19	0.028	7.194	1.197– 43.7	0.03
< 0.19	39 (65%)						
> 0.19	21 (35%)						
Tumor grade		2.815	0.784– 10.10	0.112	1.454	0.300– 7.043	0.642
G1	44 (73.33%)						
G2-G3	16 (26.67%)						
Stromal Differentiation		3.715	1.829– 16.64	0.048	1.230	1.035– 1.462	0.019
Mature	27 (45%)						
Immature	33 (55%)						
MMR		2.607	0.903– 7.526	0.076	0.456	0.0533– 3.958	0.477
Intact	46 (76.67%)						
Loss	14 (23.33%)						

Variable	Frequency	Univariate HR	CI	P Value	Multivariate HR	CI	P Value
KRAS		2.828	0.17- 47.146	0.469			
Wild type	8 (66.67%)						
Mutation	4 (33.33%)						
BRAF		1.31	0.11– 3.52	0.97			
Wild type	8 (88.89%)						
Mutation	1 (11.11%)						

Discussion

Digital pathology may finally be coming of age. Whole slide scanners are becoming ubiquitous to the practise of surgical pathology paving way for new opportunities in computational pathology and digital image analysis.

The advent of open access image analysis software has been paramount to the success of digital pathology. Qupath[13] works to meet the need for a friendly, understandable digital pathology solution. It offers a comprehensive panel for image analysis and here we demonstrate the utility of SIS for calculating PSA and myxoid stroma percentages.

Importantly, PSA is not the same as the tumoral-stromal-ratio (TSR)[9]. The PSA technique includes the entire tumor area is computational, while traditional TSR is calculated only at the invasive tumoral border and performed manually by a pathologist on glass slides, allowing for observer bias.

In colon cancer, TSR has been found to correlate with prognostic outcomes, where high stroma ratios at the invasive tumor front are associated with poor outcomes[9]. However, we found the opposite effect, with low PSA values having worse prognostic outcomes. This may be secondary to the fact that PSA encompasses the entire tumor, not just the invasive front.

If they are just looking at the invasive front in isolation, then this may overestimate the stroma to tumor ratio. DIA may allow for a more holistic assessment of the tumoral microenvironment, which in addition to showing tumor-to-stroma ratios, also quantifies the ratio (MSR) and percentage (ISP) of immature, myxoid stromal differentiation (degeneration). For the PSA, it is possible that more tumor probably means more tumor budding, which correlates with worse prognosis. Whereas for the TSR performed at

the tumor invasive front, more stroma could mean more immature SD, also associated with a with worse prognostic outcomes[17].

Regarding DIA, our findings are also consistent with what was found in luminal tumors of the breast[11] and in intrahepatic cholangiocarcinoma[10], where higher proportions of tumor (low PSA) were found to be associated with worse overall survival outcomes.

A key point to consider is the variation in DIA methods between the two mentioned studies and this publication. In the breast publication[11] a single region of interest (ROI) was identified from a WSI and tissue microarrays (TMA) were created from tumor hotspots. In this publication we performed DIA on the entire tumor from a complete WSI, also performed on the study of intrahepatic cholangiocarcinoma [10]. WSIs may be the better solution, especially considering the significant tumor heterogeneity present in CRC, we feel this is the more pragmatic approach. Here we choose to use the slide with the largest section of tumor as we felt this most accurately represented the tumoral microenvironment.

Superpixel methods[18] are gaining traction in the field of medicine and recent attempts have been made to combine SIS and DL. In dermatology the combination of superpixels and deep learning models outperformed other competing methods[19]. The advantages for SIS in the deep learning space is that it can represent the structure of an image in adaptive sizes and shapes, with the ability to improve classification performance, especially for noisy classification (corrupted labels), as well as boundary misclassification[20]. Future studies could combine applications in deep learning with superpixel methods in order to further analyze the tumoral microenvironment. This is the road to precision oncology.

Recent advances have been made in DL for CRC by Skrede et al.[21] who applied ten convolutional neural networks on supersized heterogeneous WSI. The DoMore-v.1 assay was able to differentiate prognostic groups stage independently and tested on on large patient populations, suggesting its superiority to other genomic and pathological prognostic markers. The algorithm predicted cancer-specific survival in stage II patients as (HR 2.71, 95% CI 1.25 to 5.86, $p = 0.011$) and stage III (4.09, 2.77 to 6.03, $p < 0.0001$) [21].

More practically, the differentiation of the extracellular matrix leads to characteristic immature, myxoid stroma seen on routine histologic evaluation. Understanding the peculiarities of the ECM in colorectal cancer will be necessary for future and novel therapeutic approaches [3]; it is possible that myxoid degeneration of the ECM decreases the physical barrier, improving the delivery of therapeutics, nutrients, and immune cells to solid tumors. On the other hand, mature SD may act as a barrier (FIG 11); this is supported by preclinical studies demonstrating ECM degradation to improve drug uptake and response [3].

In our study we were able to more accurately characterize SD through DIA by calculating the ISP and MSR. Through SIS we were able to predict patient outcomes and clinical profiles better than manual analysis of SD. Importantly, these techniques could be used to tailor the need for adjuvant chemotherapy, although well-designed, robust clinical studies will be needed to determine this.

Interestingly, MSR was not found to be associated with tumor stage, suggesting that it may be able to predict clinical outcomes independent of tumor stage, it may be the key to unraveling the clinical heterogeneity present in CRC. Today, the use of adjuvant treatment in stage II colon cancer has garnered much controversy[22] and recommendations options range from observation, to single agent chemotherapy, to combination regimens. We need to better tailor the need for adjuvant therapy.

For patients with colorectal cancer, a rudimentary assessment of tumoral differentiation and stage may be insufficient. Looking forward, pathologists will have to borrow concepts from biology and develop new, novel approaches to prognosticate and treat patients with CRC. The tumoral microenvironment is a prime candidate for future applications in digital pathology, and the techniques described in this body of work can be performed easily by a surgical pathologist.

Declarations

Data Availability

Pathology data and the statistical analyses for the current study are available from the corresponding author upon reasonable request.

Funding

No funding was provided for the production of this manuscript.

Acknowledgments

We thank Alexander Perry and Kathy Quinn for their role as research coordinators. Figures were constructed at Biorender.com.

Author Contributions

MN, DW, and SH developed the theoretical formalism. SH, DW, and CA contributed to the acquisition of data. DW performed the analytic calculations and performed the numerical simulations. SH, DW, CA, and MN contributed to the final version of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. K.W. Jasperson, T.M. Tuohy, D.W. Neklason, and R.W. Burt, Hereditary and familial colon cancer. *Gastroenterology* 138 (2010) 2044–2058.
2. S.M. Hacking, B. Chakraborty, R. Nasim, T. Vitkovski, and R. Thomas, A Holistic Appraisal of Stromal Differentiation in Colorectal Cancer: Biology, Histopathology, Computation, and Genomics. *Pathol*

Res Pract 220 (2021) 153378.

3. E. Henke, R. Nandigama, and S. Ergün, Extracellular Matrix in the Tumor Microenvironment and Its Impact on Cancer Therapy. *Frontiers in Molecular Biosciences* 6 (2020).
4. V.H. Koelzer, I. Zlobec, and A. Lugli, Tumor budding in colorectal cancer—ready for diagnostic practice? *Hum Pathol* 47 (2016) 4–19.
5. Q. Zhai, J. Fan, Q. Lin, X. Liu, J. Li, R. Hong, and S. Wang, Tumor stromal type is associated with stromal PD-L1 expression and predicts outcomes in breast cancer. *PLOS ONE* 14 (2019) e0223325.
6. L. Cao, P.-L. Sun, Y. He, M. Yao, and H. Gao, Desmoplastic Reaction and Tumor Budding in Cervical Squamous Cell Carcinoma are Prognostic Factors for Distant Metastasis: A Retrospective Study. *Cancer Manag Res* 12 (2020) 137–144.
7. N.A. Kemi, M. Eskuri, V.-M. Pohjanen, T.J. Karttunen, and J.H. Kauppila, Histological assessment of stromal maturity as a prognostic factor in surgically treated gastric adenocarcinoma. *Histopathology* 75 (2019) 882–889.
8. H. Ueno, M. Ishiguro, E. Nakatani, T. Ishikawa, H. Uetake, K. Murotani, S. Matsui, S. Teramukai, T. Sugai, Y. Ajioka, H. Maruo, M. Kotaka, M. Tsujie, Y. Munemoto, T. Yamaguchi, H. Kuroda, M. Fukunaga, N. Tomita, and K. Sugihara, Prognostic value of desmoplastic reaction characterisation in stage II colon cancer: prospective validation in a Phase 3 study (SACURA Trial). *British Journal of Cancer* (2021).
9. G.W. van Pelt, T.P. Sandberg, H. Morreau, H. Gelderblom, J. van Krieken, R. Tollenaar, and W.E. Mesker, The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology* 73 (2018) 197–206.
10. N. Guedj, L. Blaise, F. Cauchy, M. Albuquerque, O. Soubrane, and V. Paradis, Prognostic value of desmoplastic stroma in intrahepatic cholangiocarcinoma. *Modern Pathology* 34 (2021) 408–416.
11. E.K. Millar, L.H. Browne, J. Beretov, K. Lee, J. Lynch, A. Swarbrick, and P.H. Graham, Tumour Stroma Ratio Assessment Using Digital Image Analysis Predicts Survival in Triple Negative and Luminal Breast Cancer. *Cancers (Basel)* 12 (2020).
12. H. Lara, Z. Li, E. Abels, F. Aeffner, M.M. Bui, E.A. ElGabry, C. Kozlowski, M.C. Montalto, A.V. Parwani, M.D. Zarella, D. Bowman, D. Rimm, and L. Pantanowitz, Quantitative Image Analysis for Tissue Biomarker Use: A White Paper From the Digital Pathology Association. *Appl Immunohistochem Mol Morphol* (2021).
13. P. Bankhead, M.B. Loughrey, J.A. Fernández, Y. Dombrowski, D.G. McArt, P.D. Dunne, S. McQuaid, R.T. Gray, L.J. Murray, H.G. Coleman, J.A. James, M. Salto-Tellez, and P.W. Hamilton, QuPath: Open source software for digital pathology image analysis. *Scientific Reports* 7 (2017) 16878.
14. A. Lugli, R. Kirsch, Y. Ajioka, F. Bosman, G. Cathomas, H. Dawson, H. El Zimaity, J.F. Fléjou, T.P. Hansen, A. Hartmann, S. Kakar, C. Langner, I. Nagtegaal, G. Puppa, R. Riddell, A. Ristimäki, K. Sheahan, T. Smyrk, K. Sugihara, B. Terris, H. Ueno, M. Vieth, I. Zlobec, and P. Quirke, Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod Pathol* 30 (2017) 1299–1311.

15. K. Klintrup, J.M. Mäkinen, S. Kauppila, P.O. Väre, J. Melkko, H. Tuominen, K. Tuppurainen, J. Mäkelä, T.J. Karttunen, and M.J. Mäkinen, Inflammation and prognosis in colorectal cancer. *Eur J Cancer* 41 (2005) 2645–54.
16. H. Ueno, A.M. Jones, K.H. Wilkinson, J.R. Jass, and I.C. Talbot, Histological categorisation of fibrotic cancer stroma in advanced rectal cancer. *Gut* 53 (2004) 581–6.
17. S.M. Hacking, B. Chakraborty, R. Nasim, T. Vitkovski, and R. Thomas, A Holistic Appraisal of Stromal Differentiation in Colorectal Cancer: Biology, Histopathology, Computation, and Genomics. *Pathology - Research and Practice* (2021) 153378.
18. M. Wang, X. Liu, Y. Gao, X. Ma, and N.Q. Soomro, Superpixel segmentation: A benchmark. *Signal Processing: Image Communication* 56 (2017) 28–39.
19. G. Blanco, A.J.M. Traina, C. Traina Jr, P.M. Azevedo-Marques, A.E.S. Jorge, D. de Oliveira, and M.V.N. Bedo, A superpixel-driven deep learning approach for the analysis of dermatological wounds. *Computer Methods and Programs in Biomedicine* 183 (2020) 105079.
20. C. Shi, and C.-M. Pun, Superpixel-based 3D deep neural networks for hyperspectral image classification. *Pattern Recognition* 74 (2018) 600–616.
21. O.J. Skrede, S. De Raedt, A. Kleppe, T.S. Hveem, K. Liestol, J. Maddison, H.A. Askautrud, M. Pradhan, J.A. Nesheim, F. Albregtsen, I.N. Farstad, E. Domingo, D.N. Church, A. Nesbakken, N.A. Shepherd, I. Tomlinson, R. Kerr, M. Novelli, D.J. Kerr, and H.E. Danielsen, Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 395 (2020) 350–360.
22. A. Varghese, Chemotherapy for Stage II Colon Cancer. *Clin Colon Rectal Surg* 28 (2015) 256–261.

Figures

Figure 1

Representative virtual slide analyzed by QuPath. Two separate portion of tumor are delineated (yellow) and super-pixelated for tumor (red), non-myxoid stroma (green) and myxoid stroma (blue). Only areas selected for annotation (yellow) are included into component percentages.

Figure 2

Heatmaps and non-linear regression analyses. (a) Heatmap demonstrated the spectrum of tumoral, non-myxoid, and myxoid stromal percentages. (b) Non-linear regression for cancer-free survival data and PSA, ISP and MSR. PSA, Proportionated Stromal Area; ISP, Immature Stromal Percentage; MSR, Myxoid Stromal Ratio.

Figure 3

Hotspots of Superpixel components with associated results for PSA, IMP and MSR. PSA, Proportionated Stromal Area; ISP, Immature Stromal Percentage; MSR, Myxoid Stromal Ratio.

Figure 4

Kaplan-Meier Survival Analysis. (a) Cancer free survival between PSA low and high groups. (b) Cancer free survival between IMP low and high groups. (c) Cancer free survival between MSR low and high groups. (d) Cancer free survival between mature and immature stroma groups. (e) Cancer free survival

between low (G1-2) and high grade (G3) groups. (f) Cancer free survival between low (T1/2) and high (T3/4) tumoral stage. HR, Hazzard Ratio; CI, Confidence Interval; G, Grade; T, Stage.