



Universiteit
Leiden
The Netherlands

Towards implementation of the tumour-stroma ratio in colorectal cancer

Polack, M.

Citation

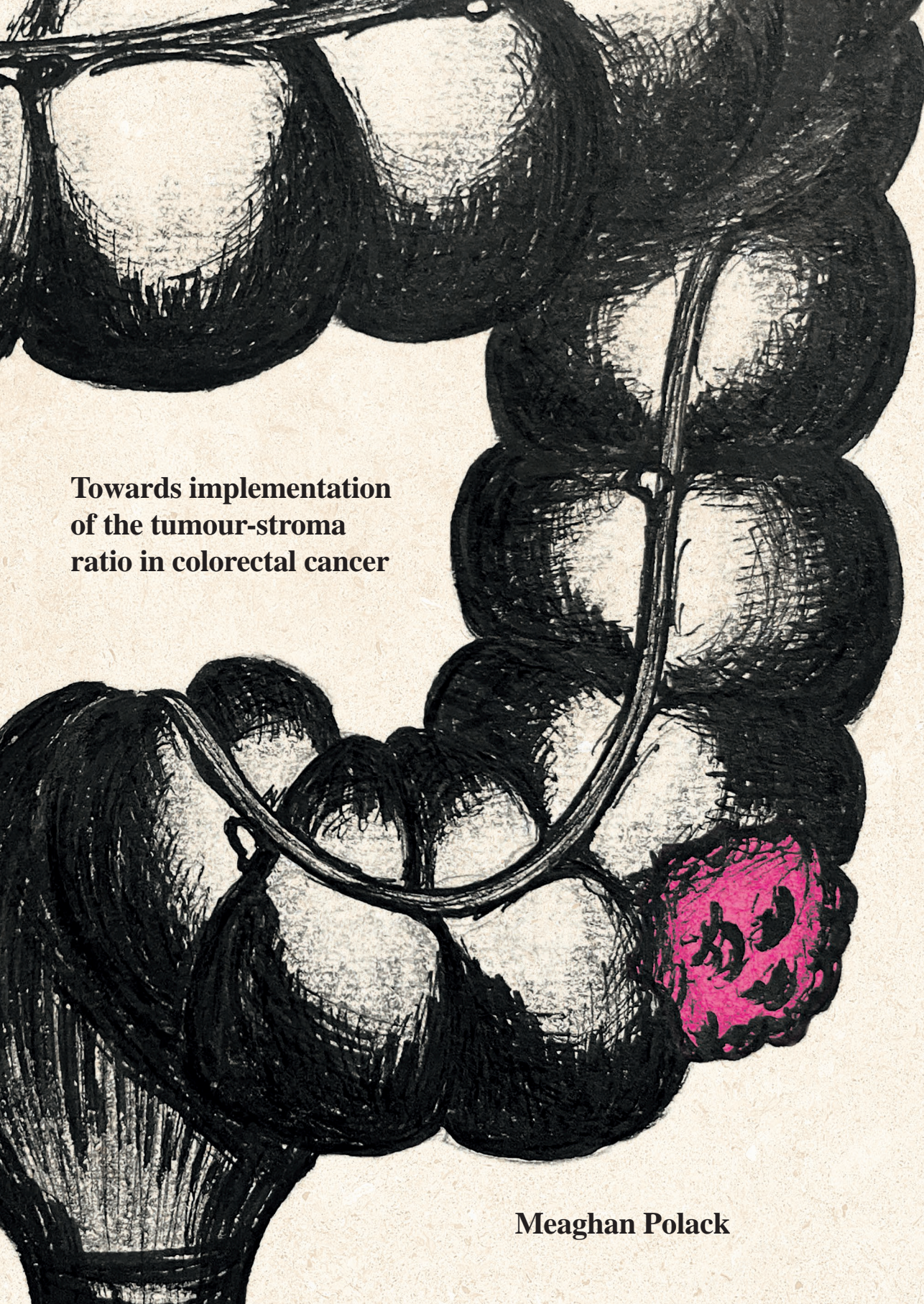
Polack, M. (2025, December 10). *Towards implementation of the tumour-stroma ratio in colorectal cancer*. Retrieved from <https://hdl.handle.net/1887/4285284>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4285284>

Note: To cite this publication please use the final published version (if applicable).



**Towards implementation
of the tumour-stroma
ratio in colorectal cancer**

Meaghan Polack

Towards implementation of the tumour-stroma ratio in colorectal cancer

Meaghan Polack

Colophon

Towards implementation of the tumour-stroma ratio in colorectal cancer

© Meaghan Polack, 2025, Leiden, the Netherlands.

ISBN: 978-94-6522-720-7

Cover: Meaghan Polack

Lay-out: Meaghan Polack

Printing: Ridderprint | www.ridderprint.nl

All rights reserved. No parts of this thesis may be reproduced, distributed, stored in a retrieval system or transmitted in any forms or by any means without prior written permission of the author.

The research in this thesis was financially supported by the Dutch Cancer Society (KWF grant number 1074), Stichting Fonds Oncologie Holland (SFOH) and the Bollenstreekfonds, Hillegom, the Netherlands.

Towards implementation of the tumour-stroma ratio in colorectal cancer

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 10 december 2025
klokke 14:30 uur

door

Meaghan Polack

geboren te Leiden
in 1994

Promotores:

Dr. W.E. Mesker

Prof. dr. R.A.E.M. Tollenaar

em. Prof. dr. J.H.J.M. van Krieken (Radboudumc)

Promotiecommissie:

Prof.dr. V.T.H.B.M. Smit

Prof.dr. M. Koopman (UMC Utrecht)

Prof.dr. G.J. Liefers

Dr. P. Snaebjornsson (Antoni van Leeuwenhoek/NKI)

Table of contents

Chapter 1	General introduction and thesis outline	7
Chapter 2	Results from the UNITED study: a Multicentre Validation Study of the Tumour-Stroma Ratio in Colon Carcinoma <i>ESMO Open. 2024 Apr;9(4):102988</i>	17
Chapter 3	The tumour-stroma ratio as predictive aid towards a biopsy-based treatment strategy in rectal carcinoma <i>Histopathology. 2025 Jul;87(1):44-57</i>	61
Chapter 4	Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT <i>Clinical Translational Oncology. 2022 Sep;24(9):1776-1784</i>	95
Chapter 5	Fibroblast activation protein-expression in colorectal carcinomas and implications for clinical application <i>Cancer Treatment and Research Communications. 2025 Jul 14;44:100964</i>	115
Chapter 6	Self-Supervised Learning Reveals Clinically Relevant Histomorphological Patterns for Therapeutic Strategies in Colon Cancer <i>Nature Communications. 2025 Mar 8;16(1):2328</i>	147
Chapter 7	Summary, general discussion and future perspectives	215
Chapter 8	Nederlandse samenvatting	225
Appendices	List of publications	240
	Curriculum Vitae	241
	Dankwoord	242



Chapter 1

General introduction
and thesis outline

General introduction and thesis outline

Colorectal cancer has remained the third highest cancer type in incidence and second cause of cancer-related deaths for years all over the world [1-3]. The cornerstone of international treatment guidelines encompasses disease extent, defined by the tumour-node-metastasis (TNM) classification [4, 5], and other assessments regarding expected patient-related outcomes, such as survival and treatment benefit [6-9]. However, recurrence rates as well as heterogeneity in survival outcomes of colorectal patients with similar TNM-stages are illustrative of a currently suboptimal therapy selection, leading to high occurrences of overtreatment and undertreatment. Moreover, even a survival paradox is observed, where patients with lower, less extensive disease and theoretically thus more favourable, TNM-stages, have a conversely worse survival outcome [6, 8, 10, 11]. Therefore, improving the prediction of patient-related outcomes and thus therapy selection is imperative.

In an attempt to attain this improvement, many histopathological parameters and biological markers (biomarkers) have been discovered and developed [12, 13]. Microsatellite instability (MSI) status, tumour deposits and tumour budding were for instance relatively recently implemented in routine pathology diagnostics [6, 14-16]. Whereas these mainly pertain to the tumour epithelial compartment, i.e. the neoplastic cells, this traditional tumour epithelium-centred approach remains therefore insufficient. It is pivotal to expand our current focus to include parameters that enable more accurate predictions and a personalized treatment strategy. Hence, the paradigm is shifting, with the surrounding tumour microenvironment increasingly emerging as an important modulator and no longer deemed an innocuous substance.

Tumour microenvironment

The past decades, the complex crosstalk between the tumour epithelial compartment and tumour microenvironment has piqued the interest of researchers worldwide [17-20]. Although the exact mechanism is yet to be elucidated, it has become apparent that this dynamic entity, predominantly comprising immune cells, extracellular matrix, vasculature and fibroblasts with various forms and functions, modulates tumour behaviour [21-23]. The tumour stroma therein, especially an abundance of stroma, has furthermore been observed to be of detrimental influence and enhances aggressive characteristics, such as tumour invasion and therapy resistance [22, 24-27]. The tumour-stroma ratio (TSR), i.e. the proportion of the tumour epithelial compartment compared to the tumour stromal compartment expressed in percentages and the main focus of this thesis, captures this effect specifically.

Tumour-stroma ratio

Since the first publication of this histological parameter in 2007 by our research group [28], the TSR has been protocolized [29, 30], promoted [31, 32] and proven [33-36], as well by other researchers [24, 37-39] and in multiple cancer types [40-42]. Iteratively, stroma-high (>50% intratumoural stroma percentages) tumours indeed lead to significantly worse patient-related outcomes than their stroma-low (\leq 50% intratumoural stroma) counterparts. A proposal to implement the TSR as additional biomarker in international guidelines and pathology diagnostics was thus presented to the TNM evaluation committee of the Union for International Cancer Control (UICC) and College of American Pathologists (CAP). While the advocated-to instances endorsed the potential of the TSR and despite this overwhelming evidence, however, a prospective study was still considered necessary. In a complying response, subsequent to an international consensus with support of the European Society of Pathology (ESP) and establishment of an official TSR-scoring E-learning [43], the multicentre UNITED study (Uniform Noting for International application of the Tumour-stroma ratio as Easy Diagnostic tool) was therefore initiated [44]. This thesis presents the results of the UNITED study in **Chapter 2**, validating the prognostic value of the TSR as biomarker on disease-free survival in colon cancer patients (DFS) [45].

Colon cancer vs. rectum cancer

Albeit often collectively termed, rectal cancer is a different entity than colon cancer [9, 46]. Treatment of rectal cancer varies thus from that of colon cancer, with a large focus on organ-sparing approaches [9, 47]. ‘Neoadjuvant therapy’ (Latin: *neo* = new, *adjuvare* = to help), first used to describe therapy intended for too extensive primary tumours for surgery [48], is a pillar in treatment in rectal cancer currently. Where it was initially implemented to enhance local control in addition to the total mesorectal excision (TME) [49, 50], research has shown that more intricate regimens can induce a complete response, potentially rendering surgery redundant [47, 51, 52]. However, this response is prone to variety and a large heterogeneity is seen between clinical and pathological responses (cCR and pCR, respectively) [53-56]. Upfront selection of patients standing to benefit from neoadjuvant therapy, and patients who should potentially not undergo a ‘watch-and-wait’ strategy [57], is thus essential.

The TSR has been shown previously to not only be a prognostic biomarker, but also to predict response to treatment [58-60]. Since neoadjuvant therapy alters the tumour microenvironment [61-63], resulting in e.g. fibrosis, this leads to inconclusive TSR scores after treatment. Scoring the TSR in the biopsies has previously proven to be representative for the TSR score of the primary tumour and, moreover, a predictive biomarker [64, 65], although large, multicentre studies evaluating the potential of the TSR in

predicting neoadjuvant response in rectal cancer are lacking [66-68]. In **Chapter 3** therefore, we integrate two large, clinical trials (RAPIDO [53] and PROCTOR-SCRIPT [69]) with our LUMC cohort and aimed to validate the predictive value of the TSR in rectal cancer.

Stromal fibroblasts and lymph node metastasis

Initial cancer diagnosis and staging, especially lymph nodes, continues to be challenging with the current CT and MRI scanning per protocol [6, 70]. PET scanning, a modality where anatomical and physiological functions are combined by using radioactive fluor-labelled fluorodeoxyglucose ([18-F]-FDG), is not commonly used in routine work-up and has pitfalls as well [6, 71, 72]. Cancer-specific tracers are increasingly researched, for which, targets pertain to the tumour stroma as well. Despite the large heterogeneity tumours are prone to, the tumour stroma generally contains quiescent fibroblasts, which can be activated through many pathways, giving rise to cancer-associated fibroblasts (CAFs) [73, 74]. Due to their tumour promoting functions, these CAFs are considered intriguing targets, and as such, a universal and high potential CAF marker has been identified: fibroblast activation protein (FAP) [75-77]. Coupled with a radioactive label, the fibroblast activation protein-inhibitor (FAPI) is exponentially subject to research in radiological imaging as tracer with FAPI-PET/CT scans, aiming to ascertain improvement to current imaging protocols [78-81]. Ultimately, tumour-targeted tracers can even be applied in a theragnostic sense [82-85]. Much research is done currently on the FAPI-tracer, although correlation studies with the golden standard, i.e. pathology, are lacking [86, 87].

As the importance of the TSR in lymph nodes metastases has been proven previously [88, 89], first, a theoretical framework for lymph node metastasis, specifically the stromal assessment, was established (**Chapter 4**). Here, we describe the tumour stroma in lymph node metastases in detail and expand on our TSR measurements, to support future correlation research between FAPI uptake on PET/CT scanning with the pathological lymph node assessments [90]. Moreover, as the CAFs had been identified as crucial modulators in the tumour behaviour of CRC, in **Chapter 5** we performed a large histological comparison between the traditional haematoxylin and eosin (H&E)-stained slides and immunohistochemical staining for FAP in CRC, biopsies and resected material with primary tumour and lymph nodes. As FAP constitutes an attractive marker for improved diagnosis through FAPI imaging, and translational studies are scarce, we aimed to describe the patterns to form a biological background for future clinical studies and ascertain correlation to the TSR.

Digitalisation and artificial intelligence

Historically, as the breakthrough that shaped the modern pathology from the mid-nineteenth century onward, the microscope has been the golden standard modality and trusted technique for the specialty [91, 92]. Glass slides containing tissue are visualized and analysed through Bright-field light microscopy in routine diagnostics, however, digitalisation is gaining interest in pathology [93]. Slides are increasingly scanned, resulting in digital whole slide images (WSIs). Initially intended to decrease workload and enable remote working for pathologists, these WSIs can also be analysed by computers, for instance by using artificial intelligence algorithms [94-96]. Often criticized and characterised as a ‘black box’, **Chapter 6** attempts to elucidate this phenomenon, where, in contrast to training on time-consuming annotations like in supervised deep learning to automate e.g. TSR scoring [97-99], we trained a state-of-art self-supervised deep learning model on unannotated WSIs [100]. Small image patches (tiles) from these WSIs were grouped based on similar histological patterns by this model, forming histomorphological phenotype clusters (HPCs). We subsequently analysed the HPCs and correlated these to patient-related outcomes, providing novel insights in potential histological aspects with the emphasis on tumour stroma, and advocate their clinical relevance.

Summary and discussion

The presented research in this thesis is ultimately summarized and followed by a discussion where future perspectives of research and more specifically, concrete and necessary steps are described towards the imperative implementation of the TSR as additional biomarker in guidelines and pathology routine diagnostics (**Chapter 7**). This thesis is concluded by the summary and discussion translated in Dutch and other appendices.

References

1. Sung, H., et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2021.
2. Siegel, R.L., et al., Cancer statistics, 2022. *CA Cancer J Clin*, 2022. 72(1): p. 7-33.
3. Bray, F., et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024. 74(3): p. 229-263.
4. Brierley, J.D., Gospodarowicz, M.K., Wittekind, C., The TNM Classification of Malignant Tumours. 8 ed. 2016: Wiley Blackwell.
5. Weiser, M.R., AJCC 8th Edition: Colorectal Cancer. *Ann Surg Oncol*, 2018. 25(6): p. 1454-1455.
6. Argiles, G., et al., Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2020. 31(10): p. 1291-1305.
7. Cervantes, A., et al., Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*, 2023. 34(1): p. 10-32.
8. Baxter, N.N., et al., Adjuvant Therapy for Stage II Colon Cancer: ASCO Guideline Update. *J Clin Oncol*, 2022. 40(8): p. 892-910.
9. Glynne-Jones, R., et al., Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2018. 29(Suppl 4): p. iv263.
10. Kim, H.S., et al., Clinicopathological and biomolecular characteristics of stage IIB/IIC and stage IIIA colon cancer: Insight into the survival paradox. *J Surg Oncol*, 2019. 120(3): p. 423-430.
11. Li, H., et al., Re-Evaluation of the Survival Paradox Between Stage IIB/IIC and Stage IIIA Colon Cancer. *Front Oncol*, 2020. 10: p. 595107.
12. Chuang, J.P., et al., Comprehensive Review of Biomarkers for the Treatment of Locally Advanced Colon Cancer. *Cells*, 2022. 11(23).
13. Chen, K., et al., Pathological Features and Prognostication in Colorectal Cancer. *Curr Oncol*, 2021. 28(6): p. 5356-5383.
14. Lugli, A., et al., Recommendations for reporting tumour budding in colorectal cancer based on the International Tumour Budding Consensus Conference (ITBCC) 2016. *Mod Pathol*, 2017. 30(9): p. 1299-1311.
15. Roth, A.D., et al., Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *J Natl Cancer Inst*, 2012. 104(21): p. 1635-46.
16. Ueno, H., et al., Optimal colorectal cancer staging criteria in TNM classification. *J Clin Oncol*, 2012. 30(13): p. 1519-26.
17. Fidler, I.J., Seed and soil revisited: contribution of the organ microenvironment to cancer metastasis. *Surg Oncol Clin N Am*, 2001. 10(2): p. 257-69, vii-viii.
18. Ouahoud, S., et al., Bidirectional tumour/stroma crosstalk promotes metastasis in mesenchymal colorectal cancer. *Oncogene*, 2020. 39(12): p. 2453-2466.
19. Sudhakar, M., H. Vignesh, and K.N. Natarajan, Crosstalk between tumour and microenvironment: Insights from spatial transcriptomics. *Adv Cancer Res*, 2024. 163: p. 187-222.
20. Werb, Z. and P. Lu, The Role of Stroma in Tumour Development. *Cancer J*, 2015. 21(4): p. 250-3.
21. Balkwill, F.R., M. Capasso, and T. Hagemann, The tumour microenvironment at a glance. *J Cell Sci*, 2012. 125(Pt 23): p. 5591-6.
22. Bremnes, R.M., et al., The role of tumour stroma in cancer progression and prognosis: emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *J Thorac Oncol*, 2011. 6(1): p. 209-17.
23. Sugai, T., et al., Microenvironmental markers are correlated with lymph node metastasis in invasive submucosal colorectal cancer. *Histopathology*, 2021. 79(4): p. 584-598.
24. Park, J.H., et al., The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Ann Oncol*, 2014. 25(3): p. 644-651.
25. Tsujino, T., et al., Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clin Cancer Res*, 2007. 13(7): p. 2082-90.
26. Zhuyan, J., et al., Critical steps to tumour metastasis: alterations of tumour microenvironment and extracellular matrix in the formation of pre-metastatic and metastatic niche. *Cell Biosci*, 2020. 10: p. 89.
27. Strous, M.T.A., et al., A high tumour-stroma ratio (TSR) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clin Transl Oncol*, 2022. 24(6): p. 1047-1058.
28. Mesker, W.E., et al., The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumour stage. *Cell Oncol*, 2007. 29(5): p. 387-98.
29. van Pelt, G.W., et al., Scoring the tumour-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*, 2018. 473(4): p. 405-412.
30. van Pelt, G.W., et al., The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 2018. 73(2): p. 197-206.
31. Smit, M.A. and W.E. Mesker, A Stromal Solution, in *The Pathologist*. 2018, Texere Publishing.
32. Mesker, W. Uniform Noting for International application of the Tumour-stroma ratio as Easy Diagnostic Tool - website. 2018 10-10-2024]; Available from: <http://watchstroma.com/>.
33. Mesker, W.E., et al., Presence of a high amount of stroma and downregulation of SMAD4 predict for worse survival for stage I-II colon cancer patients. *Cell Oncol*, 2009. 31(3): p. 169-78.
34. Huijbers, A., et al., The value of additional bevacizumab in patients with high-risk stroma-high colon cancer. A study within the QUASAR2 trial, an open-label randomized phase 3 trial. *J Surg Oncol*, 2018. 117(5): p. 1043-1048.

35. Huijbers, A., et al., The proportion of tumour-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*, 2013. 24(1): p. 179-85.
36. Zunder, S.M., et al., Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage II and stage III colon cancer. *Br J Cancer*, 2018. 119(2): p. 164-169.
37. Strous, M.T., et al., Node-negative colon cancer: histological, molecular, and stromal features predicting disease recurrence. *Molecular Medicine*, 2023. 29(1): p. 77.
38. Park, J.H., et al., Tumour invasiveness, the local and systemic environment and the basis of staging systems in colorectal cancer. *Br J Cancer*, 2017. 116(11): p. 1444-1450.
39. Sullivan, L., et al., Tumour Stroma Ratio and Its Significance in Locally Advanced Colorectal Cancer. *Curr Oncol*, 2022. 29(5): p. 3232-3241.
40. Smit, M.A., et al., The prognostic value of the tumour-stroma ratio in squamous cell lung cancer, a cohort study. *Cancer Treat Res Commun*, 2020. 25: p. 100247.
41. Vangangel, K.M.H., et al., The prognostic value of the tumour-stroma ratio is most discriminative in patients with grade III or triple-negative breast cancer. *Int J Cancer*, 2020. 146(8): p. 2296-2304.
42. Hagenaars, S.C., et al., Standardization of the tumour-stroma ratio scoring method for breast cancer research. *Breast Cancer Res Treat*, 2022. 193(3): p. 545-553.
43. Smit, M.A., et al., e-Learning for Instruction and to Improve Reproducibility of Scoring Tumour-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study. *JMIR Form Res*, 2021. 5(3): p. e19408.
44. Smit, M., et al., Uniform Noting for International Application of the Tumour-Stroma Ratio as an Easy Diagnostic Tool: Protocol for a Multicenter Prospective Cohort Study. *JMIR Res Protoc*, 2019. 8(6): p. e13464.
45. Polack, M., et al., Results from the UNITED study: a multicenter study validating the prognostic effect of the tumour-stroma ratio in colon cancer. *ESMO Open*, 2024. 9(4): p. 102988.
46. Wei, E.K., et al., Comparison of risk factors for colon and rectal cancer. *International journal of cancer*, 2004. 108(3): p. 433-442.
47. Beets, G.L., et al., A new paradigm for rectal cancer: Organ preservation: Introducing the International Watch & Wait Database (IWWD). *European Journal of Surgical Oncology (EJSO)*, 2015. 41(12): p. 1562-1564.
48. Wood, W.C., Neoadjuvant chemotherapy, in *Adjuvant Therapy of Breast Cancer*, I.C. Henderson, Editor. 1992, Springer US: Boston, MA, p. 279-291.
49. Peeters, K.C., et al., The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. *Ann Surg*, 2007. 246(5): p. 693-701.
50. Kapiteijn, E., et al., Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *New England Journal of Medicine*, 2001. 345(9): p. 638-646.
51. Garcia-Aguilar, J., et al., Organ Preservation in Patients With Rectal Adenocarcinoma Treated With Total Neoadjuvant Therapy. *J Clin Oncol*, 2022. 40(23): p. 2546-2556.
52. Custers, P.A., et al., Long-term Quality of Life and Functional Outcome of Patients With Rectal Cancer Following a Watch-and-Wait Approach. *JAMA Surgery*, 2023. 158(5): p. e230146-e230146.
53. Bahadoer, R.R., et al., Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *The Lancet Oncology*, 2021. 22(1): p. 29-42.
54. Lambregts, D.M.J., T.N. Boellaard, and R.G.H. Beets-Tan, Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights Imaging*, 2019. 10(1): p. 15.
55. Temmink, S.J., et al., Complete response rates in rectal cancer: Temporal changes over a decade in a population-based nationwide cohort. *European Journal of Surgical Oncology*, 2023. 49(11): p. 106991.
56. Nahas, S.C., et al., Diagnostic performance of magnetic resonance to assess treatment response after neoadjuvant therapy in patients with locally advanced rectal cancer. *Abdom Radiol (NY)*, 2019. 44(11): p. 3632-3640.
57. Temmink, S.J.D., et al., Watch and wait after neoadjuvant treatment in rectal cancer: comparison of outcomes in patients with and without a complete response at first reassessment in the International Watch & Wait Database (IWWD). *Br J Surg*, 2023. 110(6): p. 676-684.
58. van Pelt, G.W., et al., The value of tumour-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol*, 2020. 20: p. 39-44.
59. Hagenaars, S.C., et al., Tumour-stroma ratio is associated with Miller-Payne score and pathological response to neoadjuvant chemotherapy in HER2-negative early breast cancer. *Int J Cancer*, 2021.
60. Hansen, T.F., et al., Tumour-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. *Acta Oncol*, 2018. 57(4): p. 528-533.
61. Nagtegaal, I., et al., Morphological changes in tumour type after radiotherapy are accompanied by changes in gene expression profile but not in clinical behaviour. *J Pathol*, 2004. 204(2): p. 183-92.
62. Stone, H.B., et al., Effects of radiation on normal tissue: consequences and mechanisms. *Lancet Oncol*, 2003. 4(9): p. 529-36.
63. Miller, K.E., J.L. Ostrowski, and C.M. Quinn, Effects of chemotherapy on breast cancer tissue. *Histopathology*, 1997. 30(4): p. 397-8.
64. Park, J.H., et al., Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. *J Pathol Clin Res*, 2020. 6(1): p. 30-39.
65. Courrech Staal, E.F., et al., Reproducibility and validation of tumour stroma ratio scoring on oesophageal adenocarcinoma biopsies. *Eur J Cancer*, 2011. 47(3): p. 375-82.
66. Scheer, R., et al., Tumour-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: A retrospective cohort study. *World J Gastrointest Oncol*, 2017. 9(12): p. 466-474.

67. Strous, M.T.A., et al., Tumour-stroma ratio to predict pathological response to neo-adjuvant treatment in rectal cancer. *Surg Oncol*, 2022. 45: p. 101862.
68. Zhu, Y., et al., Prognostic Value of Tumour-Stroma Ratio in Rectal Cancer: A Systematic Review and Meta-analysis. *Front Oncol*, 2021. 11: p. 685570.
69. Breugom, A.J., et al., Adjuvant chemotherapy for rectal cancer patients treated with preoperative (chemo)radiotherapy and total mesorectal excision: a Dutch Colorectal Cancer Group (DCCG) randomized phase III trial. *Ann Oncol*, 2015. 26(4): p. 696-701.
70. Brouwer, N.P.M., et al., Clinical lymph node staging in colorectal cancer; a flip of the coin? *Eur J Surg Oncol*, 2018. 44(8): p. 1241-1246.
71. Godefroy, J., et al., Perceptual omission errors in Positron emission tomography/Computed tomography reporting. *Q J Nucl Med Mol Imaging*, 2021.
72. Lu, Y.Y., et al., A systematic review and meta-analysis of pretherapeutic lymph node staging of colorectal cancer by 18F-FDG PET or PET/CT. *Nucl Med Commun*, 2012. 33(11): p. 1127-33.
73. Chen, Y., K.M. McAndrews, and R. Kalluri, Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*, 2021. 18(12): p. 792-804.
74. Kobayashi, H., et al., Cancer-associated fibroblasts in gastrointestinal cancer. *Nat Rev Gastroenterol Hepatol*, 2019. 16(5): p. 282-295.
75. Henrich, L.M., et al., The Impact of Cancer-Associated Fibroblasts on the Biology and Progression of Colorectal Carcinomas. *Genes (Basel)*, 2024. 15(2).
76. Liu, F., et al., Fibroblast activation protein overexpression and clinical implications in solid tumours: a meta-analysis. *PLoS One*, 2015. 10(3): p. e0116683.
77. Janani, M., et al., Association of future cancer metastases with fibroblast activation protein- α : a systematic review and meta-analysis. *Front Oncol*, 2024. 14: p. 1339050.
78. Kratochwil, C., et al., (68)Ga-FAPI PET/CT: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med*, 2019. 60(6): p. 801-805.
79. Mori, Y., et al., FAPI PET: Fibroblast Activation Protein Inhibitor Use in Oncologic and Nononcologic Disease. *Radiology*, 2023. 306(2): p. e220749.
80. Dendl, K., et al., The Role of Fibroblast Activation Protein Ligands in Oncologic PET Imaging. *PET Clin*, 2021. 16(3): p. 341-351.
81. Giesel, F.L., et al., (68)Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with Various Cancers. *J Nucl Med*, 2019. 60(3): p. 386-392.
82. Parisi, A., et al., What Is Known about Theragnostic Strategies in Colorectal Cancer. *Biomedicines*, 2021. 9(2).
83. Lindner, T., et al., Targeting of activated fibroblasts for imaging and therapy. *EJNMMI Radiopharm Chem*, 2019. 4(1): p. 16.
84. Langbein, T., W.A. Weber, and M. Eiber, Future of Theranostics: An Outlook on Precision Oncology in Nuclear Medicine. *J Nucl Med*, 2019. 60(Suppl 2): p. 13s-19s.
85. Mori, Y., et al., Fibroblast Activation Protein Inhibitor Theranostics: Early Clinical Translation. *PET Clin*, 2023. 18(3): p. 419-428.
86. Mona, C.E., et al., Correlation of (68)Ga-FAPi-46 PET Biodistribution with FAP Expression by Immunohistochemistry in Patients with Solid Cancers: Interim Analysis of a Prospective Translational Exploratory Study. *J Nucl Med*, 2022. 63(7): p. 1021-1026.
87. Strating, E., et al., Fibroblast activation protein identifies Consensus Molecular Subtype 4 in colorectal cancer and allows its detection by (68)Ga-FAPI-PET imaging. *Br J Cancer*, 2022. 127(1): p. 145-155.
88. van Pelt, G.W., et al., Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *J Med Surg Pathol*, 2016. 1(02).
89. Vangangelt, K.M.H., et al., The prognostic value of tumour-stroma ratio in tumour-positive axillary lymph nodes of breast cancer patients. *Int J Cancer*, 2018. 143(12): p. 3194-3200.
90. Polack, M., et al., Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT. *Clin Transl Oncol*, 2022.
91. van den Tweel, J.G. and C.R. Taylor, A brief history of pathology: Preface to a forthcoming series that highlights milestones in the evolution of pathology as a discipline. *Virchows Arch*, 2010. 457(1): p. 3-10.
92. Loughrey, M.B., et al., Dataset for Pathology Reporting of Colorectal Cancer: Recommendations From the International Collaboration on Cancer Reporting (ICCR). *Ann Surg*, 2022. 275(3): p. e549-e561.
93. Pallua, J.D., et al., The future of pathology is digital. *Pathol Res Pract*, 2020. 216(9): p. 153040.
94. Hinton, G.E., 20 - CONNECTIONIST LEARNING PROCEDURES11This chapter appeared in Volume 40 of Artificial Intelligence in 1989, reprinted with permission of North-Holland Publishing. It is a revised version of Technical Report CMU-CS-87-115, which has the same title and was prepared in June 1987 while the author was at Carnegie Mellon University. The research was supported by contract N00014-86-K-00167 from the Office of Naval Research and by grant IST-8520359 from the National Science Foundation, in Machine Learning, Y. Kodratoff and R.S. Michalski, Editors. 1990, Morgan Kaufmann: San Francisco (CA). p. 555-610.
95. Hamet, P. and J. Tremblay, Artificial intelligence in medicine. *Metabolism*, 2017. 69s: p. S36-s40.
96. Chen, R.J., et al., Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 2024. 30(3): p. 850-862.
97. Firmbach, D., et al., Tumour-Stroma Ratio in Colorectal Cancer-Comparison between Human Estimation and Automated Assessment. *Cancers (Basel)*, 2023. 15(10).
98. Geessink, O.G.F., et al., Computer aided quantification of intratumoural stroma yields an independent prognosticator in rectal cancer. *Cell Oncol (Dordr)*, 2019. 42(3): p. 331-341.

99. Smit, M.A., et al., Deep learning based tumour-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform*, 2023. 14: p. 100191.
100. Zbontar, J., Barlow Twins: Self-Supervised Learning via Redundancy Reduction. 2021.



Chapter 2

Results from the UNITED study: a Multicentre Validation Study of the Tumour-Stroma Ratio in Colon Carcinoma

M. Polack, M.A. Smit, G.W. van Pelt, A.G.H. Roodvoets, E. Meershoek-Klein Kranenbarg, H. Putter, H. Gelderblom, A.S.L.P. Crobach, V. Terpstra, G. Petrushevska, G. Gašljević, S. Kjær-Frifeldt, E.M.V. de Cuba, N.W.J. Bulkman, G.R. Vink, R. Al Dieri, R.A.E.M. Tollenaar, J.H.J.M. van Krieken, W.E. Mesker, and the UNITED collaboration

Abstract

Background: The TNM Evaluation Committee (UICC) and College of American Pathologists (CAP) recommended to prospectively validate the cost-effective and robust tumour-stroma ratio (TSR) as independent prognostic parameter, as high intratumour stromal percentages have previously predicted poor patient-related outcomes.

Patients and methods: The UNITED study enrolled patients in 27 participating centers in 12 countries worldwide. The TSR, categorized as stroma-high (>50%) or stroma-low (\leq 50%), was scored through standardized microscopic assessment by certified pathologists, and effect on disease-free survival was evaluated with 3-year median follow-up. Secondary endpoints were benefit assessment of adjuvant chemotherapy and overall survival.

Results: A total of 1,537 patients were included, with 1,388 eligible stage II/III patients curatively operated between 2015-2021. Disease-free survival was significantly shorter in stroma-high ($N=428$) than in stroma-low patients ($N=960$) (3-year rates 70% vs. 83%; $P<0.001$). In multivariate analysis, TSR remained an independent prognosticator for disease-free survival ($P<0.001$, hazard ratio 1.49, 95% confidence interval 1.17–1.90). As secondary outcome, disease-free survival was also worse in stage II and III stroma-high patients despite adjuvant treatment (3-year rates stage II 73% vs. 92% and stage III 66% vs. 80%; $P=0.008$ and $P=0.011$, respectively). In stage II patients not receiving adjuvant chemotherapy ($N=322$), the TSR outperformed the ASCO-criteria in identifying patients at risk of events (event rate 21% vs. 9%), with a higher discriminatory 3-year disease-free survival rate (stroma-high 80% vs. ASCO high-risk 91%). A trend towards worse 5-year overall survival in stroma-high was noticeable (74% vs. 83% stroma-low; $P=0.102$).

Conclusion: The multicentre UNITED study unequivocally validates the TSR as independent prognosticator, confirming worse outcomes in stroma-high patients. The TSR improved current selection criteria for patients at risk of events, and stroma-high patients potentially experienced chemotherapy resistance. TSR implementation in international guidelines is highly recommended as aid in personalized treatment.

Trial Registration:

Dutch Trial Register NL7072; <https://clinicaltrialregister.nl/en/trial/23560>

International Registered Report Identifier (IRRID): DERR1-10.2196/13464

Introduction

Current treatment guidelines for colon cancer are traditionally based on extent of disease, expressed through the tumour-node-metastasis (TNM) classification, as well as risk assessments for patient outcome and expected benefits of adjuvant chemotherapy (ACT) [1-6]. However, the prognostic capacity of TNM staging remains suboptimal. Overtreatment, when patients do not or barely benefit from their ACT, as well as undertreatment, when patients actually could have benefited from additional treatment to prevent recurrences, therefore still occur at high rates [2, 3]. This supports the clinical need to improve individualized ACT indications upfront through additional prognostic biomarkers. Many pathological parameters have been discovered and implemented in guidelines, like tumour differentiation grade, tumour budding, or microsatellite instability (MSI) status [2, 3, 7, 8]. However, these mainly focus on the tumour epithelial compartment.

The past decades, the tumour stroma, a major component within the tumour microenvironment, emerged as an important influencer herein [9-14]. Abundance of intratumoural stroma has been demonstrated to lead to worse patient-related outcomes [15, 16]. The tumour-stroma ratio (TSR) is a histopathological parameter based on the amount of stroma expressed in percentages compared to the tumour epithelial component, and was initially developed in colon cancer, but has repeatedly been shown to be of prognostic value for almost all epithelial cancers. Patients with stroma-high tumours, i.e. >50% stroma, have a worse disease-free and overall survival (DFS and OS, respectively) than patients with stroma-low tumours, i.e. ≤50% stroma [17-21].

Implementation of the TSR in international guidelines and pathology diagnostics was advocated to the TNM Evaluation Committee, Union for International Cancer Control (UICC), and College of American Pathologists (CAP). Although these instances acknowledged the high potential of the TSR as prognostic parameter, validation was advised, including consensus on scoring of the TSR. Therefore, the present “Uniform Noting for International application of Tumour-stroma ratio as Easy Diagnostic tool” (UNITED) prospective multicentre study was initiated [22].

Herein, we hypothesized that patients with stroma-high tumours will have worse outcomes compared to patients with stroma-low tumours. Our primary endpoint was to determine the influence of the TSR on DFS, and secondary endpoints were influence of TSR on benefit of ACT, and on OS. The added value of the TSR in clinical treatment decision making could be, based on this prognostic information, to select patients with stage II stroma-high tumours for ACT, whereas the older patient with comorbidity

and a stage III stroma-low tumour could potentially be spared a burdensome and costly treatment. Validation of the TSR will result in unequivocally high-level evidence to accomplish implementation in international guidelines, aiding in shared-decision making through improved personalized treatment. Through this UNITED study, we aimed to validate the prognostic effect of the TSR in colon cancer patients.

Material and methods

Patients

The UNITED study was an investigator-driven, prospective, observational multicentre cohort study, enrolling patients in 27 centres from 12 countries. Approved and contracted centres could only start including after participating pathologists were certified through the official UNITED-study E-learning [23]. Coordination, including contract and database management, quality control and overall support, was done by the Clinical Research Center from the Leiden University Medical Centre (LUMC).

Patients ≥ 18 years of age with pathological stage II/III colon carcinoma and who had undergone a complete curative resection (R0) of their primary tumour were eligible. Patients were excluded in case of receiving neoadjuvant treatment, rectal cancer, multiple synchronous tumours, previous malignancies ≤ 10 years prior to the current cancer (except basal cell cancer or cervical cancer in situ) or any colon cancer in their medical history. Postoperative exclusion criteria were pathological stage I or IV and mortality within three months (Supplementary Table 1).

A sample size calculation was performed previously [15, 16]. For a 3-year median follow-up period, 114 and 97 recurrences were necessary for sufficient power for stage II and III colon cancer, respectively, requiring 722 and 450 evaluable patients in both groups. To obtain this minimum of 1,172 inclusions, approximately 1,500 patients (+25%) were to be registered in total, as inclusions in prospective cohort studies could ultimately be ineligible [22].

Materials and tumour-stroma analysis

Diagnostic haematoxylin-and-eosin (H&E)-stained slides of included patients, on which the T-stage of the tumour was already determined, were also used for scoring the TSR through conventional

microscopy. All participating pathologists were trained through the official UNITED study E-learning [23, 24]. This quality-controlled E-learning was supported by the European Society of Pathology (ESP) with official consensus on TSR scoring. High Cohen's interobserver agreement kappa's of >0.70 (at least substantial agreement) were previously observed, proving the reliability and efficiency in teaching the TSR scoring method, also long term [23]. The stromal percentage was scored on these H&E-stained slides according to the established method of van Pelt *et al.* [24], per 10 percent increments. Subsequently, categorization using the predefined cut-off value of 50% resulted in stroma-low ($\leq 50\%$) and stroma-high ($>50\%$) groups, similar to multiple previous studies [15, 20, 24] (Supplementary Figure S1).

Statistical analysis

DFS was defined as time between date of surgery and date of first event, i.e. recurrence (locoregional recurrence or distant metastasis) or death (by any cause). In the case of no event, DFS was calculated from date of surgery until censoring. Although accurate interpretation is only possible after five years, we also analyzed the preliminary effect of TSR on OS. OS was defined as time from date of surgery until date of death (by any cause) or until censoring. Censoring took place when patients were disease-free and/or alive at their last follow-up appointment.

Influence of TSR on benefit of ACT was assessed through comparison of TSR categories on DFS with treatment. As ACT is not routinely recommended in stage II colon cancer, the American Society of Clinical Oncology (ASCO)-criteria can be used to select those at high risk for events like recurrences, and thus could potentially benefit from this treatment: the stage II-HR. A pT4-tumour is deemed the most important ASCO risk criterium and commonly used in the Netherlands, but also sampling of <12 lymph nodes or emergency operation setting, presence of pathological risk factors like lymphovascular or perineural invasion (LVI and PnI, respectively), and poor tumour differentiation are risk factors [3]. First, for both pathological stages separately to minimize bias, influence of TSR will be determined in those receiving ACT to assess potential benefit. Subsequently, recurrence rates will be assessed for the TSR compared to ASCO criteria. To facilitate comparison and grouping of patients, all (sub)stages were recoded to the TNM-5 classification.

Statistical analysis was performed using the Chi-square test between ordinal or nominal variables. Through reversed Kaplan-Meier analysis, median follow-up time was calculated. Survival analyses were performed using Kaplan-Meier analysis with log-rank tests, and associated number needed to treat tables were added. Hazard ratio's (HR) with associated 95% confidence intervals (CI) were calculated

with the Cox proportional hazard model, using significant variables ($P < 0.05$) from the univariate analysis for the multivariate analysis.

All continuous variables were expressed in medians with interquartile ranges (IQR), whereas nominal and ordinal variables were stated as number of frequencies and corresponding percentages. Two-tailed P -values < 0.05 were considered statistically significant. Statistical analysis was performed in collaboration with the Department of Biomedical Data Sciences of the LUMC using IBM SPSS Statistics (version 29.0) and figures with GraphPad Prism (version 9.3.1).

Data storage

The Clinical Research Center coordinated data storage of the UNITED study, through the worldwide used and highly secured cloud-based platform Castor Electronic Data Capture (Castor EDC; Castor, Amsterdam, The Netherlands) [25]. Collection and supply of electronic Case Report Forms, central monitoring, quality control and generation of queries within the Castor database was performed by the Clinical Research Center, leading to high quality and reusable data. As per protocol, all data and documents are stored for a minimum of 15 years.

Ethical considerations

The UNITED study protocol was approved by the Medical Research Ethics Committee (MREC) of LUMC. All participating centres had their local MREC approve the protocol before inclusion could commence. In the Netherlands, centres were contracted through the Prospective Dutch ColoRectal Cancer cohort (PLCRC) [26, 27]. Patients from this prospective registration study, fulfilling the UNITED study eligibility criteria and treated in one of the participating centres from 2015 onward, were included. The workflow for retrieval of data through PLCRC is illustrated in Supplementary Figure S2. PLCRC explicitly included the possibility of a cohort-within-a-cohort format in their study design; patients signed broad official informed consent forms for the use of their histopathological data by other studies [27]. The UNITED study was conducted according to the Declaration of Helsinki (2013).

Results

Between January 8th, 2019 and September 9th, 2022, a total of 1,537 patients were registered. An overview of inclusion numbers per participating centre is provided in Supplementary Table 2. Baseline characteristics of all study patients are added in Supplementary Table 3. Due to the in part prospective nature of the study, 31 patients (2%) were ineligible at registration, e.g. based on medical history with another malignancy ≤ 10 years of their current colon cancer ($N=26$). Exclusion followed subsequently in 83 (5%) patients, due to presence of multiple tumours ($N=21$) or other pathology exclusions like pathological stage 0 – I ($N=42$) or IV colon cancer ($N=9$). Lastly, 35 (2%) were excluded during follow-up, mostly caused by postoperative mortality within 3 months ($N=26$).

In total, 1,388 stage II/III colon cancer patients were included in the final analysis (Figure 1). Of these, 770 patients (55%) were of male sex, and 453 patients were aged ≥ 75 years (33%). In 1,210 (87%) cases, a preoperative endoscopic biopsy was taken, highly indicative of an absence of emergency setting and indicating an elective operation. A total of 723 patients (52%) had stage II colon cancer. The tumour was categorized as stroma-high, i.e. $>50\%$ stroma, in 428 patients (31%), conform previous literature. Patient characteristics of the eligible cohort are summarized in Table 1.

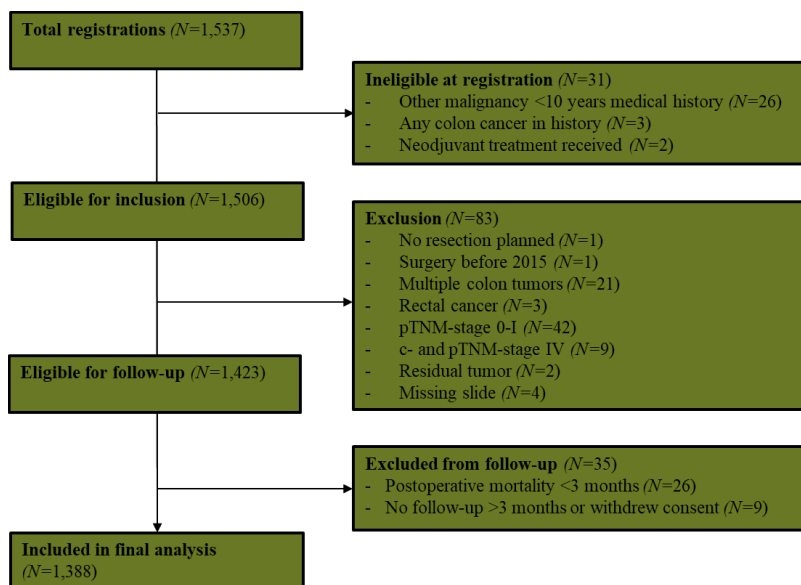


Figure 1. CONSORT diagram of the UNITED cohort. CONSORT, Consolidated Standards of Reporting Trials; UNITED, Uniform Noting for International application of Tumour-stroma ratio as Easy Diagnostic tool.

Table 1. Baseline characteristics of the eligible patients in the UNITED cohort.

Baseline characteristics	UNITED cohort (N=1,388)
Sex	
Female	618 (45)
Male	770 (55)
Age at surgery	
Median age (years)	69 (61 – 77)
≥75 years of age	453 (33)
Biopsy taken	
Yes, preoperative endoscopy	1,210 (87)
Yes, other method*	13 (1)
No**	165 (12)
Surgery	
Surgery year	2019 (2018 – 2020)
Pathological stage	
Stage II	723 (52)
Stage III	665 (48)
Lymph nodes	
Examined (<i>in total group</i>)	20 (15 – 28)
Positive (<i>in pTNM-stage III</i>)***	2 (1 – 4)
Tumour-stroma ratio	
Stroma-low (≤50%)	960 (69)
Stroma-high (>50%)	428 (31)

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding. TNM, tumour-node-metastasis stage.

*Other methods for biopsy are e.g. during surgery.

**Reasons why biopsy was not taken, is e.g. in emergency setting (obstructive ileus).

***Using the Union for International Cancer Control (UICC) TNM version 8, a tumour deposit (leading to stage N1c) will also lead to a pTNM-stage III, also when there are no positive lymph nodes.

Demographics, surgery type, tumour morphology and differentiation grade were equally distributed amongst stroma-low (N=960) and stroma-high (N=428) groups. Stroma-low tumours were more often right-sided (P=0.049) and had more often <12 lymph nodes sampled (P<0.001), however, in MSI or mismatch repair (MMR) analysis (MSI/MMR), stroma-low tumours were also more prone to MSI or MMR deficiency (MSI/dMMR; P=0.012). Stroma-high tumours ultimately had more risk factors, as these were more often stage III (P<0.001), pT4-stage (P<0.001), higher pN-stage (P<0.001), and more

often pathology risk factors like extramural venous invasion (EMVI) were present ($P<0.001$), illustrating their aggressiveness (Table 2). In merely a small subset ($N=153$, 11%), mutational status like KRAS or BRAF was determined, which was not significantly associated to TSR ($P=0.150$; Supplementary Table 4).

Table 2. Analysis of the variables of surgery, pathology and adjuvant chemotherapy in the eligible UNITED cohort, stroma-low compared to the stroma-high patients.

Variables (unit)	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Sex			0.385#
Female	420 (44)	198 (46)	
Male	540 (56)	230 (54)	
Age at surgery - older category			0.185#
<75 years of age	636 (66)	299 (70)	
≥75 years of age	324 (34)	129 (30)	
Biopsy taken			0.530#
Yes	843 (88)	380 (89)	
No*	117 (12)	47 (11)	
Surgery type			0.309†
Hemicolectomy right	435 (45)	176 (41)	
Hemicolectomy left	121 (13)	59 (14)	
Sigmoidectomy	277 (29)	136 (32)	
Other**	127 (13)	57 (13)	
Tumour-sidedness***			0.049#
Right-sided tumour	473 (49)	186 (44)	
Left-sided tumour	487 (51)	241 (56)	
Missing	0 (0)	1 (0)	
Lymph nodes			0.004#
Lymph nodes <12 examined	117 (12)	30 (7)	
Lymph nodes ≥12 examined	843 (88)	398 (93)	
Pathological tumour (pT)-stage****			<0.001#
pT1-3	810 (84)	305 (71)	
pT1	10 (1)	1 (0)	
pT2	61 (6)	6 (1)	
pT3	739 (77)	298 (70)	
pT4	150 (16)	123 (29)	
Pathological nodal (pN)-stage****			<0.001#
pN0	541 (56)	182 (43)	
pN1	292 (30)	152 (36)	
pN2	127 (13)	94 (22)	

(continued) Variables (unit)	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Tumour morphology			
Adenocarcinoma	855 (89)	390 (91)	
Mucinous carcinoma	90 (9)	33 (8)	
Other, including signet cell carcinoma	15 (2)	5 (1)	
Differentiation grade*****			0.062#
Well-Moderate	809 (84)	378 (88)	
Poor-Undifferentiated	125 (13)	41 (10)	
Grade cannot be assessed	26 (3)	9 (2)	
Pathology risk factors*****			<0.001#
No pathology risk factors present	566 (59)	197 (46)	
Presence of ≥1 pathology risk factor	394 (41)	231 (54)	
Extramural venous invasion (EMVI)			
Not reported	162 (17)	55 (13)	
Reported, of which	798 (83)	373 (87)	
Yes (EMVI+)	86 (9)	98 (23)	
No (EMVI-)	712 (74)	275 (64)	<0.001#
Venous invasion			
Not reported	99 (10)	56 (13)	
Reported, of which	861 (90)	372 (87)	
Yes (V1)	100 (12)	69 (19)	
No (V0)	761 (88)	303 (81)	<0.001#
Lymphatic invasion			
Not reported	28 (3)	20 (5)	
Reported, of which	932 (97)	408 (95)	
Yes (L1)	304 (33)	148 (36)	
No (L0)	628 (67)	260 (64)	0.193#
Perineural invasion			
Not reported	423 (44)	186 (44)	
Reported, of which	537 (56)	242 (57)	
Yes (Pn1)	68 (13)	58 (24)	
No (Pn0)	469 (87)	184 (43)	<0.001#
Microsatellite instability/Mismatch repair (MMR) status			
Not determined	439 (46)	213 (50)	
Determined, of which	521 (54)	215 (50)	
Microsatellite stable (MSS)/MMR proficient (pMMR)	403 (77)	184 (86)	
Microsatellite instable (MSI)/MMR deficient (dMMR)	118 (23)	31 (14)	0.012#
Pathological TNM-stage			
Stage II	541 (56)	182 (43)	
Stage III	419 (44)	246 (57)	<0.001#

<i>(continued)</i> Variables (unit)	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Adjuvant chemotherapy - received			<0.001#
No	539 (56)	174 (41)	
Yes	421 (44)	254 (59)	
Adjuvant chemotherapy – per pathological TNM-stage			<0.001#
Stage II + No adjuvant therapy	434 (45)	125 (29)	
Stage II + Adjuvant therapy	107 (11)	57 (13)	
Stage III + Adjuvant therapy	314 (33)	197 (46)	
Stage III + No adjuvant therapy	105 (11)	49 (11)	

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding. N/A, not applicable; TNM, tumour-node-metastasis stage.

*Reasons why biopsy was not taken, is e.g. in emergency setting (obstructive ileus).

**Other operation types include a (sub)total colectomy, high anterior resection, or transversectomy.

***A right-sided tumour is defined as a colon carcinoma in the caecum, colon ascendens, flexura hepatica or colon transversum.

****Different versions of the Union for International Cancer Control (UICC) TNM classification were used, here all variables are converted to the UICC TNM version 5 (1997).

*****Differentiation grade is variously registered as separate or combined subgroups, this is than categorized in combined grades, i.e. well – moderate or poor – undifferentiated.

*****Pathology risk factors are stated below, presence of a risk factor is defined as at least one of registered risk factors.

Absence is the absence of registered risk factors, as not all risk factors are registered.

*****Multiple mutations can occur simultaneously, hence the number of added percentages can be higher than 100.

Calculated with the Chi-square test.

† Calculated using the Chi-square test for the three most common and here presented operation types:

hemicolectomy right, hemicolectomy left and sigmoidectomy.

Median follow-up time was 36.2 months (95%CI 35.9 – 36.5) at time of database lock (January 31st, 2023), and comparable between both groups (P=0.469) (Supplementary Figure S3). Generally, follow-up was according to daily clinical practice, differing per country and centre, but approximately at 6, 12 and 24 – 36 months postoperatively. A total of 286 events occurred, of which 123 in the stroma-high group (29% of stroma-high patients; P<0.001). Mostly, this concerned distant metastases (92 stroma-high patients, 75% of events) (Supplementary Table 4). Hence, a statistically significantly worse DFS was observed in stroma-high patients with 3-year DFS rates of 70%, compared to 83% in stroma-low patients (HR 1.78, 95%CI 1.41 – 2.26; P<0.001; Figure 2).

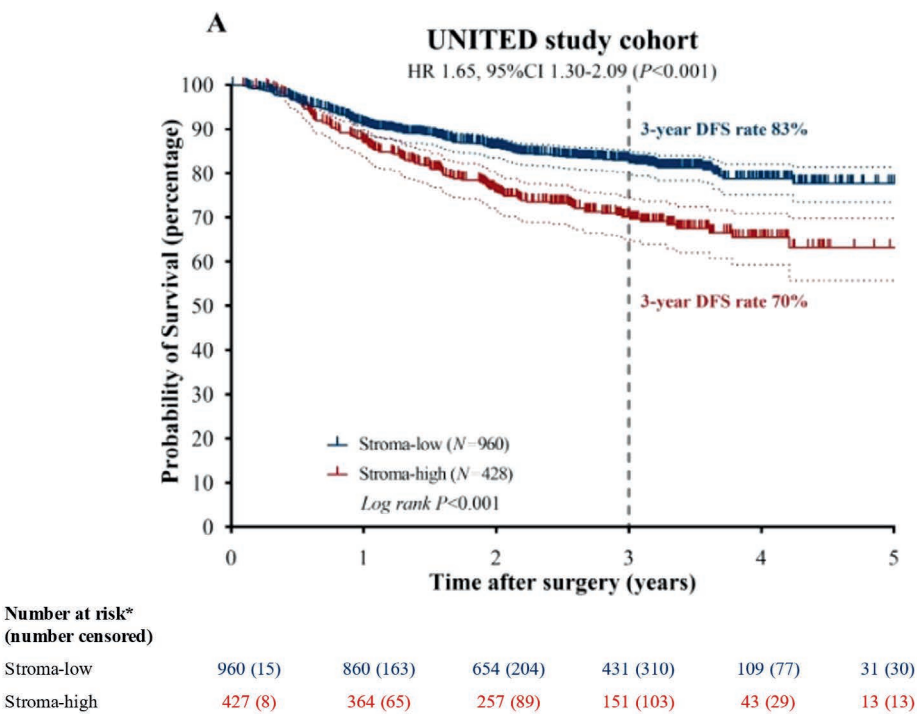
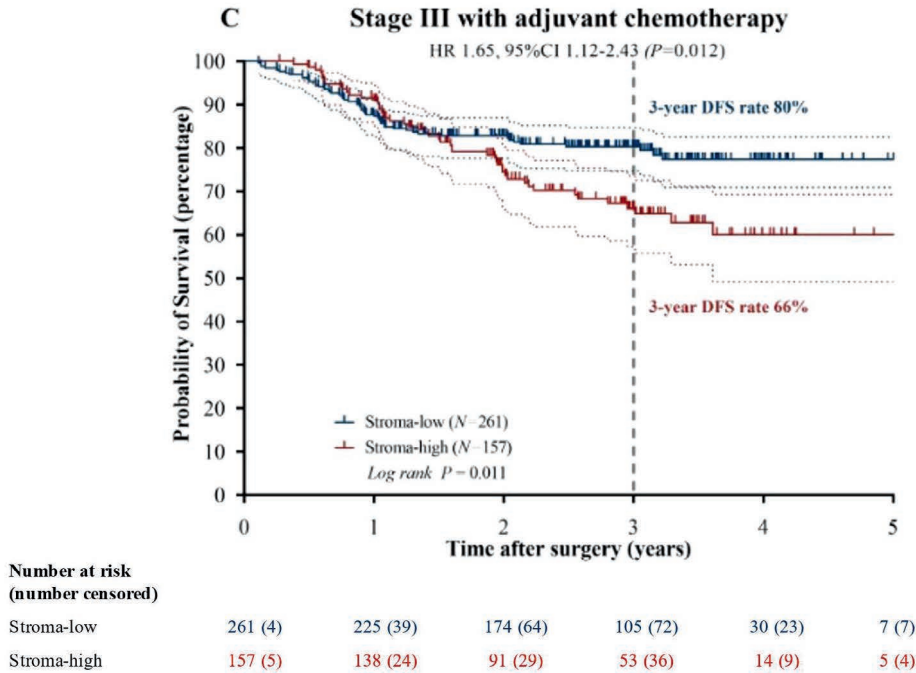
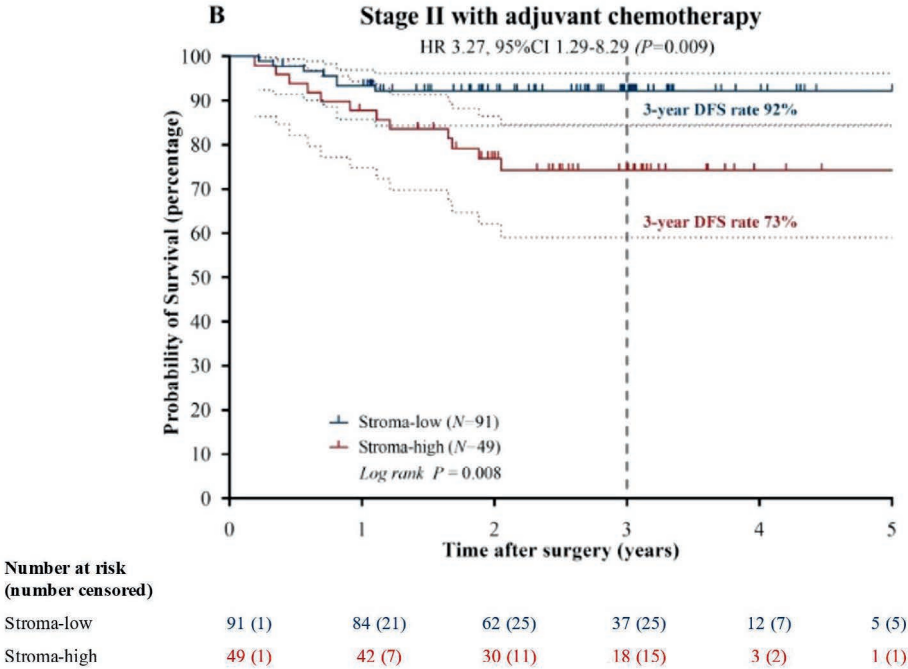
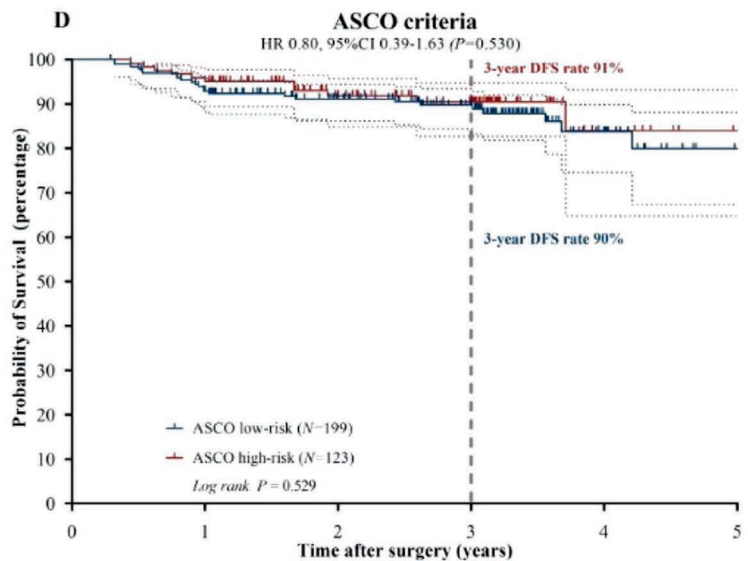


Figure 2. Disease-free survival effect of TSR in the UNITED cohort and subgroup analyses. A) Kaplan-Meier analysis and log rank test showing worse 3-year disease-free survival rates for stroma-high patients in the whole UNITED cohort (70 vs. 83%, respectively; $P<0.001$)

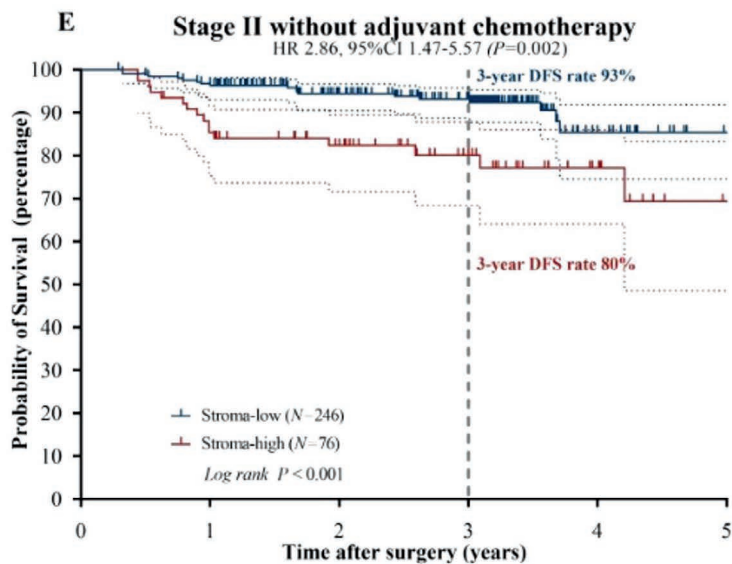
(continued on other page)





Number at risk (number censored)

ASCO low-risk	199 (4)	181 (32)	146 (40)	104 (70)	30 (19)	10 (10)
ASCO high-risk	123 (0)	117 (34)	80 (25)	54 (43)	10 (6)	4 (4)



Number at risk (number censored)

Stroma-low	246 (3)	234 (52)	178 (47)	129 (98)	27 (17)	10 (10)
Stroma-high	76 (1)	64 (14)	48 (18)	29 (15)	13 (8)	4 (4)

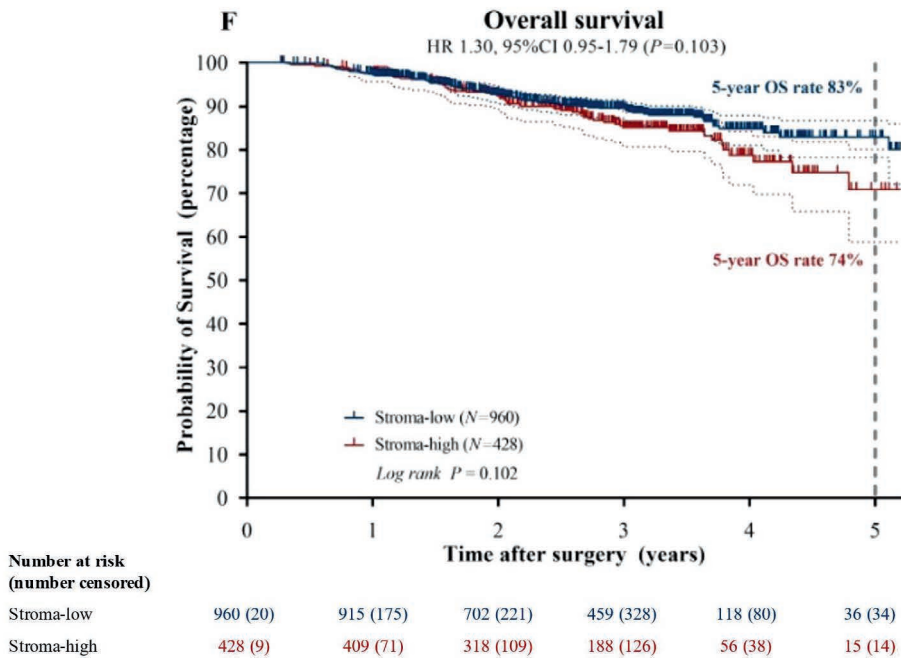


Figure 2. (continued) B) Kaplan-Meier analysis with log rank test in stage II patients receiving adjuvant chemotherapy, illustrating the worse 3-year survival rates for stroma-high patients despite treatment, indicating potential resistance to adjuvant chemotherapy (stroma-high 73% vs. stroma-low 92%; $P=0.008$). C) Kaplan-Meier analysis with log rank test in stage III patients receiving adjuvant chemotherapy, again illustrating the worse 3-year survival rates for stroma-high patients despite treatment (stroma-high 66% vs. stroma-low 80%; $P=0.011$); D) The ASCO criteria (high-risk vs. low-risk) not distinguishing any disease-free survival difference (high-risk 91% vs. low-risk 90%; $P=0.529$). E) Kaplan-Meier analysis with log rank test in stage II patients not receiving adjuvant chemotherapy, showing significant worse 3-year survival rates in stroma-high patients compared to stroma-low patients (stroma-high 80% vs. stroma-low 93%; $P<0.001$). F) Kaplan-Meier analysis with log rank test, showing overall worse survival in the stroma-high groups despite the short median follow-up of 3 years instead of 5 years, with the curves already diverging at 3 years. The 5-year overall survival rates are 74% vs. 83%, respectively ($P=0.102$). *For Disease-free survival, number of patients starting can be lower due to missing data. ASCO, American Society for Clinical Oncology; TNM, Tumour-node-metastasis stage; TSR, Tumour-stroma ratio; UNITED, Uniform Noting for International application of the Tumour-stroma ratio as Easy Diagnostic tool

Table 3. Univariate and multivariate analysis of on disease-free survival in UNITED cohort with Cox regression analysis

Variable (unit)	Univariate analysis			Multivariate analysis		
	Hazard ratio	95% Confidence interval	P-value	Hazard ratio	95% Confidence interval	P-value
Sex						
Female	1	..	0.753
Male	1.04	0.82 – 1.31
Age at surgery – older category						
<75 years of age	1	..	0.036	1	..	<0.001
≥75 years of age	1.29	1.02 – 1.63	..	1.54	1.21 – 1.96	..
Biopsy taken						
Yes	1	..	<0.001	1	..	<0.001
No*	2.60	1.96 – 3.46	..	2.33	1.72 – 3.16	..
Tumour-sidedness**						
Left-sided tumour	1	..	0.404
Right-sided tumour	0.91	0.72 – 1.14
Lymph nodes						
Lymph nodes ≥12 examined	1	..	0.009	1	..	0.067
Lymph nodes <12 examined	1.54	1.11 – 2.14	..	1.38	0.98 – 1.95	..
pT-stage***						
pT1-3	1	..	<0.001	1	..	<0.001
pT4	2.25	1.76 – 2.89	..	1.59	1.22 – 2.07	..
pN-stage***						
pN0	1	..	<0.001	1	..	<0.001
pN1	1.88	1.43 – 2.48	..	1.71	1.28 – 2.30	..
pN2	3.33	2.48 – 4.47	..	2.66	1.91 – 3.70	..

<i>(continued)</i>				
Tumour morphology	1	..	0.514	..
Adenocarcinoma	1.13	0.79 – 1.62
Other, including mucinous and signet cell carcinoma				
Differentiation grade	1	..	0.191	..
Well-moderate (high grade)	1.25	0.90 – 1.75
Poor-undifferentiated (low grade)				
Pathology risk factors	1	..	<0.001	1
No pathology risk factors present****	2.17	1.71 – 2.76	..	1.38
Presence of ≥1 pathology risk factor				1.06 – 1.81
Adjuvant chemotherapy – received	1	..	0.267	..
No	1.14	0.90 – 1.44
Yes				..
Tumour-stroma ratio	1	..	<0.001	1
Stroma-low	1.78	1.41 – 2.26	..	1.49
Stroma-high				1.17 – 1.90
Microsatellite stability*****	1	..	0.070	..
Microsatellite instable (MSI)	1.55	0.97 – 2.48
Microsatellite stable (MSS)				..

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges.

*Reasons why biopsy was not taken, is e.g. in emergency setting (obstructive ileus).

**A right-sided tumour is defined as a colon carcinoma in the caecum, colon ascendens, flexura hepatica or colon transversum.

***Different versions of the Union for International Cancer Control (UICC) classification were used. Here, all variables are converted to the UICC version 5 (1997).

****Pathology risk factors are stated below, presence of a risk factor is defined as at least one of registered risk factors. Absence is the absence of registered risk factors, as not all risk factors are registered.

*****Dependent covariate, excluded in analysis.

In multivariate analysis, after correcting for significant univariate variables, DFS remained worse for stroma-high compared to stroma-low patients, confirming the independent prognostic effect of the TSR on DFS (HR 1.49, 95%CI 1.17 – 1.90; $P<0.001$) (Table 3). Forest plots for these univariate and multivariate analyses are provided in Supplementary Figures S4 and S5. Effect of TSR on DFS per stage is illustrated in Supplementary Figure S6. In stage II, the worse DFS for stroma-high patients remained significant with 3-year DFS rates of 77% vs. 91% in stroma-low ($P<0.001$), but for stage III this was not the case (3-year DFS rates 65% vs. 72%, respectively; $P=0.055$). However, a significant bias occurred in stage III patients due to the difference in median age in stage III for patients who received ACT (65 years, IQR 57 – 72 years) and those who did not (79 years, IQR 71 – 84 years; Student's Independent T-test $P<0.001$). As this led to skewed results, stratification for age was deemed necessary here. After stratifying for age (<75 years), stroma-high stage III colon cancer led to significantly worse 3-year DFS rates as well (64% vs. 78%; $P=0.008$). The predictive potential of the TSR on benefit of ACT was investigated as secondary outcome in these groups.

A total of 676 patients (49%) of the UNITED study started with ACT, mostly intravenous oxaliplatin combined with oral capecitabine (CAPOX/XELOX; $N=394$, 58%). A detailed overview of treatment regimens is given in Supplementary Table 5. Although treatment guidelines can differ between countries, centres, and can be dependent on the decision of the physician and/or patient, we also looked within stages at those receiving ACT or not, to ascertain the benefit of patients who received additional treatment on DFS in reducing risk of recurrences.

Specifically, we analysed potential added benefit from ACT in stage II and III patients and influence of the TSR herein on DFS, after correction for age (<75 years). Within the stage II patients who did receive ACT ($N=140$), mostly stage II-HR, a significantly worse DFS was seen (3-year DFS rates stage II with ACT 73% stroma-high vs. 92% stroma-low; $P=0.008$). In the stage III group receiving standard-of-care ACT ($N=418$), 3-year DFS rates were significantly worse for stroma-high patients than their stroma-low counterparts despite their ACT, too (66% vs. 80%; $P=0.011$). (Figure 2B–C). This illustrates that stroma-low patients could benefit from ACT, but that stroma-high patients exhibit a lack of benefit or even potential resistance to ACT. Supplementary Figures S7A–B show per TSR category the different groups, with indeed worse DFS rates in all stroma-high groups not significantly increasing despite ACT ($P=0.080$) compared to stroma-low patients ($P<0.001$).

To assess which parameter could potentially have identified more patients at risk for an event, in-depth analysis on the subgroup of stage II patients <75 years not receiving ACT ($N=322$) was performed,

comparing the TSR to the ASCO criteria. According to the ASCO-criteria, in these stage II patients, 123 (38%) patients fulfilled one or more high-risk criteria. In this ASCO high-risk group, a 9% (N=11) event rate was observed, illustrating the percentage of undertreated patients. In the ASCO low-risk group however, in 24 cases (12%) an event occurred. The ASCO-criteria did thus not correctly identify patients at risk for events nor show differences in DFS rates ($P=0.383$ in Supplementary Table 6; 3-year DFS rates 90% low-risk vs. 91% high-risk, $P=0.529$ in Figure 2D, respectively). For TSR analysis, in this group of stage II patients not receiving ACT (N=322, 246 stroma-low and 76 stroma-high), a total of 35 events (11% event rate) occurred. Of these events, 16 occurred in the stroma-high (21% event rate in stroma-high group) and 19 in the stroma-low group (8% event rate in stroma-low group; $P=0.001$). DFS rates plotted in this subgroup show similar differences (80% vs. 93%; $P<0.001$) (Figure 2E). Compared to the ASCO criteria, the TSR thus identified an additional 12% patients at risk for events (21% vs. 9%) with a 91% 3-year DFS rate for ASCO high-risk patients in comparison to the 80% in stroma-high patients.

Although the UNITED study was powered specifically for DFS with a median 3-year follow-up period, as secondary endpoint the preliminary effect of the TSR on 5-year OS was estimated. A total of 163 deaths were recorded, of which 61 in the stroma-high group (14% of stroma-high patients; $P<0.001$). In the plotted Kaplan-Meier analysis and log rank analysis, effect on OS was not statistically significant, despite the relatively short median follow-up (HR 1.30, 95%CI 0.95 – 1.79; $P=0.103$). At 5 years, OS rates of 83% vs. 74%, respectively, were observed ($P=0.102$; Figure 2F).

Discussion

The UNITED study was initiated to prospectively validate the TSR as prognostic parameter in colon cancer patients. This study not only confirms that patients with stroma-high tumours indeed have significantly worse DFS, but also proves that this effect is independent from other prognostic high risk parameters, such as sampling of <12 lymph nodes and pathological T-stage or N-stage. Moreover, the TSR also outperformed the current ASCO criteria in identification of stage II colon cancer patients at risk for events. We had hypothesized that fit, stage II stroma-high patients could benefit from additional ACT and frail stage III stroma-low patients with better outcomes could perhaps be spared this treatment. However, our secondary findings contrarily indicate that all stroma-low patients benefit from ACT, whereas stroma-high patients do not and thus could actually be considered to not be selected for ACT. This study illustrates the aggressive behaviour of stroma-high tumours and the potential resistance to

(neo)adjuvant treatment of tumour stroma, which was also noticed in other studies by our research group [17, 28-30].

Despite the relatively short follow-up period of three instead of five years, a trend towards worse OS is already seen for stroma-high colon cancer. The curves diverge after three years, probably since most events, i.e. recurrences, occur within the first three years after primary diagnosis [2]. Additionally, due to the increase in sequential treatment options that may have extended survival in patients with recurrences or metastases, events mainly lead to an effect on DFS but not immediately on OS [2, 3, 31]. Therefore, we aim to collect longer follow-up data of UNITED study patients in the future, to adequately evaluate the effect of the TSR on OS after five to ten years.

In 2007, our research group was the first to describe the phenomenon of a high intratumoural stroma percentage and the associated worse patient-related outcomes in colon cancer [15]. Since then, much research has been performed regarding the role of tumour stroma, aiming to elucidate the biological mechanism. The intricate and dynamic tumour-stroma crosstalk has been observed to include the CAFs as important players, potentially also enabling the seed-and-soil principle of Paget and causing stromal metastases in lymph nodes [14, 32]. The TSR can be scored on these metastases as well, and patients with stroma-high primary tumours and stroma-high lymph node metastases have been observed to have the worst survival [33, 34]. Even small lymph nodes ≤ 5 millimetres in diameter, during routine radiologic imaging not suspected of malignancy, can contain metastases. Scoring the TSR in lymph nodes in the future, as well as more research improving positive lymph node detection, is pertinent for an even more tailored treatment [32].

Many biomarkers have emerged the past decades, as researchers are aiming to better predict tumour behaviour and patient outcomes. One such emerging biomarker is liquid biopsy, measuring circulating tumour DNA strands (ctDNA) in blood as a marker for minimal residual disease [35, 36]. Even more of interest, is the study on tumour stromal liquid biopsy panels, capturing the tumour microenvironment [37, 38]. However, not only do these increasing number of biomarkers add to existing high work load and are time consuming, often, more patient material or resources are necessary. Moreover, some biomarkers have variance in analyses, like the consensus molecular subtypes (CMS). CMS type 4, the mesenchymal type, mostly covers stroma-high tumours, but low reproducibility prohibited accurate analysis [39, 40]. Also tumour budding, now implemented in guidelines as an additional prognosticator, is known to have a less optimal interobserver agreement [41]. Although the TSR strictly does not capture the qualitative histopathological heterogeneity of the complete tumour-stroma entity, the UNITED

study shows that the quantitatively determined TSR still is an independent prognostic parameter, robust and simple, determined by pathologists during routine diagnostic microscopy assessment under two minutes without extra resources or costs, and therefore cost-effective [15, 17, 20, 24]. Potentially additional analyses, e.g. organization or maturity determination, can be done and would be of interest to perform on UNITED study material for further characterization of the tumour stroma [10, 42].

Also, the tumour immune microenvironment has been proven to affect tumour behaviour, as for instance seen in the Immunoscore [43]. High influx of tumour-infiltrating immune cells (TIICs) is often indeed correlated with decreased rates of recurrence. Hence, increasing amounts of research demonstrate the high potential of certain immunotherapeutic regimens in cancer. Although there are various potential analyses to ascertain the tumour immune microenvironment, often also requiring additional resources and patient material like with the Immunoscore method, future analysis of TIICs in all UNITED study patients would be of interest [43]. Previous research namely shows that the amount of influx of TIICs combined with TSR forms a potentially even more accurate prognosticator, as stroma-high/immune-low tumours are associated with the worst patient-related outcomes [44].

There are some limitations to the present study. MSI/MMR status, has only been obligatory in current treatment guidelines, and thus was not determined in the included older first half of patients [2]. Additionally, our analyses on ACT, as the UNITED study was not powered nor set up for this specifically, may be biased despite stratification for risk factors like age, e.g. due to the powerful a priori prognostic effect. Although the association seen in our analyses is already significant, we are currently initiating a well-balanced, matched cohort specifically powered for this endpoint to confirm the apparent resistance to ACT in stroma-high patients. Also, median three years of follow-up is relatively short, but fulfils the standard to adequately ascertain DFS, as previously taken into account in our power calculation, and can even be interpreted as a valid surrogate for OS [22, 45].

Strengths of this study include foremost the prospective nature, to minimize bias and adequately evaluate the causal effect of the TSR on survival. After the initial proposal of implementation of the TSR in TNM-based guidelines to the UICC and CAP, various supportive collaborations have been established for this study. Participating pathologists were trained with the reliable and quality-controlled UNITED study E-learning [22, 23]. Support grew, ultimately leading to the completion of this study. Not only has the TSR thus been validated, but the UNITED study also led to the foundation of an international system for pathologists to potentially implement the TSR in daily routine diagnostics. Moreover, international guidelines regarding ACT can be re-evaluated. As stroma-high patients have a

worse DFS and exhibit resistance to ACT, they could be discussed in multidisciplinary settings and potentially be spared this treatment.

As part of this study, centres were requested to send scans of the scored H&E-stained slides. This collection will be used for future development of artificial intelligence algorithms for TSR automatization. Not only can the interobserver agreement be increased even more, also analysis of difficult cases with for instance high amounts of mucin or necrosis can be facilitated, and automatization supports the increasing interest in digital pathology as well, including possibilities for even further research. Deep learning models especially can discern even more and potentially novel features in the tumour stroma, e.g. stromal organization like previously analysed by our research group, and correlations to patient-related outcomes can be assessed [42, 46].

Importantly, future studies should focus on tumour stroma-targeted therapeutic regimens or strategies, as this study shows a lack of benefit of stroma-high colon cancer to ACT, revealing a clear clinical need for new treatment options for these patients. Similar to ACT, tumours with high amounts of tumour stroma also have been observed to respond less to immunotherapeutic strategies [47]. Although confirming this potential resistance to adjuvant treatment in a powered series is necessary, fundamental research and pharmacological phase I studies should be initiated to uncover specific targets.

Conclusions

In conclusion, the UNITED study hereby unequivocally confirms the independent prognostic effect of the TSR on DFS in colon cancer patients as per request of the UICC and CAP. As stroma-high patients have worse DFS and appear to benefit less from ACT than stroma-low patients, the TSR can herein aid in clinical shared decision-making and personalized treatment. Therefore, implementation of the TSR in standard of care pathology diagnostics and reporting in addition to currently used elements as the TNM classification and ultimately in international guidelines is highly recommended.

Acknowledgements

We thank the research support staff at all participating centres, including PLCRC for providing many of the contacts and participating centres, and the Netherlands Comprehensive Cancer Organization (IKNL) for the data.

Funding

This work was supported by grants from the Dutch Cancer Society (KWF Kankerbestrijding) [grant for project number 10174] and the Stichting Fonds Oncologie Holland [no grant number], as well as the Bollenstreekfonds, Hillegom, the Netherlands [no grant number]. These funders had no role in study design, data collection and analysis, nor in the decision to publish, nor in the preparation of the manuscript.

References

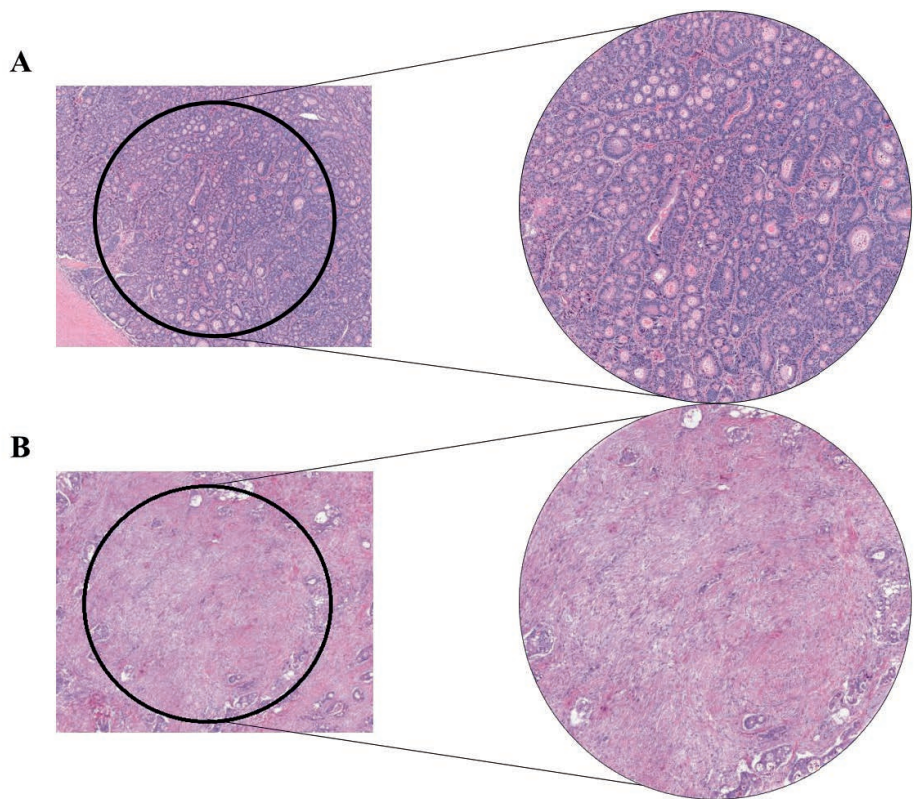
1. Brierley, J.D., Gospodarowicz, M.K., Wittekind, C., The TNM Classification of Malignant Tumours. 8 ed. 2016: Wiley Blackwell.
2. Argiles, G., et al., Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2020. 31(10): p. 1291-1305.
3. Baxter, N.N., et al., Adjuvant Therapy for Stage II Colon Cancer: ASCO Guideline Update. *J Clin Oncol*, 2022. 40(8): p. 892-910.
4. Cervantes, A., et al., Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*, 2023. 34(1): p. 10-32.
5. Weiser, M.R., AJCC 8th Edition: Colorectal Cancer. *Ann Surg Oncol*, 2018. 25(6): p. 1454-1455.
6. Costas-Chavarri, A., et al., Treatment of Patients With Early-Stage Colorectal Cancer: ASCO Resource-Stratified Guideline. *Journal of Global Oncology*, 2019(5): p. 1-19.
7. Lugli, A., et al., Recommendations for reporting tumour budding in colorectal cancer based on the International Tumour Budding Consensus Conference (ITBCC) 2016. *Mod Pathol*, 2017. 30(9): p. 1299-1311.
8. Roth, A.D., et al., Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *J Natl Cancer Inst*, 2012. 104(21): p. 1635-46.
9. Sarkar, M., et al., Cancer-associated fibroblasts: The chief architect in the tumour microenvironment. *Front Cell Dev Biol*, 2023. 11: p. 1089068.
10. Ueno, H., et al., Histologic categorization of desmoplastic reaction: its relevance to the colorectal cancer microenvironment and prognosis. *Ann Surg Oncol*, 2015. 22(5): p. 1504-12.
11. El Sissy, C., et al., Therapeutic Implications of the Immunoscore in Patients with Colorectal Cancer. *Cancers (Basel)*, 2021. 13(6).
12. Joshi, R.S., et al., The Role of Cancer-Associated Fibroblasts in Tumour Progression. *Cancers (Basel)*, 2021. 13(6).
13. Park, J.H., et al., The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Ann Oncol*, 2014. 25(3): p. 644-651.
14. Sandberg, T.P., et al., Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumour-stroma ratio in colorectal cancer. *BMC Cancer*, 2019. 19(1): p. 284.
15. Mesker, W.E., et al., The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumour stage. *Cell Oncol*, 2007. 29(5): p. 387-98.
16. Mesker, W.E., et al., Presence of a high amount of stroma and downregulation of SMAD4 predict for worse survival for stage I-II colon cancer patients. *Cell Oncol*, 2009. 31(3): p. 169-78.
17. Huijbers, A., et al., The proportion of tumour-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*, 2013. 24(1): p. 179-85.
18. Vangangelt, K.M.H., et al., The prognostic value of the tumour-stroma ratio is most discriminative in patients with grade III or triple-negative breast cancer. *Int J Cancer*, 2020. 146(8): p. 2296-2304.
19. Smit, M.A., et al., The prognostic value of the tumour-stroma ratio in squamous cell lung cancer, a cohort study. *Cancer Treat Res Commun*, 2020. 25: p. 100247.
20. van Pelt, G.W., et al., The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 2018. 73(2): p. 197-206.
21. Zunder, S.M., et al., Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage II and stage III colon cancer. *Br J Cancer*, 2018. 119(2): p. 164-169.
22. Smit, M., et al., Uniform Noting for International Application of the Tumour-Stroma Ratio as an Easy Diagnostic Tool: Protocol for a Multicenter Prospective Cohort Study. *JMIR Res Protoc*, 2019. 8(6): p. e13464.
23. Smit, M.A., et al., e-Learning for Instruction and to Improve Reproducibility of Scoring Tumour-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study. *JMIR Form Res*, 2021. 5(3): p. e19408.
24. van Pelt, G.W., et al., Scoring the tumour-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*, 2018. 473(4): p. 405-412.
25. EDC, C. Castor Electronic Data Capture. 2019 [cited 2019; Available from: <https://www.castoredc.com/>].
26. Derksen, J.W.G., et al., The Prospective Dutch Colorectal Cancer (PLCRC) cohort: real-world data facilitating research and clinical care. *Sci Rep*, 2021. 11(1): p. 3923.
27. Burbach, J.P., et al., Prospective Dutch colorectal cancer cohort: an infrastructure for long-term observational, prognostic, predictive and (randomized) intervention research. *Acta Oncol*, 2016. 55(11): p. 1273-1280.
28. Hagensaars, S.C., et al., Tumour-stroma ratio is associated with Miller-Payne score and pathological response to neoadjuvant chemotherapy in HER2-negative early breast cancer. *Int J Cancer*, 2021.
29. van Pelt, G.W., et al., The value of tumour-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol*, 2020. 20: p. 39-44.
30. Strous, M.T.A., et al., A high tumour-stroma ratio (TSR) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clin Transl Oncol*, 2022. 24(6): p. 1047-1058.
31. Iveson, T.J., et al., 3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial. *Lancet Oncol*, 2018. 19(4): p. 562-578.
32. Polack, M., et al., Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT. *Clin Transl Oncol*, 2022.
33. van Pelt, G.W., et al., Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *J Med Surg Pathol*, 2016. 1(02).

34. Vangangel, K.M.H., et al., The prognostic value of tumour-stroma ratio in tumour-positive axillary lymph nodes of breast cancer patients. *Int J Cancer*, 2018. 143(12): p. 3194-3200.
35. Schraa, S.J., et al., Cell-Free Circulating (Tumour) DNA before Surgery as a Prognostic Factor in Non-Metastatic Colorectal Cancer: A Systematic Review. *Cancers (Basel)*, 2022. 14(9).
36. Parikh, A.R., et al., Minimal Residual Disease Detection using a Plasma-only Circulating Tumour DNA Assay in Patients with Colorectal Cancer. *Clin Cancer Res*, 2021. 27(20): p. 5586-5594.
37. Kuruc, M., et al., New Strategies to Categorize Blood for Proteomic Biomarker Discovery. *J. Proteom. Bioinform*, 2020. 2: p. 90-107.
38. Ravensbergen, C.J., et al., The Stroma Liquid Biopsy Panel Contains a Stromal-Epithelial Gene Signature Ratio That Is Associated with the Histologic Tumour-Stroma Ratio and Predicts Survival in Colon Cancer. *Cancers (Basel)*, 2021. 14(1).
39. Guinney, J., et al., The consensus molecular subtypes of colorectal cancer. *Nat Med*, 2015. 21(11): p. 1350-6.
40. Sandberg, T.P., et al., Molecular profiling of colorectal tumours stratified by the histological tumour-stroma ratio - Increased expression of galectin-1 in tumours with high stromal content. *Oncotarget*, 2018. 9(59): p. 31502-31515.
41. Smit, M.A., et al., Tumour-stroma ratio outperforms tumour budding as biomarker in colon cancer: a cohort study. *Int J Colorectal Dis*, 2021. 36(12): p. 2729-2737.
42. Zunder, S., et al., Stromal organization as predictive biomarker for the treatment of colon cancer with adjuvant bevacizumab; a post-hoc analysis of the AVANT trial. *Cell Oncol (Dordr)*, 2019. 42(5): p. 717-725.
43. Pagès, F., et al., International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 2018. 391(10135): p. 2128-2139.
44. Ravensbergen, C.J., et al., Combined Assessment of the Tumour-Stroma Ratio and Tumour Immune Cell Infiltrate for Immune Checkpoint Inhibitor Therapy Response Prediction in Colon Cancer. *Cells*, 2021. 10(11).
45. Yin, J., et al., Reevaluating Disease-Free Survival as an Endpoint vs Overall Survival in Stage III Adjuvant Colon Cancer Trials. *J Natl Cancer Inst*, 2022. 114(1): p. 60-67.
46. Smit, M.A., et al., Deep learning based tumour-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform*, 2023. 14: p. 100191.
47. Valkenburg, K.C., A.E. de Groot, and K.J. Pienta, Targeting the tumour stroma to improve cancer therapy. *Nat Rev Clin Oncol*, 2018. 15(6): p. 366-381.

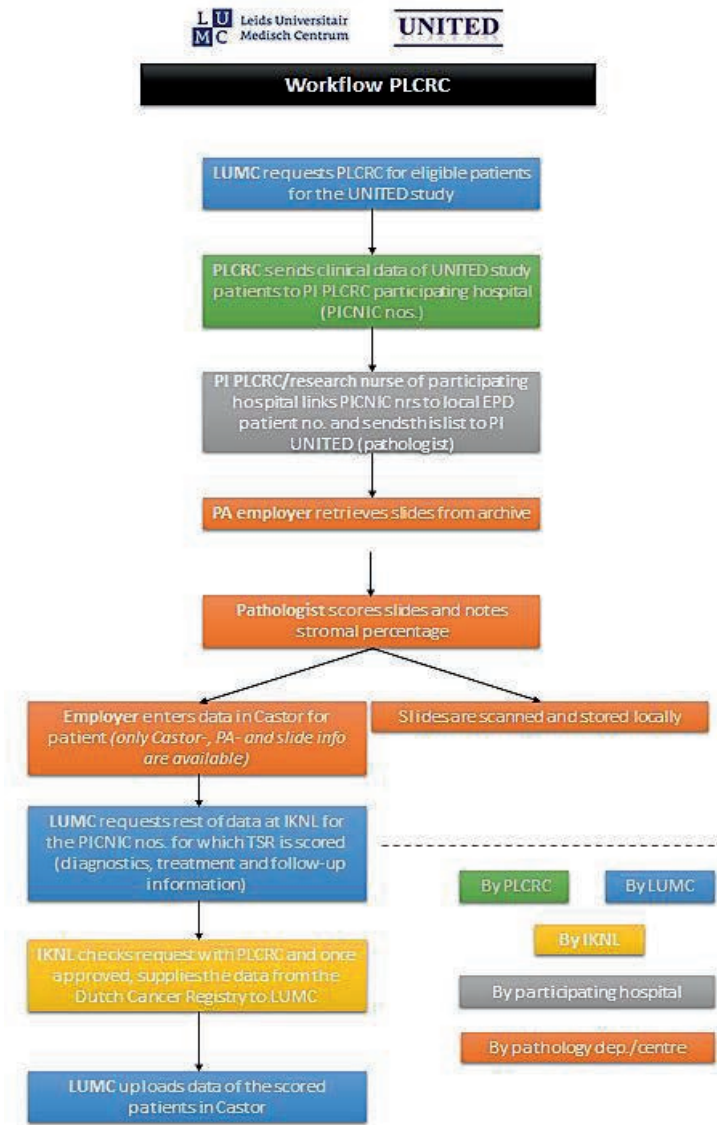
Supplementary Material

Supplementary Table 1. Final inclusion and exclusion criteria of the UNITED study

UNITED study	Criteria
Inclusion criteria (<i>eligible for inclusion and follow-up</i>)	<ul style="list-style-type: none"> • Operation date after 2015 • Histologically proven colon cancer • Pathological stage II (T3-4, N0, M0) or III (every T, N1-2, M0) • Tumour-stroma ratio score • Age ≥ 18 years and signed informed consent
Exclusion criteria (<i>ineligible</i>)	<ul style="list-style-type: none"> • Neoadjuvant therapy • Other malignancy 10 year prior to current colon cancer (except basal cell carcinoma or cervix carcinoma in situ) or in the complete medical history a colon carcinoma • Rectal cancer • Multiple synchronous malignant colon cancer • No complete curative resection (R1 or R2 resection) • Postoperative mortality within 3 months of operation



Supplementary Figure S1. Zoomed in examples of the scoring of the tumour-stroma ratio on haematoxylin and eosin-stained slides. (A) Stroma-low ($\leq 50\%$) colon cancer; (B) Stroma-high ($> 50\%$) colon cancer. 10x magnification



Supplementary Figure S2. Workflow patient inclusion through PLCRC collaboration.

PA, Pathology; PI, Principle investigator; PLCRC, Prospective Dutch ColoRectal Cancer cohort; LUMC, Leiden University Medical Centre; IKNL, Institute for Cancer in the Netherlands.

Supplementary Table 2. Participating centres and collaborative investigators with associated inclusion rates

No. inclusions	Location (country)	Institute	Investigator (department)
230	Skopje (Macedonia)	Medical Faculty of Ss. Cyril and Methodius University	Gordana Petrushevska# (pathology) Magdalena Bogdanovska (pathology) Panche Zdravkoski (pathology) Svetozar Antovic (surgery) Darko Dzambaz (surgery) Panche Karagjovov (surgery)
133	Rotterdam (Netherlands) region	PATHAN Laboratories\$ (Franciscus Gasthuis & Vlietland; Admiraal De Ruyter Ziekenhuis; Usselland Ziekenhuis)	Erienne M. V. de Cuba## (pathology) Frédérique Beverdam# (surgery; Franciscus Gasthuis & Vlietland) Jan Jansen# (surgery; Admiraal de Ruyter Ziekenhuis) Maarten Vermaas# (surgery; Usselland Ziekenhuis)
117	Ljubljana (Slovenia)	Onkološki Inštitut	Gorana Gašljević# (pathology)
114	Vejle (Denmark)	Vejle Sygehus – Sygehus Lillebælt	Sanne Kjer-Frifeld# (pathology) Jan Lindebjerg (pathology)
111	Venlo (Netherlands)	VieCuri Medisch Centrum	Maud Strous# (pathology) Jeroen F. Vogelaar (surgery)
101	Hoofddorp and Haarlem (Netherlands)	Spaarne Gasthuis\$	Nicole W.J. Bulkman## (pathology)
83	Enschede region (Netherlands)	LabPON\$ (Medisch Spectrum Twente; Ziekenhuisgroep Twente; ZorgSaam Terneuzen)	Joop van Baarlen# (pathology; retired)

			Leonie Mekenkamp (medical oncology; Medisch Spectrum Twente) Ronald Hoekstra (medical oncology; Ziekenhuisgroep Twente) Mark Sie (medical oncology; ZorgSaam Terneuzen)
80	Barcelona (Spain)	Hospital Clinic	Miriam Cuatrecasas# (pathology) Sara Simonetti (pathology) María Teresa Rodrigo (pathology) Iván Archilla Sanz (pathology) Jose Guerrero Pineda (pathology)
75	Deventer (Netherlands)	Deventer Ziekenhuis\$	Natalja E. Leeuwis-Fedorovich# (pathology) Koen A. Talsma (surgery)
70	João Pessoa (Brazil)	Napoleão Laureano Hospital	Ricella M. Souza da Silva# (pathology)
57	Utrecht (Netherlands)	Universitair Medisch Centrum Utrecht\$	Miangela M. Lacle# (pathology) Miriam Koopman (medical oncology)
55	Delft (Netherlands)	Reinier de Graaf Gasthuis\$	Jan Willem T. Dekker# (surgery) Arjan van Tilburg (pathology)
53	Barcelona (Spain)	Vall d'Hebron Institute of Oncology	Paolo Nuciforo# (pathology) Xenia Villalobos Alberú (pathology) Stefania Landolfi (pathology) Adriana Zucchiatti (pathology)
42	Alkmaar (Netherlands)	Symbiant Laboratoires\$ (Noordwest Ziekenhuisgroep Alkmaar)	Emma Witteveen# (pathology) Arad Bordbar (pathology) Mathijs P. Hendriks (medical oncology)

41	Amersfoort (Netherlands)	Meander Medisch Centrum\$	René Arensman# (pathology)
38	Hardwick (United Kingdom)	University Hospital of North Tees	Shonali Natu# (pathology)
34	Glasgow regio (United Kingdom)	NHS Greater Glasgow and Clyde	Noori Maka# (pathology)
28	Leiden (Netherlands)	Leids Universitair Medisch Centrum	Wilma E. Mesker# (surgery) Rob A.E.M. Tollenaar (surgery) Meaghan Polack (surgery) Marloes A. Smit (surgery) Gabi W. van Pelt (surgery) Hein Putter (biomedical data sciences) Elma Meershoek-Kleinenbarg (clinical research center, surgery) Annet G.H. Roodvoets (clinical research center, surgery) Augustinus S.L.P. Crobach (pathology) Hans Gelderblom (medical oncology)
28	Lisbon (Portugal)	Hospital CUF Tejo	Mário Fontes e Sousa# (medical oncology) Paula Borralho Nunes (pathology) João Cruz (pathology) Ana Raimundo (medical oncology) Nelson Silva (surgery)
24	Almada (Portugal)	Hospital Garcia de Orta	Maria J. Brito# (pathology)
19	The Hague (Netherlands)	Haaglanden Medisch Centrum	Valeska Terpstra# (pathology)
4	Kiev (Ukraine)	Bogomolets - Kyiv Oncology Center	L.M. Zakhartseva (pathology)

N/A	Brussels (Belgium)	European Society for Pathology	Raed Al Dieri (pathology) Jean-François Fléjou (pathology) Roger Feakins (pathology) Els Dequeker (pathology)
N/A	Utrecht (Netherlands)	Netherlands Comprehensive Cancer Organisation (IKNL)\$	Geraldine R. Vink (research and development)
N/A	Nijmegen (Netherlands)	Radboud University Medical Center	J. Han J.M. van Krieken (pathology)

N/A, Not applicable.

Local principal investigator.

\$ Part of the PLCRC collaboration.

* Currently employed elsewhere.

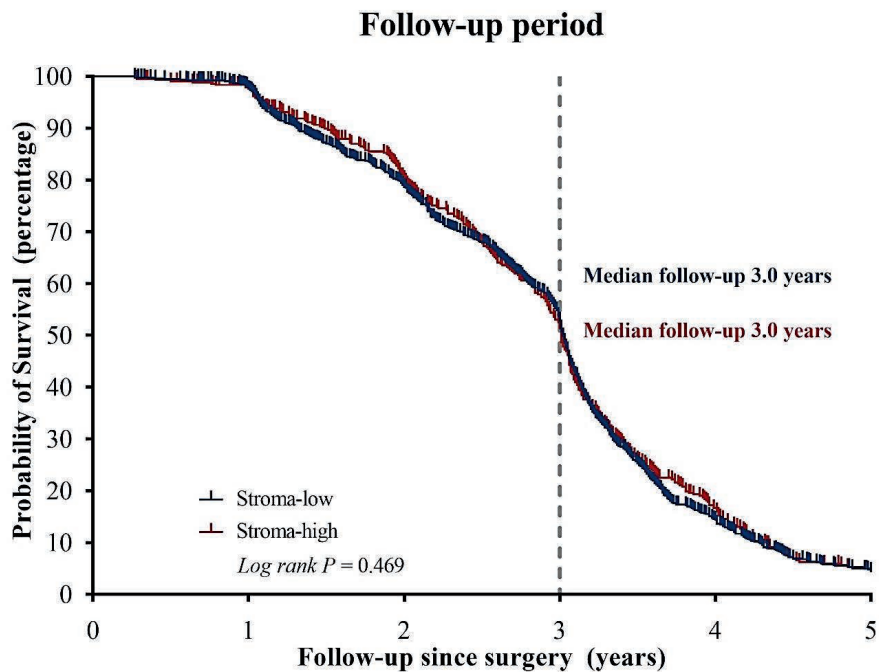
Supplementary Table 3. Baseline characteristics of the total UNITED cohort

Baseline characteristics	Total UNITED cohort (N=1,537)
Sex	
Female	686 (45)
Male	851 (55)
Age at surgery – years	
Median age	70 (61 – 77)
≥75 years of age	509 (33)
Ineligible at registration	31 (2)
Biopsy taken	
Yes	1,353 (88)
No*	179 (12)
Rectal cancer	3 (0)
Surgery	
Surgery year	2019 (2018 – 2020)
Surgery before 2015	1 (0)
No surgery	1 (0)
Pathological TNM-stage	
Stage 0 - I	42 (3)
Stage II	756 (49)
Stage III	694 (45)
Stage IV	9 (1)
Multiple colon tumours	21 (1)
Lymph nodes - number	
Examined (<i>in total group</i>)	20 (14 – 28)
Positive (<i>in pTNM-stage III</i>)**	2 (0 – 33)
Tumour-stroma ratio	
Stroma-low (≤50%)	969 (63)
Stroma-high (>50%)	433 (28)
Residual tumour	2 (0)
Missing slide	4 (0)
Follow-up	
Postoperative mortality <3 months	26 (2)
Not started follow-up	5 (0)
Withdrew consent	4 (0)
Included in final analysis	1,388 (90)

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding. TNM, tumour-node-metastasis stage.

* Reasons why biopsy was not taken, is almost always in emergency setting (obstructive ileus).

**Although there are no positive lymph nodes, using the UICC version 8, a tumour deposit (leading to stage N1c) will also lead to a pathological TNM-stage III.



Supplementary Figure S3. Median follow-up period.

Follow-up times calculated with reverse Kaplan-Meier analysis and log rank test, showing similar follow-up curves for the stroma-high group (median follow-up 3.0 years, 95% confidence interval 2.9 – 3.1) and stroma-low (median follow-up 3.0 years, 95% confidence interval 3.0 – 3.1) ($P=0.469$).

Supplementary Table 4. General overview of patient outcomes

Patient outcome	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Follow-up time – years			0.469\$
Median follow-up time	3.0 (3.0 – 3.1)	3.0 (2.9 – 3.1)	
Mutational status determined			
Not determined	864 (90)	371 (87)	
Determined, of which	96 (10)	57 (13)	0.068#
No mutations	38 (40)	16 (28)	
Mutations present, of which*	58 (60)	41 (72)	0.150#
<i>KRAS</i>	21 (36)	21 (51)	
<i>BRAF</i>	31 (53)	16 (39)	
<i>NRAS</i>	1 (2)	2 (5)	
<i>PIK3A</i>	4 (7)	2 (5)	
<i>TP53</i>	6 (10)	8 (20)	
Other	8 (14)	7 (17)	
Disease-free survival – years			
Disease-free survival time	5.2 (5.0 – 5.3)	4.8 (4.4 – 5.1)	<0.001‡
Stage and treatment group			<0.001#
Stage II - No adjuvant therapy	434 (45)	125 (29)	
<i>Number of events</i>	47 (11)	27 (22)	
Stage II + Adjuvant therapy	107 (11)	57 (13)	
<i>Number of events</i>	9 (8)	14 (25)	
Stage III + Adjuvant therapy	314 (33)	197 (46)	
<i>Number of events</i>	63 (21)	60 (31)	
Stage III - No adjuvant therapy	105 (11)	49 (11)	
<i>Number of events</i>	44 (42)	22 (45)	
Disease-free status			
No event	797 (83)	305 (71)	<0.001#
Event	163 (17)	123 (29)	
Type of event			<0.001#
No event	797 (83)	305 (71)	
Death by any cause	39 (4)	15 (4)	
Distant metastasis	105 (11)	92 (22)	
Locoregional recurrence	10 (1)	7 (2)	
Simultaneous distant metastasis and locoregional recurrence	9 (1)	9 (2)	

<i>(continued)</i> Patient outcome	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Location of distant metastasis			0.007#
Liver	37 (35)	26 (29)	
Lung	22 (21)	6 (7)	
Liver and lung	14 (13)	6 (7)	
Bone	2 (2)	1 (1)	
Brain	1 (1)	3 (3)	
Peritoneal metastases	10 (10)	12 (13)	
Abdominal lymph nodes	2 (2)	2 (2)	
Two or more locations	16 (15)	29 (32)	
Other	1 (1)	5 (6)	
Overall survival – years			
Overall survival time	5.5 (5.3 – 5.7)	5.6 (5.3 – 5.9)	0.102‡
Overall survival status			0.053#
Alive	858 (89)	367 (86)	
Died	102 (11)	61 (14)	
Cause of death (Metastases of) current colon cancer			0.020#
	58 (57)	48 (79)	
Second primary malignancy	5 (5)	4 (7)	
Other, including pre-existing comorbidity	27 (26)	6 (10)	
Missing	12 (12)	3 (5)	

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. N/A, not applicable.

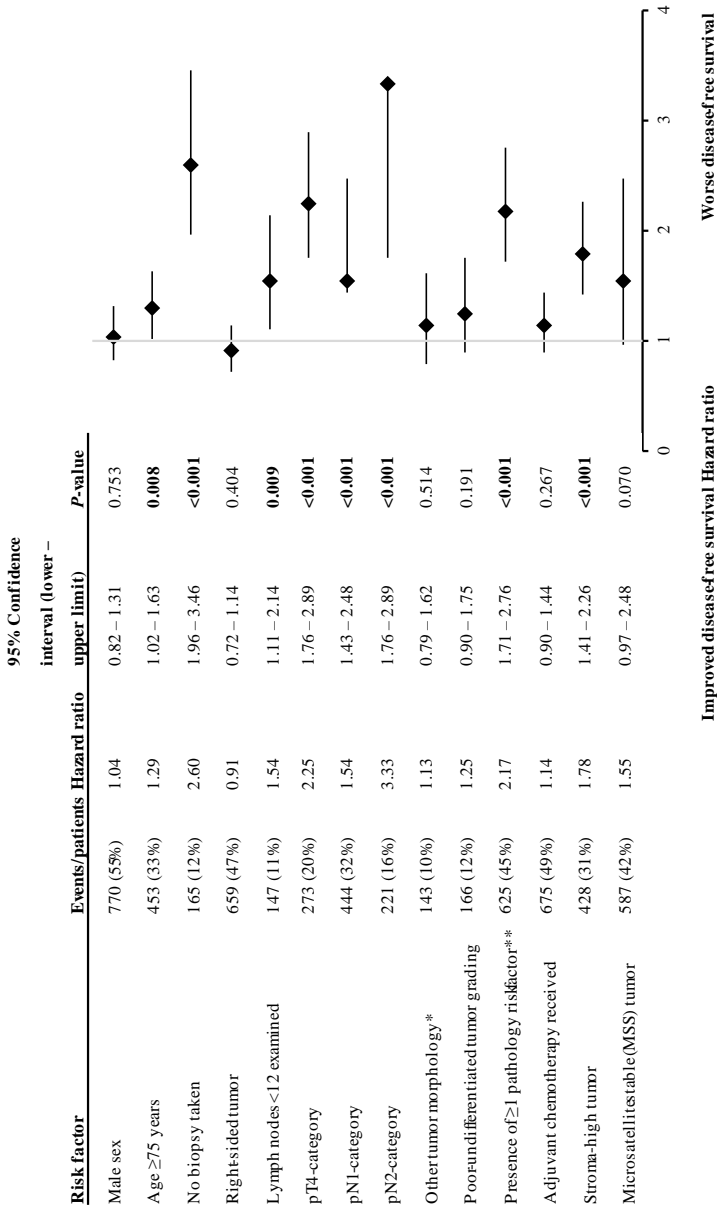
*Multiple mutations can occur simultaneously, hence the number of added percentages can be higher than 100 and no analysis is performed.

Calculated with the Chi-square test.

\$ Calculated with a reverse Kaplan-Meier analysis and log rank test.

‡ Calculated with Kaplan-Meier analysis and log rank test.

Supplementary Figure S4. Forest plot of the effect of risk factors on disease-free survival (univariate Cox regression analysis).

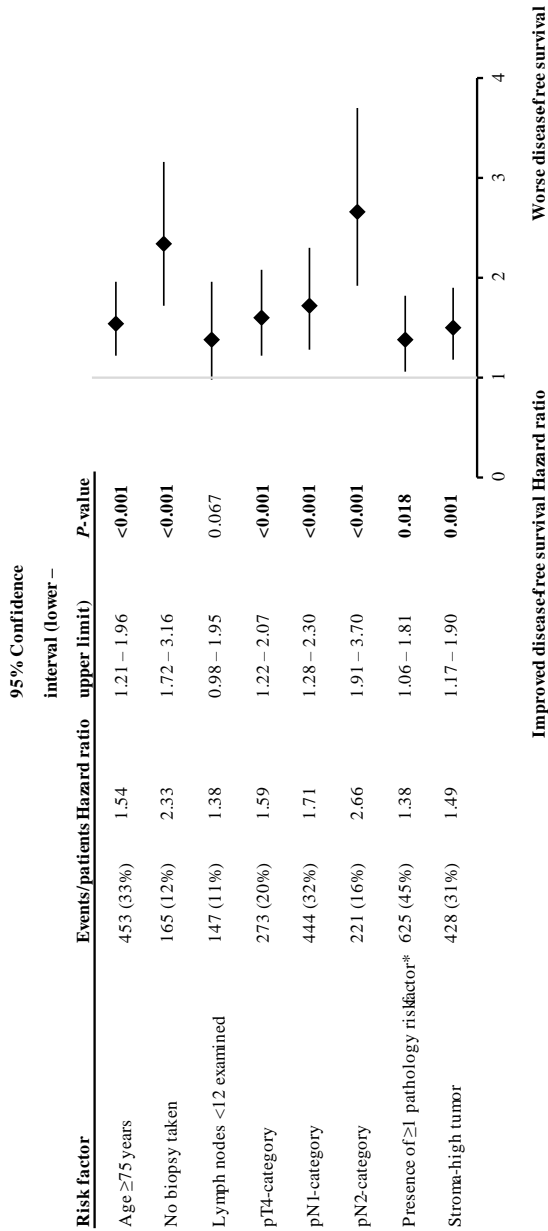


* Other tumour morphology include signet cell carcinoma or medullary carcinoma.

** Pathology risk factors are stated below, presence of a risk factor is defined as at least one of registered risk factors.

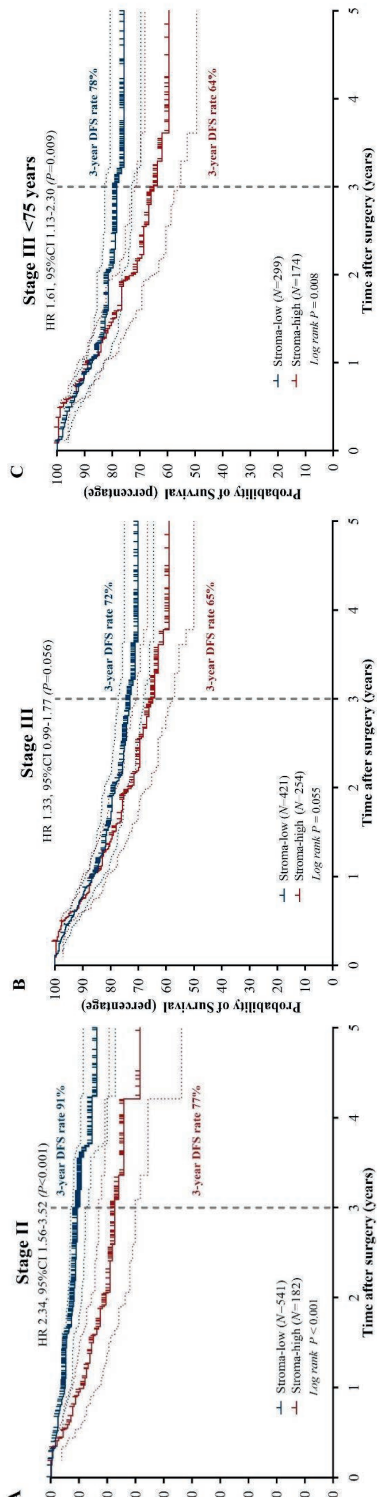
Absence is the absence of registered risk factors, as not all risk factors are registered. Risk factors include extramural vascular invasion (EMVI), perineural invasion (PnI), etc.

Supplementary Figure S5. Forest plot of the effect of risk factors on disease-free survival using significant variables from the univariate analysis (multivariate Cox regression analysis).



* Pathology risk factors are stated below, presence of a risk factor is defined as at least one of registered risk factors.

Absence is the absence of registered risk factors, as not all risk factors are registered. Risk factors include extramural vascular invasion (EMVI), perineural invasion (PnI), etc.



Supplementary Figure S6. TSR on Disease-free survival per TNM-stage.

Kaplan-Meier curve and log rank analysis of TSR category and plotted 95% confidence intervals. A) Stage II with 3-year survival rates of 77% vs. 91%, respectively ($P < 0.001$); B) Stage III, with 3-year survival rates of 65% vs. 72%, respectively ($P = 0.055$). TNM-stage III is only nearly significant due to bias through high number of patients with comorbidities or high age and not treated with standard adjuvant chemotherapy. C) After stratifying for age <75 years, stage III stroma-high colon cancer also leads to significantly worse disease-free survival, with 3-year rates of 64% vs. 78% ($P = 0.008$).

TNM, tumour-node-metastasis stage; TSR, tumour-stroma ratio.

Supplementary Table 5. Detailed overview of the adjuvant chemotherapy regimens and durations

Variables	Total eligible (N=1,388)	Stroma-low (N=960)	Stroma-high (N=428)	P-value
Adjuvant chemotherapy – not started				
Total not started adjuvant treatment	713 (51)	539 (56)	174 (41)	<0.001#
Stage II	559 (78)	434 (80)	125 (72)	<0.001#
Stage III	154 (22)	105 (20)	49 (28)	<0.001#
Reasons not started adjuvant treatment*				N/A
Not indicated	509 (71)	401 (76)	108 (62)	
Comorbidity, age	100 (14)	70 (14)	30 (17)	
Other, including patients wish	88 (16)	56 (10)	32 (18)	
Missing	20 (3)	15 (3)	5 (3)	
Adjuvant chemotherapy – started				
Total started treatment	675 (49)	421 (44)	254 (59)	<0.001#
Stage II	164 (24)	107 (26)	57 (22)	<0.001#
Stage III	511 (76)	314 (74)	197 (78)	<0.001#
Regimen started adjuvant chemotherapy**				N/A
CAPOX/XELOX***	394 (58)	238 (57)	156 (61)	
Capecitabine monotherapy	173 (26)	112 (27)	61 (24)	
FOLFOX	33 (5)	26 (6)	7 (3)	
Other, including 5FU monotherapy	32 (5)	18 (4)	14 (6)	
Missing	43 (6)	27 (6)	16 (6)	
Duration adjuvant chemotherapy – months				0.118\$
Median duration	3 (2 - 5)	3 (2 - 5)	2 (2 - 5)	
Missing	52 (8)	32 (8)	20 (8)	

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. 5FU, 5-fluorouracil intravenous chemotherapy; CAPOX/XELOX, oral capecitabine with intravenous oxaliplatin chemotherapy; FOLFOX, intravenous 5-fluorouracil and oxaliplatin; TNM, tumour-node-metastasis stage.

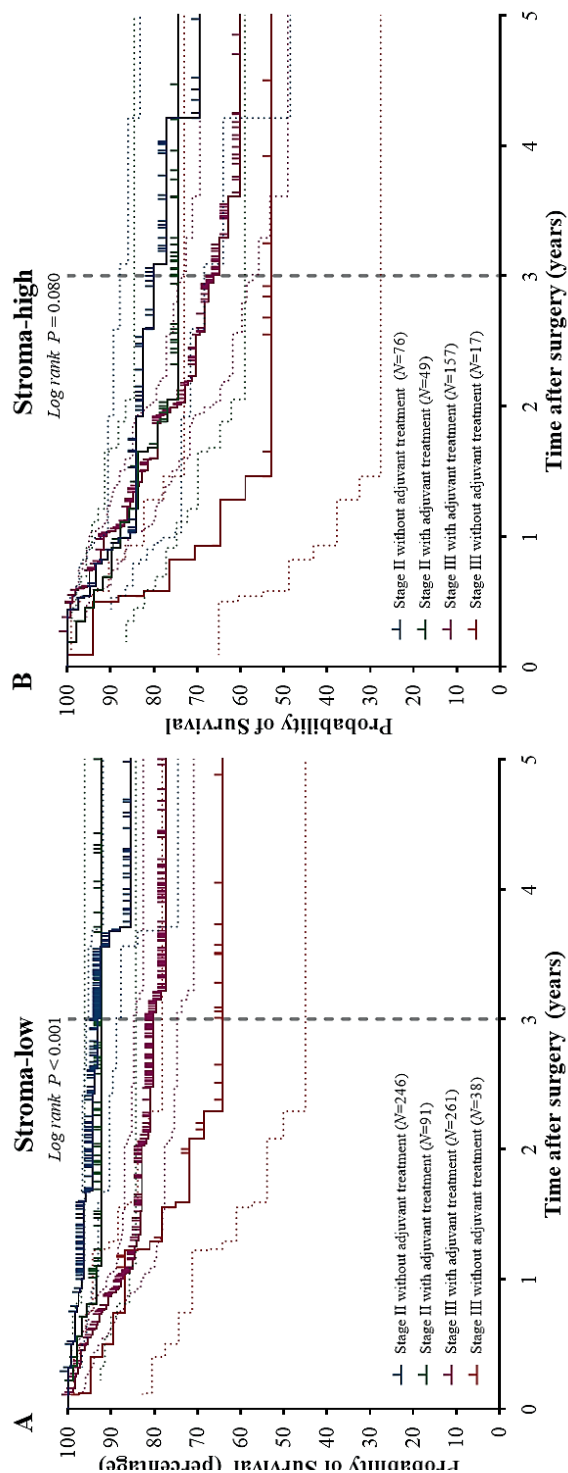
*Multiple reasons can occur, hence the number of added percentages can be higher than 100. The first given reason is shown here.

**Switch in regimens can occur, hence the number of added percentages can be higher than 100. The initially started regimen is mostly shown here.

***CAPOX/XELOX regimens here are both the 3 and 6 cycles.

Calculated with the Chi-square test.

\$ Calculated with an Independent Student's T-test.



Supplementary Figure S7. Disease-free survival per TSR category and benefit of adjuvant treatment per TNM-stage.

Kaplan-Meier curve and log rank analysis of TSR category and plotted 95% confidence intervals. A) Stroma-low patient groups showing an significant influence of adjuvant treatment, with 3-year survival rates of 93% vs. 80% vs. 64%, respectively ($P < 0.001$); B) Stroma-high patient groups showing worse outcomes than stroma-low groups in A but also within groups no significant difference despite adjuvant treatment, with 3-year survival rates of 80% vs. 73% vs. 66% vs. 52% ($P = 0.080$). In TNM-stage III patients not receiving adjuvant treatment, bias occurs e.g. due to small numbers.

TNM, tumour-node-metastasis stage; TSR, tumour-stroma ratio.

Supplementary Table 6. Overview of ASCO criteria and TSR categories

Variables	Total no. (%)	No. events (%)	P-value
Stage II - No adjuvant chemotherapy			<0.001#
Patients <75 years of age*	322 (100)	35 (11)	
Tumour-stroma ratio			0.001#
Stroma-low	246 (76)	19 (8)	
Stroma-high	76 (24)	16 (21)	
ASCO-criteria**			0.383#
ASCO low-risk	199 (62)	24 (12)	
ASCO high-risk	123 (38)	11 (9)	

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. ASCO, American Society for Clinical Oncology; TSR, tumour-stroma ratio.

*Compared to the category as defined in Supplementary Table 4.

**ASCO-criteria include a pT4 tumour, sampling of <12 lymph nodes or emergency setting of surgery, presence of pathological risk factors like lymphovascular or perineural invasion, and poor tumour differentiation as risk factors. If one is present, categorization as ASCO high-risk followed.

Calculated with the Chi-square test.



Chapter 3

The tumour-stroma ratio as predictive aid towards a biopsy-based treatment strategy in rectal carcinoma

Meaghan Polack, Gabi W. van Pelt, Davita H. van den
Heuvel, Elma Meershoek Klein-Kranenburg, Annet G.H.
Roodvoets, Hein Putter, Augustinus S.L.P. Crobach,
Iris D. Nagtegaal, Koen C.M.J. Peeters, Rob A.E.M.
Tollenaar, J. Han J.M. van Krieken, Wilma E. Mesker

Abstract

Aims: Tumour-stroma ratio (TSR) scores of biopsy material in rectal carcinoma (RC) could aid a biomarker-based, upfront and personalized treatment strategy selection for RC patients. In a large retrospective, multicentre cohort, we aimed to validate the predictive value of biopsy-scored TSR on neoadjuvant therapy response, and secondary, disease-free and overall survival (DFS, OS).

Methods and results: Scanned haematoxylin-and-eosin-stained RC biopsy slides were collected from LUMC (N=116) and from the clinical PROCTOR-SCRIPT (N=142) and RAPIDO (N=271) trials. TSR was scored per protocol and categorized as stroma-low ($\leq 50\%$) or stroma-high ($> 50\%$). Major response was defined as tumour regression grade (TRG) 1+2 by Mandard, including pathological complete response. Ultimately, a large and varied cohort with 373 RC patients was established. Locally advanced RC were more often stroma-high ($P < 0.001$). We subsequently observed significantly less major response rates in the stroma-high RC after a neoadjuvant treatment approach (hazard ratio 0.63, 95% confidence interval 0.41 – 0.99; $P = 0.044$). Despite correction for well-known risk factors in Cox hazard regression analysis, such as (y)pTNM-substages or residual tumour status, the TSR had no singular significant influence on DFS nor OS in multivariate analysis ($P = 0.438$; $P = 0.934$, respectively).

Conclusions: Biopsy-scored TSR can predict neoadjuvant therapy efficacy, as RC patients with stroma-high biopsies achieve less major response. Patient survival, however, is multifactorial, although response is an important predictor, influenced by TSR. Scoring TSR on RC biopsy material is a reliable histological parameter, implementation of which in treatment guidelines could aid to improve upfront selection for a watch-and-wait strategy.

Introduction

Optimalisation of therapeutic strategies in rectal carcinoma (RC) has been subject to many clinical trials over the years [1]. Management of RC has evolved, shifting the paradigm from initially aiming for enhanced locoregional control to whole organ preservation, rapidly improving patient-related outcomes like disease-free and overall survival (DFS and OS, respectively) [1, 2]. Currently, the cornerstone of international treatment guidelines encompasses risk stratification, based on disease extent as defined by the tumour-node-metastasis (TNM) classification [3], and clinical imaging factors like mesorectal fascia (MRF) involvement and/or extramural vascular invasion (EMVI) [4, 5]. Implementation of total mesorectal excision (TME) surgery [6] and preoperative treatment regimens [7-11] have led to optimal local control.

Starting with short course radiotherapy (SCRT) for improvement of locoregional control and survival [7], now, regimens including preoperative radiotherapy and chemotherapy, i.e. neoadjuvant therapy [10-12], has given rise to the watch-and-wait strategy [13], delaying and even potentially sparing patients burdensome surgery. However, response rates are prone to variation [14]. Moreover, heterogeneity is observed in reached clinical and pathological complete response (cCR and pCR, respectively) [10, 15-17]. With increasing RC incidence [4, 18], and high rates of treatment complications [4, 19], it is thus pivotal to improve upfront treatment selection. Current pathological risk factor parameters however focus mainly on the tumour epithelial compartment, i.e. neoplastic cells [4].

Convincing evidence is emerging that elements of the tumour microenvironment, especially the tumour stroma, are of detrimental influence of tumour behaviour, promoting tumour invasion and metastasis [20, 21]. Capturing this effect, the tumour-stroma ratio (TSR) is a robust and cost-effective histopathological parameter based on intratumoural stromal percentages [22]. The TSR has been validated as an independent biomarker in multiple tumour types: indeed, stroma-high (>50% stroma) gastrointestinal tract carcinomas not only have worse OS and DFS [23-25], as observed in colon carcinoma in the recently published prospective international UNITED study [26, 27], but also predicts worse response to (neo)adjuvant therapy [28-31]. Albeit often collectively termed, RC is a different entity than colon carcinoma [4]. Literature on TSR in RC specifically is scarce and mostly consists of relatively older, single centred RC series with limited patients and/or treatment types [25, 32-34].

To address this knowledge gap, we integrated two prominent clinical trials, i.e. PROCTOR-SCRIPT [9] and RAPIDO [10], with our local cohort. This collaboration enabled us to create an extensive overview

of the TSR in varied RC patient populations. A more biomarker-based approach on biopsies is crucial to improve future selection of responders and the TSR could aid in this prediction as literature has shown [32, 33]. As primary endpoint, we assessed the correlation between TSR and neoadjuvant therapy response. This study analysed the predictive potential of biopsy-scored TSR on DFS and OS as secondary endpoints. We hypothesized that the more aggressive and resistant stroma-high tumours would reach a response less often, and would have worse DFS and OS compared to their stroma-low counterparts, potentially influencing patient selection for a watch-and-wait strategy in the future.

Methods

Patient cohorts

Our local cohort (Leiden University Medical Centre, LUMC; N=116) was combined with available material from two well-established, independent clinical validation cohorts, i.e. subgroups of the PROCTOR-SCRIPT [9] (N=142) and RAPIDO [10] (N=271) trials. Hence, a large series comprising various TNM-stages and treatment regimens was analysed. All cohorts included patients ≥ 18 years with given informed consent. Additional inclusion and exclusion criteria for the clinical trials are mentioned in previous reports [9, 10]. Summarizing, the PROCTOR-SCRIPT was a combined study assessing the role of adjuvant chemotherapy compared to observation in RC patients treated with neoadjuvant therapy consisting of (chemo)radiotherapy and TME, whereas the RAPIDO analysed different neoadjuvant regimes in locally advanced RC (LARC). The LUMC cohort consisted of consecutive patients with a variety of (neo)adjuvant therapy types, of which available material was collected from patients operated after year 2000 with stage I-III RC and no previous malignancy <10 years prior to current RC.

Of note, PROCTOR-SCRIPT patients were included postoperatively, hence clinical TNM-stage nor clinical risk factors were not registered in the study database, although one inclusion criterium was pathological TNM-stage II/III. Stage II/III was given in case no imaging is performed and locoregional extent of disease was uncertain, e.g. in earliest included patients. Moreover, as different versions of the TNM classification were used, for optimal grouping and comparison, all variables are converted to TNM version 5 (1997). Age was registered at randomisation (PROCTOR-SCRIPT and RAPIDO) or diagnosis (LUMC). Supplementary Table 1 gives a detailed overview of treatment types and regimens per cohort.

Materials and tumour-stroma analysis

Scanned haematoxylin-and-eosin (H&E)-stained slides of diagnostic biopsies were collected at LUMC (LUMC, PROCTOR-SCRIPT) or requested from Radboud University Medical Centre (Radboudumc; RAPIDO). At LUMC, slides were scanned with the Panoramic-250 scanner, Radboudumc used the Panoramic-1000 (3DHistech, Hungary; 20x magnification). Analysis was performed with 3DHistech CaseViewer software (v2.7). Two independent observers (MP-GvP: LUMC, PROCTOR-SCRIPT; MP-DH: RAPIDO) scored the TSR on biopsies according to van Pelt *et al.* [35], blinded for clinical data. Subsequently, categorisation in stroma-low ($\leq 50\%$ stroma) and stroma-high ($> 50\%$ stroma) followed (Figure 1, created in BioRender.com). Neoadjuvant therapy response was assessed on resection material through the tumour regression grade (TRG) in five categories as defined by Mandard [36] by local pathologists (LUMC, PROCTOR-SCRIPT) or in three groups (no-partial-complete response; RAPIDO). To ascertain the predictive correlation to TSR, TRG was dichotomized in clinically relevant and previously also maintained groups of TRG1+2 (including pCR and major responders) and TRG3-5 (non-major responders) [37].

Ethical considerations

Use of anonymised material from the studies PROCTOR-SCRIPT and RAPIDO were approved by the steering committees. These trials were conducted according to the Declaration of Helsinki (2013). All already available archival material and data of the LUMC cohort were coded and handled according to the Dutch National Ethical Guidelines (“Code of proper secondary use of human tissue”). No informed consent was necessary under the legislation for this retrospective analysis.

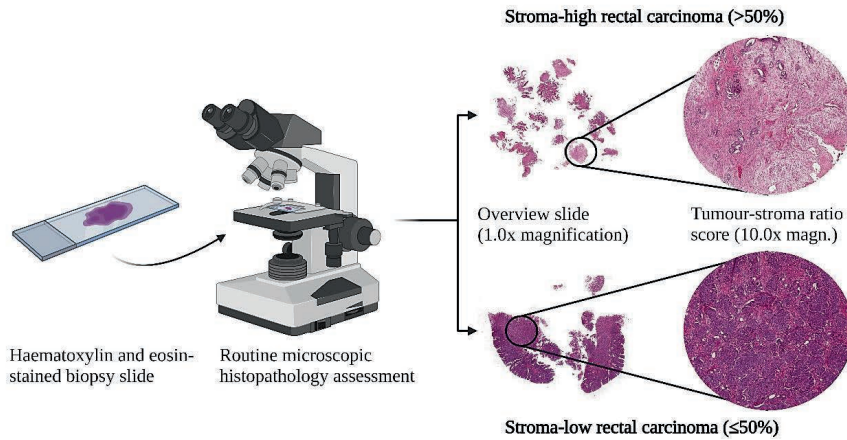


Figure 1. The process of scoring the tumour-stroma ratio (TSR) on haematoxylin and eosin-stained biopsy material using light microscopy.

First, using 1.0 – 2.5x magnification for a general overview of the complete slide, the area with the highest amount of tumour stroma is selected. Subsequently, the TSR is scored on a 10.0x magnifying objective, as per protocol of van Pelt *et al.* Finally, the biopsy is categorised as stroma-high (>50% intratumoural stroma; example shown above) or stroma-low (≤50% intratumoural stroma; example shown below). Created in BioRender.com.

Statistical analysis

Interobserver agreement Cohens kappa's were calculated between TSR biopsy scores. Assessment of prediction of TSR to neoadjuvant therapy response was done per TSR category and per therapy type subsequently. DFS was defined as period between date of surgery until any first event, i.e. recurrence (locoregional recurrence or distant metastasis), death (any cause) or until censoring. OS pertained to the period between date of surgery until death (any cause) or censoring. Censoring occurred when patients were disease-free and/or alive at last registration or after 10 years of follow-up.

Chi-square tests for nominal, Goodman Kruskal gamma statistics for ordinal, and Student's Independent T-tests for continuous variables were performed. Median follow-up time was calculated with the reversed Kaplan-Meier method, survival analyses were performed with Kaplan-Meier analyses and associated log-rank tests. Cox regression analysis for hazard ratio's (HR) with 95% confidence intervals

(CI) were calculated in univariate analysis for major response (event defined as pCR/TRG1+2), for the period between surgery and diagnosis (LUMC), first radiotherapy dose (PROCTOR-SCRIPT), or randomisation (RAPIDO). The variables significant of influence ($P < 0.05$) in univariate analysis, were included in the multivariate analysis.

Continuous variables were expressed in means with standard deviations (SD), whereas nominal and ordinal variables were stated as number of frequencies and corresponding percentages. Two-tailed P-values < 0.05 were considered statistically significant. Statistical analysis was performed using IBM SPSS Statistics (v29.0).

Results

Patient cohorts

Establishing the final patient population, exclusions followed from the three cohorts, e.g. per(i)operative pathological stage IV (N=18) or absence of pathological data (N=68). In total, 373 RC patients were ultimately included in this study (Figure 2). Baseline characteristics of the combined cohort and those separately are presented in Table 1. There are differences between cohorts, inherently correlating to the used studies, including treatment type ($P < 0.001$), clinical risk factors ($P < 0.001$) and clinical TNM-stage ($P < 0.001$). Most importantly, as could be expected due to the more aggressive nature of the tumours involved, the TSR is also already higher in the LARC patients of the RAPIDO, where 57% of patients were stroma-high, compared to approximately one-third in other cohorts and literature ($P < 0.001$). Hereby, the full spectrum of presentations of RC is covered in our study.

Overview TSR analyses

The TSR scores had high Cohen's interobserver agreement kappa's of 0.84 (MP – DH; RAPIDO biopsies) and 0.77 (MP – GP; LUMC and PROCTOR-SCRIPT biopsies). Supplementary Table 2 presents an overview of stroma-low compared to stroma-high clinical variables in the total patient population and per cohort separately. Overall, although more clinical risk factors were found in stroma-high patients as expected ($P = 0.014$), stroma-low patients had more often not undergone neoadjuvant therapy ($P < 0.001$) and were operated on in earlier years (2009 vs. 2012 in stroma-high patients; $P < 0.001$).

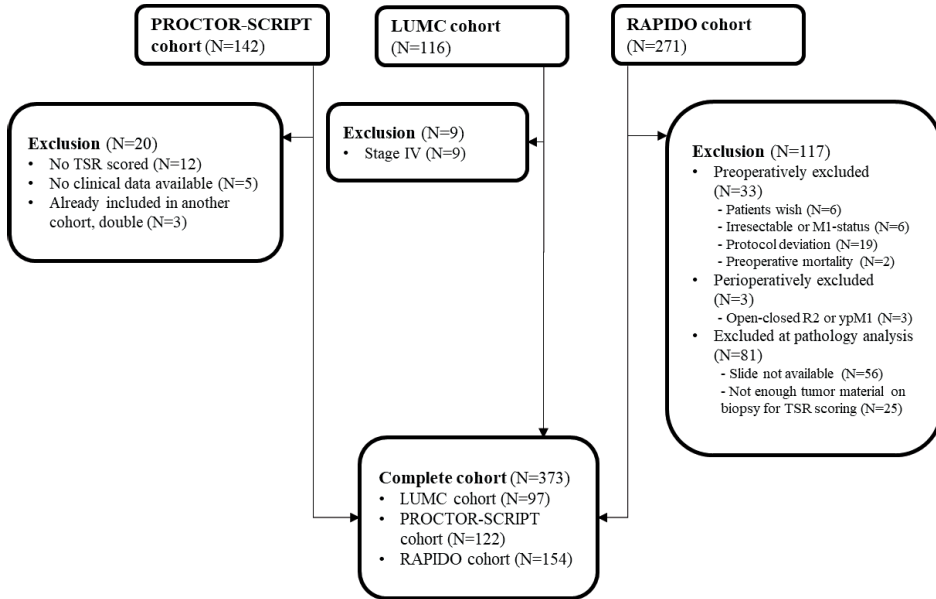


Figure 2. Flowchart showing the patient population with initial inclusion rates per separate cohort (PROCTOR-SCRIPT, LUMC or RAPIDO) and exclusion numbers and reasons, leading to the ultimately included final patient cohort (N=373). TSR, tumour-stroma ratio.

Table 1. Baseline characteristics of the eligible patients in the complete cohort and separate cohorts.

Baseline characteristics	Complete cohort (N=373)	LUMC (N=97)	PROCTOR-SCRIPT (N=122)	RAPIDO (N=154)	P-value
Participating centres	28	1	22	13	N/A
Total no. participating centres					
Sex					0.520#
Female	117 (31)	26 (27)	41 (34)	50 (33)	
Male	256 (69)	71 (73)	81 (66)	104 (68)	
Age					0.017\$ 0.018#
Median age (years)	62 (55-69)	65 (57-73)	60 (55-68)	64 (54-69)	
Age of >70 years	75 (20)	29 (30)	19 (16)	27 (18)	
Treatment type					<0.001#
Neoadjuvant treatment and surgery	289 (78)	74 (76)	64 (52)	152 (99)	
Neoadjuvant treatment, surgery and adjuvant treatment	64 (17)	3 (3)	58 (48)	2 (1)	
Surgery and adjuvant treatment	4 (1)	4 (4)	0 (0)	0 (0)	
Surgery alone	16 (4)	16 (17)	0 (0)	0 (0)	
Clinical TNM-stage					<0.001#
II/III	4 (1)	4 (4)	N/A	0 (0)	
II	42 (11)	28 (29)	N/A	14 (9)	
III	202 (54)	62 (64)	N/A	140 (91)	
Unknown	125 (34)	3 (3)	122 (100)	0 (0)	
Clinical Locally Advanced Rectal Carcinoma (LARC)					<0.001#
No, no clinical LARC	23 (6)	23 (24)	N/A	0 (0)	
Yes, clinical LARC	221 (59)	67 (69)	N/A	154 (0)	
Unknown	129 (35)	7 (7)	122 (100)	0 (0)	

<i>(continued)</i> Baseline characteristics	Complete cohort (N=373)	LUMC (N=97)	PROCTOR-SCRIPT (N=122)	RAPIDO (N=154)	P-value
Tumour location					0.463#
Low rectum (<5cm anal verge)	133 (36)	42 (43)	40 (34)	51 (33)	
Mid rectum (5-10 cm anal verge)	108 (29)	27 (28)	34 (28)	46 (31)	
High rectum (>10 cm anal verge)	126 (34)	27 (28)	44 (36)	55 (36)	
Unknown	6 (2)	1 (1)	4 (3)	2 (1)	
Clinical risk factors					<0.001#
No, no additional risk factors	24 (10)	24 (26)	N/A	0 (0)	
Yes, 1 clinical risk factor present	48 (20)	32 (35)	N/A	16 (10)	
Yes, 2 clinical risk factors present	85 (35)	19 (21)	N/A	66 (43)	
Yes, 3 or more clinical risk factors present	88 (36)	16 (18)	N/A	72 (47)	
Not (enough) determined, unknown	128 (34)	6 (6)	122 (100)	0 (0)	
Clinical risk factors					
Extramural venous invasion	19 (5)	1 (1)	N/A	18 (12)	
Mesorectal fascia involvement	95 (26)	16 (17)	N/A	79 (51)	
Clinical lateral lymph nodes	42 (11)	17 (18)	N/A	25 (16)	
Clinical T4-stage	38 (10)	7 (7)	N/A	31 (20)	
Clinical N+-stage	202 (54)	62 (64)	N/A	140 (91)	
Clinical N2-stage	122 (33)	22 (23)	N/A	100 (65)	
Tumour-stroma ratio biopsy					<0.001#
Stroma-low (≤50%)	215 (58)	69 (71)	80 (66)	66 (43)	
Stroma-high (>50%)	158 (42)	28 (29)	42 (34)	88 (57)	

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding.

N/A, not applicable; TNM, tumour-node-metastasis

Calculated with Chi-square test. \$ Calculated with one-way ANOVA analysis

TSR predictor of major response to neoadjuvant therapy

Pathological outcomes per TSR category of the total patient population and per therapy type separately are summarized in Supplementary Table 3. The CRT and RAPIDO regimens are combined in a large neoadjuvant treatment (NAT) group (N=182). Of note, SCRT was not intended to be used to reach pCR and only after more than 7 weeks after neoadjuvant therapy any significant downsizing can be seen[7, 38]. Hence, it almost mimics the situation in treatment-naïve patients: in stroma-high patients, more often a higher ypT-stage 2-4 (P=0.043) and less response was seen (P=0.044). Analysing major response rates per treatment type in-depth subsequently, the biopsy-scored TSR emerged as a valuable predictor, as stroma-high patients reached significantly less major response to NAT than stroma-low patients (HR 0.63, 95%CI 0.41 – 0.99; P=0.044) (Table 2).

Table 2. Univariate Cox regression analysis on the hazard of major response of TSR in total cohort and per treatment type.

Therapy type	Variable	Number major responders (%)	Major response - Univariate analysis*		
			Hazard ratio	95% Confidence interval	P-value
Complete cohort (N=337)**	Stroma-low (N=188)	39 (21)	1		0.071
	Stroma-high (N=149)	41 (28)	0.666	0.428 – 1.035	
NAT (CRT + RAPIDO; N=182)	Stroma-low (N=81)	38 (47)	1		0.044
	Stroma-high (N=101)	40 (40)	0.632	0.405 – 0.989	
CRT (N=107)	Stroma-low (N=56)	26 (46)	1		0.308
	Stroma-high (N=51)	17 (33)	0.726	0.393 – 1.343	
RAPIDO (N=75)	Stroma-low (N=25)	12 (48)	1		0.839
	Stroma-high (N=50)	23 (42)	0.930	0.461 – 1.877	

CRT, chemoradiation (25x1.8-2 Gray and capecitabine monotherapy); N/A, not applicable; NAT, neoadjuvant treatment; RAPIDO (5x5 Gray followed by 6 cycles capecitabine and oxaliplatin).

*The period between surgery (for pathology) and diagnosis (LUMC), first radiotherapy dose (PROCTOR-SCRIPT), or randomisation (RAPIDO). The event is defined as major response (TRG1+2/pathological complete response).

**Complete cohort here pertains to those patients who had undergone neoadjuvant therapy and with a known and determined response

Survival analyses

To assess the predictive value of the TSR on DFS and OS, Kaplan-Meier analysis with log rank tests were first performed for the complete cohort and separate therapy groups (Supplementary Table 4; Supplementary Figure 1-2). No significant influence of the TSR was observed here. Subsequently, we used Cox hazard regression in the complete cohort. In univariate analysis, higher (y)pT and/or (y)pN stages, as well as residual tumour status or not reaching a major response were so significantly of influence, it potentially introduced bias and the TSR reached no significance (DFS $P=0.800$; OS $P=0.856$) (Table 3). Aiming to analyse the effect of the TSR relatively to the other, well-known, risk factors, we added the TSR as variable in the multivariate analysis. However, even correcting for these variables, the TSR did not assert an additional influence here either (DFS $P=0.438$; OS $P=0.934$).

<i>(continued)</i>	Residual tumour (Wittekind 2009)	1 3.082	1.874-5.069	<0.001	1 1.354	0.569-3.220	0.493	1 3.367	1.907-5.946	<0.001	1 3.401	1.782-6.490	<0.001
	R0 resection R+ resection	1 2.449	1.466-4.092	<0.001	1 3.407	0.377-30.83	0.275	1 1.968	1.129-3.432	0.017	1 1.183	0.501-2.793	0.701
	Major response TRG1+2 TRG3-5	1 2.748	1.520-4.966	<0.001	1 N/A	N/A	0.973	1 2.065	1.105-3.860	0.023	1 1.521	0.575-4.020	0.398
	y(p)T-category y(p)T-stage 0+1 y(p)T-stage 2-4	1 2.748	1.520-4.966	<0.001	1 N/A	N/A	0.973	1 2.065	1.105-3.860	0.023	1 1.521	0.575-4.020	0.398
	Lymph nodes (LN) LN examined ≥12 LN examined <12	1 0.941	0.676-1.308	0.716	N/A			1 1.001	0.682-1.471	0.995	N/A		
	y(p)N-category y(p)N-stage 0 y(p)N-stage +	1 2.694	1.900-3.819	<0.001	1 1.678	0.748-3.764	0.210	1 2.152	1.452-3.187	<0.001	1 2.117	1.274-3.517	0.004
	Differentiation grade tumour Low grade (well-moderate) High grade (poor-undifferentiated)	1 1.124	0.631-2.002	0.691	N/A			1 1.564	0.849-2.882	0.152	N/A		
	Pathology risk factors No risk factors present Yes, 1 or more risk factors present	1 2.301	1.290-4.107	0.005	1 1.889	0.983-3.623	0.056	1 1.746	0.868-3.513	0.118	N/A		
	TSR biopsy Stroma-low Stroma-high	1 1.044	0.750-1.453	0.800	1 0.798	0.451-1.412	0.438	1 1.036	0.705-1.523	0.856	1 0.982	0.637-1.512	0.934

N/A, not applicable. NAT, neoadjuvant treatment. SCRT, short course radiotherapy. TNM, tumour-node-metastasis classification. TRG, tumour regression grade. TSR, tumour-stroma ratio. Disease-free survival and overall survival are censored at 10 years. Variables with a significant ($P \leq 0.050$) influence on outcome in univariate analysis are included in multivariate analysis.

Discussion

The present study set out to determine the predictive effect of biopsy-scored TSR on response to neoadjuvant therapy. In a large and varied multicentre patient population, we first observed that LARC, characterised through imaging as a more aggressive and invading tumour, was significantly more often categorised as stroma-high. We subsequently validated the TSR's predictive value on response. In RC patients undergoing a NAT approach, significantly less major response rates were achieved in those with a stroma-high biopsy. Identifying potential major therapy responders can ultimately aid upfront selection of patients for treatment strategies, e.g. watch-and-wait.

As secondary endpoint, the prognostic effect on DFS and OS of the TSR in RC patients was assessed. Despite correction for the significant influence of well-known risk factors like (y)pTNM-substages and residual tumour status, the TSR did not significantly assert a singular predictive influence on DFS or OS in our cohort. Response however, remained an important prognosticator, to which the TSR is a major contributor. This is most likely due to the presence of large multifactorial causal relations of various risk factors on survival, such as increasing sequential treatment options, underlining the need for a multidisciplinary and patient-tailored approach [4]. Moreover, underlying biological processes influencing response and tumour behaviour such as mutational status, are merely determined in a minority of included patients but increasingly analysed in current practices [39].

The past decades, the traditional neoplastic cell-centred view has incrementally been expanded to include the surrounding tumour microenvironment [40, 41]. The complex interaction between these entities has hence been subject to increasing research [20, 42]. In 2007, our research group first explored the absolute ratio of the tumour epithelium compartment compared to the stromal compartment: the TSR [23]. Since then, the field of the TSR has exponentially gained interest, worldwide [22, 24, 25, 28, 30, 43]. Recently, the detrimental influence of tumour stromal abundance on patient-related outcomes was proven in our UNITED study, where stroma-high colon carcinomas indeed led to worse DFS [26, 27]. This parameter has therefore been proposed [27] to be implemented in international guidelines, for instance in the TNM-classification.

Although the TSR had already previously proven to be of predictive value in biopsies and RC [32, 33, 44], large, novel and multicentre studies were lacking [34, 45]. As upfront selection of RC patients for personalised therapeutic strategies is gaining importance with the shift towards a watch-and-wait strategy [2, 4, 46], implementing biopsy-based biomarkers to aid this selection is pivotal. Potential

implications of our results, proving that biopsy-scored stroma-high RC patients are less prone to reach major response in a NAT approach, could be far-reaching. Including the TSR in the panel with all other important clinicopathological variables currently used by multidisciplinary meeting, could lead to an improved tailored approach, where in a shared-decision making setting, patients could even be advised to be spared the potentially less efficient though burdensome treatment and instead be recommended for immediate surgery, especially in frail elderly with comorbidities.

The largest strength but simultaneously also limitation of this study was inherent to the choice for a large and varied patient population; the differences between the cohorts prohibited large pooling as this could lead to skewed results. Of interest, it was for example observed that the RAPIDO trial, including only LARC patients, had significantly more stroma-high patients than the other cohorts with less advanced tumours. While this also emphasizes the importance of the stromal compartment and influence of its abundance on tumour invasiveness, it also gave rise to potential selection bias. Therefore, we aimed to answer our research questions in specific, though smaller, subgroups. Another limitation pertains to the retrospective aspect of this study. Moreover, despite the use of large clinical trials, these are still somewhat dated. With the rise of various neoadjuvant approaches [47], scoring more recent trials or even aiming to include the TSR scored on biopsies as parameter in prospective or randomized controlled trials in the future would therefore be imperative.

The current ESMO guidelines in management of RC state that there are several regression grade classifications including the Mandard [36] and the Dworak [48] classifications without indicating an optimal method for scoring tumour regression. Despite the fact that the Mandard classification may be not the most widely used for determining tumour regression and TRG has been described to potentially underestimate tumour shrinkage compared to e.g. a fragmentation pattern [49], many studies, including ours, use the Mandard score [32], which enhances potential comparison. In any case, a reliable score should include a 'complete-partial-no response' category [4]. Although the International Collaboration on Cancer Reporting (ICCR) colorectal pathology guidelines [50] mention another, the Ryan four-tier system [51], ultimately, all classifications for scoring result in similar endpoints.

Future studies including more parameters of the tumour microenvironment as well could enhance our understanding on the dynamic interplay between the various components of this entity. The immune cell component [52-54], tumour budding [55], or biomarkers such as circulating tumour DNA (ctDNA) and other liquid biopsy methods show promise in colon [56, 57] as well as RC [58], and could provide valuable insights. TSR is, in contrast to most biomarkers, uncomplicated and cost-effective [35], taught

by E-learning [59] and hence standardized with high interobserver agreement [60]. Furthermore, in this study, we collected scanned H&E stained biopsy slides. Artificial intelligence is exponentially being researched in pathology, nowadays successfully completing high-performance tasks like classification [61]. Future studies could use this collection for developing algorithms, e.g. supervised deep learning models trained by pathologists' annotations to characterise tissue types [62], resulting in automatised TSR quantification in biopsies, supporting pathologists' workload. Currently, some studies show promise in various TSR scoring methods, mostly performed on primary tumours in colon [63-65]. However, for RC, studies are still scarce [66], even less aiming to predict response from TSR in biopsies [67]. Moreover, unsupervised learning could potentially discover novel histological patterns on unannotated slides, improving personalized predictions on therapy response and survival [68].

We have hereby presented a large and multicentre study validating the predictive significance of biopsy-scored TSR on neoadjuvant therapy response in RC. Using well-established clinical trials, we conclude that patients with stroma-high RC biopsies will reach less major response in a NAT approach. The TSR should thus be implemented in routine pathology diagnostics and the current clinicopathological panel of parameters, which multidisciplinary meetings consider for RC patients in personalized selection of treatment strategies.

Acknowledgements

We thank the PROCTOR-SCRIPT and RAPIDO steering committees for use of their data and support. We also want to thank pathologist Arantza Farina Sarasqueta for aiding in the initial scoring procedures. We thank Sonay Kuş Öztürk (Radboudumc) for scanning of the RAPIDO slides, and Ronald van Vlierberghe and Geeske Dekker-Ensink (both LUMC) for scanning of the PROCTOR-SCRIPT and LUMC slides, as well as laboratory support.

Funding

This work was supported by the Bollenstreekfonds, Hillegom, the Netherlands (no grant number). This funder had no role in study design, data collection and analysis, nor in the decision to publish, nor in the preparation of the manuscript. The original PROCTOR-SCRIPT trial was supported by the Dutch Cancer Society (KWF 1999-03 and KWF 2003-16), the Dutch Colorectal Cancer Group, and the Swedish Cancer Society. Furthermore, Roche has provided an unrestricted educational grant. The original RAPIDO trial was funded by the Dutch Cancer Foundation, Swedish Cancer Society, Spanish Ministry of Economy and Competitiveness, and Spanish Clinical Research Network.

Clinical trial registration

PROCTOR-SCRIPT: Dutch Colorectal Cancer group, CKTO 2003-16, ISRCTN36266738

RAPIDO: EudraCT 2010-023957-12, and ClinicalTrials.gov NCT01558921

References

- Iv AA, Koprowski MA, Nabavizadeh N, Tsikitis VL. The evolution of rectal cancer treatment: The journey to total neoadjuvant therapy and organ preservation. *Ann Gastroenterol* 2022;35:226-233.
- Beets GL, Figueiredo NL, Habr-Gama A, van de Velde CJH. A new paradigm for rectal cancer: Organ preservation: Introducing the international watch & wait database (iwwd). *European Journal of Surgical Oncology (EJSO)* 2015;41:1562-1564.
- Weiser MR. Ajcc 8th edition: Colorectal cancer. *Ann Surg Oncol* 2018;25:1454-1455.
- Glynn-Jones R, Wyrwicz L, Tiet E et al. Rectal cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2018;29:iv263.
- Beets-Tan RGH, Lambregts DMJ, Maas M et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 european society of gastrointestinal and abdominal radiology (esgar) consensus meeting. *European Radiology* 2018;28:1465-1475.
- Kapiteijn E, Putter H, van de Velde CJ. Impact of the introduction and training of total mesorectal excision on recurrence and survival in rectal cancer in the netherlands. *Br J Surg* 2002;89:1142-1149.
- Peeters KC, Marijnen CA, Nagtegaal ID et al. The time trial after a median follow-up of 6 years: Increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. *Ann Surg* 2007;246:693-701.
- Kapiteijn E, Marijnen CA, Nagtegaal ID et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *New England Journal of Medicine* 2001;345:638-646.
- Breugom AJ, van Gijn W, Muller EW et al. Adjuvant chemotherapy for rectal cancer patients treated with preoperative (chemo)radiotherapy and total mesorectal excision: A dutch colorectal cancer group (dccg) randomized phase iii trial. *Ann Oncol* 2015;26:696-701.
- Bahadoer RR, Dijkstra EA, van Etten B et al. Short-course radiotherapy followed by chemotherapy before total mesorectal excision (tme) versus preoperative chemoradiotherapy, tme, and optional adjuvant chemotherapy in locally advanced rectal cancer (rapido): A randomised, open-label, phase 3 trial. *The Lancet Oncology* 2021;22:29-42.
- Garcia-Aguilar J, Patil S, Gollub MJ et al. Organ preservation in patients with rectal adenocarcinoma treated with total neoadjuvant therapy. *J Clin Oncol* 2022;40:2546-2556.
- Conroy T, Bosset JF, Etienne PL et al. Neoadjuvant chemotherapy with folfoxir and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (unicancer-prodige 23): A multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* 2021;22:702-715.
- van der Valk MJ, Hilling DE, Bastiaannet E et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the international watch & wait database (iwwd): An international multicentre registry study. *The Lancet* 2018;391:2537-2545.
- Greenbaum A, Martin DR, Bocklage T, Lee J-H, Ness SA, Rajput A. Tumor heterogeneity as a predictor of response to neoadjuvant chemotherapy in locally advanced rectal cancer. *Clinical colorectal cancer* 2019;18:102-109.
- Lambregts DMJ, Boellaard TN, Beets-Tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern mr imaging: A pictorial review. *Insights Imaging* 2019;10:15.
- Temmink SJ, Martling A, Angenete E, Nilsson PJ. Complete response rates in rectal cancer: Temporal changes over a decade in a population-based nationwide cohort. *European Journal of Surgical Oncology* 2023;49:106991.
- Nahas SC, Nahas CSR, Cama GM et al. Diagnostic performance of magnetic resonance to assess treatment response after neoadjuvant therapy in patients with locally advanced rectal cancer. *Abdom Radiol (NY)* 2019;44:3632-3640.
- Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA Cancer J Clin* 2023;73:233-254.
- Custers PA, van der Sande ME, Grotenhuis BA et al. Long-term quality of life and functional outcome of patients with rectal cancer following a watch-and-wait approach. *JAMA Surgery* 2023;158:e230146-e230146.
- Werb Z, Lu P. The role of stroma in tumor development. *Cancer J* 2015;21:250-253.
- Bremnes RM, Dønnem T, Al-Saad S et al. The role of tumor stroma in cancer progression and prognosis: Emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *Journal of Thoracic Oncology* 2011;6:209-217.
- van Pelt GW, Sandberg TP, Morreau H et al. The tumour-stroma ratio in colon cancer: The biological role and its prognostic impact. *Histopathology* 2018;73:197-206.
- Mesker WE, Junggeburst JM, Szuhai K et al. The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Cell Oncol* 2007;29:387-398.
- Huijbers A, Tollenaar RA, v Pelt GW et al. The proportion of tumor-stroma as a strong prognosticator for stage ii and iii colon cancer patients: Validation in the victor trial. *Ann Oncol* 2013;24:179-185.
- Park JH, Richards CR, McMillan DC, Horgan PG, Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Ann Oncol* 2014;25:644-651.
- Smit M, van Pelt G, Roodvoets A et al. Uniform noting for international application of the tumor-stroma ratio as an easy diagnostic tool: Protocol for a multicenter prospective cohort study. *JMIR Res Protoc* 2019;8:e13464.
- Polack M, Smit MA, van Pelt GW et al. Results from the united study: A multicenter study validating the prognostic effect of the tumor-stroma ratio in colon cancer. *ESMO Open* 2024;9:102988.
- Zunder SM, van Pelt GW, Gelderblom HJ et al. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage ii and stage iii colon cancer. *Br J Cancer* 2018;119:164-169.
- van Pelt GW, Krol JA, Lips IM et al. The value of tumor-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol* 2020;20:39-44.

30. Hansen TF, Kjær-Frifeldt S, Lindebjerg J et al. Tumor-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. *Acta Oncol* 2018;57:528-533.
31. Strous MTA, Faes TKE, Gubbels A et al. A high tumour-stroma ratio (tsr) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clin Transl Oncol* 2022;24:1047-1058.
32. Strous MTA, Faes TKE, Heemskerk J et al. Tumour-stroma ratio to predict pathological response to neo-adjuvant treatment in rectal cancer. *Surg Oncol* 2022;45:101862.
33. Park JH, van Wyk H, McMillan DC et al. Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. *J Pathol Clin Res* 2020;6:30-39.
34. Scheer R, Baidoshvili A, Zoidze S et al. Tumor-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: A retrospective cohort study. *World J Gastrointest Oncol* 2017;9:466-474.
35. van Pelt GW, Kjaer-Frifeldt S, van Krieken J et al. Scoring the tumor-stroma ratio in colon cancer: Procedure and recommendations. *Virchows Arch* 2018;473:405-412.
36. Mandard AM, Dalibard F, Mandard JC et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 1994;73:2680-2686.
37. Nagtegaal ID, Glynne-Jones R. How to measure tumour response in rectal cancer? An explanation of discrepancies and suggestions for improvement. *Cancer Treatment Reviews* 2020;84:101964.
38. Tulchinsky H, Shmueli E, Figer A, Klausner JM, Rabau M. An interval > 7 weeks between neoadjuvant therapy and surgery improves pathologic complete response and disease-free survival in patients with locally advanced rectal cancer. *Annals of surgical oncology* 2008;15:2661-2667.
39. Glynne-Jones R, Wyrwicz L, Tiret E et al. Rectal cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2017;28:iv22-iv40.
40. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. *J Cell Sci* 2012;125:5591-5596.
41. Ueno H, Shinto E, Shimazaki H et al. Histologic categorization of desmoplastic reaction: Its relevance to the colorectal cancer microenvironment and prognosis. *Ann Surg Oncol* 2015;22:1504-1512.
42. Valkenburg KC, de Groot AE, Pienta KJ. Targeting the tumour stroma to improve cancer therapy. *Nature Reviews Clinical Oncology* 2018;15:366-381.
43. Strous MT, van der Linden RL, Gubbels AL et al. Node-negative colon cancer: Histological, molecular, and stromal features predicting disease recurrence. *Molecular Medicine* 2023;29:77.
44. Zhu Y, Jin Z, Qian Y, Shen Y, Wang Z. Prognostic value of tumor-stroma ratio in rectal cancer: A systematic review and meta-analysis. *Front Oncol* 2021;11:685570.
45. Tian W, Yang Y, Qin Q et al. Vimentin and tumor-stroma ratio for neoadjuvant chemoradiotherapy response prediction in locally advanced rectal cancer. *Cancer Sci* 2023;114:619-629.
46. Temmink SJD, Peeters K, Bahadoer RR et al. Watch and wait after neoadjuvant treatment in rectal cancer: Comparison of outcomes in patients with and without a complete response at first reassessment in the international watch & wait database (iwwd). *Br J Surg* 2023;110:676-684.
47. Kagawa Y, Smith JJ, Fokas E et al. Future direction of total neoadjuvant therapy for locally advanced rectal cancer. *Nat Rev Gastroenterol Hepatol* 2024;21:444-455.
48. Dworak O, Keilholz L, Hoffmann A. Pathological features of rectal cancer after preoperative radiochemotherapy. *International journal of colorectal disease* 1997;12:19-23.
49. Kus Ozturk S, Graham Martinez C, Sheahan K et al. Relevance of shrinkage versus fragmented response patterns in rectal cancer. *Histopathology* 2023;83:870-879.
50. Loughrey MB, Webster F, Arends MJ et al. Dataset for pathology reporting of colorectal cancer: Recommendations from the international collaboration on cancer reporting (iccr). *Ann Surg* 2022;275:e549-e561.
51. Ryan R, Gibbons D, Hyland J et al. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *Histopathology* 2005;47:141-146.
52. Ravensbergen CJ, Polack M, Roelands J et al. Combined assessment of the tumor-stroma ratio and tumor immune cell infiltrate for immune checkpoint inhibitor therapy response prediction in colon cancer. *Cells* 2021;10.
53. Bagaev A, Kotlov N, Nomie K et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer cell* 2021;39:845-865. e847.
54. Pagès F, Mlecnik B, Marliot F et al. International validation of the consensus immunoscore for the classification of colon cancer: A prognostic and accuracy study. *Lancet* 2018;391:2128-2139.
55. Aherne S, Donnelly M, Ryan É J et al. Tumour budding as a prognostic biomarker in biopsies and resections of neoadjuvant-treated rectal adenocarcinoma. *Histopathology* 2024;85:224-243.
56. Kramer A, Greuter MJE, Schraa SJ et al. Early evaluation of the effectiveness and cost-effectiveness of ctDNA-guided selection for adjuvant chemotherapy in stage II colon cancer. *Ther Adv Med Oncol* 2024;16:17588359241266164.
57. Ravensbergen CJ, Kuruc M, Polack M et al. The stroma liquid biopsy panel contains a stromal-epithelial gene signature ratio that is associated with the histologic tumor-stroma ratio and predicts survival in colon cancer. *Cancers (Basel)* 2021;14.
58. Sahin IH, Yanes R, Saridogan T, Holder-Murray J, Dasari AN. The role of circulating tumor DNA for management of patients with rectal cancer: Challenges and opportunities. *Cancer J* 2024;30:290-296.
59. Smit MA, van Pelt GW, Dequeker EM et al. E-learning for instruction and to improve reproducibility of scoring tumor-stroma ratio in colon carcinoma: Performance and reproducibility assessment in the united study. *JMIR Form Res* 2021;5:e19408.
60. Smit MA, van Pelt GW, Terpstra V et al. Tumour-stroma ratio outperforms tumour budding as biomarker in colon cancer: A cohort study. *Int J Colorectal Dis* 2021;36:2729-2737.
61. Chen RJ, Ding T, Lu MY et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine* 2024;30:850-862.

62. Coudray N, Ocampo PS, Sakellaropoulos T et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018;24;1559-1567.
63. Smit MA, Ciompi F, Bokhorst JM et al. Deep learning based tumor-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform* 2023;14;100191.
64. Firmbach D, Benz M, Kuritsyn P et al. Tumor-stroma ratio in colorectal cancer-comparison between human estimation and automated assessment. *Cancers (Basel)* 2023;15.
65. Sinicrope FA, Nelson GD, Saberzadeh-Ardestani B et al. Use of deep learning to evaluate tumor microenvironmental features for prediction of colon cancer recurrence. *Cancer Res Commun* 2024;4;1344-1350.
66. Geessink OGF, Baidoshvili A, Klaase JM et al. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. *Cell Oncol (Dordr)* 2019;42;331-341.
67. Liang Y, Zhu Y, Lin H et al. The value of the tumour-stroma ratio for predicting neoadjuvant chemoradiotherapy response in locally advanced rectal cancer: A case control study. *BMC Cancer* 2021;21;729.
68. Liu B, Polack M, Coudray N et al. Self-supervised learning reveals clinically relevant histomorphological patterns for therapeutic strategies in colon cancer. *bioRxiv* 2024;2024.2002.2026.582106.

Supplementary Table 1. Detailed overview of the treatment types within the three separate cohorts.

Cohort	Treatment group	Neoadjuvant therapy	Adjuvant therapy
LUMC (N=97)	Neoadjuvant + surgery (N=74)	SCRT (N=52); CRT (N=17); Other (N=5)	N/A
	Neoadjuvant + surgery + adjuvant (N=3)	SCRT (N=3)	PROCTOR (N=2); SCRIPT (N=1)
	Surgery + adjuvant (N=4)	None	PROCTOR (N=1); Radiotherapy (N=2); CAPOX (N=1)
	Surgery alone (N=16)	None	N/A
PROCTOR-SCRIPT (N=122)	Neoadjuvant + surgery (N=64)	SCRT (N=59); CRT (N=5)	N/A
	Neoadjuvant + surgery + adjuvant (N=58)	SCRT (N=52); CRT (N=6)	
	Surgery + adjuvant (N=0)	N/A	
	Surgery alone (N=0)	N/A	N/A
RAPIDO (N=154)	Neoadjuvant + surgery (N=152)	RAPIDO (N=75); CRT (N=77)	N/A
	Neoadjuvant + surgery + adjuvant (N=2)	CRT (N=2)	Unknown (N=2)
	Surgery + adjuvant (N=0)	N/A	N/A
	Surgery alone (N=0)	N/A	N/A

CAPOX, capecitabine and oxaliplatin; CRT, chemoradiation (25x1.8-2 Gray and capecitabine monotherapy); N/A, not applicable; PROCTOR (5-FU + Leucovorin); RAPIDO (5x5 Gray followed by 6 cycles of CAPOX); SCRIPT (capecitabine monotherapy); SCRT, short course radiotherapy (5x5 Gray).

Supplementary Table 2. Characteristics of the biopsy-scored stroma-low compared to stroma-high patients in total cohort and per cohort.

Variables (unit)		Total cohort (N=373)		LUMC (N=97)		PROCTOR-SCRIPT (N=122)		RAPIDO (N=154)				
		Stroma-low (N=215)	Stroma-high (N=158)	P-value	Stroma-low (N=69)	Stroma-high (N=28)	P-value	Stroma-low (N=80)	Stroma-high (N=42)	Stroma-low (N=66)	Stroma-high (N=88)	P-value
Sex				0.209#			0.798#					0.112#
Female		73 (34)	44 (28)		19 (28)	7 (25)		28 (35)	13 (31)	26 (39)	24 (27)	
Male		142 (66)	114 (72)		50 (73)	21 (75)		52 (65)	29 (69)	40 (61)	64 (73)	
Age*												
Mean age (in years; SD)		62 (10)	62 (9)	0.978\$	64 (13)	66 (9)	0.372\$	61 (9)	61 (9)	62 (10)	61 (9)	0.605\$
Age, category**				0.841#			0.856#					0.854#
≤70 years of age		171 (80)	127 (80)		48 (70)	20 (71)		69 (86)	34 (81)	54 (82)	73 (83)	
>70 years of age		44 (21)	31 (20)		21 (30)	8 (29)		11 (14)	8 (19)	12 (18)	15 (17)	
Clinical TNM-stage***				0.103#			0.091#					0.257#
II		27 (21)	15 (13)		23 (33)	5 (18)		N/A****		4 (6)	10 (11)	
III		102 (79)	100 (87)		40 (58)	22 (79)				62 (94)	78 (89)	
Unknown and II/III***** (excluded from analysis)		86 (40)	43 (27)		6 (9)	1 (4)				0 (0)	0 (0)	
Tumour location				0.186#			0.611#					0.827#
Low rectum (<5cm anal verge)		79 (37)	54 (35)		28 (41)	14 (50)		28 (35)	12 (29)	23 (66)	28 (32)	
Mid rectum (5-10 cm anal verge)		54 (26)	53 (34)		19 (28)	8 (29)		17 (21)	17 (40)	18 (27)	28 (32)	
High rectum (>10 cm anal verge)		78 (37)	48 (31)		21 (30)	6 (21)		33 (41)	11 (26)	24 (36)	31 (35)	
Unknown (excluded from analysis)		4 (2)	3 (1)		1 (1)	0 (0)		2 (3)	2 (5)	1 (2)	1 (1)	
Clinical risk factors*****				0.014#			0.229#					0.111#
No, no additional risk factors		20 (16)	4 (3)		20 (29)	4 (14)		N/A*****		N/A	N/A	
Yes, 1 clinical risk factor present		23 (18)	25 (22)		20 (29)	12 (43)				3 (5)	13 (15)	
Yes, 2 clinical risk factors present		40 (31)	45 (39)		11 (16)	8 (29)				29 (44)	37 (42)	
Yes, 3 or more clinical risk factors present		46 (36)	42 (36)		12 (17)	4 (14)				34 (52)	38 (43)	
Not (enough) determined, unknown (excluded from analysis)		85 (40)	42 (27)		6 (9)	0 (0)				0 (0)	0 (0)	
Treatment type				<0.001#			0.074#					0.100#
Neoadjuvant treatment and surgery		151 (70)	139 (88)		49 (71)	25 (89)		38 (48)	26 (62)	64 (97)	88 (100)	
Neoadjuvant treatment, surgery and adjuvant treatment		47 (22)	16 (10)		3 (4)	0 (0)		42 (53)	16 (38)	2 (3)	0 (0)	
Surgery and adjuvant treatment		2 (1)	2 (1)		2 (3)	2 (7)		N/A	N/A	N/A	N/A	
Surgery alone		15 (7)	1 (1)		15 (22)	1 (4)		N/A	N/A	N/A	N/A	

<i>(continued)</i>										
Neoadjuvant treatment received	17 (8)	3 (2)	0.011#	17 (25)	3 (11)	0.125#	N/A	N/A	N/A	N/A
No, no neoadjuvant treatment	198 (92)	155 (98)	<0.001#	52 (75)	25 (89)	0.350#	80 (100)	42 (100)	66 (100)	88 (100)
Yes, neoadjuvant treatment received, of which										0.020#
Short course radiotherapy (SCRT; 5x5 Gray)	114 (58)	52 (34)		39 (80)	16 (70)		75 (94)	36 (86)	N/A	N/A
Chemoradiotherapy (CRT; 28x1.2 Gray and capecitabine monotherapy)	56 (28)	51 (33)		10 (20)	7 (30)		5 (6)	6 (14)	41 (62)	38 (43)
RAPIDO-regime (5x5 Gray followed by 6 cycles capecitabine and oxaliplatin [CAPOX])	25 (13)	50 (32)		N/A	N/A		N/A	N/A	25 (38)	50 (57)
Other (excluded from analysis)	3 (2)	2 (1)		2 (2)	2 (2)		0 (0)	0 (0)	0 (0)	0 (0)
Waiting period before surgery										
Mean duration period (in weeks; SD)	10 (9)	15 (10)	<0.001\$	10 (7)	11 (7)	0.883\$	2 (3)	3 (4)	20 (6)	22 (6)
Period shorter <12 weeks	129 (60)	58 (37)	<0.001#	51 (74)	20 (71)	0.802#	78 (98)	38 (91)	N/A	N/A
Period longer ≥12 weeks	86 (40)	100 (63)		18 (26)	8 (29)		2 (3)	4 (10)	66 (100)	88 (100)
Surgery year										
Mean surgery year (SD)	2009 (5)	2012 (4)	<0.001\$	2007 (5)	2010 (5)	0.027\$	2006 (3)	2007 (3)	2015 (1)	2015 (1)
Surgery type, grouped			0.174#			0.710#				0.500\$
Low anterior resection (LAR)	122 (57)	84 (53)		35 (51)	13 (46)		49 (61)	22 (52)	38 (58)	49 (56)
Abdominoperineal resection (APR)	85 (40)	60 (38)		32 (46)	15 (54)		28 (35)	17 (40)	25 (38)	28 (32)
Hartmann	5 (2)	11 (7)		1 (1)	0 (0)		3 (4)	2 (5)	1 (2)	9 (10)
Other (excluded from analysis)	3 (1)	3 (2)		1 (1)	0 (0)		0 (0)	1 (2)	2 (3)	2 (2)
Adjuvant treatment received			0.005#			0.986#				0.100#
No, no adjuvant treatment	166 (77)	140 (89)		64 (93)	26 (93)		38 (48)	26 (62)	64 (97)	88 (100)
Yes, adjuvant treatment received, of which	49 (23)	18 (11)	0.125#	5 (7)	2 (7)	0.208#	42 (53)	16 (38)	2 (3)	0 (0)
5-FU and Leucovorin (PROCTOR)	19 (9)	3 (2)		3 (60)	0 (0)		16 (38)	3 (19)	0 (0)	0 (0)
Capecitabine monotherapy (SCRIPT)	27 (13)	13 (8)		1 (20)	0 (0)		26 (62)	13 (81)	0 (0)	0 (0)
Radiotherapy	1 (1)	1 (1)		1 (20)	1 (50)		0 (0)	0 (0)	0 (0)	0 (0)
Capecitabine and oxaliplatin (CAPOX)	0 (0)	1 (1)		0 (0)	1 (50)		0 (0)	0 (0)	0 (0)	0 (0)
Missing	2 (1)	0 (0)		0 (0)	0 (0)		0 (0)	0 (0)	2 (100)	0 (0)

All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding. N/A, not applicable. SD, standard deviation. TNM, tumour-node-metastasis.

* Age at randomisation (PROCTOR-SCRIPT and RAPIDO) or diagnosis (LUMC).

** Age cutoff of 70 years is used here, averaging most trials.

***Different versions of the TNM classification were used, for optimal grouping and comparison, all variables are converted to TNM version 5 (1997).

****Stage II/III is given in case no imaging is performed and locoregional extent of disease is uncertain.

PROCTOR-SCRIPT patients were included after surgery, hence clinical TNM-stage or risk factors were not registered in the study database, although inclusion criteria was pathological TNM-stage II/III

*****Clinical risk factors as determined by guidelines to assess locally advanced rectal carcinoma (e.g. ESMO guidelines), see below overview of all.

Clinical risk factors were not always determined on imaging, especially in patients operated before 2010.

*****Multiple risk factors may exist simultaneously, hence percentages do not add up to 100. Not always determined on imaging, especially in patients operated before 2010.

Calculated with Chi-square test

\$ Calculated with the Student's Independent Samples T-test

Supplementary Table 3. Overview of the pathology outcomes between biopsy-scored stroma-low compared to stroma-high patients in the total cohort and per treatment type.

Variables (unit)	Total cohort (N=373)		No neoadjuvant therapy (N=20)		SCRT (N=166)		NAT (N=182)				
	Stroma-low (N=215)	Stroma-high (N=158)	P-value	Stroma-low (N=17)	Stroma-high (N=3)	Stroma-low (N=114)	Stroma-high (N=52)	Stroma-low (N=81)	Stroma-high (N=101)	P-value	
Residual tumour* R0 resection R+ resection <i>Missing, unknown (excluded from analysis)</i>	202 (94) 12 (6) 1 (0)	146 (92) 12 (8) 0 (0)	0.441#	17 (100) 0 (0) 0 (0)	3 (100) 0 (0) 0 (0)	N/A		106 (94) 7 (6) 1 (1)	48 (92) 4 (8) 0 (0)	0.720#	
Minimal distance to circumferential margin (CRM)			0.280\$			0.802\$				0.936\$	
Mean CRM distance (in millimetres; SD)	11 (12)	10 (9)		16 (13)	14 (11)		13 (14)	13 (12)	8 (5)		
(y)pTNM-stage** (y)pTNM-stage 0 (y)pTNM-stage I (y)pTNM-stage II (y)pTNM-stage III	22 (10) 23 (11) 53 (25) 117 (54)	22 (14) 21 (13) 43 (27) 72 (46)	0.058†	N/A 7 (41) 5 (29) 5 (29)	N/A 1 (33) 1 (33) 1 (33)	0.815†	N/A 10 (9) 23 (20) 81 (71)	N/A 3 (6) 5 (10) 44 (85)	22 (27) 10 (12) 19 (24) 30 (37)	22 (22) 23 (23) 30 (30) 26 (26)	0.441†
Maximum diameter tumour Mean tumour diameter (in millimetres; SD)	30 (18)	38 (111)	0.280\$	39 (27)	25 (10)	0.146\$	36 (12)	40 (13)	18 (17)	37 (128)	
All y(p)T-categories** y(p)T-stage 0 y(p)T-stage I y(p)T-stage II y(p)T-stage III y(p)T-stage IV	26 (12) 8 (4) 52 (24) 114 (53) 15 (7)	23 (15) 5 (3) 40 (25) 83 (53) 7 (4)	0.384†	N/A 1 (6) 7 (41) 8 (47) 1 (6)	N/A 0 (0) 1 (33) 2 (67) 0 (0)	0.678†	N/A 4 (4) 30 (26) 77 (68) 3 (3)	N/A 0 (0) 11 (21) 39 (75) 2 (4)	26 (32) 3 (4) 13 (16) 29 (36) 10 (12)	23 (23) 5 (5) 27 (27) 41 (41) 5 (5)	0.998†
Combined y(p)T-categories** y(p)T-stage 0+I y(p)T-stage II-IV	34 (16) 181 (84)	28 (18) 130 (82)	0.627†	1 (6) 16 (94)	0 (0) 3 (100)	0.348†	4 (4) 110 (97)	0 (0) 52 (100)	29 (36) 52 (64)	28 (28) 73 (72)	0.245†
Lymph nodes (LN) No. patients ≥ LN 12 examined No. patients < LN 12 examined <i>Missing (excluded from analysis)</i>	93 (44) 121 (57) 1 (0)	79 (50) 78 (50) 1 (0)	0.190#	11 (65) 6 (35) 0 (0)	2 (67) 1 (33) 0 (0)	0.948#	49 (43) 64 (57) 1 (1)	21 (41) 30 (59) 1 (2)	33 (41) 48 (59) 0 (0)	55 (55) 46 (46) 0 (0)	0.066#

<i>(continued)</i>	LN per patient	12 (6)	12 (7)	0.269\$	15 (6)	16 (9)	0.470\$	11 (6)	11 (5)	0.438\$	11 (5)	13 (8)	0.086\$
Mean LN examined (SD)		2 (3)	1 (2)	0.042\$	2 (4)	1 (2)	0.801\$	2 (4)	3 (3)	0.659\$	1 (2)	1 (1)	0.014\$
Mean LN positive (SD)				0.030†			0.876†			0.121†			0.051†
y(pN)-category**													
y(pN)-stage 0	98 (46)	86 (54)			12 (71)	2 (67)		33 (29)	8 (15)		51 (63)	75 (74)	
y(pN)-stage I	79 (37)	57 (36)			1 (6)	1 (33)		58 (51)	32 (62)		19 (24)	23 (23)	
y(pN)-stage II	38 (18)	15 (10)			4 (24)	0 (0)		23 (20)	12 (23)		11 (14)	3 (3)	
Differentiation grade tumour				0.135#			0.814#			0.629#			0.198#
Well to moderate (low grade)	140 (73)	110 (77)			11 (73)	2 (67)		88 (89)	43 (92)		41 (85)	64 (93)	
Poor to undifferentiated (high grade)	23 (12)	10 (7)			4 (27)	1 (33)		11 (11)	4 (9)		7 (15)	5 (7)	
<i>Missing (excluded from analysis)</i>	25 (12)	15 (10)			2 (12)	0 (0)		15 (13)	5 (10)		47 (58)	71 (70)	
<i>N/A (no tumour; excluded from analysis)</i>	27 (13)	23 (15)			N/A	N/A		N/A	N/A		25 (31)	23 (23)	
Pathology risk factors***													
Not determined	85 (40)	83 (53)			8 (47)	2 (67)		28 (25)	8 (15)		47 (58)	71 (70)	
Determined, and	105 (49)	52 (33)		0.757#	9 (53)	1 (33)	0.598#	86 (75)	44 (85)		9 (11)	7 (7)	0.949#
No risk factors present	39 (37)	18 (35)			7 (78)	1 (100)		28 (33)	14 (32)		4 (44)	3 (43)	
Yes, 1 or more risk factors present	66 (63)	34 (65)			2 (22)	0 (0)		58 (67)	30 (68)		5 (56)	4 (57)	
<i>N/A (no tumour; excluded from analysis)</i>	25 (12)	23 (15)			N/A	N/A		N/A	N/A		25 (31)	23 (23)	
Tumour response****													
Determined, and	188 (87)	149 (94)		0.074†	N/A			105 (92)	47 (90)		22 (27)	22 (22)	0.265†
Complete pathological response (pCR; TRG 1)	22 (12)	22 (15)						0 (0)	0 (0)				
Partial pathological response (pPR; TRG 2+3)	73 (39)	68 (46)						21 (20)	4 (9)		50 (62)	63 (63)	
No or nearly no pathological response (TRG 4+5)	93 (50)	59 (40)						84 (80)	43 (92)		9 (11)	16 (16)	
<i>N/A (no neoadjuvant therapy; excluded from analysis)</i>	17 (8)	3 (2)						N/A	N/A		N/A	N/A	
<i>Not determined, unknown (excluded from analysis)</i>	10 (5)	6 (4)						9 (8)	5 (10)		N/A	N/A	
Major response in those determined				0.150†	N/A					0.316†			0.322†
No, no major response (TRG3-5)	149 (79)	108 (73)						104 (99)	47 (100)		43 (53)	61 (60)	
Yes, major response (TRG1+2)	39 (21)	41 (28)						1 (1)	0 (0)		38 (47)	40 (40)	0.322†

Other type of NAT (N=5) than above is excluded at category for analysis purposes. All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding.

N/A, not applicable; SD, standard deviation; SCRT, short course radiotherapy (5x5 Gray); TNM, tumour-node metastasis; NAT, neoadjuvant treatment, including chemoradiation (25x1.8-2 Gray and capecitabine monotherapy); and RAPIDO (5x5 Gray followed by 6 cycles capecitabine and oxaliplatin); TRG, tumour regression grade.

*Residual tumour according to Wittekind (2009).

**Different versions of the TNM classification were used, here all variables are converted to TNM version 5 (1997).

***Pathology risk factors include: extramural venous invasion, lymphatic invasion, (intramural) venous invasion, perineural invasion or combinations of these.

****Tumour response is categorized in three; no/partial/complete response, or based on TRG groups as categorized by Mandard.

Calculated using the Chi-square test. \$ Calculated using the Student's Independent Samples T-test. † Calculated using the Goodman and Kruskal's test for gamma.

Supplementary Table 4. Overview of the clinical outcomes between biopsy-scored stroma-low compared to stroma-high patients in the total cohort and per treatment type.

Variables (unit)	Total cohort (N=373)		No neoadjuvant therapy (N=20)		SCRT (N=166)		NAT (N=182)	
	Stroma-low (N=215)	Stroma-high (N=158)	Stroma-low (N=17)	Stroma-high (N=3)	Stroma-low (N=114)	Stroma-high (N=52)	Stroma-low (N=81)	Stroma-high (N=101)
Follow-up time - months								
Mean follow-up time (in years; SE)	8.5 (0.4)	7.7 (0.3)	11.1 (1.8)	13.5 (4.9)	8.7 (0.6)	7.8 (0.7)	6.9 (0.3)	7.4 (0.2)
Disease-free survival								
Mean disease-free survival time (in years; SE)	7.0 (0.3)	6.9 (0.3)	7.7 (0.9)	3.3 (0.9)	6.7 (0.4)	6.4 (0.6)	6.8 (0.4)	7.3 (0.4)
Events until censoring	81 (38)	62 (39)	5 (29)	2 (67)	48 (42)	21 (40)	27 (33)	37 (37)
Event type								
Locoregional recurrence	11 (14)	11 (18)	0 (0)	0 (0)	4 (8)	2 (10)	6 (22)	9 (24)
Distant metastasis	52 (64)	40 (65)	3 (60)	2 (100)	33 (69)	17 (81)	16 (59)	20 (54)
Death, any cause	18 (22)	11 (18)	2 (40)	0 (0)	11 (23)	2 (10)	5 (19)	8 (22)
Overall survival								
Mean overall survival time (in years; SE)	8.1 (0.2)	8.1 (0.2)	8.8 (0.6)	8.1 (1.6)	8.0 (0.3)	7.7 (0.5)	8.1 (0.4)	8.4 (0.3)
Deaths until censoring	60 (28)	46 (29)	4 (24)	1 (33)	34 (30)	16 (31)	21 (26)	27 (27)
Cause of death								
(Metastases of) current cancer	38 (63)	30 (65)	2 (33)	1 (33)	22 (65)	12 (75)	14 (67)	16 (59)
Other, including preexisting comorbidity	15 (25)	11 (24)	1 (25)	0 (0)	8 (24)	2 (13)	6 (29)	9 (33)
Missing, unknown	7 (12)	5 (11)	1 (25)	0 (0)	4 (12)	2 (13)	1 (5)	2 (7)

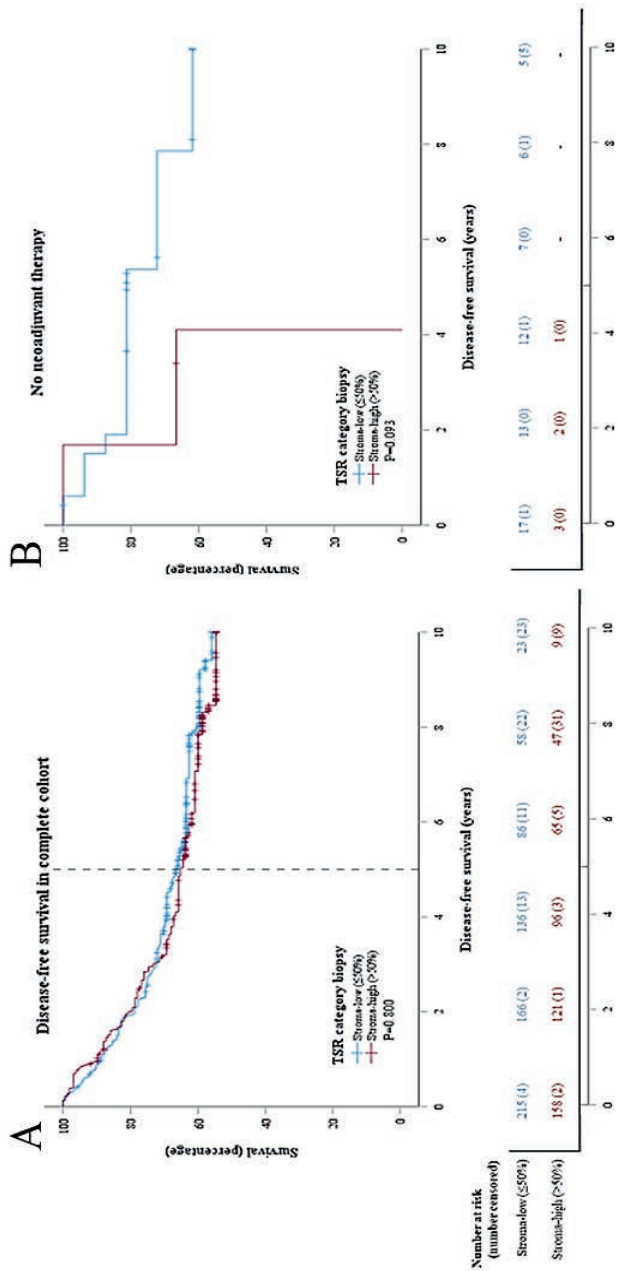
Other type of NAT (N=5) than above is excluded at category for analysis purposes. Disease-free and overall survival are censored after 10 years. All variables are given as absolute numbers with associated percentages or medians with interquartile ranges. Sum of percentages can be less or more than 100 due to rounding.

N/A, not applicable; SE, standard error; SCRT, short course radiotherapy (5x5 Gray); TNM, tumour-node metastasis; NAT, neoadjuvant treatment, including chemoradiation (25x1.8-2 Gray and capecitabine monotherapy) and RAPIDO (5x5 Gray followed by 6 cycles capecitabine and oxaliplatin).

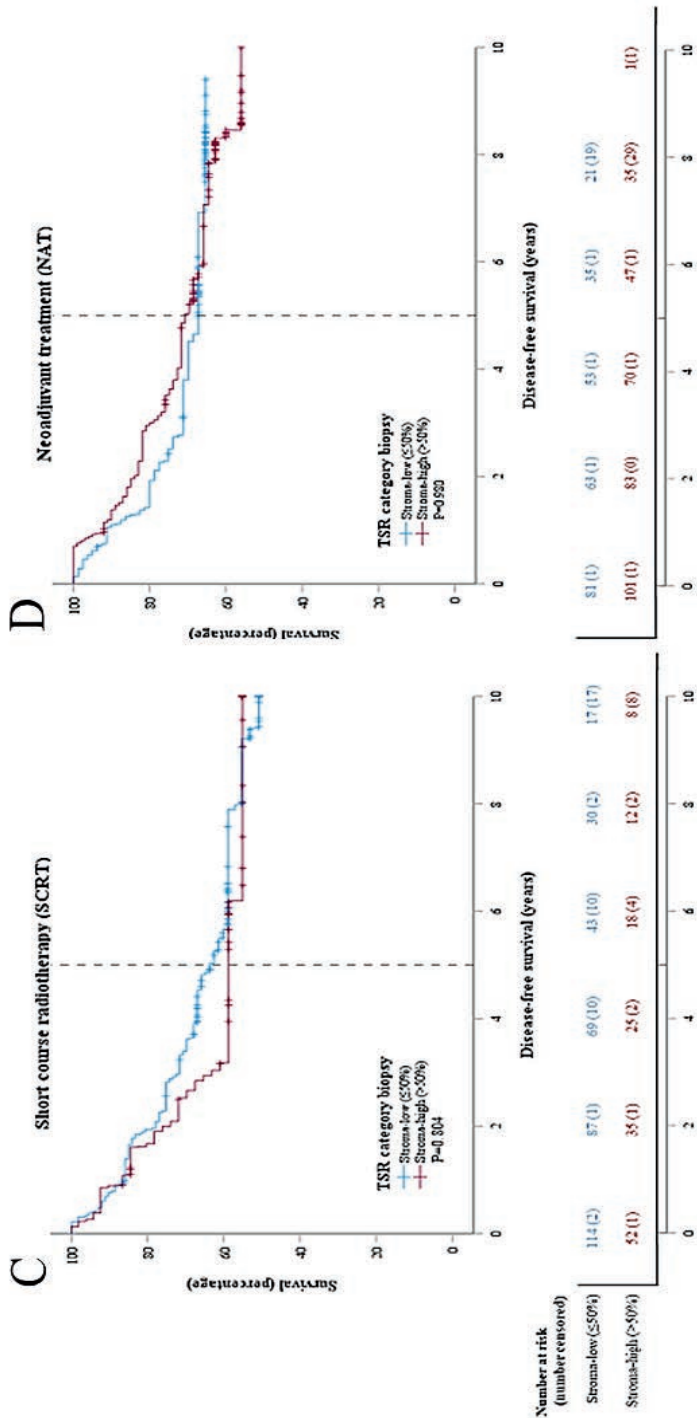
† Calculated with a reverse Kaplan-Meier analysis and log rank test.

§ Calculated with Kaplan-Meier analysis and log rank test.

Calculated using the Chi-square test.

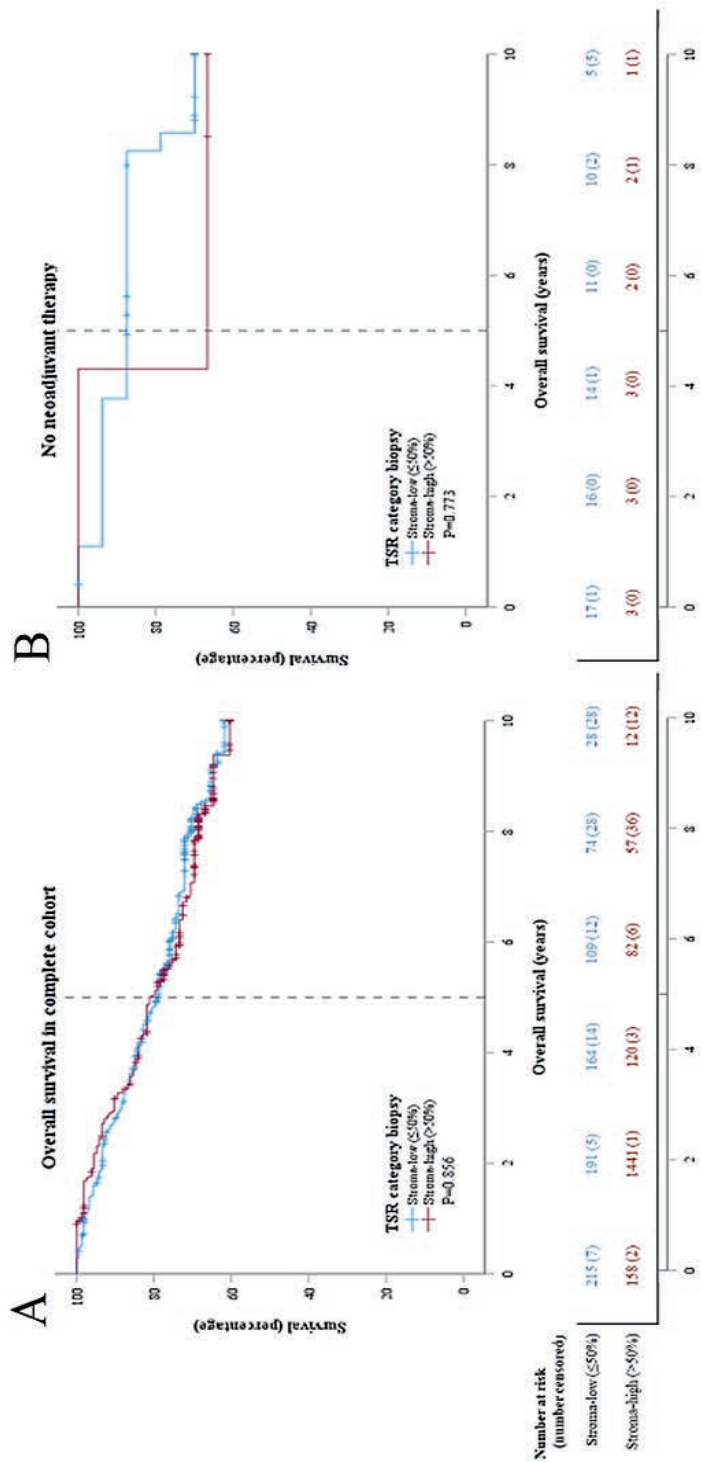


Supplementary Figure 1. Kaplan-Meier analysis and log rank test for disease-free survival (DFS) in stroma-low rectal carcinoma patients compared to stroma-high rectal carcinoma patients. (A) In the complete cohort (5-year DFS 66% vs. 65%; $P=0.800$). (B) In the group without neoadjuvant therapy (5-years DFS 81% vs. none; $P=0.093$).
(continued on next page)

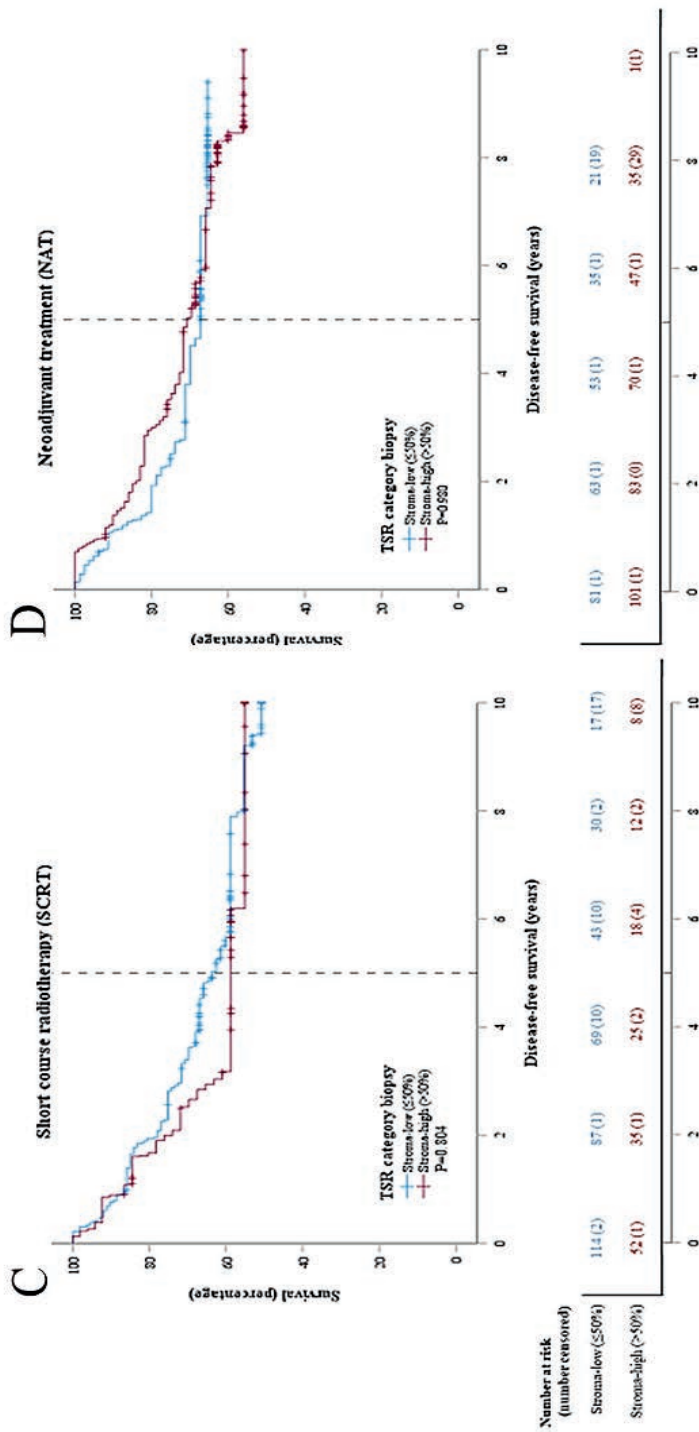


(continued) **Supplementary Figure 1.** Kaplan-Meier analysis and log rank test for disease-free survival (DFS) in stroma-low rectal carcinoma patients compared to stroma-high rectal carcinoma patients. (C) In the group with SCRT (5-year DFS 64% vs. 59%; P=0.804). (D) In the NAT group, including the chemoradiation and RAPIDO-treated patients (5-year DFS 67% vs. 71%; P=0.980). The TSR as single parameter does not reach significance in any analysis.

TSR, tumour-stroma ratio.

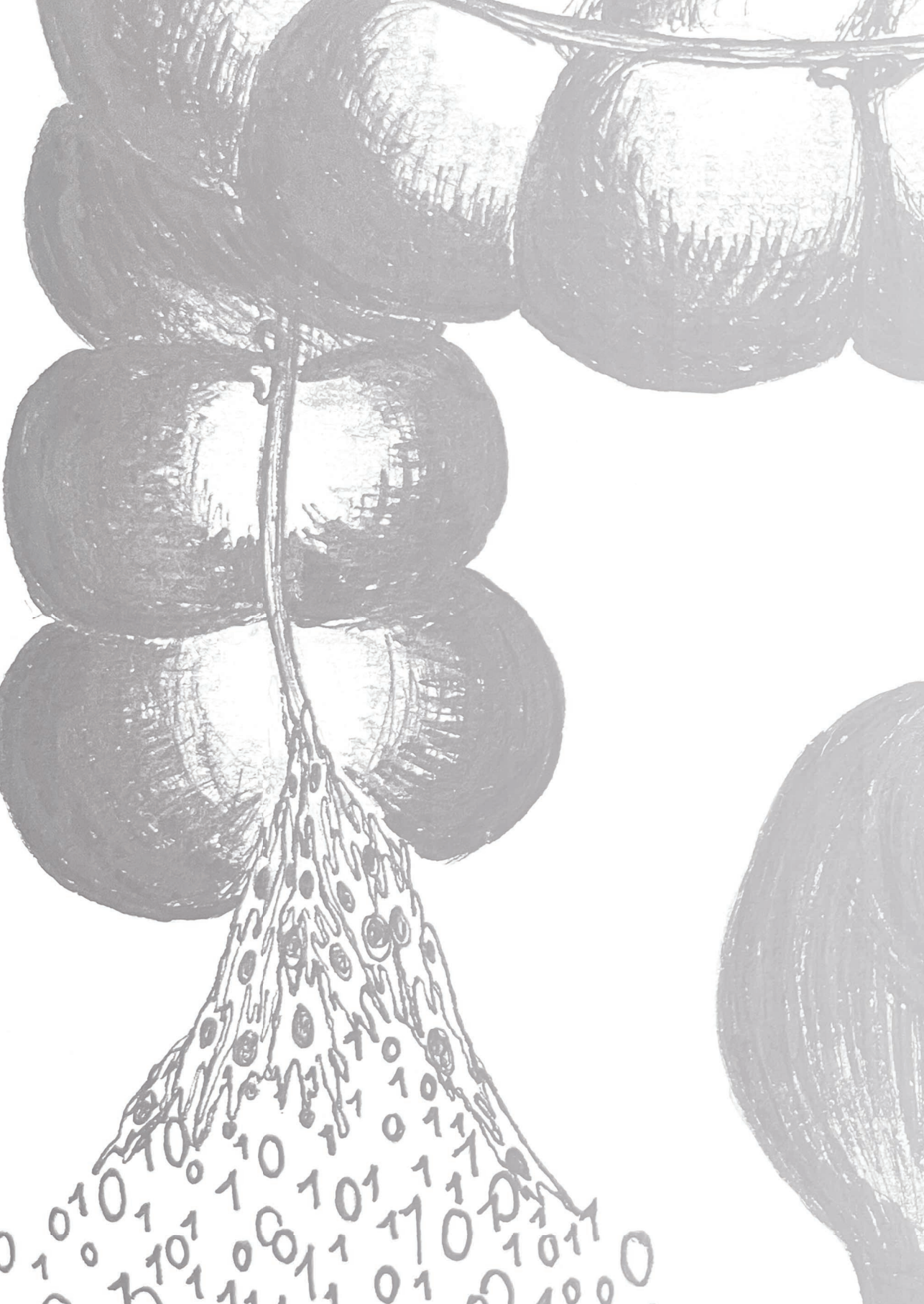


Supplementary Figure 2. Kaplan-Meier analysis and log rank test for overall survival (OS) in stroma-low rectal carcinoma patients compared to stroma-high rectal carcinoma patients. (A) In the complete cohort (5-year OS 79% vs. 81%; $P=0.856$). (B) In the group without neoadjuvant therapy (5-years OS 88% vs. 67%; $P=0.773$). (continued on next page)



(continued) **Supplementary Figure 2.** Kaplan-Meier analysis and log rank test for overall survival (OS) in stroma-low rectal carcinoma patients compared to stroma-high rectal carcinoma patients. (C) In the group with SCRT (5-year OS 77% vs. 72%; $P=0.696$). (D) In the NAT group, including the chemoradiation and RAPIDO-treated patients (5-year OS 80% vs. 88%; $P=0.764$). The TSR as single parameter does not reach significance in any analysis.

TSR, tumour-stroma ratio.



Chapter 4

Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT

Meaghan Polack, Sophie C. Hagenaars, Alice Couwenberg,
Walter Kool, Rob A.E.M. Tollenaar, Wouter V. Vogel, Petur
Snaebjornsson*, Wilma E. Mesker*

* shared last author

Abstract

Purpose: The recently developed fibroblast activation protein inhibitor (FAPI) tracer for PET/CT, binding tumour-stromal cancer-associated fibroblasts, is a promising tool for detection of positive lymph nodes. This study provides an overview of features, including sizes and tumour-stromal content, of lymph nodes and their respective lymph node metastases (LNM) in colorectal cancer (CRC), since literature lacks on whether LNMs contain sufficient stroma to potentially allow FAPI-based tumour detection.

Methods: Haematoxylin and eosin stained tissue slides from 73 stage III colon cancer patients were included. Diameters and areas of all lymph nodes and their LNMs were assessed, the amount of stroma by measuring the stromal compartment area, the conventional and total tumour-stroma ratios (TSR-c and TSR-t, respectively), as well as correlations between these parameters. Also, subgroup analysis using a minimal diameter cut off of 5.0 mm was performed.

Results: In total, 126 lymph nodes were analysed. Although positive correlations were observed between node and LNM for diameter and area ($r=0.852$, $P<0.001$ and $r=0.960$, $P<0.001$, respectively), and also between the LNM stromal compartment area and nodal diameter ($r=0.612$, $P<0.001$), nodal area ($r=0.747$, $P<0.001$) and LNM area ($r=0.746$, $P<0.001$), novel insight was that nearly all (98%) LNMs contained stroma, with median TSR-c scores of 35% (IQR 20–60%) and TSR-t of 20% (IQR 10–30%). Moreover, a total of 32 (25%) positive lymph nodes had a diameter of <5.0 mm.

Conclusion: In LNMs, stroma is abundantly present, independent of size, suggesting a role for FAPI PET/CT in improved lymph node detection in CRC.

Introduction

Colorectal cancer (CRC) is the third most common cancer type worldwide, with high morbidity and mortality rates [1-3]. In the diagnostic and preoperative workup, imaging with CT and/or MRI is performed for local, i.e. the primary tumour and regional lymph nodes, and distant extent of the disease. For clinical lymph node staging, positive lymph nodes are identified using size cut-off values and other morphological characteristics, such as roundness or irregular borders, according to radiological guidelines using these standard anatomical imaging techniques [4-8]. The detection of actual lymph node metastases (LNMs) remains suboptimal however, since false negative outcomes are prevalent, affecting treatment selection [9]. Therefore, other imaging modalities, e.g. FDG PET/CT or functional MRI, are being studied, but have not yet led to improved accuracy for detecting LNMs in CRC patients [10, 11].

Based on recent developments in the field of the tumour microenvironment, a novel PET/CT tracer has been developed: the fibroblast activation protein inhibitor (FAPI), which can be conjugated to a radioactive isotope such as Gallium-68 or Fluorine-18. The FAPI tracer binds to a part of the tumour microenvironment, or tumour stroma, which plays an essential role in tumour behaviour and the metastasis process. This tumour stroma is mainly composed of lymphocytes, vasculature, the extracellular matrix and fibroblasts. A desmoplastic reaction of constant remodelling by collagenases and matrix metalloproteases is associated with an epithelial-to-mesenchymal transition [12, 13]. This process not only promotes tumorigenesis, but also activates the quiescent fibroblasts in various ways, giving rise to the so-called cancer-associated fibroblasts (CAFs). These CAFs express the fibroblast activation protein (FAP), to which the FAPI can bind specifically [13-15].

So far, literature has shown promising results with FAPI PET/CT in cancer patients, suggesting improved sensitivity and specificity compared to the FDG PET/CT [16-21]. In addition to good uptake in the primary tumour, a high tumour-to-background ratio is seen with the FAPI tracer, leading to better imaging of regional and distant extent of disease. Furthermore, no specific patient protocols are necessary to prevent errors in imaging, e.g. fasting or regulation in medicine, due to the characteristics in biodistribution and the pharmacokinetics of FAPI [20, 22-26].

Over the past decades, the tumour stroma has been subject to extensive research. The tumour-stroma ratio (TSR), standardized and validated for primary tumours in multiple retrospective studies, expresses the proportional size of the stromal compartment relative to the epithelial tumour cell compartment and

has proven to be of significant prognostic and predictive value in various types of carcinomas [13, 27-30]. Stromal abundance in primary tumours is herein associated with a worse overall survival (OS) and disease-free survival (DFS), including in colon cancer [31-35]. In addition to the metastasized tumour epithelial cells, a LNM often also contains tumour stroma. Patients with a high amount of stroma in their LNMs, especially when combined with a stroma-high primary tumour, were observed to have an even worse OS and DFS than patients with stroma-low tumours and/or LNMs. Studies have shown that the TSR in biopsies correlate to the TSR score in resection material [36]. However, no information between the TSR scores between the primary tumour and their LNMs exists, necessitating examination of all positive lymph nodes for accurate diagnosis and prognosis [37-39].

In theory, the FAPI tracer binds to CAFs in the primary tumour as well as to CAFs in the LNMs, which is potentially visualized by a PET/CT scan. Although the FAPI PET/CT scanning is being studied more and more in depth, and imaging recently has distinguished different pathological cancer types based on uptake, however, it is still unknown what the quantification is of tumour stroma in the LNMs in CRC patients, and thus whether small positive lymph nodes and their LNM contain enough stroma and thus uptake to potentially be detected with imaging [40]. The aim of this work is therefore to determine the histopathological features of the positive lymph nodes, their respective LNMs and especially the tumour stroma within, in stage III CRC patients, to be used as a theoretical framework. Furthermore, the correlation between the lymph node and LNM size and stromal amount will be assessed, as well as a subgroup analysis in small positive lymph nodes with a diameter of less than 5.0 mm.

Materials and methods

Study population

The study cohort included 73 patients diagnosed with colon carcinoma between 1996 and 2011 at the Leiden University Medical Center, the Netherlands, described in more detail in a previous study [37]. Patients underwent complete surgical resection of the primary tumour including regional lymph nodes, and all had a histologically proven stage III adenocarcinoma or mucinous adenocarcinoma. Of these 73 patients, 5 µm haematoxylin and eosin (H&E) stained tissue sections of all the resected associated regional lymph nodes were selected. The final slides were scanned with a 3D-Histech scanner and analysed digitally using CaseViewer (version 2.4). Exclusion followed in the case of negative lymph nodes, fragmented tissue or poor quality of the digital slides. Patient samples were previously handled in a coded fashion, according to national ethical guidelines (“Code for Proper Secondary Use of Human Tissue”, Dutch Federation of Medical Scientific Societies) [37].

Histopathological analysis

Positive lymph nodes were separately analysed by two investigators (MP, SH). Several histopathological measurements retaining to the lymph nodes and the LNM themselves were assessed, listed and described in detail in Table 1. The amount of tumour stroma within the LNM was assessed in two ways, i.e. by determining proportion, expressed in percentages with the TSR, and by measuring the absolute size of the stromal compartment, expressed in mm². The TSR was analysed using the conventional TSR scoring method (TSR-c) first, according to the requirements as described by van Pelt et al. [29], after which the positive lymph nodes were subsequently categorized as stroma-high (>50%) or stroma-low (≤50%). Then, the total stromal percentage of this whole lymph node (TSR-t) was scored by eyeballing. The TSR-t was further categorized as <5% stroma in the whole LNM, 5% stroma, and then per tenfold (i.e. 10%, 20% etc.). Consensus was reached that for the TSR-t, the cut-off value would be 5%, as this was deemed sufficient to potentially be detected by pathologists and imaging. Stromal percentages were also independently scored by the two investigators (MP, SH), maintaining a maximum of 10% difference in score for optimal interobserver agreement. If scores differed more than 10%, the case was discussed until agreement was reached. If necessary, a third observer, an experienced pathologist (PS), was decisive.

In evaluating the correlation between size of the positive lymph node and the corresponding LNM and stromal presence, subgroup analysis was performed using a cut-off value for lymph node diameter of 5.0 mm and a corresponding area of 19.6 mm² ($\frac{1}{4} \times \pi \times \text{diameter}^2$). This cut off corresponds with the smallest radiological cut-off value that is clinically applied; in accordance with the Dutch CRC guideline and previous histopathological studies, imaging criteria for LNM include a short axis diameter of ≥5.0 mm and/or different morphological features, like roundness [41, 42].

Statistical analysis

Analyses are reported for the total group of analysed lymph nodes. All continuous variables were not normally distributed and therefore shown as medians with interquartile ranges (IQR). Categorical variables are presented as absolute numbers with the respective percentages. Cohen's kappa was used for interobserver agreement in stromal scores. Correlations were evaluated between lymph node and their LNM size and stromal percentages, using the Pearson correlation coefficient. Two-tailed P-values of less than 0.05 were considered statistically significant. All statistical analyses were conducted using the IBM SPSS Statistics version 25.0.

Table 1. Measurements of the positive lymph nodes.

Variable (unit)	Description
Diameter positive lymph node (mm)	The largest diameter of the complete positive lymph node
Diameter LNM (mm)	The largest diameter of the metastatic lesion within the positive lymph node
Area positive lymph node (mm ²)	The area of the complete positive lymph node, including the capsule and all directly adjacent lymphoid tissue and vessels within the contour of the lymph node
Area LNM (mm ²)	The area of the largest continuous metastatic lesion in the positive lymph node, or, in case of multiple smaller lesions, including the adjacent metastatic lesions (i.e. lesions for which the distance between two lesions is less than the size of the lesion themselves)
Ratio area LNM : lymph node (%)	The ratio of the area of the metastatic lesion and the complete positive lymph node, portraying the percentage that the metastasis occupies of the lymph node
TSR-c (%)	The ratio of the stromal percentage in comparison to the tumour epithelial percentage within the metastatic lesion, scored in a 3.1 mm ² annotation, and meeting the requirements of van Pelt <i>et al</i>
TSR-t (%)	The absolute amount of stroma in the metastatic lesion, taking in account all tissue types within the metastasis, e.g. necrosis or mucin, and categorized in percentages starting at 0-1%, 1-5% and then per tenfold (i.e. 10%, 20%, etc.)
TSR-c category (no. – perc.) <ul style="list-style-type: none"> - Stroma-low - Stroma-high - Not applicable 	<p>≤50% stroma</p> <p>>50% stroma</p> <p>In the case the official 3.1 mm² annotation area could not meet the requirements and an adjusted annotation size is used</p>
Area adjusted TSR-c annotation (mm ²)	The area of the annotation used for scoring the TSR-c in the case the official 3.1 mm ² annotation area could not meet the requirements
Area of stroma in LNM (mm ²)	A feature portraying the area of the stromal compartment in the metastasis, calculated through the formula: Area stroma in metastasis = (Percentage stroma / 100) * Area metastasis)

LNM, Lymph node metastasis; TSR-c, Tumour-stroma ratio – conventional scoring method; TSR-t, Tumour-stroma ratio – total percentage of stroma.

Results

Patient population and lymph nodes

Patient characteristics are summarized in Table 2. The cohort of 73 patients had a total 350 available resected lymph nodes, of which 208 (59%) were positive and thus contained a LNM. Of these positive lymph nodes, 82 (23%) were fragmented and consequently excluded. Finally, 45 patients with 126 evaluable positive lymph nodes were included in the analysis. A flowchart of the selection of analysed nodes is shown in Figure 1, and an overview of the measured features was provided previously in Table 1.

Table 2. Patient characteristics

Variable (unit)	Result (N=45)
Age (years)	70 (55–78)
Sex (male, number – percentage)	23 (51)
Total lymph nodes per patient (no.)	3 (2–5)
Total positive lymph nodes per patient (no.)	2 (1–3)
Histology (number – percentage)	
- Adenocarcinoma	38 (85)
- Mucinous adenocarcinoma	7 (15)

IQR = interquartile range; Baseline characteristics are presented in medians with interquartile range or numbers with frequencies

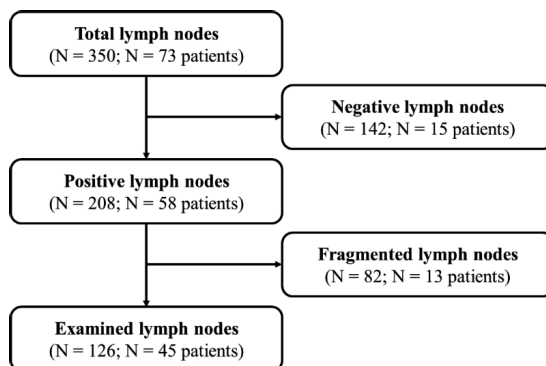


Figure 1. Flowchart showing the selection of analysable lymph nodes.

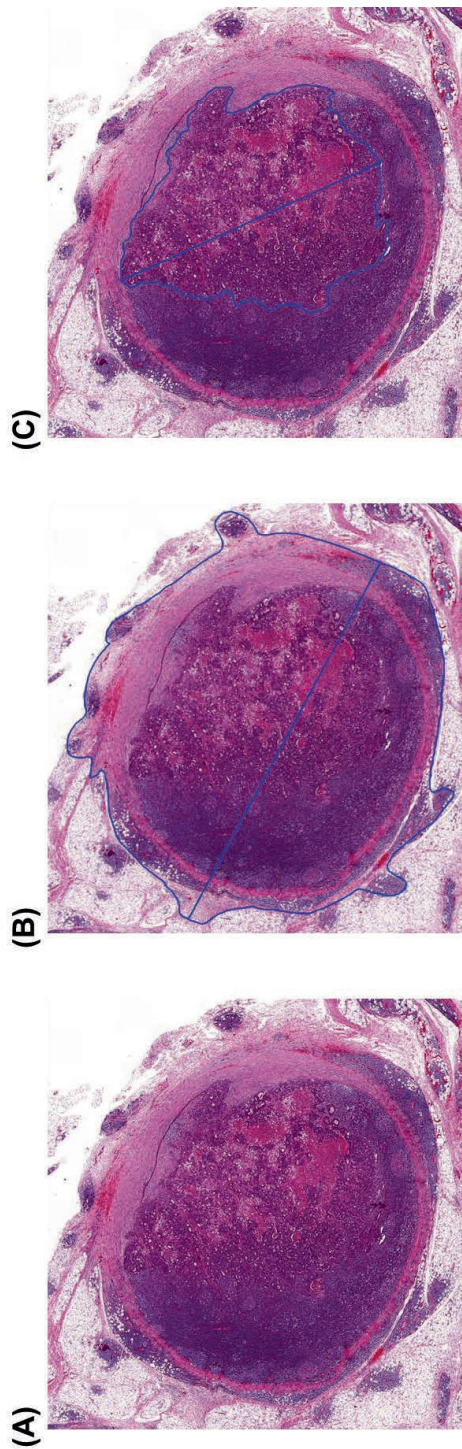


Figure 2. Positive lymph node containing a lymph node metastasis: (A) General overview. (B) Measurement of the diameter and area of the complete lymph node. (C) Measurement of the diameter and area of the lymph node metastasis.

In general, it was noticeable that the LNM were heterogeneous in size, shape and composition, i.e. with regard to stromal percentages and amount of tumour epithelial cells, necrosis and/or mucin, both between as well as within individual patients. Figure 2 illustrates all of the measurements performed for each positive lymph node and their LNM.

Tumour-stroma analysis

Lymph node characteristics are summarized in Table 3. For the TSR-c score, the third, independent observer (PS) was decisive in 5 of the positive lymph nodes (4%). In the rest of the cases and the TSR-t, the scores were similar or agreement was reached, leading to an excellent overall Cohens' kappa of >0.9.

Table 3. Lymph node characteristics

Variable (unit)	Result (N=126)
Diameter positive lymph node (millimetres)	6.4 (4.9–9.6)
Diameter LNM (mm)	5.2 (2.8–8.2)
- Micrometastases (number – percentage)	17 (13.5)
Area positive lymph node (mm ²)	24.5 (12.8–52.6)
Area LNM (mm ²)	13.3 (4.1–34.1)
Ratio area LNM : lymph node (%)	63.4 (31.5–79.0)
TSR-c (%)	35 (20–60)
TSR-t (%)	20 (10–30)
TSR-c category (no.–perc.)	
- Stroma-low	56 (44.4)
- Stroma-high	35 (27.8)
- Not applicable	35 (27.8)
Area adjusted TSR-c annotation (mm ²)	1.0 (0.4–1.7)
Area of stroma in LNM (mm ²)	2.1 (0.5–5.2)

LNM, Lymph node metastasis; TSR-c, Tumour-stroma ratio – conventional scoring method; TSR-t, Tumour-stroma ratio – total percentage of stroma; TSR categories: stroma-low ($\leq 50\%$ stroma) or stroma-high ($> 50\%$ stroma). Baseline characteristics are presented in medians with interquartile range or numbers with frequencies.

The median TSR-c percentage of all the LNMs was 35% (IQR 20–60%). A total of 56 of the LNMs was defined as stroma-low (44%) and 35 (28%) were stroma-high, while in 35 cases (28%), the TSR-c method could not be applied according to the requirements of the guideline for scoring as described by van Pelt et al [29], e.g. due to the limited size of tissue sample. In the latter group of LNMs, a median adjusted TSR-c annotation area of 1.0 mm² (IQR 0.4–1.7 mm²) was used. Figure 3 shows a stroma-high and a stroma-low LNM and the corresponding TSR-c annotations. The median TSR-t score of all the lymph nodes was 20% (IQR 10–30%). A total of 3 (2%) of all analysed LNMs contained <5% tumour stroma, the rest all were scored with a sufficient amount of stroma, i.e. TSR-t \geq 5%.

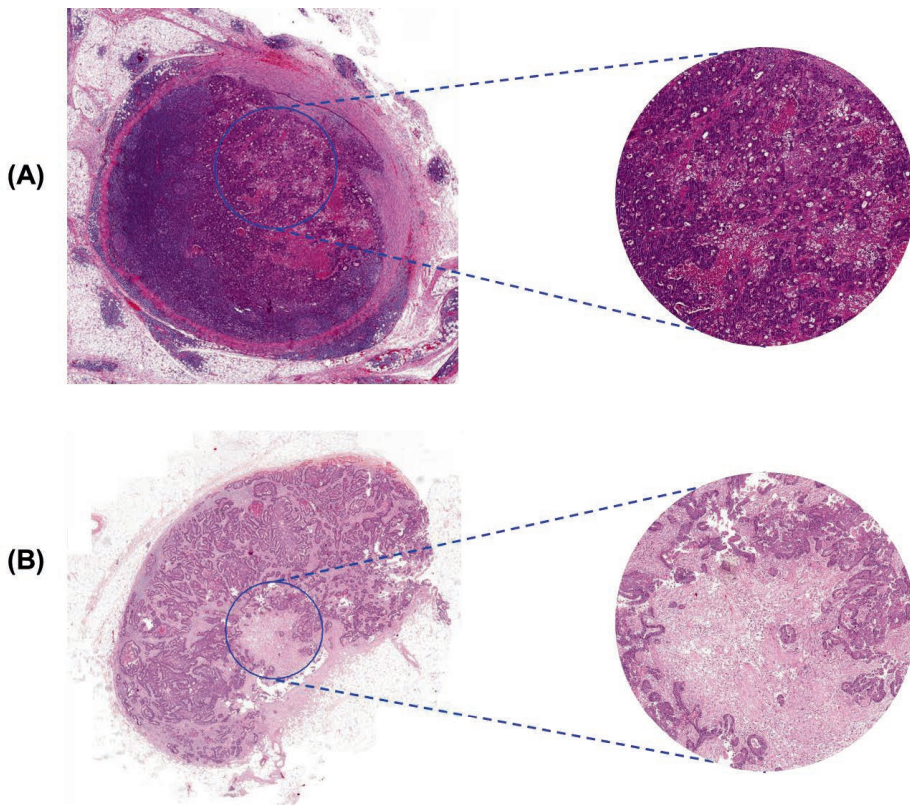


Figure 3. Positive lymph nodes containing a lymph node metastasis with conventional tumour-stroma ratio annotations and corresponding 3.14mm² highlighted areas. (A) Stroma-low metastasis. (B) Stroma-high metastasis.

Subgroup analysis of positive lymph nodes using the radiology cut off

A total of 32 of the 126 positive lymph nodes (25%) had a diameter smaller than the 5.0 mm radiology cut-off value. In these nodes, the median TSR-c of the LNM was 30% (IQR 10–40%), and the median TSR-t was 20% (IQR 10–30%). Furthermore, a total of 52 of the 126 positive lymph nodes (41%) had an area of less than 19.6 mm². A median TSR-c of 30% (IQR 12.50–60%) and TSR-t of 20% (IQR 10–30%) was observed in this subgroup. In all 32 small but positive lymph nodes with the criterion of a diameter of less than 5.0 mm, and in 51 positive lymph nodes covering an area of less than 19.6 mm² (98%), a clear metastatic lesion of $\geq 5\%$ total stroma was observed.

Correlations between positive lymph nodes and the LNM

Correlations between the positive lymph nodes and their respective LNMs were assessed, regarding size and stromal amount, to ascertain whether on current and future imaging, assumptions regarding lymph node size and potential malignancy could be made. Here, correlations were observed between the sizes of the positive lymph node and their LNM; diameters and areas of the lymph node and that of the LNM were significantly positively correlated ($r=0.852$, $P<0.001$ and $r=0.960$, $P<0.001$, respectively), i.e. that the larger the size of the positive lymph node was, a larger LNM was observed. The area of the amount of stroma in the LNM and the total area of the LNM showed a significant positive correlation as well ($r=0.746$, $P<0.001$). Moreover, the LNM stromal compartment area was correlated to the total positive lymph node area ($r=0.747$, $P<0.001$) and diameter ($r=0.612$, $P<0.001$). See Figure 4 for an overview of the scatterplots showing these correlations.

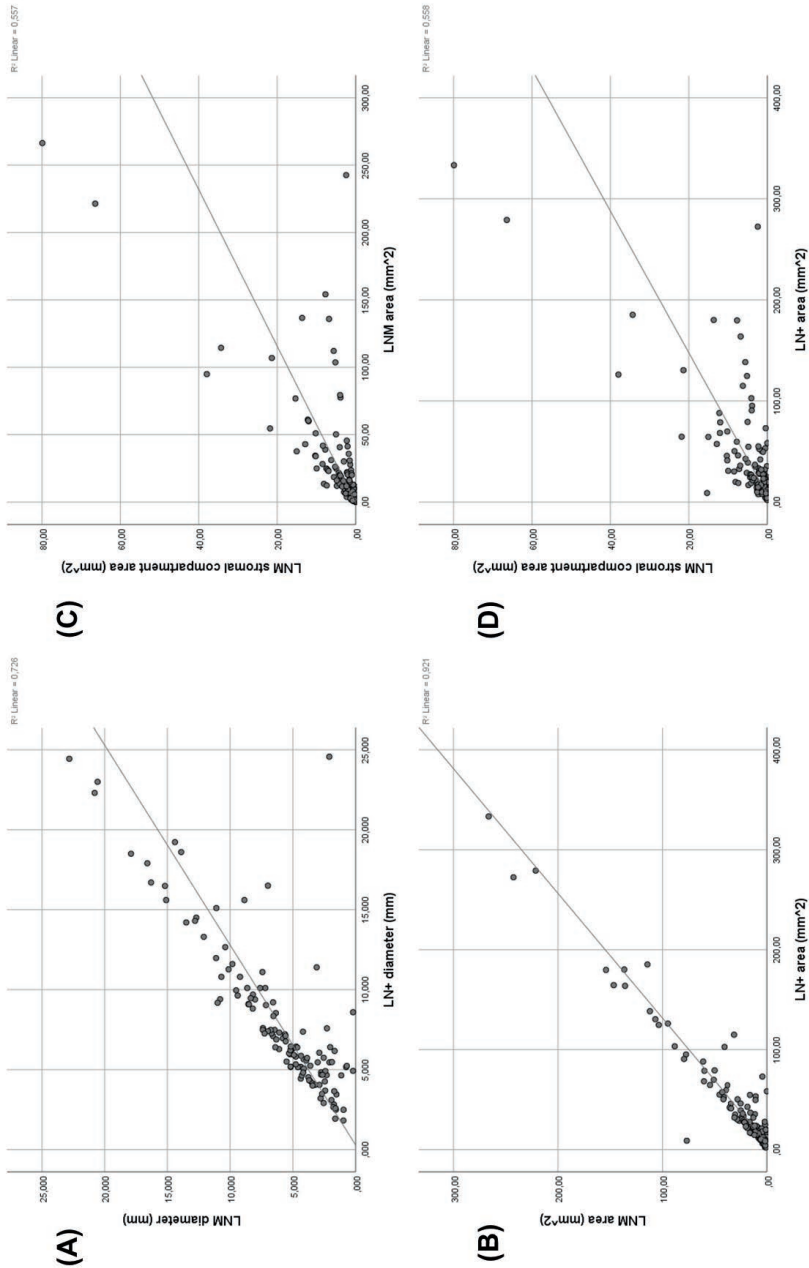


Figure 4. Overview of the different correlations, portrayed by scatterplot diagrams. (continued on next page)

(continued) (A) LNM diameter and positive lymph node diameter ($r=0.852$, $P<0.001$). (B) LNM area and positive lymph node area ($r=0.960$, $P<0.001$). (C) LNM stromal compartment area and LNM area ($r=0.746$, $P<0.001$). (D) LNM stromal compartment area and positive lymph node area ($r=0.747$, $P<0.001$)

Discussion

This study provides an overview of histopathological features of positive lymph nodes, their respective LNMs and especially the tumour stroma within, in stage III CRC patients. First, in line with previous literature as well, we found that nearly all LNMs contained tumour stroma, despite morphological heterogeneity regarding size and proportional composition of non-stromal components [42-44]. Interestingly, all the positive lymph nodes smaller in diameter than 5.0 mm (32 out of 126, 25%), which could potentially have been missed with conventional imaging using the cut-off value as stated by guidelines, contained a sufficient amount of tumour stroma as well, i.e. a TSR-t score $\geq 5\%$.

Furthermore, we have shown that positive lymph node size positively correlates with the size of their respective LNMs, in diameter as well as in area, i.e. that larger positive lymph nodes contain larger LNMs. Both the lymph node and LNM area correlate with the absolute amount of stroma in the LNM, measured in mm² as area. This means that, as positive lymph nodes are larger, a higher absolute amount of stroma can be expected. We have also assessed the proportional composition of LNM, i.e. the proportion of stroma versus other LNM components, e.g. tumour epithelial cells, mucin and/or necrosis, and have shown that the TSR does neither correlate with lymph node diameter nor with LNM diameter or area. In other words, both small and large positive lymph node or LNM may be proportionally stroma-high or stroma-low. These data may at first glance seem contradicting, but this means that, as LNM get larger, the absolute amount of stroma increases, independent of their proportional composition (stroma vs. other components), which is supported by other research [45].

It remains unknown whether the high proportion of stroma in positive lymph nodes will lead to improved detection with FAPI PET/CT, since literature lacks prospective studies with imaging and pathology correlation. Still, in molecular studies, FAP is strongly expressed in the CAFs, highly prevalent in stroma-high tumours, which could be an indication of a high number of binding spots for FAPI [13, 19]. Therefore, positive lymph nodes containing CAFs could potentially be visualized better, which correlates with previous findings, where the novel FAPI PET/CT scan has shown to improve

detection of positive lymph nodes using this more targeted tracer in comparison to the standard FDG PET/CT, even leading to upstaging [21-25].

However, the CAF group makes up a heterogeneous cell population. FAP expression is mainly used as marker to localize and determine the CAFs, but there are many different subtypes of CAFs involved in the process of metastasizing, with various markers. Moreover, although minimally, FAP is expressed by more cells than the CAFs exclusively, e.g. the quiescent stromal cells in other tissue types, leading to some background uptake [46-48]. In addition, it has been demonstrated that inflammation or radiation- and surgery-induced fibrosis can show uptake of FAPI as well [18, 22, 24]. FAPI PET/CT scans may thus give false positive results, necessitating some precaution in evaluating previous study results. Moreover, although recent research distinguishes pathological substrate due to differences in uptake of FAPI, more prospective clinical studies for corroboration of imaging to histopathology need to be initiated as well [18, 22, 23, 40].

There are some limitations to this study. First, the radiological cut off of 5.0 mm in diameter used for analysis is not the cut off used in pathology. In pathology standards, lymph nodes containing only isolated tumour cells (<0.2 mm) are seen as negative nodes, whereas lymph nodes with a LNM larger than 0.2 mm are classified as positive [8]. Moreover, this cut-off value used in radiology likely does not correspond exactly to the pathological 5.0 mm, since the short axis diameter on imaging is fully visible in a three-dimensional plane, whereas on a pathology diagnostic slide, the two-dimensional tissue is arbitrarily prepared. More prospective studies on correlations with radiological and pathological findings are however currently initiated for improved assessment. Second, in almost a third of cases, the TSR-c method could not be applied, e.g. due to limited size of the tissue sample or insufficient viable tissue in the LNM. The standard annotation size was adjusted until the rules were met as stated by the guideline by van Pelt *et al.* [29]. Moreover, although the TSR has also been proven to be of prognostic value in mucinous carcinomas, which have the potential to make as much stroma as adenocarcinomas without a mucinous component, it is not known how the mucin reacts to the FAPI tracer [28, 29]. Lastly, no additional immunohistochemistry was performed on the diagnostic slides, thus the actual presence of FAP on the CAFs was not verified in this work. Since there are many different CAF subtypes with various markers, this potentially could have led to an overestimated number of CAFs and thus FAPI uptake. For optimal corroboration of results, future studies should hence include immunohistochemistry with specific FAP staining.

Only patients with stage III colon cancer who underwent a curative resection without neoadjuvant therapy were included in this study. Therefore, the outcomes cannot be applied to patients receiving neoadjuvant chemotherapy or (chemo)radiotherapy, most commonly applied in rectal cancer patients [49, 50]. Neoadjuvant treatment alters the stromal component, potentially leading to fibrosis with aberrant, less detectable CAFs and thus a disrupted image of the tumour in resection material [51, 52]. In turn, this could cause discrepancies between imaging and histopathology results. Previously, it was found that the TSR is of significant value in predicting therapeutic response of neoadjuvant treatment in oesophageal and breast cancer, with a worse response in stroma-high tumours than stroma-low tumours [34, 35]. It would be interesting to evaluate these treated patients primary tumours and positive lymph nodes, including FAPI PET/CT imaging and biopsy beforehand and subsequent analyse the associated histological slides and the tumour stroma with additional immunohistochemistry, as current literature is lacking herein.

This is, to the best of our knowledge, the first article describing histopathological tumour stromal features of LNMs in CRC patients in detail. In LNMs, stroma is abundantly present, serving as a potentially good substrate for FAPI PET/CT, even in small positive lymph nodes. Hence, the first step for a theoretical framework is taken for new standardization of imaging in cancer. The next step is to correlate tracer uptake in positive lymph nodes on imaging with FAPI PET/CT scanning to the histopathology results in prospective, clinical studies.

Acknowledgements

We would like to thank A.S.L.P. Crobach for his aid and expertise in pathology for the initiation of the project. This work was financially supported from the Bollenstreekfonds, Hillegom, the Netherlands.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021. doi:10.3322/caac.21660.
2. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol.* 2019;14:89-103. doi:10.5114/pg.2018.81072.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394-424. doi:10.3322/caac.21492.
4. Tudyka V, Blomqvist L, Beets-Tan RG, Boelens PG, Valentini V, van de Velde CJ, et al. EURECCA consensus conference highlights about colon & rectal cancer multidisciplinary management: the radiology experts review. *Eur J Surg Oncol.* 2014;40:469-75. doi:10.1016/j.ejso.2013.10.029.
5. Dighe S, Purkayastha S, Swift I, Tekkis PP, Darzi A, A'Hern R, et al. Diagnostic precision of CT in local staging of colon cancers: a meta-analysis. *Clin Radiol.* 2010;65:708-19. doi:10.1016/j.crad.2010.01.024.
6. Schaap DP, Boogerd LSF, Konishi T, Cunningham C, Ogura A, Garcia-Aguilar J, et al. Rectal cancer lateral lymph nodes: multicentre study of the impact of obturator and internal iliac nodes on oncological outcomes. *Br J Surg.* 2021;108:205-13. doi:10.1093/bjs/znaa009.
7. Ogura A, Konishi T, Beets GL, Cunningham C, Garcia-Aguilar J, Iversen H, et al. Lateral Nodal Features on Restaging Magnetic Resonance Imaging Associated With Lateral Local Recurrence in Low Rectal Cancer After Neoadjuvant Chemoradiotherapy or Radiotherapy. *JAMA Surg.* 2019;154:e192172. doi:10.1001/jamasurg.2019.2172.
8. Weiser MR. AJCC 8th Edition: Colorectal Cancer. *Ann Surg Oncol.* 2018;25:1454-5. doi:10.1245/s10434-018-6462-1.
9. Brouwer NPM, Stijns RCH, Lemmens V, Nagtegaal ID, Beets-Tan RGH, Fütterer JJ, et al. Clinical lymph node staging in colorectal cancer; a flip of the coin? *Eur J Surg Oncol.* 2018;44:1241-6. doi:10.1016/j.ejso.2018.04.008.
10. Godefroy J, Ben Haim S, Rosenbach E, Meital AN, Levy A, Chicheportiche A, et al. Perceptual omission errors in Positron emission tomography/Computed tomography reporting. *Q J Nucl Med Mol Imaging.* 2021. doi:10.23736/s1824-4785.21.03339-2.
11. Lu YY, Chen JH, Ding HJ, Chien CR, Lin WY, Kao CH. A systematic review and meta-analysis of pretherapeutic lymph node staging of colorectal cancer by 18F-FDG PET or PET/CT. *Nucl Med Commun.* 2012;33:1127-33. doi:10.1097/MNM.0b013e328357b2d9.
12. Ueno H, Shinto E, Shimazaki H, Kajiura Y, Sueyama T, Yamamoto J, et al. Histologic categorization of desmoplastic reaction: its relevance to the colorectal cancer microenvironment and prognosis. *Ann Surg Oncol.* 2015;22:1504-12. doi:10.1245/s10434-014-4149-9.
13. Sandberg TP, Stuart M, Oosting J, Tollenaar R, Sier CFM, Mesker WE. Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumour-stroma ratio in colorectal cancer. *BMC Cancer.* 2019;19:284. doi:10.1186/s12885-019-5462-2.
14. Joshi RS, Kanugula SS, Sudhir S, Pereira MP, Jain S, Aghi MK. The Role of Cancer-Associated Fibroblasts in Tumour Progression. *Cancers (Basel).* 2021;13. doi:10.3390/cancers13061399.
15. Koustoulidou S, Hoorens MWH, Dalm SU, Mahajan S, Debets R, Seimille Y, et al. Cancer-Associated Fibroblasts as Players in Cancer Development and Progression and Their Role in Targeted Radionuclide Imaging and Therapy. *Cancers (Basel).* 2021;13. doi:10.3390/cancers13051100.
16. Giesel FL, Kratochwil C, Lindner T, Marschalek MM, Loktev A, Lehnert W, et al. (68)Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with Various Cancers. *J Nucl Med.* 2019;60:386-92. doi:10.2967/jnumed.118.215913.
17. Koerber SA, Staudinger F, Kratochwil C, Adeberg S, Haefner MF, Ungerechts G, et al. The Role of (68)Ga-FAPI PET/CT for Patients with Malignancies of the Lower Gastrointestinal Tract: First Clinical Experience. *J Nucl Med.* 2020;61:1331-6. doi:10.2967/jnumed.119.237016.
18. Kratochwil C, Flechsig P, Lindner T, Abderrahim L, Altmann A, Mier W, et al. (68)Ga-FAPI PET/CT: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med.* 2019;60:801-5. doi:10.2967/jnumed.119.227967.
19. Sharma P, Singh SS, Gayana S. Fibroblast Activation Protein Inhibitor PET/CT: A Promising Molecular Imaging Tool. *Clin Nucl Med.* 2021;46:e141-e50. doi:10.1097/rlu.0000000000003489.
20. Ballal S, Yadav MP, Moon ES, Kramer VS, Roesch F, Kumari S, et al. Biodistribution, pharmacokinetics, dosimetry of [68Ga]Ga-DOTA-SA-FAPi, and the head-to-head comparison with [18F]F-FDG PET/CT in patients with various cancers. *European Journal of Nuclear Medicine and Molecular Imaging.* 2020;48:1915-31. doi:10.1007/s00259-020-05132-y.
21. Dendl K, Schlittenhardt J, Staudinger F, Kratochwil C, Altmann A, Haberkorn U, et al. The Role of Fibroblast Activation Protein Ligands in Oncologic PET Imaging. *PET Clin.* 2021;16:341-51. doi:10.1016/j.cpet.2021.03.012.
22. Chen H, Zhao L, Ruan D, Pang Y, Hao B, Dai Y, et al. Usefulness of [(68)Ga]Ga-DOTA-FAPi-04 PET/CT in patients presenting with inconclusive [(18)F]FDG PET/CT findings. *Eur J Nucl Med Mol Imaging.* 2021;48:73-86. doi:10.1007/s00259-020-04940-6.
23. Chen H, Pang Y, Wu J, Zhao L, Hao B, Wu J, et al. Comparison of [(68)Ga]Ga-DOTA-FAPi-04 and [(18)F] FDG PET/CT for the diagnosis of primary and metastatic lesions in patients with various types of cancer. *Eur J Nucl Med Mol Imaging.* 2020;47:1820-32. doi:10.1007/s00259-020-04769-z.
24. Pang Y, Zhao L, Luo Z, Hao B, Wu H, Lin Q, et al. Comparison of (68)Ga-FAPI and (18)F-FDG Uptake in Gastric, Duodenal, and Colorectal Cancers. *Radiology.* 2021;298:393-402. doi:10.1148/radiol.2020203275.
25. Hicks RJ, Roselt PJ, Kallur KG, Tothill RW, Mileskin L. FAPI PET/CT: Will It End the Hegemony of (18)F-FDG in Oncology? *J Nucl Med.* 2021;62:296-302. doi:10.2967/jnumed.120.256271.

26. Ito J, Nogami M, Morita Y, Sakaguchi K, Komada H, Hirota Y, et al. Dose-dependent accumulation of glucose in the intestinal wall and lumen induced by metformin as revealed by (18) F-labelled fluorodeoxyglucose positron emission tomography-MRI. *Diabetes Obes Metab*. 2021;23:692-9. doi:10.1111/dom.14262.
27. Mesker WE, Junggeburst JM, Suzhai K, de Heer P, Morreau H, Tanke HJ, et al. The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumour stage. *Cell Oncol*. 2007;29:387-98. doi:10.1155/2007/175276.
28. van Pelt GW, Sandberg TP, Morreau H, Gelderblom H, van Krieken J, Tollenaar R, et al. The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*. 2018;73:197-206. doi:10.1111/his.13489.
29. van Pelt GW, Kjaer-Frifeldt S, van Krieken J, Al Dieri R, Morreau H, Tollenaar R, et al. Scoring the tumour-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*. 2018;473:405-12. doi:10.1007/s00428-018-2408-z.
30. Smit MA, Philipsen MW, Postmus PE, Putter H, Tollenaar RA, Cohen D, et al. The prognostic value of the tumour-stroma ratio in squamous cell lung cancer, a cohort study. *Cancer Treat Res Commun*. 2020;25:100247. doi:10.1016/j.ctarc.2020.100247.
31. Zunder SM, van Pelt GW, Gelderblom HJ, Mancao C, Putter H, Tollenaar RA, et al. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage II and stage III colon cancer. *Br J Cancer*. 2018;119:164-9. doi:10.1038/s41416-018-0083-0.
32. Huijbers A, Tollenaar RA, v Pelt GW, Zeestraten EC, Dutton S, McConkey CC, et al. The proportion of tumour-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*. 2013;24:179-85. doi:10.1093/annonc/mds246.
33. Huijbers A, van Pelt GW, Kerr RS, Johnstone EC, Tollenaar R, Kerr DJ, et al. The value of additional bevacizumab in patients with high-risk stroma-high colour cancer. A study within the QUASAR2 trial, an open-label randomized phase 3 trial. *J Surg Oncol*. 2018;117:1043-8. doi:10.1002/jso.24998.
34. Hagenaars SC, de Groot S, Cohen D, Dekker TJA, Charehbili A, Meershoek-Klein Kranenbarg E, et al. Tumour-stroma ratio is associated with Miller-Payne score and pathological response to neoadjuvant chemotherapy in HER2-negative early breast cancer. *Int J Cancer*. 2021. doi:10.1002/ijc.33700.
35. van Pelt GW, Krol JA, Lips IM, Peters FP, van Klaveren D, Boonstra JJ, et al. The value of tumour-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol*. 2020;20:39-44. doi:10.1016/j.ctro.2019.11.003.
36. Courrech Staal EF, Smit VT, van Velthuysen ML, Spitzer-Naaykens JM, Wouters MW, Mesker WE, et al. Reproducibility and validation of tumour stroma ratio scoring on oesophageal adenocarcinoma biopsies. *Eur J Cancer*. 2011;47:375-82. doi:10.1016/j.ejca.2010.09.043.
37. van Pelt GW, Hansen TF, Bastiaannet E, van Krieken J. Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *J Med Surg Pathol*. 2016;1.
38. Vangangelst KMH, Tollenaar LSA, van Pelt GW, de Kruif EM, Dekker TJA, Kuppen PJK, et al. The prognostic value of tumour-stroma ratio in tumour-positive axillary lymph nodes of breast cancer patients. *Int J Cancer*. 2018;143:3194-200. doi:10.1002/ijc.31658.
39. Mesker WE, van Pelt GW, Tollenaar R. Tumour stroma as contributing factor in the lymph node metastases process? *Oncotarget*. 2019;10:922-3. doi:10.18632/oncotarget.26644.
40. Wei Y, Cheng K, Fu Z, Zheng J, Mu Z, Zhao C, et al. [(18)F]AIF-NOTA-FAPI-04 PET/CT uptake in metastatic lesions on PET/CT imaging might distinguish different pathological types of lung cancer. *Eur J Nucl Med Mol Imaging*. 2021. doi:10.1007/s00259-021-05638-z.
41. (LWGIT) LWG-IT. Dutch Guideline Colorectal Carcinoma and Colorectal Liver Metastases. Integraal Kankercentrum Neder Land. 2014.
42. Märkl B, Röble J, Arnholdt HM, Schaller T, Krammer I, Cacchi C, et al. The clinical significance of lymph node size in colon cancer. *Mod Pathol*. 2012;25:1413-22. doi:10.1038/modpathol.2012.92.
43. Rodriguez-Bigas MA, Maamoun S, Weber TK, Penetrante RB, Blumenon LE, Petrelli NJ. Clinical significance of colorectal cancer: metastases in lymph nodes < 5 mm in size. *Ann Surg Oncol*. 1996;3:124-30. doi:10.1007/bf02305790.
44. Diaz-Cano SJ. Tumour heterogeneity: mechanisms and bases for a reliable application of molecular marker design. *Int J Mol Sci*. 2012;13:1951-2011. doi:10.3390/ijms13021951.
45. Kotanagi H, Fukuoka T, Shibata Y, Yoshioka T, Aizawa O, Saito Y, et al. The size of regional lymph nodes does not correlate with the presence or absence of metastasis in lymph nodes in rectal cancer. *J Surg Oncol*. 1993;54:252-4. doi:10.1002/jso.2930540414.
46. Costa A, Kieffer Y, Scholer-Dahirel A, Pelon F, Bourachot B, Cardon M, et al. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell*. 2018;33:463-79.e10. doi:10.1016/j.ccell.2018.01.011.
47. Pelon F, Bourachot B, Kieffer Y, Magagna I, Mermet-Meillon F, Bonnet I, et al. Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat Commun*. 2020;11:404. doi:10.1038/s41467-019-14134-w.
48. Ohlund D, Elyada E, Tuveson D. Fibroblast heterogeneity in the cancer wound. *J Exp Med*. 2014;211:1503-23. doi:10.1084/jem.20140692.
49. Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rodel C, Cervantes A, et al. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2018;29:iv263. doi:10.1093/annonc/mdy161.
50. Beets-Tan RGH, Lambregts DMJ, Maas M, Bipat S, Barbaro B, Curvo-Semedo L, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *European Radiology*. 2018;28:1465-75. doi:10.1007/s00330-017-5026-2.

- Chapter 4

51. Prall F, Wöhlke M, Klautke G, Schiffmann L, Fietkau R, Barten M. Tumour regression and mesorectal lymph node changes after intensified neoadjuvant chemoradiation for carcinoma of the rectum. *Apmis*. 2006;114:201-10. doi:10.1111/j.1600-0463.2006.apm_304.x.
52. Hellevik T, Martinez-Zubiaurre I. Radiotherapy and the tumour stroma: the importance of dose and fractionation. *Front Oncol*. 2014;4:1. doi:10.3389/fonc.2014.00001.



Chapter 5

Fibroblast activation protein-expression in colorectal carcinomas and implications for clinical application

Meaghan Polack, Gabi W. van Pelt, Augustinus S.L.P.
Crobach, Lioe-Fee de Geus-Oei, Rob A.E.M. Tollenaar,
J. Han J.M. van Krieken, Wilma E. Mesker

Abstract

Background: Colorectal cancer is highly prevalent. The stromal tumour microenvironment significantly influences tumour behaviour, and cancer-associated fibroblasts (CAFs), as major component of tumour stroma, are increasingly studied. Specifically, CAF-marker fibroblast activation protein (FAP) is gaining interest as tracer for imaging using radiolabelled FAP-inhibitor (FAPI). We describe patterns of FAP-expression, and associations to intratumoural stroma amount, establishing a biological background and potential future reference to pathology assessment.

Materials and methods: Archival histological material from 125 stage-II/III CRC patients was collected. Haematoxylin-and-eosin staining was performed to determine the tumour-stroma ratio (TSR), indicating stromal percentages in primary tumours, lymph nodes (LNs) and biopsies. On immunohistochemistry stains, FAP-expression was semiquantitatively scored as little-no, heterogeneous, or moderate-high expression. Correlation of TSR and FAP-expression with Chi-square testing was assessed. Other patterns were also described, e.g. tumour epithelial FAP-expression.

Results: In total, 93 patients (40 stage-III colon [CC], 53 stage-II/III RC) were included. Correlation between 41 (44%) stroma-high CRC (18/40 CC, 45%; 23/53 RC, 43%) with high FAP-expression was not significant ($P=0.428$ CRC; $P=0.470$ CC; $P=0.615$ RC). The majority of CRC had any FAP-expression (78 CRC, 84%), mostly the invasive front, and in most associated LN metastases (87% CRC tumour-positive LN). However, FAP-expression was often heterogeneous, even staining healthy colon and lymphoid tissue.

Conclusions: CRCs and LN metastases generally express FAP, but levels vary significantly between and within tumours and have no direct correlation with TSR. Care has to be taken translating FAPI-PET/CT results with e.g. disease extent and activity, emphasizing the importance of multidisciplinary approach.

Introduction

Colorectal cancer (CRC) is a multifactorial and multifaceted disease [1]. Although the focus is increasing on prevention through life style improvements and early detection, such as nationwide screening programs, it still remains a highly prevalent disease, accounting for nearly 2 million new cases in 2022 worldwide [2-4]. Therapeutic guidelines are mainly based on clinicopathological variables, such as extent of disease as described by the tumour-node-metastasis (TNM) classification [5-9]. However, the last decades, a paradigm shift has occurred in oncology research. The traditional focus on tumour epithelial cells is increasingly redirected to the tumour microenvironment as a crucial influencer of tumour behaviour [10, 11]. Mainly composed of immune cells, including tumour-infiltrating T-cells and tumour-associated macrophages, vasculature and the angiogenesis process, the extracellular matrix including collagen, and activated quiescent fibroblasts, the tumour stroma is subject to extensive study [11-16]. Crosstalk between this dynamic entity and tumour epithelial cells disrupts the homeostasis and promotes tumorigenesis, in which the quiescent fibroblasts are activated. These cancer-associated fibroblasts (CAFs) have been identified as important contributors [11, 16-18].

CAFs comprise a significant proportion of cells within the tumour stroma and have a proliferative phenotype. Furthermore, they actively induce neoplastic cell growth and through e.g. epithelial-to-mesenchymal transition, recruit more CAFs [11, 16-18]. Studies moreover revealed the pivotal role of CAFs in the development and progression of CRC [19-21]. High levels of accumulated CAFs as well as a high amount of intratumoural stroma are known to predict the course of CRC [22-25]. Indeed, the tumour-stroma ratio (TSR), capturing this phenomenon, has iteratively proven to of predictive value to patient-related outcomes, with stroma-high tumours leading to worse survival [23, 26-28] and poor response to (neo)adjuvant therapy [29-31].

Hence, tumour stroma and especially CAFs constitute attractive targets for improved CRC diagnosis and prognosis. One approach currently being explored in nuclear medicine is the targeting of fibroblast activation protein (FAP), a universal CAF-marker, with the developed fibroblast activation protein-inhibitor (FAPI) [32, 33]. Coupled with a radioisotope label like [18F]Fluor or [68Ga]Gallium, FAPI is increasingly researched using PET/CT scanning to optimise tumour detection, tumour characterisation and even treatment strategies in CRC [34-36]. However, translational studies on the correlating histopathology level, determining FAP expression of CAFs in CRC, are scarce [37-40].

The aim of this study was therefore primarily to assess the levels and patterns of expression of FAP in CRC specimens, establishing a biological background for clinical implementation, which could potentially serve for future reference in pathology assessment. Secondly, we determine the association of the prognostic TSR parameter and the amount of FAP expression of CAFs in CRC. We hypothesized that FAP levels would correlate to TSR, as stroma-high tumours contain more CAFs, theoretically leading to a higher FAP expression. TSR combined with FAP expression could potentially then aid upfront selection of CRC patients for therapeutic strategies, e.g. identifying more aggressive and therapy-resistant stroma-high patients on imaging.

Materials and methods

Patients and material

For this retrospective study, we included 125 CRC patients operated between 2000-2016 at the Leiden University Medical Centre (LUMC): 50 stage III colon carcinoma (CC) and 75 stage II/III rectal carcinoma (RC). Paraffin blocks were collected of resection material from the most invasive part of the primary tumour (PT) of all CRC, of all resected lymph nodes (LNs) of stage III CRC, and of biopsy material from RC patients, as well as all associated clinicopathological data. Exclusion followed when blocks were missing, staining artefacts occurred, or if there was not enough tumour material present. All already available archival material and data were coded and handled according to the National Ethical Guidelines (“Code of proper secondary use of human tissue”). No informed consent was necessary under the legislation for this retrospective analysis. In short, a varied patient population was created to grant insight in potential FAP expression correlation to different histopathology patterns: Supplementary Figure 1 shows a flowchart of ultimately included CRC patients (N=93). Baseline characteristics, treatment types and patient-related outcomes of the cohorts are shown in Supplementary Table 1. General pathology characteristics of the cohorts are further summarized in Table 1.

Table 1. Overview of the general pathology variables of the CC and RC cohorts

Characteristic (unit)	CC cohort (N=40)	RC cohort (N=53)
Tumour diameter – millimetres		
Median maximum diameter (range)	50 (15-125)	30 (0-120)
Residual tumour*		
R0 resection	40 (100)	51 (96)
R1 resection	0 (0)	1 (2)
R2 resection	0 (0)	1 (2)
y(p)T-category**		
y(p)T-stage 0	0 (0)	3 (6)
y(p)T-stage I	0 (0)	2 (4)
y(p)T-stage II	2 (5)	19 (36)
y(p)T-stage III	32 (80)	27 (51)
y(p)T-stage IV	6 (15)	2 (4)
LN (median; range)		
LN examined	12 (4-50)	13 (3-27)
LN positive	2 (1-13)	0 (0-6)
LN – hurdle scoring		
N/A (no LN positive)	0 (0)	32 (61)
Positive LN	40 (100)	21 (40)
No hurdles, normal metastasis and scoring possible	15 (38)	12 (57)
Mucin abundantly present	7 (18)	2 (10)
Necrosis abundantly present	18 (45)	4 (19)
Fibrosis present	N/A	3 (14)
y(p)N-category**		
y(p)N-stage 0	0 (0)	30 (57)
y(p)N-stage I	27 (68)	20 (38)
y(p)N-stage II	13 (33)	3 (6)
(y)pTNM-stage**		
(y)pTNM-stage 0	0 (0)	3 (6)
(y)pTNM-stage I	0 (0)	16 (30)
(y)pTNM-stage II	0 (0)	11 (21)
(y)pTNM-stage III	40 (100)	23 (43)
Tumour morphology (PT or biopsy)		
Adenocarcinoma	31 (78)	51 (96)
Mucinous adenocarcinoma	9 (23)	2 (4)
Differentiation grade tumour		
Well to moderate	4 (10)	9 (17)
Moderate to poorly	22 (55)	31 (59)
Poor to undifferentiated	7 (18)	2 (4)
N/A (no tumour)	0 (0)	4 (8)
Not determined	7 (18)	7 (13)

<i>(continued)</i> Characteristic (unit)	CC cohort (N=40)	RC cohort (N=53)
No risk factors present	3 (8)	15 (28)
Lymphatic invasion	1 (3)	1 (2)
Vascular invasion	3 (8)	2 (4)
Lymphangial invasion	2 (5)	1 (2)
Perineural invasion	1 (3)	5 (9)
Lymphatic and perineural invasion	0 (0)	1 (2)
All of the above	1 (3)	0 (0)
Not determined	29 (73)	29 (53)
Tumour response***		
N/A (no neoadjuvant therapy)	40 (100)	12 (23)
No or nearly no pathological response	N/A	28 (53)
Partial pathological response (pPR)	N/A	19 (18)
Complete pathological response (pCR)	N/A	3 (6)
Tumour regression grade (TRG)****		
N/A (no neoadjuvant therapy)	40 (100)	12 (23)
Absence of regressive changes (TRG5)	N/A	12 (23)
Residual tumour cells outgrowing fibrosis (TRG4)	N/A	16 (30)
More residual tumour cells but fibrosis predominates (TRG3)	N/A	5 (9)
Rare viable tumour cells (TRG2)	N/A	5 (9)
Complete regression, no viable tumour cells (TRG1)	N/A	3 (6)
Microsatellite status		
Not determined	31 (78)	36 (68)
Determined, and		
Microsatellite instable (MSI) and/or deficient mismatch repair (dMMR) enzymes	0 (0)	0 (0)
Microsatellite stable (MSS) and/or proficient mismatch repair (pMMR) enzymes	9 (22)	17 (32)
Mutational burden		
No mutational analysis performed	32 (80)	45 (85)
Performed, and		
No mutations found	4 (10)	2 (4)
KRAS mutation	4 (10)	5 (9)
BRAF mutation	0 (0)	0 (0)
APC mutation	0 (0)	1 (2)

All variables are given as absolute numbers with associated percentages or medians with ranges (minimum-maximum). Sum of percentages can be less or more than 100 due to rounding. CC, colon carcinoma; LN, lymph node; PT, primary tumour; RC, rectal carcinoma; TNM, tumour-node metastasis.

*Residual tumour according to Wittekind (2009). **Different versions of the TNM classification were used, here all variables are converted to the AJCC/UICC TNM version 5 (1997). ***Tumour response is roughly categorized in three; no/partial/complete response. ****Tumour regression grade as categorized by Mandard.

Staining protocols

Additional 5 µm thick paraffin sections were cut from all collected archival blocks, on which standard haematoxylin and eosin (H&E) staining and immunohistochemical staining (IHC) using a FAP antibody (rabbit monoclonal antibody IgG, ERP20021, ab207178, Abcam - dilution 1:200) was performed, according to Sandberg et al [20]. After deparaffinization and rehydration, antigen retrieval was performed using the PreTreatment Link Module (Dako), following mounting with antifade reagent, and lastly, counterstaining (DAPI, 2µg/ml Sigma-Aldrich).

Tumour stroma and pathology analysis

The H&E-stained slides were used for the TSR, scored by two trained independent researchers blinded for clinical data (MP and GWvP). Subsequently, categorisation into stroma-low ($\leq 50\%$ stroma) or stroma-high ($>50\%$ stroma) followed according to van Pelt et al.[41] (Supplementary Figure 2). Cohen's interobserver agreement kappa's were 0.87 for CC (almost perfect agreement; scored on PT), and 0.67 for RC (substantial agreement; scored on biopsy). As neoadjuvant treatment alters the tumour stromal compartment, e.g. creating fibrosis, the TSR in PT material in neoadjuvant treated RC is not representative [42]. However, as was proven in literature, the TSR of biopsy material is deemed representative for the TSR of the associated PTs [31, 43], hence, this was used to categorize RC. In CC, as these patients did not receive neoadjuvant treatment, the complete PT material was available, and no biopsy material was collected. Moreover, we had previously observed that the LN containing the highest stromal percentage was significantly associated to patient outcomes, thus this LN was deemed representative for all LNs, as well [30, 44].

FAP expression analysis

FAP expression was assessed by two observers, also blinded for clinical data (pathologist JHJMvK, researcher MP). Biopsies, PTs and associated LNs were all semi quantitatively scored on IHC slides, in stroma as well as tumour epithelium: (0) little to no, (1) heterogeneous, and (2) moderate to high expression. In LNs, distinction was made for assessments of the staining of metastases as well as healthy lymphoid tissue. For comparison and association to TSR for future implications, only moderate-high expression of FAP was deemed sufficient for potential outcomes.

Slide and statistical analysis

Slides were scanned with the Panoramic 250 scanner (3DHistech, Hungary) (tissue level pixel size ~0.33 $\mu\text{m}/\text{pixel}$; 20x magnification). Annotations and figures were made using the 3DHistech SlideViewer 2.7 software. Continuous variables were expressed in medians with interquartile ranges (IQR), and nominal and ordinal variables in number of frequencies with corresponding percentages. Statistical analysis was performed using Chi-square analyses with significance determined at two-tailed P-values <0.05, using IBM SPSS Statistics 29.0.

Results

TSR findings

TSR characteristics per cohort are summarized in Table 2. A majority of both CC (scored on PT; N=22, 55%) and RC (scored on biopsy; N=30, 57%) were stroma-low, similar to previous studies [23, 26-28]. The TSR of biopsies in all RC and their associated PT did not completely correlate (P=0.053; Supplementary Table 2), caused by a large insignificant effect on neoadjuvant treated RC PTs (P=0.785; Supplementary Table 2), illustrating that the TSR can not be scored in resected material due to extensive fibrosis induced by any neoadjuvant treatment [45]. The TSR of treatment-naïve RC biopsies did however significantly correlate to the TSR of the PT (P=0.007), which confirms the reliability and representativeness of biopsy-assessed TSR scores (Supplementary Table 2).

Table 2. Overview of the TSR and FAP variables of the CRC and separate CC and RC cohorts.

Characteristic (unit)	CRC cohort (N=93)	CC cohort (N=40)	RC cohort (N=53)
TSR – biopsy			
Stroma-low	30 (32)	N/A	30 (57)
Stroma-high	23 (25)	N/A	23 (43)
N/A (no biopsy)	40 (43)	40 (100)	N/A
TSR – PT*			
Stroma-low	38 (41)	22 (55)	16 (30)
Stroma-high	52 (56)	18 (45)	34 (64)
N/A (no tumour)	3 (3)	N/A	3 (6)
TSR – representative LN**			
Stroma-low	51 (55)	22 (55)	29 (55)
Stroma-high	20 (22)	18 (45)	2 (4)
N/A (no LN positive)	22 (24)	0 (0)	22 (42)
FAP overall expression – biopsy			
Little to no expression	6 (6)	N/A	6 (11)
Heterogeneous expression	26 (28)	N/A	26 (49)
Moderate to high expression	21 (23)	N/A	21 (40)
N/A (no biopsy)	40 (43)	40 (100)	N/A
FAP expression tumour epithelium – biopsy			
No, only tumour stroma	23 (25)	N/A	23 (43)
Heterogeneous expression	18 (19)	N/A	18 (34)
Yes, high expression in tumour epithelium	12 (13)	N/A	12 (23)
N/A (no biopsy)	40 (43)	40 (100)	N/A
FAP overall expression – PT			
Little to no expression	11 (12)	7 (18)	4 (8)
<i>Any expression, of which</i>	82 (88)	33 (83)	49 (92)
Heterogeneous expression	47 (61)	25 (63)	32 (60)
Moderate to high expression	25 (27)	8 (20)	17 (32)
FAP expression tumour epithelium – PT			
No, only tumour stroma	43 (46)	33 (83)	10 (19)
Yes, any FAP expression, of which	47 (51)	7 (18)	40 (75)
Heterogeneous expression	26 (28)	5 (13)	21 (40)
High expression in tumour epithelium	21 (23)	2 (5)	19 (36)
N/A (no tumour)	3 (3)	N/A	3 (6)

<i>(continued)</i> Characteristic (unit)	CRC cohort (N=93)	CC cohort (N=40)	RC cohort (N=53)
N/A (no LN positive)	32 (34)	0 (0)	32 (60)
Positive LN, of which	61 (66)	40 (100)	21 (40)
Little to no expression	15 (25)	14 (35)	1 (2)
Heterogeneous expression	21 (32)	17 (43)	4 (8)
Moderate to high expression	25 (41)	9 (23)	16 (30)
FAP expression tumour epithelium – representative LN**			
N/A (no LN positive)	32 (34)	0 (0)	32 (60)
Expression, of which	61 (66)	40 (100)	21 (40)
No, only expression in tumour stroma	35 (57)	32 (80)	3 (14)
Heterogeneous expression	15 (25)	8 (20)	7 (33)
Yes, moderate to high expression in tumour epithelium	11 (18)	0 (0)	11 (52)
FAP expression – general			
Equal distribution of FAP expression	15 (16)	7 (18)	8 (15)
High expression in remodelling areas	78 (84)	33 (83)	45 (85)
FAP discrepancy in expression – biopsy and PT			
No discrepancy, similar expression	25 (27)	N/A	25 (47)
Slight discrepancy, biopsy higher	14 (15)	N/A	14 (26)
Slight discrepancy, PT higher	12 (13)	N/A	12 (23)
Major discrepancy	2 (2)	N/A	2 (4)
N/A (no biopsy)	40 (43)	40 (100)	N/A
FAP discrepancy in expression – PT and representative LN**			
No discrepancy, similar expression	27 (29)	16 (40)	11 (21)
Slight discrepancy, PT higher	9 (10)	9 (23)	0 (0)
Slight discrepancy, LN higher	19 (20)	11 (28)	8 (15)
Major discrepancy	6 (6)	4 (10)	2 (4)
N/A (no LN positive)	32 (34)	0 (0)	32 (60)

All variables are given as absolute numbers with associated percentages or medians with IQR. Sum of percentages can be less or more than 100 due to rounding. CC, colon carcinoma; FAP, fibroblast activation protein; IQR, interquartile range; LN, lymph node; N/A, not applicable; PT, primary tumour; RC, rectal carcinoma; TSR, tumour-stroma ratio.

*All primary tumours are shown here, including those with neoadjuvant treatment.

**Representable lymph node is the lymph node with most tumour stroma.

In total, we analysed 288 CRC LNs, of which 125 (43%) were positive for metastasis, accounting for 61 patients (66%; 40 CC and 21 RC patients). The TSR of LNs with metastases was not associated with the TSR score in the PT ($P=0.529$ in CRC; Table 3), emphasising their heterogeneity. A trend was seen in stroma-high CC patients that their LNs were more often difficult to score due to e.g. necrosis ($P=0.071$), whereas in RC merely 2 patients had stroma-high tumour-positive LNs, both potentially skewing results (Supplementary Table 3).

FAP expression

General FAP characteristics are given in Table 1. During microscopic analysis of FAP expression, various patterns were noted and scored accordingly, illustrated by Figure 1. FAP was expressed in the majority of CRC patients (78 in CRC, 33 in CC, 83%; and 47 in RC, 89%) (Figure 1A-B), as well as in associated LN metastases (87% of all CRC tumour-positive LN) (Figure 1C-D). However, levels of FAP expression varied and could also be heterogeneous throughout the tumour (61% in CRC PT) (Figure 1E-F). In some cases, also some FAP expression was seen in tumour epithelium of PT as well (51% in CRC, mostly in RC after neoadjuvant treatment [75%]) (Figure 1G-H). FAP was expressed mostly in areas of active remodelling, i.e. the invasive front (78 in CRC; 84%). Overall, there was often a discrepancy in intensity of FAP expressed by biopsy and PT, as well as between PT and LN (Table 1).

Correlations FAP expression and TSR and other findings

In ascertaining the correlation between TSR and FAP expression in CRC, although stroma-high CRC more often expressed higher levels of FAP, this did not reach significance ($P=0.428$ in CRC; $P=0.470$ in CC and $P=0.615$ in RC, respectively) in their counterpart nor in their LNs (Table 3). Despite the fact that the majority of PTs and LN metastases express FAP, still 22% of all analysed CRC tumour-positive LN did not stain for FAP (Supplementary Table 3). Moreover, in 22% of tumour-negative CRC LNs, FAP conversely did get expressed, e.g. in overactive germinal centres (Figure 2A-B). A mean percentage of 69% of all analysed CRC LNs per patient was found that was tumour-positive and had corresponding FAP-positivity, whereas 31% of tumour-positive LNs did not express FAP, and 14% of tumour-negative LNs were actually FAP-positive within the patient (Supplementary Table 3).

Table 3. Comparison of stroma-low and stroma-high tumours within the CC and RC cohorts

Characteristic (unit)	CRC cohort (N=93)		CC cohort (N=40)		RC cohort (N=53)	
	Stroma-low (N=52)	Stroma-high (N=41)	P-value	Stroma-low (PT; N=22)	Stroma-high (PT; N=18)	P-value
Clinical TNM-stage*			0.705#			0.332#
cTNM-stage I	1 (2)	0 (0)		0 (0)	0 (0)	
cTNM-stage II	17 (33)	12 (29)		6 (27)	7 (39)	
cTNM-stage II/III**	12 (23)	8 (20)		11 (50)	8 (44)	
cTNM-stage III	22 (42)	21 (51)		5 (23)	3 (17)	
Treatment type, grouped			0.141#			0.022#
Neoadjuvant therapy and surgery	20 (39)	21 (51)		0 (0)	0 (0)	
Surgery and adjuvant therapy	12 (23)	12 (29)		12 (55)	11 (61)	
Surgery alone	20 (39)	8 (20)		10 (45)	7 (39)	
Pathology risk factors***						
Not determined	31 (60)	26 (63)		18 (82)	11 (61)	
Determined, and	21 (40)	15 (37)	0.735#	4 (18)	7 (39)	0.125#
No pathology risk factors present	10 (48)	8 (53)		0 (0)	3 (43)	
Yes, 1 or more risk factors present	11 (52)	7 (47)		4 (100)	4 (57)	
TSR						
TSR in PT			N/A			0.053#
Stroma-low	N/A	N/A		N/A	N/A	
Stroma-high	N/A	N/A		N/A	N/A	
N/A (no tumour)	N/A	N/A		N/A	N/A	

<i>(continued)</i> TSR in representative LN**** N/A (missing, no LN positive) Positive LN, of which Stroma-low Stroma-high	18 (35)	14 (34)	0 (0)	0 (0)	18 (60)	14 (61)
	34 (65)	27 (66)	0.529#	18 (100)	12 (40)	9 (39)
	24 (71)	17 (63)		9 (50)	11 (92)	8 (89)
	10 (29)	10 (37)		9 (41)	1 (8)	1 (11)
			0.428#			0.615#
FAP expression – with counterpart						
Little expression (no - heterogeneous expression)	37 (71)	26 (63)		18 (82)	19 (63)	13 (57)
Moderate to high expression	15 (29)	15 (37)		4 (18)	11 (37)	10 (44)
FAP in representative LN****						
N/A (missing, no LN positive)	18 (35)	14 (34)		0 (0)	18 (60)	14 (61)
Positive LN, of which	34 (65)	27 (66)	0.191#	18 (100)	12 (40)	9 (39)
Little expression (no - heterogeneous expression)	17 (50)	18 (67)		16 (73)	1 (8)	4 (44)
Moderate to high expression	17 (50)	9 (33)		6 (27)	11 (92)	5 (56)
Any FAP expression in tumour epithelium – with counterpart			0.576#			0.259#
No, only in tumour stroma	30 (58)	26 (63)		19 (86)	11 (37)	12 (52)
Yes, also some in tumour epithelium	22 (42)	15 (37)		3 (14)	19 (63)	11 (48)
FAP expression in remodelling areas			0.431#			0.683#
Equal distribution of FAP expression	7 (14)	8 (20)		3 (14)	4 (13)	4 (17)
High expression in remodelling areas	45 (87)	33 (81)		19 (86)	26 (87)	19 (83)
FAP discrepancy – biopsy-PT			N/A			0.657#
No discrepancy, similar expression	N/A	N/A		N/A	15 (52)	10 (46)
Yes, any discrepancy	N/A	N/A		N/A	14 (48)	12 (55)

<i>(continued)</i> FAP discrepancy – PT-representative LN**** N/A (missing, no LN positive)	18 (35)	14 (34)	0 (0)	0 (0)	18 (60)	14 (61)
Positive LN, of which	34 (65)	27 (66)	0.586#	22 (100)	18 (100)	0.604#
No discrepancy, similar expression	14 (41)	13 (48)		8 (36)	8 (44)	5 (56)
Yes, any discrepancy	20 (59)	14 (52)		14 (64)	10 (56)	4 (44)
Tumour response*****			N/A			0.864#
No or nearly no pathological response	N/A	N/A		N/A	N/A	9 (45)
Partial pathological response (pPR)	N/A	N/A		N/A	N/A	10 (50)
Complete pathological response (pCR)	N/A	N/A		N/A	N/A	1 (5)
Pathological TNM-stage*			0.828#			0.829#
(y)pTNM-stage 0	1 (2)	2 (5)		0 (0)	0 (0)	1 (3)
(y)pTNM-stage I	9 (17)	7 (17)		0 (0)	0 (0)	9 (30)
(y)pTNM-stage II	7 (14)	4 (10)		0 (0)	0 (0)	7 (23)
(y)pTNM-stage III	35 (67)	28 (68)		22 (100)	18 (100)	13 (43)

All variables are given as absolute numbers with associated percentages. Sum of percentages can be less or more than 100 due to rounding. CC, colon carcinoma; FAP, fibroblast activation protein; LN, lymph node; N/A, not applicable; PT, primary tumour; RC, rectal carcinoma; TNM, tumour-node-metastasis; TSR, tumour-stroma ratio.

*Different versions of the TNM classification were used, here all variables are converted to the AJCC/UICC TNM version 5 (1997).

**Stage II/III is given in case no imaging is performed and locoregional extent of disease is uncertain.

*** Pathology risk factors include lymphatic invasion, venous invasion, perineural invasion.

****Representable lymph node is the lymph node with most tumour stroma.

*****Tumour response is roughly categorized in three; no/partial/complete response.

Calculated with the Chi-square test.

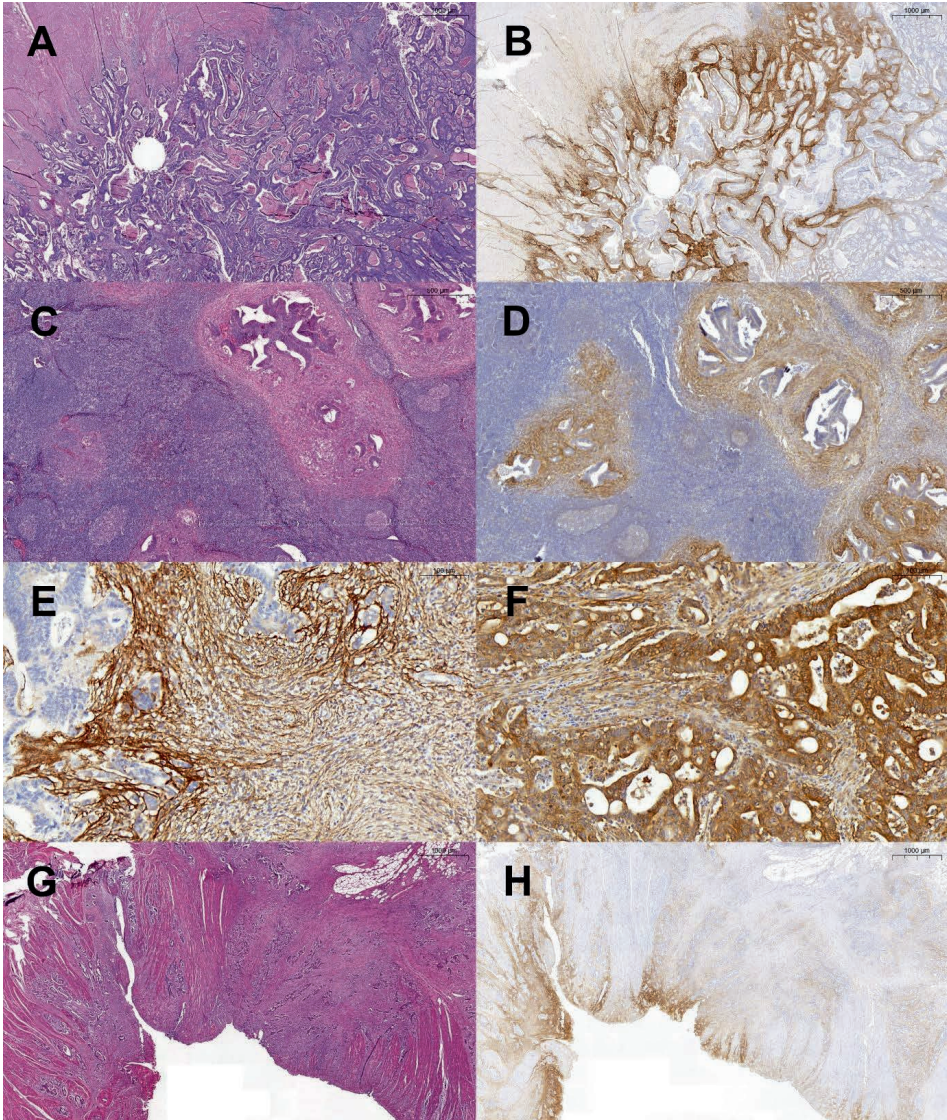


Figure 1. (A-B) Haematoxylin-and-eosin (H&E)-stained slide and immunohistochemistry-stained fibroblast activation protein (FAP) slide examples of a primary tumour and expression towards the invasive front. Punch hole for localisation (2.0x magnification). (C-D) H&E and FAP examples of lymph node metastases (5.0x magnification). (E-F) FAP expression in stroma with stained cancer-associated fibroblasts, compared to strong staining of FAP in tumour epithelial cells more than in stroma (20.0x magnification). (G-H) H&E and FAP examples of a primary tumour with heterogeneous expression (2.0x magnification).

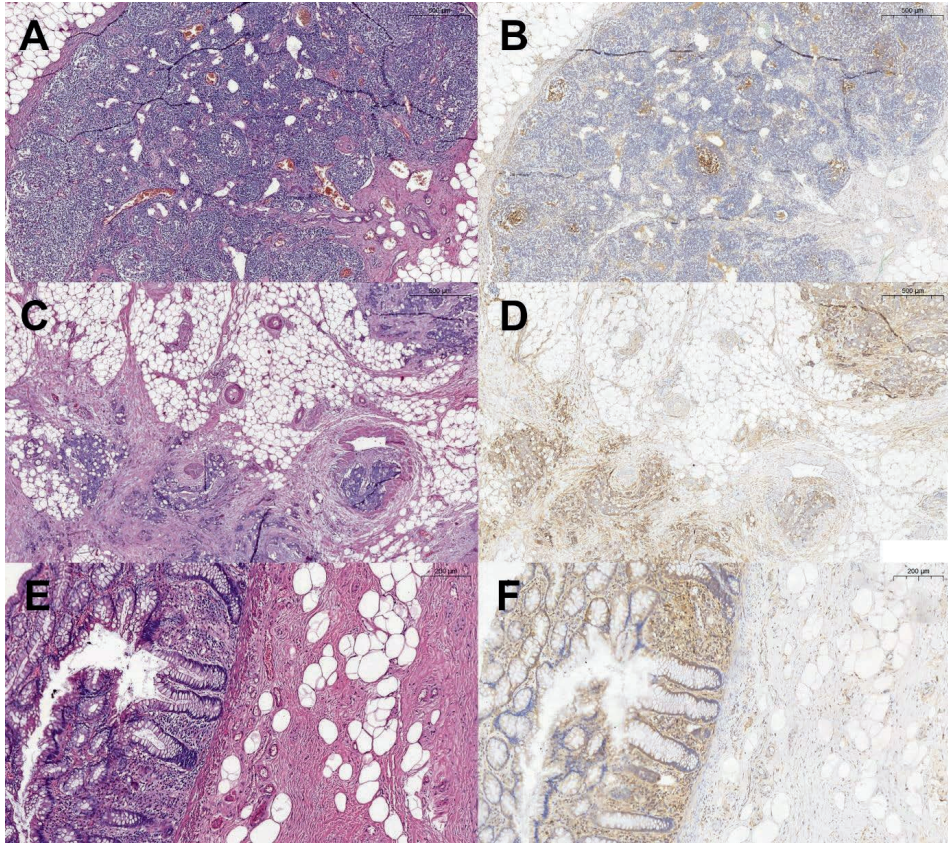


Figure 2. (A-B) H&E and FAP of a tumour-negative lymph node metastasis but positive FAP expression (10.0x magnification). (C-D) H&E and FAP examples of a tumour deposit and extramural venous invasion, strongly staining for FAP, illustrating aggressiveness (5.0x magnification). (E-F) H&E orientation for FAP expression in post-radiation mucosa in a rectal cancer patient with a pathological complete response (10.0x magnification). FAP, fibroblast activation protein; H&E, haematoxylin and eosin staining.

Interestingly, the sporadically present tumour deposits often strongly expressed FAP, potentially illustrating tumour aggressiveness (Figure 2C-D). In this study, no correlation could be found between TSR scores of biopsies and predicted tumour regression grade or response of pathological TNM stage (ypTNM) in RC, although a large bias exists in treatment type ($P=0.022$) and a trend towards higher a priori stages in stroma-high RC ($P=0.332$) (Table 2). We also observed, of important note, in mucosa, for instance post-radiation, expression of FAP as well, i.e. in one RC patient with a pathological complete response (Figure 2E-F).

Discussion

In this study, we hypothesized that stroma-high CRC, composed of more CAFs, would have higher FAP-marker expression. Theoretically, this could aid upfront selection of patients for treatment strategies, where for instance on a future FAPI PET/CT scan, more tracer uptake would indicate more aggressive and therapy-resistant stroma-high tumours. However, no direct correlation between the well-established TSR parameter and FAP expression was observed. Moreover, although FAP was expressed in the majority of CRC patient material, this was not ubiquitous and more importantly, also non-tumourous tissue stained positive for FAP. In conclusion, our results elucidate a biological background in contrast to current literature, necessitating care on the direct translation of FAPI as radiopharmaceutical for CRC detection and staging, which is increasingly researched, emphasizing the importance of a multidisciplinary approach.

FAP is a proteolytic enzyme on the cell surface of CAFs, promoting tumour invasion through processes like active extracellular matrix degradation and tissue remodelling [46]. Most tumour types contain CAFs expressing FAP, including gastrointestinal cancers as also described by Mona et al [37]. Our results confirm that the majority of CRC indeed expressed FAP, however, that this was more intense towards the invasive front. This most likely explains the discrepancy with TSR scores, which are generally acquired from the tumour centre [20, 41, 47]. Interestingly, we observed strong staining for FAP in the sporadically present tumour deposits and extramural venous invasion. This is corroborated with other literature, stating that higher FAP expression or FAP expression in tumour epithelium was associated with more invasive tumours [46, 48], emphasizing the correlation of FAP expression and tumour aggressiveness.

Traditionally, after an initial diagnosis with an endoscopic biopsy, determining extent of disease as major characteristic of tumour behaviour, is important for diagnosis and prognosis prediction, also in CRC [6, 7]. This assessment of PT, LNs and potential distant metastases, using the TNM stage classification, is concurrently performed through imaging [5]. However, especially accurate nodal staging remains challenging despite current high-standard imaging modalities and techniques [49-51]. Since FAPI PET/CT scanning could improve staging as described in literature [52, 53], and as we had previously seen that most LN metastases, independent from size, contain tumour stroma [54], we aimed to determine in this present study whether FAP, as a proven derivative of FAPI PET/CT imaging, might indeed theoretically, biologically, improve diagnostic accuracy [37].

However, our results illustrate that FAP expression was prone to variation. Not all CRC expressed FAP, and moreover, non-cancerous tissue could also exhibit FAP expression. Although this could in part be due to general intratumoural heterogeneity [55], the clinical consequences are far-reaching. In a few cases, when directly translating FAP expression to future FAPI tracer uptake in LNs for instance, this could have potentially lead to understaging in a patient. As an observed FAP-negative but tumour-positive LN could in turn lead to a clinical stage II but pathological stage III tumour, this could have necessitated other (neo)treatment strategies [6, 7]. Examples from literature also mention specifically FAP staining in post-radiation mucosa, dysplasia or fibrosis [39, 56], which our results corroborated, e.g. FAP expression in some patients with pathological complete response. Consequently, RC watch-and-wait strategies [57] could be rendered insufficient with mere FAPI PET/CT in follow-up, as these patients could be subjected to surgeries while potentially have reached a complete regression of their tumour. Especially with the rise of newer neoadjuvant chemoradiotherapy combination regimens, like the total neoadjuvant treatment strategy or RAPIDO trial [58], this could influence diagnostic results. Also, non-oncologic diseases leading to fibrosis, scarring or inflammation like Crohn's disease have been found to influence FAPI-uptake [59, 60], rendering assessment in patients with chronic illnesses inaccurate and requiring further research [61].

Importantly, while FAP is widely researched and deemed the universal CAF-marker, there are however, many CAF subtypes. Derived from a variety of origins through direct recruitment or an epithelial-to-mesenchymal transition, a concrete molecular consensus is thus lacking for this heterogeneous group and merely distinguishable from tumour epithelial cells by using panels with different markers [20, 62]. CAFs are generally associated with tumour invasion and other hallmarks of tumour aggressiveness, such as treatment resistance, angiogenesis and immunosuppression, e.g. by producing growth factors and inflammatory cytokines (inflammatory CAFs; iCAFs). Contrarily, some exhibit antitumoural properties, for instance in antigen-presenting (apCAF) subtypes, or even dual functions, like alpha-

smooth muscle actin-expressing CAFs, identifying a more myofibroblast-like (myCAF) cell [63, 64]. Solely focussing on FAP could thus possibly lead to underrepresentation of fibroblast presence or even misinterpretation of tumour behaviour. Furthermore, CAF subtypes in metastases have been shown to vary from those remaining in the PT, more actively inducing cell invasion and are of deleterious influence on patient survival [65, 66]. Crucial research is therefore currently being initiated to unravel this complex entity [67].

This study has some limitations. We performed a single-centre study with a relatively small patient cohort, also hindering patient-related outcome analyses, with inclusions up until 2016. Novel treatment strategies are hence not taken into account, and moreover, data on currently per protocol analysed microsatellite or mutational status [6, 7], was scarce or not available. Additionally, a limitation also pertained to our semi quantitative scoring methodology. Although these semi quantitative measures are often used [68], e.g. in Mona et al. [29], visual assessment could potentially lead to suboptimal scores [69]. Standardisation or automatic analysis using artificial intelligence could overcome this hurdle in the future and support the increasing workload for pathologists, as is already being endeavoured with the TSR [70]. Lastly, despite the possibility of some variations of the chemical compounds within the FAPI-tracer, which could in theory lead to a minor discrepancy with standard IHC staining with the FAP antibody [34], the biodistribution did previously correlate, granting a representative insight in CRC [37, 38]. Results of future studies on the correlation between FAPI uptake on imaging and pathology assessment are awaited, like the upcoming clinical trials FAPI-CRC1 (NCT05209750) and FoCus (NCT06191120).

However, this study has strengths as well, including the in-depth analysis of a wide variety of CRC patients, hence establishing a literal biological framework for current trends in research. Although our study did not find a direct correlation between biopsy-scored TSR and the response on treatment and/or tumour regression grade in RC, most likely due to bias, we aim to analyse this specifically in a larger future study. Moreover, as our research group had previously found that the TSR in LN metastases formed an additional prognostic parameter, we assessed whether FAP expression in LNs could be predicted by the PT [44]. The TSR scores, however, like FAP expression, show high heterogeneity between the PT and LNs, highlighting once again the importance of analysing all LNs in pathology, as stroma-high LN metastases indicate more aggressive tumours and lead to worse patient-related outcomes [30, 44].

In conclusion, CRC and associated positive LNs can express FAP, though this is subject to variation, necessitating care in translating future intensity of FAPI-PET/CT directly into clinical results, e.g. tumour characterisation. No clear agreement between TSR and FAP expression exists, hindering prediction of a TSR score in patients from mere imaging of uptake through FAPI PET/CT scanning and underlining the importance of collaboration for pathology assessment. In light of the increasing interest and research the past years, these findings need to be considered when FAPI as radiopharmaceutical in PET/CT scanning is implemented to improve CRC staging, emphasizing the need for a multidisciplinary approach.

Acknowledgements

We thank Cindy Cano Ramos, Shriya Boedhai and Lois Stokkers for their work in the lab, along with Ronald van Vlierberghe and Geeske Dekker-Ensink for laboratory support and scanning.

Funding

This work was supported by the Bollenstreekfonds, Hillegom, the Netherlands [no grant number]. This funder had no role in study design, data collection and analysis, nor in the decision to publish, nor in the preparation of the manuscript.

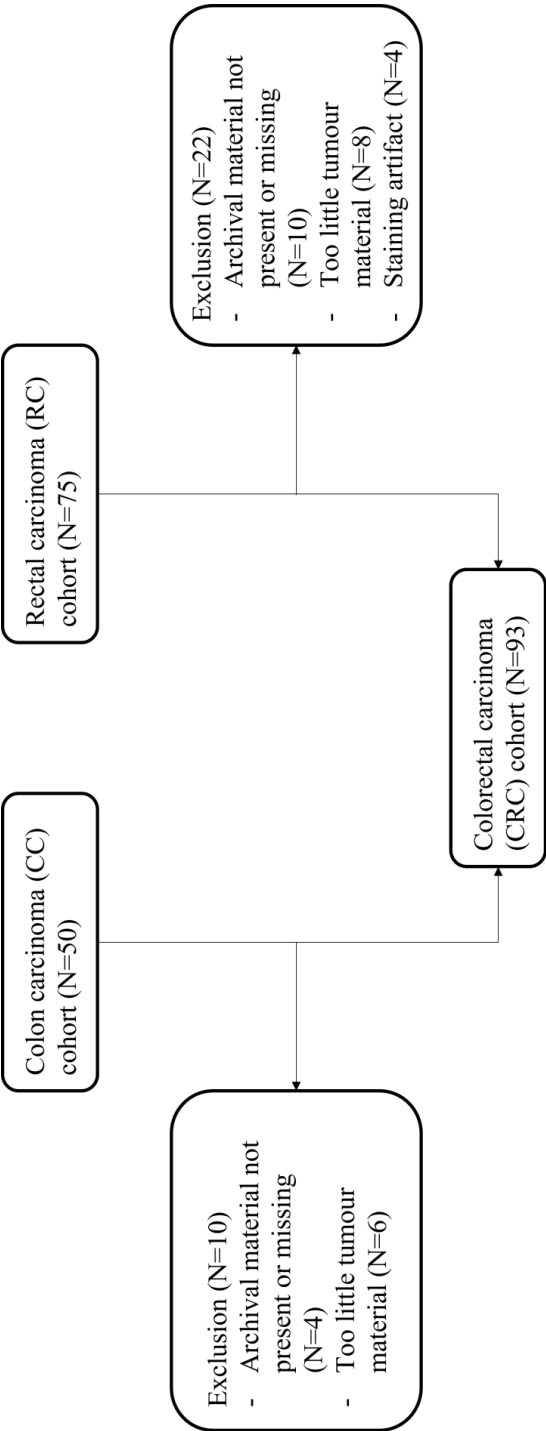
References

1. Dekker, E., et al., Colorectal cancer. *Lancet*, 2019. 394(10207): p. 1467-1480.
2. Bray, F., et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024. 74(3): p. 229-263.
3. Ahmadpour, S., et al., Early Detection of Colorectal Cancer: Conventional Techniques and Current Biomarkers. *Current Cancer Therapy Reviews*, 2024. 20(4): p. 327-340.
4. Modarres, M.P., et al., Alterations of the Gut Microbiome and Metabolome Following: A Opportunity for Early Detection of Colorectal Cancer. *Current Cancer Therapy Reviews*, 2025. 21(1): p. 10-22.
5. Brierley, J.D., Gospodarowicz, M.K., Wittekind, C., *The TNM Classification of Malignant Tumours*. 8 ed. 2016: Wiley Blackwell.
6. Argiles, G., et al., Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2020. 31(10): p. 1291-1305.
7. Glynn-Jones, R., et al., Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2018. 29(Suppl 4): p. iv263.
8. Cervantes, A., et al., Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*, 2023. 34(1): p. 10-32.
9. Weiser, M.R., *AJCC 8th Edition: Colorectal Cancer*. *Ann Surg Oncol*, 2018. 25(6): p. 1454-1455.
10. Balkwill, F.R., M. Capasso, and T. Hagemann, The tumor microenvironment at a glance. *J Cell Sci*, 2012. 125(Pt 23): p. 5591-6.
11. Bremnes, R.M., et al., The role of tumor stroma in cancer progression and prognosis: emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *J Thorac Oncol*, 2011. 6(1): p. 209-17.
12. Pagès, F., et al., International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 2018. 391(10135): p. 2128-2139.
13. Ahmadpour, S., et al., The effects of tumor-derived supernatants (TDS) on cancer cell progression: A review and update on carcinogenesis and immunotherapy. *Cancer Treatment and Research Communications*, 2024. 40: p. 100823.
14. Lorestani, P., et al., The complex role of macrophages in pancreatic cancer tumor microenvironment: a review on cancer progression and potential therapeutic targets. *Discover Oncology*, 2024. 15(1): p. 369.
15. Ravensbergen, C., et al., Whole-Slide ECM Imaging Reveals Dense Fibrous Matrix as a High-Risk Factor for Recurrence in Stage II Colon Cancer. *bioRxiv*, 2024: p. 2024.11. 14.622985.
16. D'Arcangelo, E., et al., The life cycle of cancer-associated fibroblasts within the tumour stroma and its importance in disease outcome. *Br J Cancer*, 2020. 122(7): p. 931-942.
17. Henrich, L.M., et al., The Impact of Cancer-Associated Fibroblasts on the Biology and Progression of Colorectal Carcinomas. *Genes (Basel)*, 2024. 15(2).
18. Kalluri, R., The biology and function of fibroblasts in cancer. *Nat Rev Cancer*, 2016. 16(9): p. 582-98.
19. Koliarakis, V., et al., Mesenchymal Cells in Colon Cancer. *Gastroenterology*, 2017. 152(5): p. 964-979.
20. Sandberg, T.P., et al., Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumor-stroma ratio in colorectal cancer. *BMC Cancer*, 2019. 19(1): p. 284.
21. Tsujino, T., et al., Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clin Cancer Res*, 2007. 13(7): p. 2082-90.
22. Wikberg, M.L., et al., High intratumoral expression of fibroblast activation protein (FAP) in colon cancer is associated with poorer patient prognosis. *Tumour Biol*, 2013. 34(2): p. 1013-20.
23. Mesker, W.E., et al., The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Cell Oncol*, 2007. 29(5): p. 387-98.
24. Nissen, N.I., M. Karsdal, and N. Willumsen, Collagens and Cancer associated fibroblasts in the reactive stroma and its relation to Cancer biology. *J Exp Clin Cancer Res*, 2019. 38(1): p. 115.
25. Coto-Llerena, M., et al., High Expression of FAP in Colorectal Cancer Is Associated With Angiogenesis and Immunoregulation Processes. *Front Oncol*, 2020. 10: p. 979.
26. Huijbers, A., et al., The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*, 2013. 24(1): p. 179-85.
27. Zunder, S.M., et al., Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage II and stage III colon cancer. *Br J Cancer*, 2018. 119(2): p. 164-169.
28. Polack, M., et al., Results from the UNITED study: a multicenter study validating the prognostic effect of the tumor-stroma ratio in colon cancer. *ESMO Open*, 2024. 9(4): p. 102988.
29. van Pelt, G.W., et al., The value of tumor-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol*, 2020. 20: p. 39-44.
30. Strous, M.T.A., et al., A high tumour-stroma ratio (TSR) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clin Transl Oncol*, 2022. 24(6): p. 1047-1058.
31. Park, J.H., et al., Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. *J Pathol Clin Res*, 2020. 6(1): p. 30-39.
32. Dendl, K., et al., The Role of Fibroblast Activation Protein Ligands in Oncologic PET Imaging. *PET Clin*, 2021. 16(3): p. 341-351.
33. Koustoulidou, S., et al., Cancer-Associated Fibroblasts as Players in Cancer Development and Progression and Their Role in Targeted Radionuclide Imaging and Therapy. *Cancers (Basel)*, 2021. 13(5).
34. Giesel, F.L., et al., (68)Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with Various Cancers. *J Nucl Med*, 2019. 60(3): p. 386-392.

35. Kratochwil, C., et al., (68)Ga-FAPI PET/CT: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med*, 2019. 60(6): p. 801-805.
36. Koerber, S.A., et al., The Role of (68)Ga-FAPI PET/CT for Patients with Malignancies of the Lower Gastrointestinal Tract: First Clinical Experience. *J Nucl Med*, 2020. 61(9): p. 1331-1336.
37. Mona, C.E., et al., Correlation of (68)Ga-FAPI-46 PET Biodistribution with FAP Expression by Immunohistochemistry in Patients with Solid Cancers: Interim Analysis of a Prospective Translational Exploratory Study. *J Nucl Med*, 2022. 63(7): p. 1021-1026.
38. Strating, E., et al., Fibroblast activation protein identifies Consensus Molecular Subtype 4 in colorectal cancer and allows its detection by (68)Ga-FAPI-PET imaging. *Br J Cancer*, 2022. 127(1): p. 145-155.
39. Dadgar, H., et al., Exploring the efficacy of FAPI PET/CT in the diagnosis and treatment management of colorectal cancer: a comprehensive literature review and initial experience. *Clinical and Translational Imaging*, 2024.
40. Zboralski, D., et al., Preclinical evaluation of FAP-2286 for fibroblast activation protein targeted radionuclide imaging and therapy. *European Journal of Nuclear Medicine and Molecular Imaging*, 2022. 49(11): p. 3651-3667.
41. van Pelt, G.W., et al., Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*, 2018. 473(4): p. 405-412.
42. Nagtegaal, I., et al., Morphological changes in tumour type after radiotherapy are accompanied by changes in gene expression profile but not in clinical behaviour. *J Pathol*, 2004. 204(2): p. 183-92.
43. Strous, M.T.A., et al., Tumour-stroma ratio to predict pathological response to neo-adjuvant treatment in rectal cancer. *Surg Oncol*, 2022. 45: p. 101862.
44. van Pelt, G.W., et al., Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *J Med Surg Pathol*, 2016. 1(02).
45. Prall, F., et al., Tumour regression and mesorectal lymph node changes after intensified neoadjuvant chemoradiation for carcinoma of the rectum. *Apmis*, 2006. 114(3): p. 201-10.
46. Basalova, N., et al., Fibroblast Activation Protein Alpha (FAP α) in Fibrosis: Beyond a Perspective Marker for Activated Stromal Cells? *Biomolecules*, 2023. 13(12).
47. van Pelt, G.W., et al., The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 2018. 73(2): p. 197-206.
48. Janani, M., et al., Association of future cancer metastases with fibroblast activation protein- α : a systematic review and meta-analysis. *Front Oncol*, 2024. 14: p. 1339050.
49. Tudyka, V., et al., EURECCA consensus conference highlights about colon & rectal cancer multidisciplinary management: the radiology experts review. *Eur J Surg Oncol*, 2014. 40(4): p. 469-75.
50. Brouwer, N.P.M., et al., Clinical lymph node staging in colorectal cancer; a flip of the coin? *Eur J Surg Oncol*, 2018. 44(8): p. 1241-1246.
51. Beets-Tan, R.G.H., et al., Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *European Radiology*, 2018. 28(4): p. 1465-1475.
52. Kömek, H., et al., Comparison of [68 Ga]Ga-DOTA-FAPI-04 PET/CT and [18F]FDG PET/CT in colorectal cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, 2022. 49(11): p. 3898-3909.
53. Yang, L., et al., [18F] AIF-NOTA-FAPI-04 PET/CT as a promising tool for imaging fibroblast activation protein in gastrointestinal system cancers: a prospective investigation of comparative analysis with 18F-FDG. *European Journal of Nuclear Medicine and Molecular Imaging*, 2023. 50(13): p. 4051-4063.
54. Polack, M., et al., Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT. *Clin Transl Oncol*, 2022.
55. Zheng, Z., et al., Intratumor heterogeneity: A new perspective on colorectal cancer research. *Cancer Medicine*, 2020. 9(20): p. 7637-7645.
56. Mori, Y., et al., FAPI PET: Fibroblast Activation Protein Inhibitor Use in Oncologic and Nononcologic Disease. *Radiology*, 2023. 306(2): p. e220749.
57. Temmink, S.J.D., et al., Watch and wait after neoadjuvant treatment in rectal cancer: comparison of outcomes in patients with and without a complete response at first reassessment in the International Watch & Wait Database (IWWd). *Br J Surg*, 2023. 110(6): p. 676-684.
58. Bahadoer, R.R., et al., Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *The Lancet Oncology*, 2021. 22(1): p. 29-42.
59. Chen, L., et al., [68Ga]Ga-FAPI-04 PET/CT on assessing Crohn's disease intestinal lesions. *European Journal of Nuclear Medicine and Molecular Imaging*, 2023. 50(5): p. 1360-1370.
60. Hotta, M., et al., Non-oncologic incidental uptake on FAPI PET/CT imaging. *Br J Radiol*, 2023. 96(1142): p. 20220463.
61. Langbein, T., W.A. Weber, and M. Eiber, Future of Theranostics: An Outlook on Precision Oncology in Nuclear Medicine. *J Nucl Med*, 2019. 60(Suppl 2): p. 13s-19s.
62. Ohlund, D., E. Elyada, and D. Tuveson, Fibroblast heterogeneity in the cancer wound. *J Exp Med*, 2014. 211(8): p. 1503-23.
63. Chen, Y., K.M. McAndrews, and R. Kalluri, Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*, 2021. 18(12): p. 792-804.
64. Yang, D., et al., Cancer-associated fibroblasts: from basic science to anticancer therapy. *Experimental & Molecular Medicine*, 2023. 55(7): p. 1322-1332.
65. Kou, Z., et al., Heterogeneity of primary and metastatic CAFs: From differential treatment outcomes to treatment opportunities (Review). *Int J Oncol*, 2024. 64(5).

66. Pelon, F., et al., Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat Commun*, 2020. 11(1): p. 404.
67. Abudukelimu, S., N. de Miranda, and L. Hawinkels, Fibroblasts in Orchestrating Colorectal Tumorigenesis and Progression. *Cell Mol Gastroenterol Hepatol*, 2024. 17(5): p. 821-826.
68. Meyerholz, D.K. and A.P. Beck, Principles and approaches for reproducible scoring of tissue stains in research. *Lab Invest*, 2018. 98(7): p. 844-855.
69. Meyerholz, D.K. and A.P. Beck, Fundamental Concepts for Semiquantitative Tissue Scoring in Translational Research. *Ilar j*, 2018. 59(1): p. 13-17.
70. Smit, M.A., et al., Deep learning based tumor-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform*, 2023. 14: p. 100191.

Supplementary Material



Supplementary Figure 1. Flowchart of ultimately included patients in separate and combined cohorts.

Supplementary Table 1. Baseline characteristics of the complete CRC cohort, and the CC and RC cohorts shown separately

Characteristic (unit)	Total CRC cohort (N=93)	CC cohort (N=40)	RC cohort (N=53)
Sex			
Male	57 (61)	22 (55)	35 (66)
Female	36 (39)	18 (45)	18 (34)
Age at diagnosis – years			
Median age (range)	67 (39-95)	70 (39-91)	67 (45-95)
Age ≥75 years old	26 (28)	13 (33)	13 (25)
Diagnosis year			
Median (range)	2008 (2000-2016)	2005 (2000-2011)	2013 (2005-2016)
Median period until surgery (weeks)	6 (0-29)	2 (0-11)	7 (0-29)
Clinical TNM-stage*			
cTNM-stage I	1 (1)	0 (0)	1 (2)
cTNM-stage II	29 (31)	13 (33)	16 (30)
cTNM-stage II/III**	20 (22)	19 (48)	1 (2)
cTNM-stage III	43 (46)	8 (20)	35 (66)
cTNM-stage IV	0 (0)	0 (0)	0 (0)
Treatment			
Neoadjuvant therapy and surgery	41 (44)	0 (0)	41 (77)
Surgery alone	28 (30)	17 (43)	11 (21)
Neoadjuvant, surgery and adjuvant therapy	0 (0)	0 (0)	0 (0)
Surgery and adjuvant therapy	24 (26)	23 (57)	1 (2)
Operation setting			
Elective	81 (87)	28 (70)	53 (0)
Emergency***	12 (13)	12 (30)	0 (0)
Location tumour****			
Right-sided colon	24 (26)	24 (60)	N/A
Left-sided colon	15 (16)	15 (38)	N/A
Rectosigmoid	2 (2)	1 (2)	1 (2)
High rectum (>10cm anal verge)	15 (16)	N/A	15 (28)
Mid rectum (5-10cm anal verge)	16 (17)	N/A	16 (30)
Low rectum (<5cm anal verge)	21 (23)	N/A	21 (40)
Operation type			
Subtotal colectomy	3 (3)	3 (8)	0 (0)
Hemicolectomy right (also extended)	21 (21)	21 (53)	0 (0)
Transversectomy	2 (2)	2 (5)	0 (0)
Hemicolectomy left	5 (5)	5 (13)	0 (0)
Sigmoidectomy, Hartmann	9 (10)	8 (20)	1 (2)
Low anterior resection	33 (36)	1 (3)	32 (60)
Abdominal perineal resection	20 (22)	0 (0)	20 (38)
Disease-free survival – years			
Median survival (range)	5.1 (0.1-21.5)	4.6 (0.2-21.5)	5.3 (0.1-17.8)

<i>(continued)</i> Characteristic (unit)	Total CRC cohort (N=93)	CC cohort (N=40)	RC cohort (N=53)
No event occurred	43 (46)	14 (35)	29 (55)
Events, of which type	50 (54)	26 (65)	24 (45)
Locoregional recurrence	3 (6)	1 (4)	2 (8)
Distant metastasis	23 (46)	12 (46)	11 (46)
Death	24 (48)	13 (50)	11 (46)
Overall survival – years			
Median survival (range)	7.4 (0.4-23.1)	7.3 (0.5-23.1)	7.4 (0.4-18.0)
Overall survival status			
Alive	46 (50)	15 (38)	31 (59)
Dead, due to the cause of	47 (51)	25 (63)	22 (42)
Current colorectal carcinoma	22 (47)	11 (46)	11 (48)
Other, e.g. comorbidity	18 (38)	9 (38)	9 (39)
Unknown	7 (15)	4 (17)	3 (13)

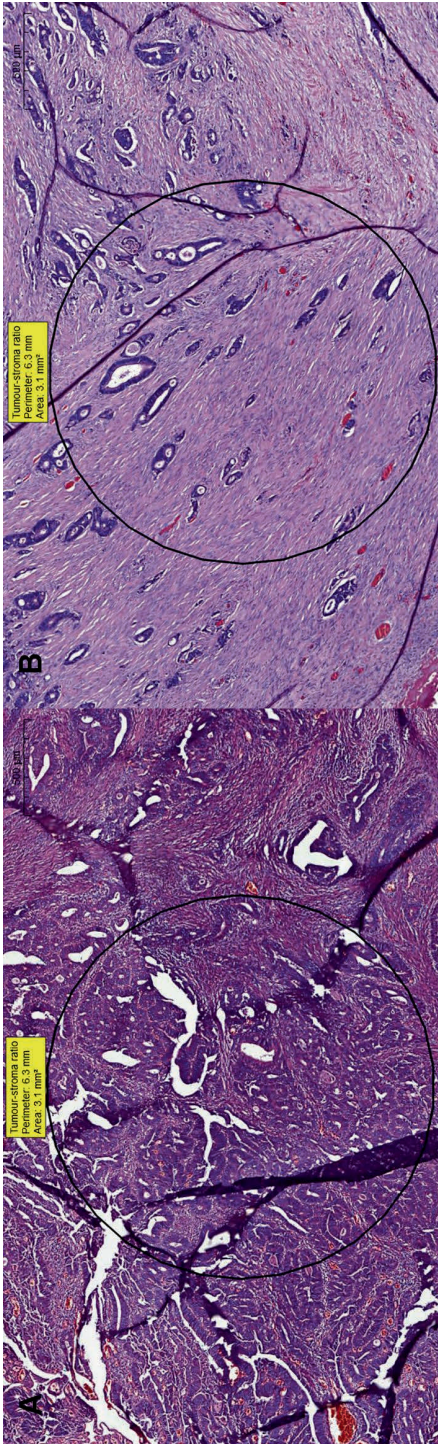
All variables are given as absolute numbers with associated percentages or medians with ranges (minimum-maximum). Sum of percentages can be less or more than 100 due to rounding. CC, colon carcinoma; N/A, not applicable; RC, rectal carcinoma; TNM, tumour-node metastasis.

*Different versions of the TNM classification were used, here all variables are converted to the AJCC/UICC TNM version 5 (1997).

**Stage II/III is given in case no imaging is performed and locoregional extent of disease is uncertain.

***Emergency setting includes the presenting of patients with obstructive ileus and e.g. during operation a tumour is found.

****Location is determined through imaging primarily. A right-sided tumour is defined as a colon carcinoma in the caecum, colon ascendens, flexura hepatica or colon transversum until the flexura lienalis.



Supplementary Figure 2. Zoomed in examples of the tumour-stroma ratio annotations on haematoxylin and eosin-stained slides in colorectal cancer patients. (A) Stroma-low ($\leq 50\%$); (B) Stroma-high ($> 50\%$). General overview 5.0x magnification, annotation 10.0x magnification.

Supplementary Table 2. Comparison of biopsy-scored stroma-low and stroma-high tumours within the RC cohort

Characteristic (unit)	RC cohort (N=53)		
	Stroma-low (biopsy; N=30)	Stroma-high (biopsy; N=23)	P-value
FAP expression – biopsy			
No expression	4 (13)	2 (9)	0.814#
Any expression (heterogeneous - high expression)			
Heterogeneous expression	15 (50)	11 (48)	
Little expression (no - heterogeneous expression)	19 (63)	13 (57)	
Moderate to high expression	11 (37)	10 (44)	
Treatment type – all treatments			
TSR – PT			0.053#
Stroma-low	13 (43)	3 (13)	
Stroma-high	16 (53)	18 (78)	
N/A (no tumour)	1 (3)	2 (9)	
FAP expression – PT			0.712#
Little expression (no - heterogeneous expression)	21 (70)	15 (65)	
Moderate to high expression	9 (30)	8 (35)	
Representative LN*			
N/A (missing, no LN positive)	18 (60)	14 (61)	0.830#
Positive LN, of which	12 (40)	9 (39)	
TSR			
Stroma-low	11 (92)	8 (89)	
Stroma-high	1 (8)	1 (11)	
FAP expression – LN*			0.055#
Little expression (no - heterogeneous expression)	1 (8)	4 (44)	
Moderate to high expression	11 (92)	5 (56)	
Treatment type – no neoadjuvant treatment			
No neoadjuvant treatment received	10 (33)	2 (9)	0.007#
TSR – PT**			
Stroma-low	9 (90)	0 (0)	
Stroma-high	1 (10)	2 (100)	
FAP expression – PT			0.584#
Little expression (no - heterogeneous expression)	7 (70)	1 (50)	
Moderate to high expression	3 (30)	1 (50)	
Representative LN*			
N/A (missing, no LN positive)	9 (90)	1 (50)	N/A
Positive LN, of which	1 (10)	1 (50)	
TSR			
Stroma-low	0 (0)	0 (0)	
Stroma-high	1 (100)	1 (100)	

(continued)			N/A
FAP expression – LN*			
Little expression (no - heterogeneous expression)	1 (100)	1 (100)	
Moderate to high expression	0 (0)	0 (0)	
Treatment type – neoadjuvant treatment***			
Neoadjuvant treatment received	20 (67)	21 (91)	
TSR – PT			0.785#
Stroma-low	4 (20)	3 (14)	
Stroma-high	15 (75)	16 (76)	
N/A (no tumour)	1 (5)	2 (10)	
FAP expression – PT			0.819#
Little expression (no - heterogeneous expression)	14 (70)	14 (67)	
Moderate to high expression	6 (30)	7 (33)	
Representative LN*			
N/A (missing, no positive LN)	9 (45)	13 (62)	
LN positive, of which	11 (55)	8 (38)	
TSR			N/A
Stroma-low	11 (100)	8 (100)	
Stroma-high	0 (0)	0 (0)	
FAP expression – representative LN*			0.027#
Little expression (no - heterogeneous expression)	0 (0)	3 (38)	
Moderate to high expression	11 (100)	5 (63)	

All variables are given as absolute numbers with associated percentages. Sum of percentages can be less or more than 100 due to rounding. FAP, fibroblast activation protein; LN, lymph node; N/A, not applicable; PT, primary tumour; RC, rectal carcinoma; TSR, tumour-stroma ratio.

*Representable lymph node is the lymph node with most tumour stroma.

**The TSR scored in biopsies is representable for the primary tumour TSR, as seen in literature and here proven in not neoadjuvantly treated tumours.

***Neoadjuvant treatment includes either short course radiotherapy (SCRT, 5x5 Gray) (N=30) or chemoradiation (CRT, 25x2/28x1.2 Gray + capecitabine monotherapy) (N=11).

#Calculated with the Chi-square test.

Supplementary Table 3. Comparison of the tumours with stroma-low and stroma-high lymph node metastases within the CRC and separate CC and RC cohorts.

Characteristic (unit)	CRC cohort (N=61)				CC cohort (N=40)				RC LN cohort (N=21)			
	All stage III patients (N=61)	Stroma-low LN patients* (LN, N=41)	Stroma- high LN patients* (N=20)	P- value	All stage III patients (N=40)	Stroma-low LN patients* (N=22)	Stroma- high LN patients* (N=18)	P- value	All stage III patients (N=21)	Stroma-low LN patients* (N=19)	Stroma- high LN patients* (N=2)	P- value
All LN analysed												
Total LN and tumour deposits scored, of which <i>Mean total examined LN</i> (SD)	288 (100) 5 (4)	169 (100) 4 (3)	119 (100) 6 (5)	0.151\$	189 (66) 5 (4)	76 (40) 4 (3)	113 (60) 6 (5)	0.049\$	99 (34) 2 (3)	93 (94) 5 (4)	6 (6) 3 (0)	0.048\$
Total LN with metastasis (tumour-positive)	125 (43)	78 (46)	47 (39)		77 (41)	34 (45)	44 (39)		47 (47)	44 (47)	3 (50)	
<i>Mean LN tumour-positive</i> (SD)	2 (2)	2 (1)	2 (2)	0.397\$	2 (2)	2 (1)	2 (2)	0.122\$	2 (2)	2 (2)	2 (1)	0.304\$
Total LN without metastasis (tumour-negative)	163 (57)	91 (54)	72 (61)		112 (59)	42 (55)	69 (61)		52 (53)	49 (53)	3 (50)	
<i>Mean LN tumour-negative</i> (SD)	3 (3)	2 (3)	4 (3)	0.109\$	3 (3)	2 (3)	4 (3)	0.051\$	3 (3)	3 (3)	2 (1)	0.243\$
FAP expression in LN												
All LNs analysed, of which												
Total LN with any FAP expression (FAP-positive)	288 (100) 112 (39)	169 (100) 75 (46)	119 (100) 37 (31)		189 (100) 53 (28)	76 (100) 19 (25)	113 (100) 34 (30)		99 (100) 59 (60)	93 (100) 56 (60)	6 (100) 3 (50)	
<i>Mean all LNs FAP-positive</i> (SD)	2 (2) 87 (78)	2 (2) 56 (75)	2 (2) 31 (84)	0.968\$	1 (2) 45 (85)	1 (1) 17 (89)	2 (2) 28 (82)	0.050\$	3 (2) 42 (71)	3 (2) 39 (70)	2 (1) 3 (50)	0.114\$
LN tumour-positive and FAP-positive	1 (1)	1 (1)	2 (2)	0.649\$	1 (1)	1 (1)	2 (2)	0.074\$	2 (1)	2 (2)	2 (1)	0.453\$
Mean LN tumour-positive and FAP-positive (SD)	25 (22)	19 (25)	6 (16)		8 (15)	2 (11)	6 (18)		17 (29)	17 (30)	0 (0)	
LN tumour-negative, but FAP-positive (e.g. FAP expression in healthy lymphoid tissue)	0 (1)	0 (1)	0 (1)	0.442\$	0 (0)	0 (0)	0 (0)	0.127\$	1 (1)	1 (2)	0 (0)	0.015\$
Mean LN tumour-negative, but FAP-positive (SD)												

<i>(continued)</i>										
Total LN without FAP expression (FAP-negative)	176 (61)	94 (56)	82 (69)	136 (72)	57 (75)	99 (88)	40 (40)	37 (40)	3 (50)	
Mean all LNs FAP-negative (SD)	3 (3)	2 (2)	4 (4)	3 (3)	3 (3)	4 (4)	2 (2)	2 (2)	2 (1)	0.571\$
LN tumour-negative and FAP-negative	138 (78)	72 (77)	66 (80)	103 (76)	40 (70)	63 (64)	35 (88)	32 (86)	3 (50)	
Mean LN tumour-negative and FAP-negative (SD)	2 (3)	2 (2)	3 (3)	3 (3)	2 (2)	4 (3)	2 (2)	2 (2)	2 (1)	0.811\$
LN tumour-positive, but FAP-negative	38 (22)	22 (23)	16 (20)	33 (24)	17 (30)	16 (16)	5 (13)	5 (14)	0 (0)	
Mean LN tumour-positive, but FAP-negative (SD)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	0 (1)	0 (1)	0 (0)	0.056\$
Intra-individual FAP expression										
Tumour-positive LN analysed, patients**	61 (100)	41 (67)	20 (33)	40 (66)	22 (55)	18 (45)	21 (34)	19 (90)	2 (10)	0.113\$
Mean percentage LNs tumour-positive and FAP-positive (SD)	69 (42)	70 (43)	67 (42)	57 (46)	52 (48)	63 (43)	91 (23)	91 (24)	100 (0)	
Mean percentage LNs tumour-positive and FAP-negative (SD)	31 (42)	30 (43)	33 (42)	43 (46)	48 (48)	37 (43)	8 (23)	9 (24)	0 (0)	
Tumour-negative LN analysed, patients**	43 (72)	26 (60)	17 (40)	27 (63)	12 (44)	15 (66)	16 (37)	14 (87)	2 (13)	0.013\$
Mean percentage LNs tumour-negative and FAP-positive (SD)	14 (29)	19 (35)	7 (13)	6 (12)	4 (10)	7 (14)	29 (41)	33 (42)	0 (0)	
Mean percentage LNs tumour-negative and FAP-negative (SD)	87 (28)	81 (35)	93 (13)	94 (12)	96 (10)	93 (14)	71 (41)	67 (42)	100 (0)	

All variables are given as absolute numbers with associated percentages or means with SD. Sum of percentages can be less or more than 100 due to rounding. CC, colon carcinoma; FAP, fibroblast activation protein; LN, lymph node; N/A, not applicable; PT, primary tumour; RC, rectal carcinoma; SD, standard deviation; TSR, tumour-stroma ratio.

*Representable lymph node is the lymph node with most tumour stroma.

**Not all patients had negative lymph nodes which were examinable or available.

#Calculated with the Chi-square test.

\$Calculated with Independent Samples T-test, equal variances not assumed.



Chapter 6

Self-Supervised Learning Reveals Clinically Relevant Histomorphological Patterns for Therapeutic Strategies in Colon Cancer

Meaghan Polack+, Bojing Liu+, Nicolas Coudray#, Adalberto
Claudio Quiros#, Theodore Sakellaropoulos, Hortense Le, Afreen
Karimkhan, Augustinus S.L.P. Crobach, J. Han J.M. van Krieken,
Ke Yuan, Rob A.E.M. Tollenaar, Wilma E. Mesker, and Aristotelis
Tsirigos

+, # author with equal contribution

•

Abstract

Self-supervised learning (SSL) automates the extraction and interpretation of histopathology features on unannotated hematoxylin-eosin-stained whole slide images (WSIs). We train an SSL Barlow Twins encoder on 435 colon adenocarcinoma WSIs from The Cancer Genome Atlas to extract features from small image patches (tiles). Leiden community detection groups tiles into histomorphological phenotype clusters (HPCs). HPC reproducibility and predictive ability for overall survival is confirmed in an independent clinical trial (N=1213 WSIs). This unbiased atlas results in 47 HPCs displaying unique and shared clinically significant histomorphological traits, highlighting tissue type, quantity, and architecture, especially in the context of tumor stroma. Through in-depth analysis of these HPCs, including immune landscape and gene set enrichment analysis, and associations to clinical outcomes, we shine light on the factors influencing survival and responses to treatments of standard adjuvant chemotherapy and experimental therapies. Further exploration of HPCs may unveil additional insights and aid decision-making and personalized treatments for colon cancer patients.

Introduction

Traditionally, the diagnosis of colon cancer is confirmed by microscopic assessment of resection specimens on hematoxylin and eosin-stained (H&E) slides by pathologists. For each patient, a personalized treatment strategy is tailored through a multidisciplinary meeting, following guidelines consisting of risk assessments based on clinicopathological characteristics, including the tumor-node-metastasis (TNM) classification and additional biomarkers [1–4]. However, due to an aging population and increasing amount of biomarker research, diagnosing and predicting the prognosis of colon cancer patients can be time consuming, or complicated and resource-demanding, especially when incorporating screening for mutational variants [2,3,5,6].

In modern digital pathology, scanning H&E slides into high-resolution whole slide images (WSIs) has enabled the applications of deep learning (DL) [7]. Deep convolutional neural networks, in particular, have benefited the diagnostic process, initially by minimizing inter-rater disagreement and workload [7–9]. In colorectal cancer, supervised DL models also showed the ability to predict molecular pathways (i.e. mutation density, microsatellite instability [MSI], chromosomal instability) and key mutations like BRAF and KRAS [10,11]. DL even has intriguing potential in predicting complicated prognostic outcomes such as patient survival [12–14]. Moreover, integrating multi-omic data with the associated H&E slides, i.e. multimodal data integration, led to improvements in prognostic prediction on overall 49 survival (OS) for most cancer types [15,16].

Previous DL studies primarily focused on training models to extract features from WSIs under supervision of potentially extensive and time-consuming human-derived annotations on slide or pixel level, i.e. supervised learning [10,12,15,17]. Self-supervised learning (SSL) on the other hand, has gained significantly increasing attention for its capacity to automatically capture image features from unlabeled data [11]. Applications of SSL models, including general-purpose foundation models [18,19], have demonstrated superior performance in various downstream cancer classification, survival, and molecular phenotype prediction tasks compared to traditional supervised learning models [20–24]. Barlow Twins, an SSL model designed to learn non-redundant image features, has several advantages over other SSL learning models (e.g. contrastive learning models), including not requiring extensive batch sizes nor asymmetry between the network "twins" [25].

Despite the efficacy in decision-making, DL models are often labeled as "black boxes", posing significant challenges in terms of interpretability. Supervised attention-based multiple instance learning

is a common interpretive method, enabling DL models to concentrate on informative segments within WSI according to predefined training labels [26]. Another approach involves employing unsupervised clustering algorithms to organize extracted features into clinically relevant and interpretable clusters which can be subsequently linked to diverse patient-related outcomes [23,27,28]. These clustering methods offer significant advantages, including the prediction of various clinical outcomes, intuitive visualization for pathologists, and interpretation and correlation with a range of molecular data.

In this work, our objective is twofold: first, to automatically and reliably extract clinically relevant histologic patterns from WSIs, which can be interpreted by pathologists, and second, to investigate the connections between these patterns and patient outcomes as well as molecular phenotypes across different treatment groups within a large clinical trial for colon cancer. To achieve this, we apply an SSL pipeline [23] involving the Barlow Twins encoder for feature extraction, followed by a community detection algorithm to construct an unbiased atlas of histomorphologic phenotype clusters (HPCs). This algorithm is exclusively trained on public data from the colon adenocarcinoma cohort within The Cancer Genome Atlas (TCGA) multi-institutional database [29]. Remarkably, the identified HPCs generalizes well in unseen WSIs obtained from the clinical Bevacizumab-Avastin® adjuvant (AVANT) trial [30]. Subsequently, HPCs are linked to patient OS. An HPC-based classifier trained using TCGA data for OS, demonstrates prognostic significance in the external validation of the AVANT study, even when considering key clinical and demographic factors typically employed in clinical settings. Notably, by conducting comprehensive analyses of the distinct histomorphologic features of each HPC and their associations with immune and genetic profiles, we provide insight into morphological and molecular determinants of patient survival upon different treatments (e.g. standard-of-care adjuvant chemotherapy and experimental targeted therapies). Exploring these features further could yield additional insights into other histopathology diagnostics, supporting shared decision-making and advancing personalized treatment options for colon cancer patients in the future.

Results

Self-supervised learning of WSI features using the multi-institutional TCGA dataset

We trained the self-supervised algorithm using data exclusively from the TCGA colon adenocarcinoma (TCGA-COAD) set, eliminating the need for annotations by pathologists (Figure 1). A total of 435 WSIs (428 patients) obtained from the TCGA-COAD dataset (see Methods: Study population for details) was first divided into smaller image patches (224-by-224 pixels), also known as image tiles, at a magnification level of 10x (Figure 1a) (see Methods: Data pre-processing for details). To identify features on these patches, we trained an SSL Barlow Twins feature extractor using a random subset of tiles (Figure 1a) (see Methods: Extracting image features using Barlow Twins for details). The Barlow Twins was trained with the objective function to evaluate the cross-correlation matrix between the embeddings (feature vector z) of two identical backbone networks, which were fed distorted variants of a batch of image tiles. The objective function was optimized by minimizing the deviation of the crosscorrelation matrix from the identity matrix. This led to increased similarity among the embedding of z vectors of the distorted sample versions, while reducing redundancy among the individual components of these vectors. As a result, each tile was described as a vector of 128 extracted features that can subsequently be used to group tiles into clusters by similarity.

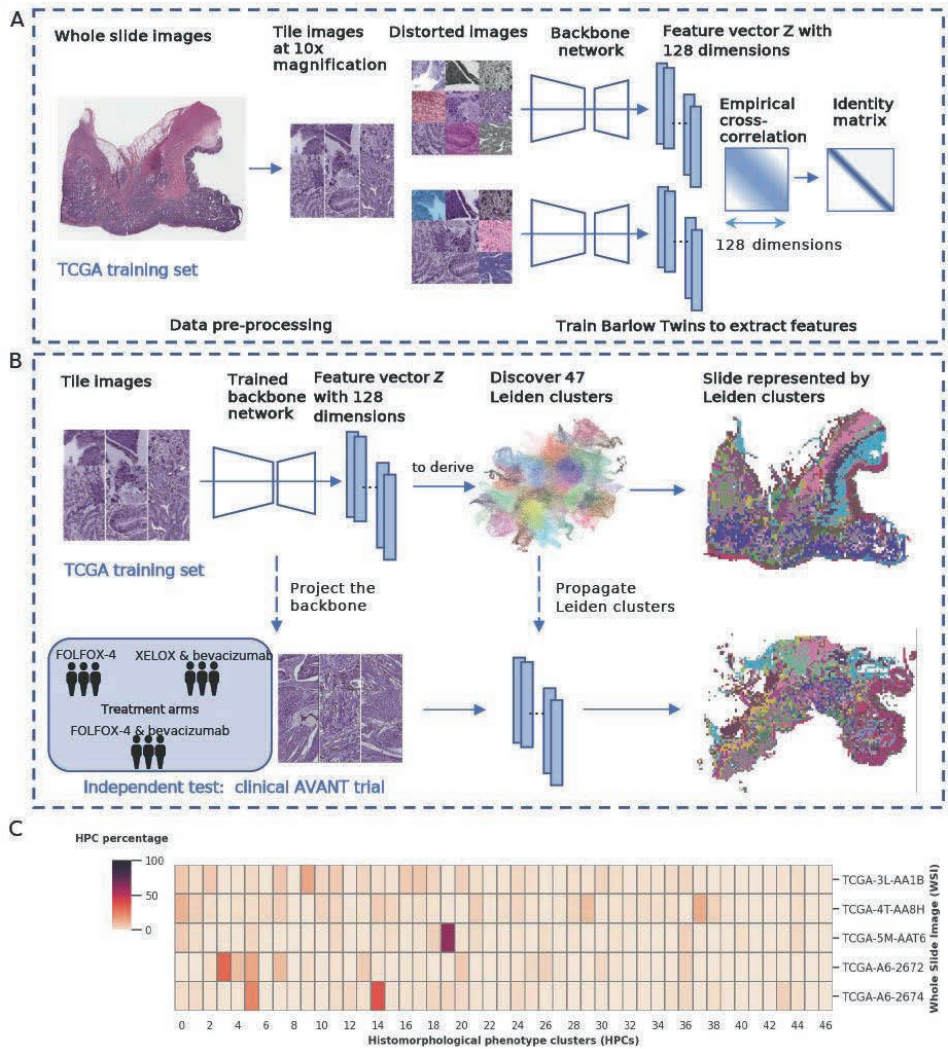


Figure 1. Overview of the Model Architecture: Training Barlow Twins and deriving Histomorphological Phenotype Clusters. (A) Training Barlow Twins with TCGA. WSIs from TCGA were processed to extract image tiles and normalize stain colours. The Barlow Twins network was employed to learn 128-dimensional z vectors from these image tiles. (B) Deriving HPCs. The tiles from TCGA were projected into z vector representations obtained from the trained Barlow Twins network. HPCs were defined by applying Leiden community detection to the nearest neighbour graph of z tile vector representations. Each WSI was represented by a compositional vector of the derived HPCs, indicating the percentage of each HPC with respect to the total tissue area. The Barlow Twins model and HPCs were then projected and integrated into the external AVANT trial. (C) Whole Slide Image Representation. The compositional HPC data represented the WSIs in the study. AVANT, Bevacizumab-Avastin® adjuvant trial. HPC, histomorphological phenotype cluster. TCGA, The Cancer Genome Atlas. WSI, whole slide image.

Construction of an unbiased atlas of histologic patterns through community detection

We applied the Leiden community detection algorithm to derive HPCs, i.e. the clusters with similar histologic patterns (Figure 1b) (see Methods: Identification of HPCs for details). The process began by first projecting the trained Barlow Twins onto the entire TCGA-COAD dataset, extracting 128 dimensional feature representations for each image tile. Subsequently, we utilized Leiden community detection on a nearest neighbor graph constructed from these tile vector representations (Figure 1b). Tiles with similar vector representations were clustered into a group and assigned a specific HPC ID number. The optimization of the Leiden configuration was achieved through an unsupervised process (see Methods: Identification of HPCs for details, Supplementary Figure 1a), resulting in the identification of a total of 47 HPCs, visually represented in a dimensionality reduction plot (Uniform Manifold Approximation and Projection; UMAP plot) (Figure 2a).

As external dataset, we analyzed a total of 1213 colon cancer patients with diagnostic pathology H&E WSIs (one WSI per patient), a subset of the clinical AVANT trial [30,31] (see Methods: Study population for details). We harnessed the optimized SSL Barlow Twins model to generate embeddings of the unseen AVANT WSI tiles. The assignment of identified HPCs to the unseen AVANT was achieved using the K-nearest neighbors approach. The HPC label of each tile in the AVANT data was determined based on majority votes from its K-nearest neighbors in the TCGA training set (Figure 1b). As a result, we obtained comprehensive visual representations of the WSIs where the WSI tiles are colored by their corresponding HPC (Figure 1b). Additionally, we were able to capture the characteristics and heterogeneity of WSIs using the compositional data derived from the HPCs, i.e. the percentage of the area on a WSI covered by each HPC, thus facilitating downstream analyses and modeling (Figure 1c).

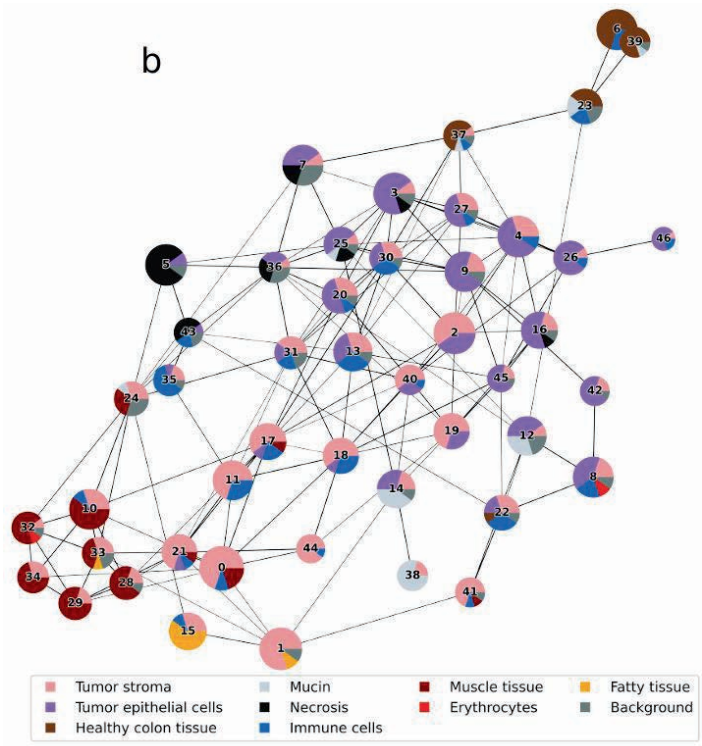
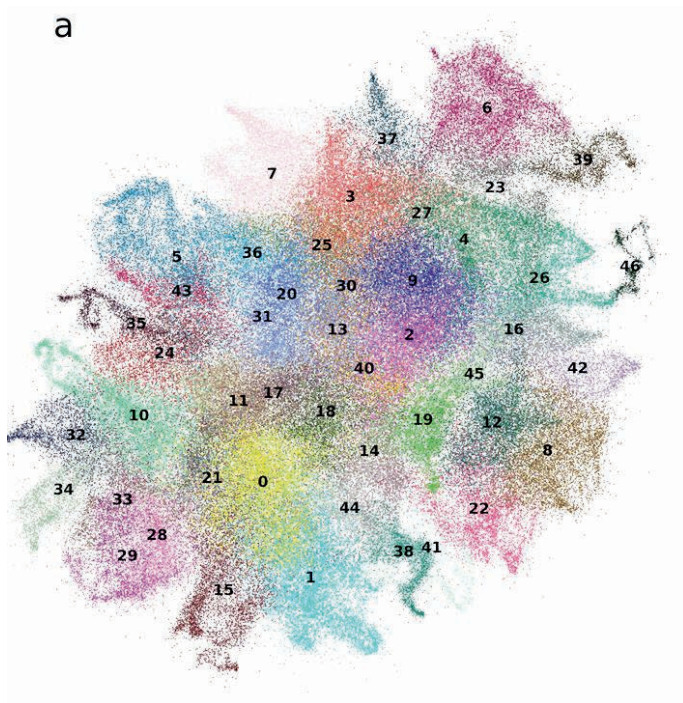
Histopathological assessment and characterization of HPCs

Each HPC underwent histopathological analysis on a randomly selected set of 32 tiles per cluster within TCGA, independently evaluated by two pathologists (ASLPC and JHJMvK) and a researcher (MP)(see Methods: Interpretation of HPCs for details). Tissue types, as observed on the tiles, were described with specific attention to tumor epithelium, tumor stroma and immune cells. Other unique histopathological features or patterns, such as tumor differentiation grade and stromal organization, were noted as well. All present tissue types were scored in percentages (Supplementary Table 1) and depicted using pie charts (Figure 2b). We plotted the interconnections of 47 HPCs using partition-based graph abstraction (PAGA) [32], with the pie charts reflecting their tissue compositions (Figure 2b). Interestingly, distinct larger groups of clusters, or "super-clusters", could be observed based on the similarity of tissue composition, interconnectedness, and topology of HPCs in the PAGA plot (Figure 2b).

In total, we identified eight super-clusters: (1) healthy and dysplastic colon tissue, (2) necrosis, (3) mucinous areas, (4) immune cells, (5) muscle tissue (longitudinal/axial), (6) fatty tissue, (7) tumor stroma, and, (8) tumor epithelium, in no particular order, formed by groups of HPCs shown in Figure 2c. Common histopathological characteristics were noted among HPCs within each designated super-cluster, while HPCs encompassing varied tissue types across multiple super-clusters were often situated at their intersections. For instance, HPC 12, containing not only tumor epithelium but also mucinous tumor, was located between HPCs belonging to the two super-clusters of mucinous tumor and tumor epithelium in the PAGA plot. Moreover, HPC 23, marked by dysplastic colon tissue, formed a bridge between the healthy colon tissue HPC 39 and the tumor epithelium-containing super-cluster HPCs (e.g. HPCs 4, 26, 46), suggesting a potential chronological pathogenesis. In summary, the derived HPCs displayed distinctive histopathologic characteristics. Moreover, HPCs located in close proximity on the UMAP and PAGA plot, based on the extracted features, demonstrated common traits, hinting at potential pattern relationships, mixed phenotypes, or pathogenic trajectories.

Assessment of HPC consistency within and across TCGA and AVANT cohorts

Although SSL methods have been applied recently in histopathology, there is usually no systematic analysis of consistency of the histologic patterns discovered by these methods within and across datasets [21,27]. Here, we address this potential pitfall by incorporating several qualitative and quantitative assessments.



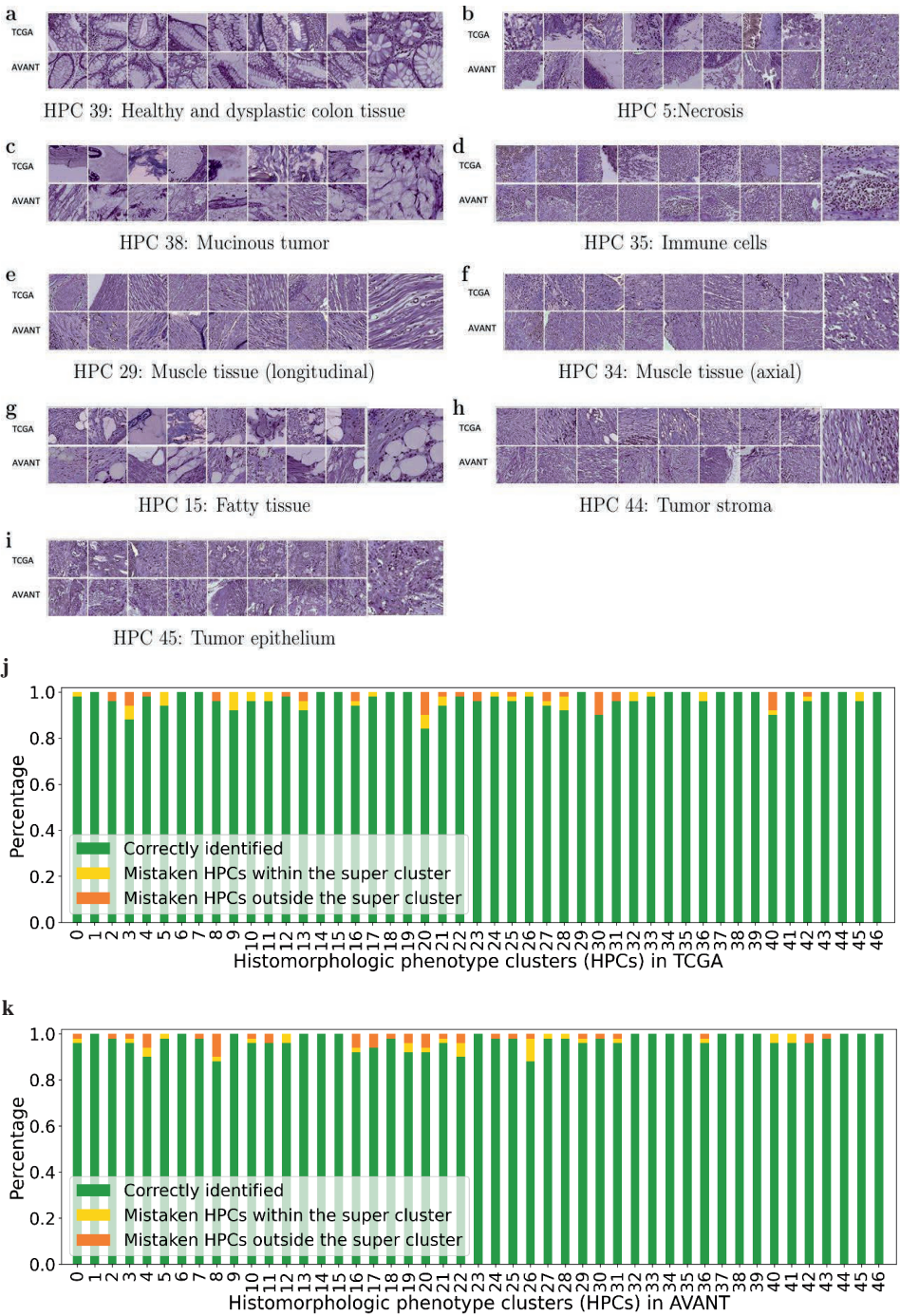
C

Healthy and dysplastic colon tissue	<div>39 Healthy colon tissue</div> <div>6 Inflamed healthy colon tissue</div> <div>23 Dysplastic colon tissue</div> <div>37 High grade dysplasia-adenoma</div> <div>224mm</div>	<div>5 Necrotic</div> <div>43 Necrotic</div> <div>7 Avital tumor with necrosis</div> <div>36 Avital tumor with necrosis</div> <div>25 Moderately differentiated tumor necrotic</div>
Mucinous	<div>14 Mucinous tumor</div> <div>38 Mucinous tumor stroma</div> <div>12 Mucinous tumor</div>	<div>Immune cells</div> <div>13 Infiltrated stroma</div> <div>18 Infiltrated stroma</div> <div>35 Immune cells</div> <div>30 Mixed, immune-high</div> <div>22 Infiltrated stroma</div> <div>31 Mixed, infiltrated stroma</div>
Muscle tissue (longitudinal/axial fibres)	<div>29 Muscle (longitudinal)</div> <div>10. Muscle (longitudinal)</div> <div>33 Muscle (longitudinal)</div> <div>28 Muscle (longitudinal)</div> <div>34 Muscle (axial)</div> <div>24 Vessel-like stroma/muscle</div> <div>32 Muscle (axial)</div>	<div>Fatty tissue</div> <div>15 Fatty tissue</div>
Tumor stroma	<div>0 Disorganized stroma</div> <div>1 Edematous stroma</div> <div>41 Disorganized and loose stroma</div> <div>19 Edematous disorganized stroma</div> <div>11 Vascularized stroma-immune cells</div> <div>21 Vascularized stroma</div> <div>17 Infiltrated aligned stroma-high</div> <div>44 Disorganized, immune-low stroma (desert-type)</div> <div>2 Aligned stroma - well differentiated tumor</div> <div>40 Disorganized stroma, growing infiltratively</div>	<div>Tumor epithelium</div> <div>4 Moderately differentiated tumor - aligned stroma</div> <div>27 Moderately differentiated tumor - aligned stroma</div> <div>3 Moderately differentiated tumor</div> <div>9 Moderately epithelial-infiltrated stroma</div> <div>8. Mixed tumor epithelial-infiltrated stroma</div> <div>16 Well-differentiated tumor</div> <div>45 Poor-differentiated tumor</div> <div>20 Disorganized and vascularized stroma-low</div> <div>42 Moderately differentiated tumor</div> <div>46 Infiltrated tumor</div> <div>26 Infiltrated poor-undifferentiated</div>

Figure 2. Identification of HPCs in TCGA and subsequent classification into super-clusters. (A) UMAP showing 47 HPCs identified from the TCGA dataset, each scatter representing an image tile. (B) PAGA plot of HPCs. Each node represented an HPC with edges representing connections between HPCs based on their vector representation similarity. The pie chart of each node represented the tissue composition for each HPC. (C) Grouping of HPCs into super-clusters according to histopathology tissue similarities. Representative tiles for each HPC were labelled with ID and a brief description. HPC, histomorphological phenotype cluster. PAGA, partition-based abstraction graph. TCGA, The Cancer Genome Atlas. UMAP, uniform manifold approximation and projection plot.

First, qualitative assessment was conducted to evaluate the within-cluster and between-cluster heterogeneity of 47 HPCs derived in TCGA-COAD. Based on 32 randomly selected tiles from each HPC in the TCGA-COAD, three experts (ASLPC, JHJMvK, and MP) independently assessed each HPC by comparing tissue type quantity and architectures (histopathological assessment procedure stated above). Overall, all raters reached the general consensus that there was a noteworthy level of within-cluster morphological similarity and a significant diversity among the 47 HPCs (Figure 3 [a-i]), although phenotypic similarities varied across HPCs, implying that some HPCs may appear more similar than others. To delve deeper into the within-cluster and between-cluster heterogeneity, we carried out quantitative objective blinded tests within TCGA and AVANT tiles separately. This was to ascertain whether the morphological patterns identified by each HPC could also be recognized by human experts.

In this test, the assessor (MP) was shown three groups of image tiles, each containing five tiles. Two groups were from the same HPC, and the third was from a randomly selected other HPC, also called the "odd HPC". The assessor was required to identify the "odd HPC" (Supplementary Figure 2a, see Methods: Pathologist assessment of HPCs for details). Each of the 47 HPCs underwent 50 tests to determine the success rate. Within TCGA, we found that 17 out of the 47 HPCs achieved 100% identification rate, while the remaining 30 HPCs had a correct identification percentage ranging from 84% to 98%. Similarly, within in AVANT, 17 out of 47 HPCs achieved a perfect accuracy, while the rest had an accuracy ranging from 88% to 98% (Figure 3j). In general, HPCs in close proximity to each other in the PAGA plot or belonging to the same super-cluster were more prone to erroneous assignment.



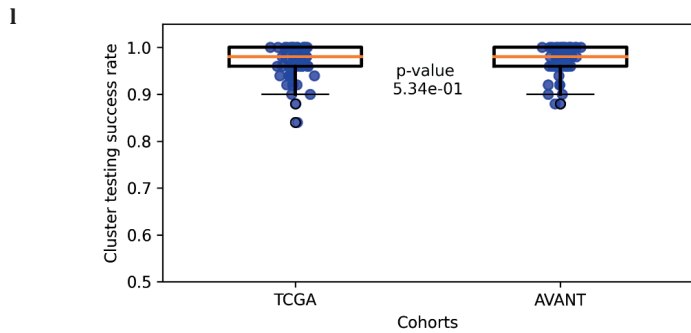


Figure 3. Verification of HPCs in the TCGA training set and the external clinical AVANT trial (a-i) Example tiles from TCGA (upper row) and AVANT (lower row) showcase the eight super-clusters with a zoomed-in representative tile. The muscle tissue super-cluster is further divided into longitudinal and axial subgroups. (j-k) Stacked bar plots illustrate instances of misclassification for each HPC in TCGA training set and AVANT external test set. Green bars represent the percentage of correctly identified odd clusters, yellow bars indicate misclassifications within the tested HPC’s super-cluster, orange bars show misclassifications outside the super-cluster. (l) Box plots display similar distributions of success test rates (corresponding to the green bars in panels j and k) for HPCs in TCGA and AVANT cohorts (two-sided Wilcoxon signed-rank test, $p = 0.534$). Each blue point within each box plot represents the success test rate for a single HPC, calculated based on 50 tests per HPC. The central orange line within each box represents the median, while the bounds of the box indicate the 25th and 75th percentiles (interquartile range). Whiskers extend to the minima and maxima within 1.5 times the interquartile range. Points outside the whiskers represent outliers. HPC, histomorphological phenotype cluster. TCGA, The Cancer Genome Atlas. AVANT, Bevacizumab-Avastin® adjuvant trial.

To evaluate the effectiveness of transferring morphological patterns from the TCGA to the external AVANT test set, three experts (ASLPC, JHJMvK, and MP) independently reviewed a randomly chosen set of 32 tiles from the TCGA-COAD subset and another 32 tiles randomly selected from the AVANT trial. This qualitative comparison concluded a remarkable resemblance between the TCGA and AVANT tiles within their respective HPCs (Figure 3 [a-i]). In comparing the objective test results from TCGA and AVANT, we found an 80% overlap in the misclassified HPCs between the two datasets and 65% overlap in correctly classified HPCs (Figure 3j, 3k). These results indicate that the robust morphological features extracted from the training set can be effectively transferred to an independent unseen test set.

HPC-based classifier was associated to OS in patients treated with standard-of-care and AVANT-experimental treatment

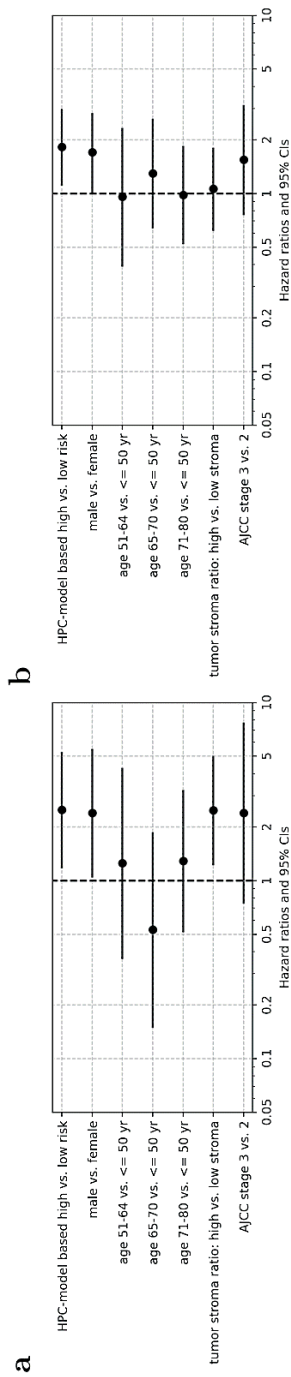
We explored the prognostic significance of HPCs on OS. The OS prediction model was developed within TCGA-COAD (see Methods: Optimization of HPCs and the prediction of OS using Cox regressions with L2 regularization for details). For external validation, we utilized the control group from the AVANT trial who had only received standard adjuvant chemotherapy (i.e. FOLFOX-4). The AVANT trial aimed to investigate whether combining bevacizumab, a humanized anti-vascular endothelial growth factor (VEGF) monoclonal antibody, with standard chemotherapy would improve survival among colon cancer patients [30,33]. The trial had three treatment arms: FOLFOX-4, bevacizumab+FOLFOX-4, bevacizumab+XELOX. The study was prematurely terminated due to the adverse effect in patient survival associated with bevacizumab [30]. Given the unique bevacizumab experimental treatment and its adverse effects, we hypothesized that the survival model trained on TCGA-COAD patients may not generalize well in the AVANT bevacizumab-treated group. We referred to our source population, represented by the TCGA-COAD, as the "standard-of-care group", in contrast to the unique bevacizumab treatment in the AVANT trial. We then opted to validate the OS prediction model, trained on TCGA-COAD data, primarily on AVANT patients who exclusively received standard FOLFOX-4 chemotherapy. This subset is henceforth referred to as the "AVANT control group", serving as an independent test set for validation.

We modeled HPCs on OS using Cox regression with L2 regularization trained within the TCGA-COAD incorporated all 47 HPCs as predictors. The model was optimized through five-fold cross validation (CV) on the TCGA training set (Supplementary Figure 3a). The optimized regularized Cox model was then tested in the independent test set of the AVANT control group. We observed a test set c-index of 0.65 (bootstrap 95% confidence interval [CI]=0.55-0.74) (Supplementary Figure 3b). The HPC-based classifier (i.e. high risk versus low risk) was determined by the median predicted hazard obtained in the TCGA-COAD. The HPC-based model also outperformed a clinical baseline model trained on age, sex, tumor-stroma ratio (TSR), and AJCC TNM stage (Supplementary Figure 3c). To investigate whether our HPC-based classifier provides additional prognostic value to the existing important clinical predictors, a regular multivariable Cox regression was fitted within the external AVANT control test set. The model included HPC-based classifier as well as important clinical and demographic variables (Figure 4a). Notably, the HPC-based risk classifier demonstrated significance as an independent prognostic factor (hazard ratio [HR] = 2.50, 95% CI=1.18 – 5.31), along with male sex 213 (HR = 2.42, 95% CI = 1.07 – 5.47) and the TSR (HR = 2.49, 95% CI = 1.23 – 5.04). The 20 most important HPCs associated to OS were summarized using the interpretable SHapley Additive exPlanations (SHAP) (Figure 4c).

Harnessing the well-defined experimental protocols in AVANT, we were granted the unique opportunity to examine the influence of HPCs on colon cancer OS in the bevacizumab treatment groups (i.e. bevacizumab+FOLFOX-4 or bevacizumab+XELOX). Given the proven comparable therapeutical efficacy of FOLFOX-4 and XELOX [33–35], we consolidated the bevacizumab+FOLFOX-4 and bevacizumab+XELOX cohorts into a unified "AVANT-experimental group". Similar to the analysis stated above, we trained Cox regressions with L2 regularization encompassing all 47 HPCs within the AVANT-experimental patients using 5-fold CV (Supplementary Figure 3d). The HPC-based classifier remained of independent prognostic value ($HR = 1.82$, 95% $CI = 1.11 - 2.99$) after adjusting for age, sex, TNM tumor staging, and TSR (Figure 4b). The top 20 most influential HPCs on OS prediction were shown in the SHAP summary plot (Figure 4d).

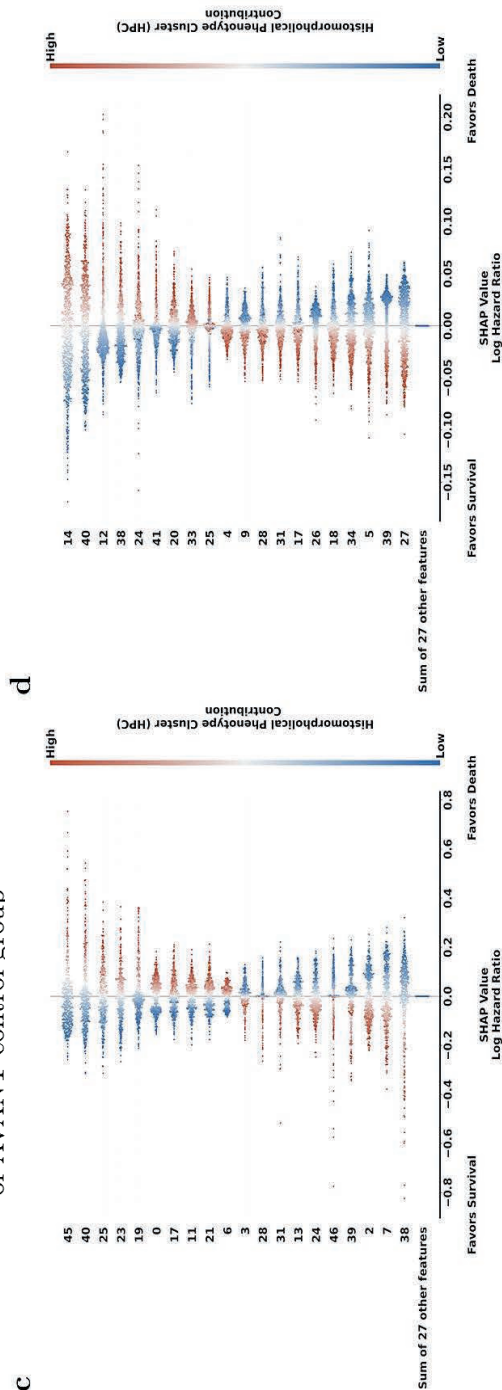
Pathological assessment of OS-associated HPCs in the standard-of-care and AVANT-experimental groups

We highlighted the top 20 most influential HPCs on OS prediction reflecting the SHAP summary plot (Figure 4c, 4d) in the PAGA plots for the standard-of-care and AVANT-experimental groups (Figure 5). The survival-favorable HPCs were highlighted in a shade of red and the survival-unfavorable HPCs were highlighted in blue. Preliminary findings showed that HPCs containing proportionally more healthy colon tissue or immune cells appeared associated with an improved OS, and HPCs comprising mucinous tumor, tumor stroma, and poor-to-undifferentiated (high-grade) tumor epithelium were associated with worse OS in both groups. An overview of the HPCs and their associations to OS, including pathological explanations and references, were given in Supplementary Table 2. We will highlight some OS-associated HPCs in depth here.



Standard-of-care group: external test set
of AVANT control group

AVANT-experimental group



Top 20 OS-associated HPCs
trained in standard-of-care TCGA

Top 20 OS-associated HPCs
trained in AVANT-experimental groups

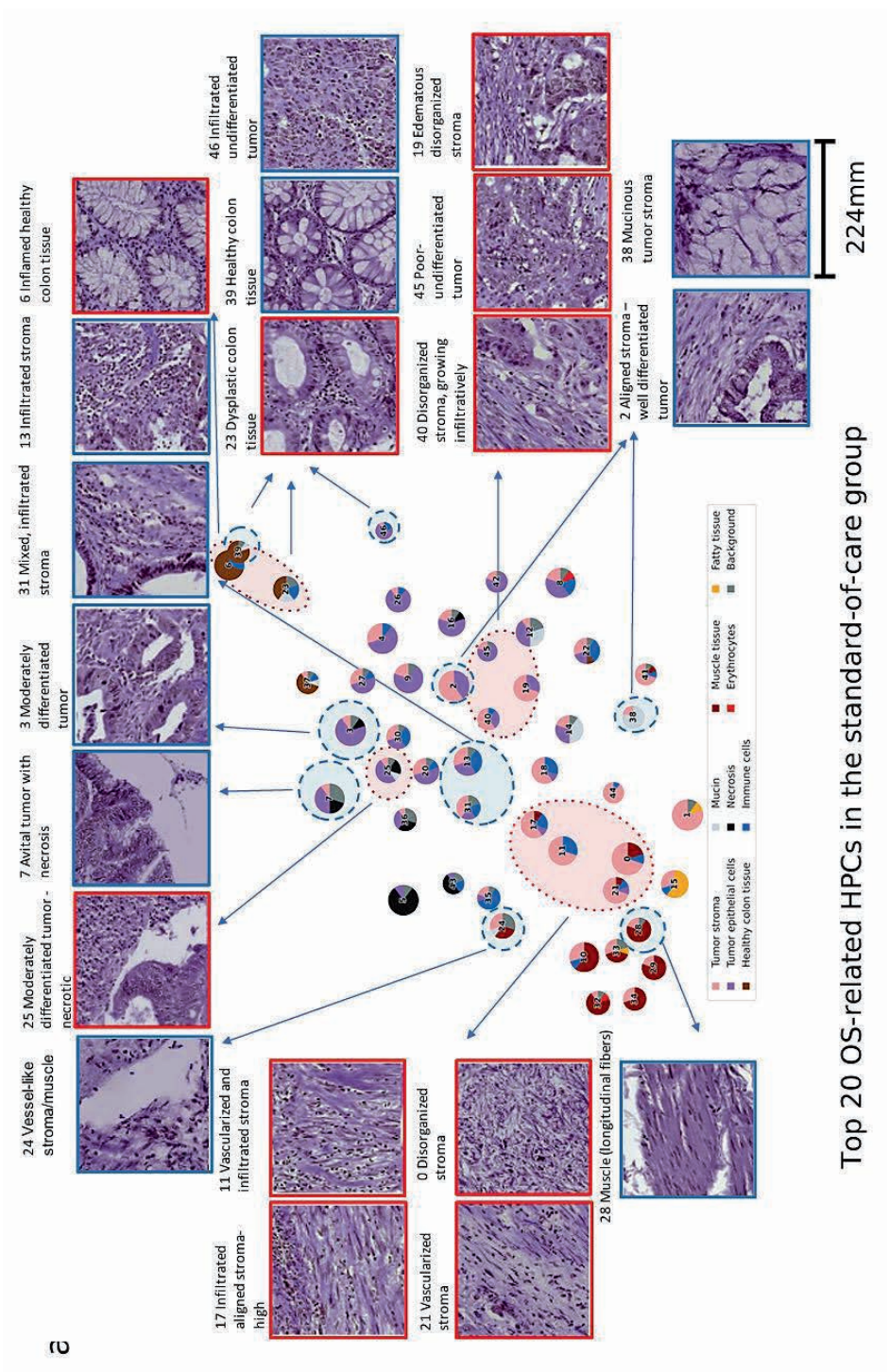
(previous page) **Figure 4.** HPC-based classifier was associated with OS in patients treated with standard-of-care and AVANT-experimental treatment. (a) Ordinary Cox regression for OS, incorporating the HPC-based risk classifier, along with sex, age categories, tumor-stroma ratio, and AJCC TNM staging, was conducted within the external test set of the AVANT control group (N=379 patients after excluding those with missing clinical information). Each point represents the point estimate of HR, and the horizontal whiskers depict the 95% CI. The HPC model-based classifier stands as an independent prognostic factor (HR = 2.50, 95% CI = 1.18 – 5.31) for OS. (b) Ordinary Cox regression for OS, incorporating the HPC-based risk classifier, along with sex, age categories, tumor-stroma ratio, and AJCC TNM staging, was conducted within the AVANT experimental group (N=751 patients after excluding those with missing clinical information). Each point represents the point estimate of HR, and the horizontal whiskers depict the 95% CI. The HPC model-based classifier stands as an independent prognostic factor (HR = 1.82, 95% CI = 1.11 – 2.99) for OS. (c and d) The SHAP summary plots depict the relationship between the center-log-transformed compositional value of an HPC and its impact on death hazard prediction. Statistics estimated from AVANT control (N=405 after excluding those with missing survival data) (c) and experimental groups (N=780 patients after excluding those with missing survival data) (d). The color bar indicates the relative compositional value of an HPC, with red indicating higher and blue indicating lower composition. Higher compositions of the top 10 HPCs were associated with worse OS, while higher compositions of the bottom 10 HPCs were linked to improved OS. AJCC TNM, American Joint Committee on Cancer tumor-node-metastasis classification. AVANT, Bevacizumab-Avastin® adjuvant trial. HPC, histomorphological phenotype cluster. OS, overall survival. SHAP, SHapley Additive exPlanations. TCGA, The Cancer Genome Atlas.

The healthy colon tissue-containing HPC 39 was among the top survival-favorable HPCs as indicated by larger SHAP values in both groups (Figure 4c, 4d). In the standard-of-care group, HPC 23 characterized by dysplastic and low-grade tumor epithelium and HPC 6 marked by stroma-infiltrated healthy colon tissue indicating inflammation, were associated with a worse OS. HPC 31, predominately composed of immune cells, was associated with an improved OS in both treatment groups. A similar association was also noted for the infiltrated-stroma HPC 13, which is part of the immune super-cluster, in the standard-of-care group (Figure 5a) as well as HPC 18 (infiltrated stroma, containing aligned tumor stroma and immune cells) and HPC 17 (infiltrated aligned stroma-high; organized tumor stroma with an immune component) in the AVANT-experimental group (Figure 5b). In the AVANT-experimental group, the mucinous super-cluster (HPCs 12, 14 and 38) was associated with a poor survival (Figure 5b), while conversely, HPC 38 (mucinous tumor stroma) led to a better survival in the standard-of-care group (Figure 5a).

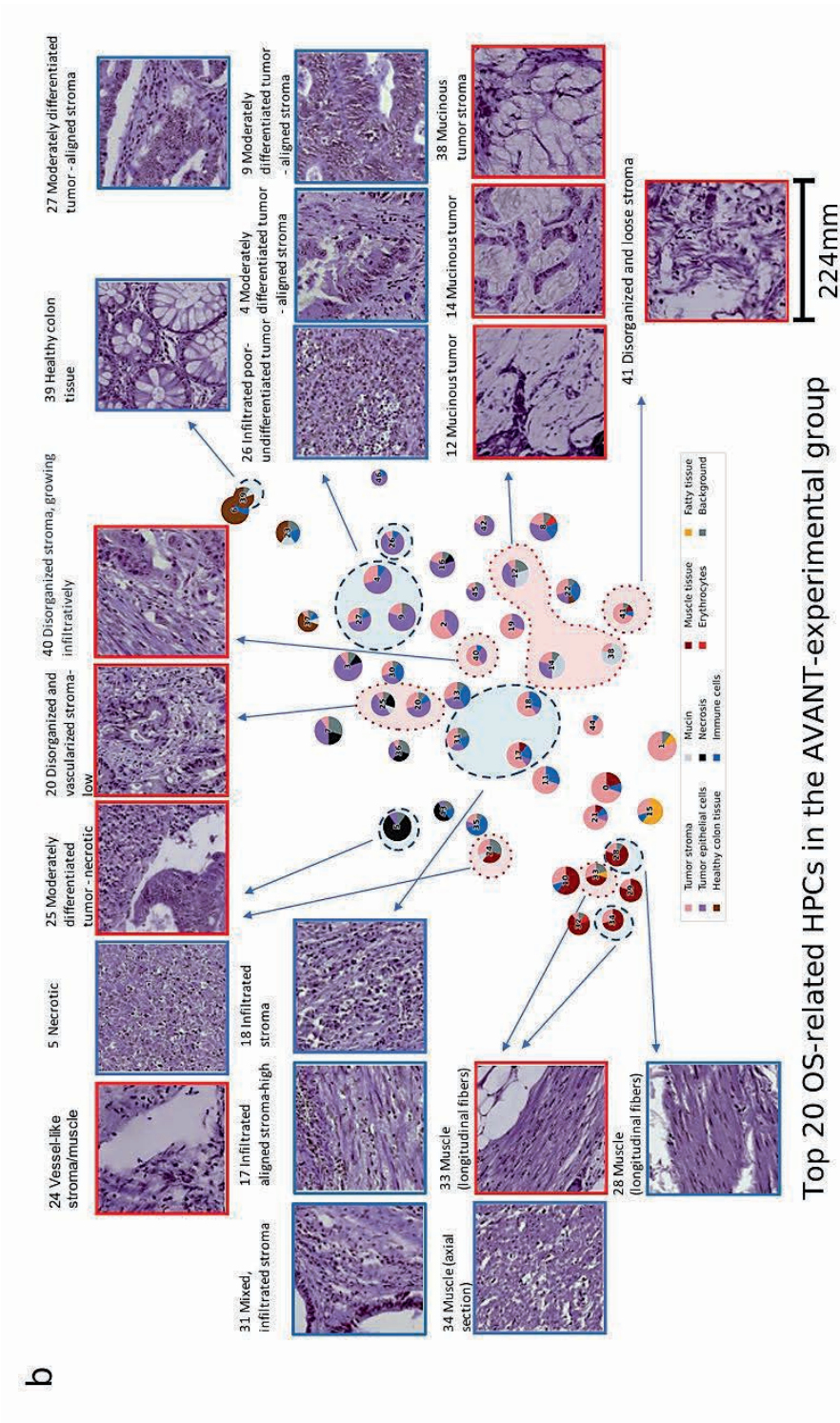
Upon closer examination, we noted that the algorithm captured a distinction in stroma organization within the stromal tissue presented in the tumor stroma and tumor epithelium super-clusters. HPCs in these two super-clusters all contained some component of stromal tissue. However, HPCs with disorganized or heterogeneous tumor stroma with neovascularization (HPC 40 in both groups; 0, 11 and 21 in standard-of-care, and 41 in the AVANT-experimental group) were associated with a poor survival, whereas aligned and organized tumor stromal "strands" were often observed among the top survival-favorable HPCs (HPC 2 in the standard-of-care and HPC 27 in the AVANT-experimental group).

Moreover, the tumor differentiation grade was another parameter correlating with HPCs and their impact on OS. Poor-to-undifferentiated (high-grade) tumor epithelium, e.g. HPC 45, in the standard-of-care group, was linked to a worse survival, while well-to-moderately differentiated (low-grade) tumor led to an improved survival (e.g. HPCs 3 in the standard-of-care group and HPC 4 in the AVANT-experimental group). A unique pattern was furthermore observed in survival-favorable HPCs 46 (standard-of-care group) and 26 (AVANT-experimental group), which was characterized by a predominance of poor-to-undifferentiated tumor epithelium but accompanied by a notable influx of immune cells.

Additional findings were related to the survival-favorable HPC 7 (avital tumor with necrosis) in the standard-of-care group (Figure 5a) and the positive association between necrosis-dominated HPC 5 and better survival in the AVANT-experimental group (Figure 5b). The associations between HPCs containing muscle tissue (e.g. HPC 24, 33 and 34) and survival were generally inconsistent, possibly due to the similarity of muscle fibers and tumor stroma, caused for instance by the organization of muscle fibers and vascularization.



Top 20 OS-related HPCs in the standard-of-care group



(previous pages) **Figure 5.** PAGA plots highlighted with important HPCs related to OS in the standard-of-care and experimental treated group (a) Standard treated group: HPCs colored in the red are linked to worse survival and HPCs colored in blue are linked to better survival. (b) AVANT-experimental treated group: HPCs colored in the red are linked to worse survival and HPCs colored in blue are linked to better survival. AVANT, Bevacizumab-Avastin® adjuvant trial. HPC, histomorphological phenotype cluster. PAGA, partition-based graph abstraction.

Additional findings were related to the survival-favorable HPC 7 (avital tumor with necrosis) in the standard-of-care group (Figure 5a) and the positive association between necrosis-dominated HPC 5 and better survival in the AVANT-experimental group (Figure 5b). The associations between HPCs containing muscle tissue (e.g. HPC 24, 33 and 34) and survival were generally inconsistent, possibly due to the similarity of muscle fibers and tumor stroma, caused for instance by the organization of muscle fibers and vascularization.

Outcome-associated HPCs are linked to diverse immune features in the tumor microenvironment

Spearman correlation coefficients were calculated between the top 20 survival-related HPCs on one hand, and the TCGA immune landscape on the other hand (see Methods: Linkage between survival-related HPCs, immune landscape, and gene expression data for details). The correlation heatmap was plotted with bi-directional hierarchical clustering (Figure 6a). Interestingly, in the standard-of-care group (Figure 6a), HPCs 0, 11, 17 and 21, all part of the phenotypic tumor stroma super-cluster, were identified by the genotypic immune analysis as having a positive correlation with the stroma-high category. Moreover, their SHAP values indicated an positive association with worse survival. Conversely, HPCs 13 and 31 in the immune cell super-cluster were correlated with higher leukocyte fraction. Survival prediction model suggested the immune cell super-cluster was in general associated with a better survival. Further anticipated findings included the validation of HPCs 2 and 24, marked by stroma-high tiles, through immune feature correlation within the stroma-high category.

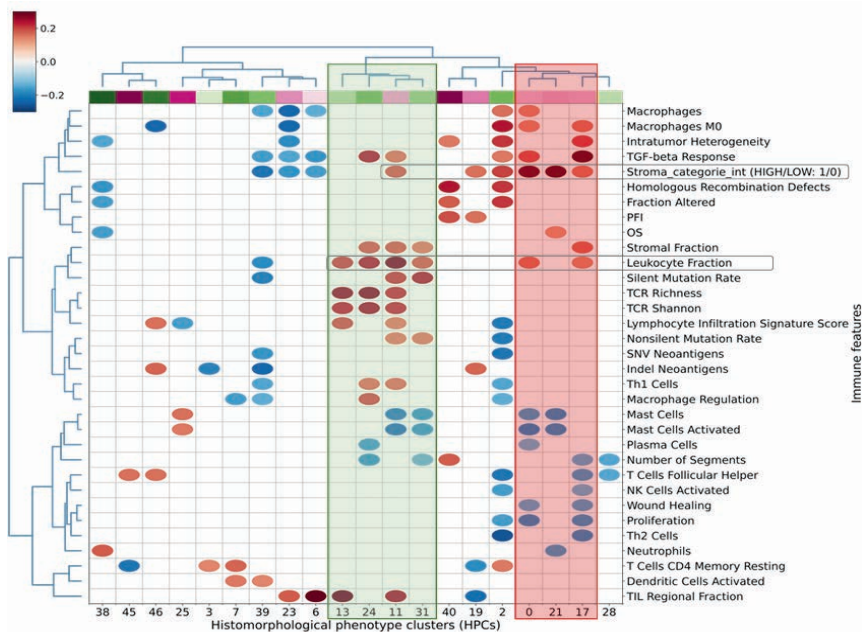
In the AVANT-experimental group (Figure 6b), survival-favorable HPCs 5, 17, 31, 18 correlated with an increased leukocyte fraction. Notably, HPCs 17 and 31 aligned with an elevated immune cell composition (Figure 5b). Additionally, HPCs 17 and 18 were also associated with a higher expression of stroma, which was consistent with the observed stroma-high morphology (Figure 5b). On the contrary, survival-unfavorable HPCs 14, 40 and 20 exhibited higher genomic instability. For instance,

HPCs 40 and 20 were positively correlated with homologous recombination defects, intratumor heterogeneity, and HPC 14 was positively linked to nonsilent mutation rate, single nucleotide variants, and indel neoantigens. Although HPC 14 also showed a positive correlation with the leukocyte fraction, a more pronounced association with genomic instability through multiple pathways seemed to play a greater role in its negative impact on OS.

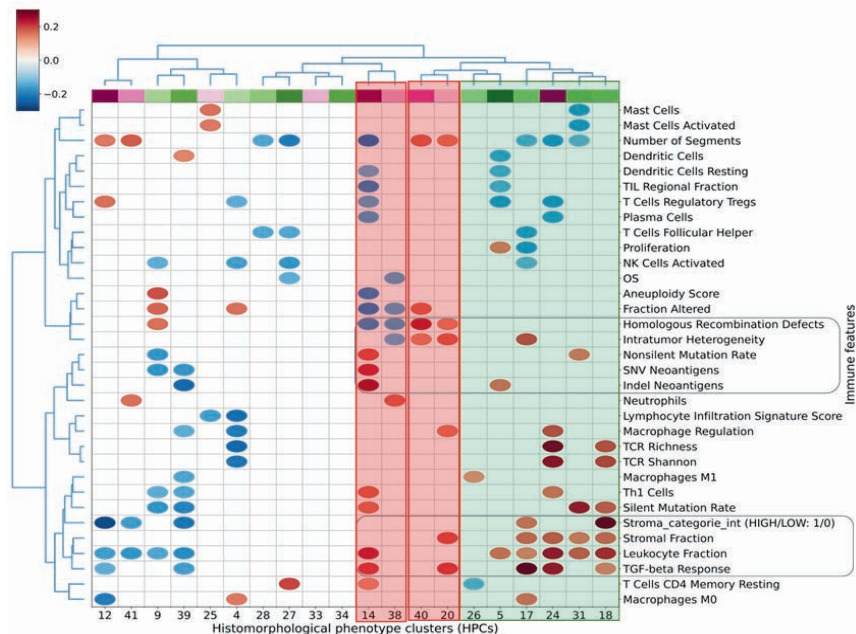
The resulting correlations with immune landscape data aligned with the observed morphologies of HPCs, particularly in stromal and immune features. Furthermore, the data suggest a potential role in genomic instabilities within the AVANT-experimental group. Taken together, these results demonstrate that HPCs can capture the remarkable heterogeneity of the tumor microenvironment.

Outcome-associated HPCs are linked to oncogenic pathways and bevacizumab's mechanism of action

Next, we performed gene set enrichment analysis (GSEA) to discern associations between top OS-related HPCs from both the standard-of-care and AVANT-experimental groups and key cancer hallmark pathways (Figure 6c and d). In the AVANT-experimental cohort, survival-related HPCs showed striking alignment with the enrichment observed in oncogenic hallmark pathways (Figure 6d), which such alignment was overall much less pronounced in standard-of-care group (Figure 6c). Still, in the standard-of-care group, pathways encoding epithelial-to-mesenchymal transition, leading to an increased tumor-stromal percentage, were enriched in survival-unfavorable HPCs 11, 17, 40 and 19. HPCs related to inflammatory response pathways showed primarily positive associations with OS, e.g. with HPCs 13 and 31 related to better survival.



Standard-of-care group: immune landscape



AVANT-experimental group: immune landscape

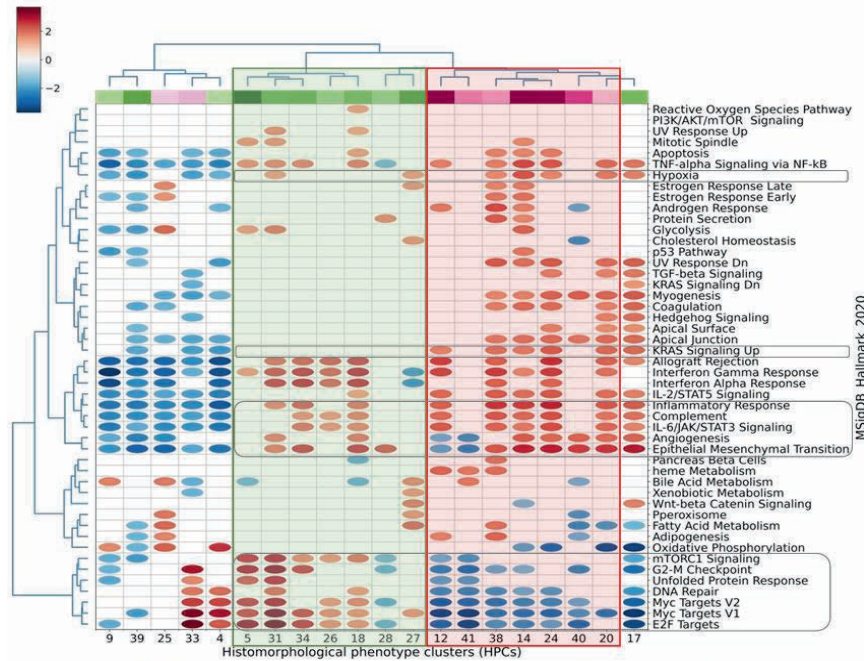
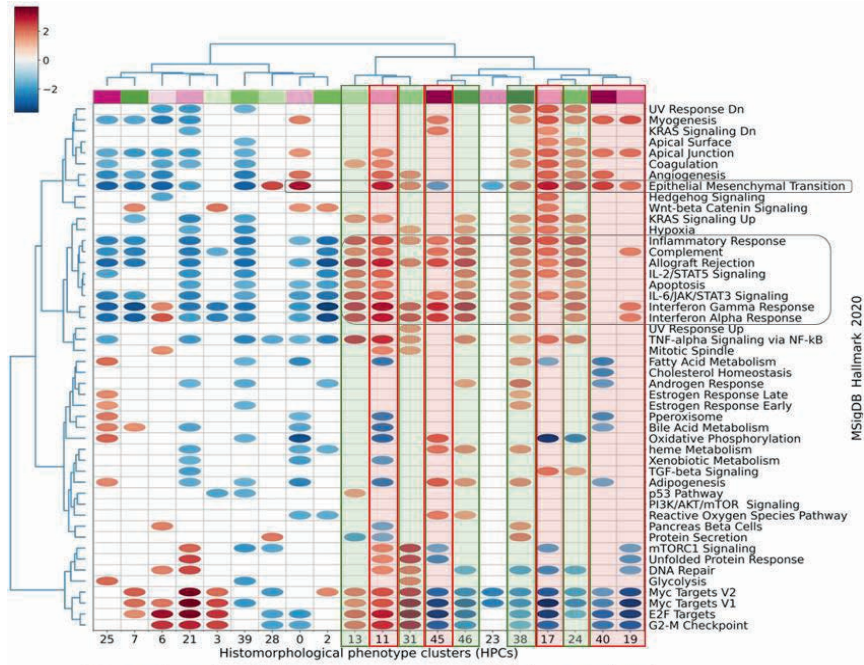


Figure 6. Survival-associated HPCs in relation to immune and genetic profile. (a) Standard-of-care group: Spearman's correlations between top 20 OS-related HPCs and immune landscape features. HPCs (columns of the matrix) were colored according to the beta-coefficients estimated from the optimized regularised Cox regression, with red indicating HPCs related to worse survival and green indicating HPCs related to better survival. The color bar at the upper left corner indicates the value of correlation coefficients with red denoting positive and blue denoting negative correlations. (b) AVANT-experimental treated group: Spearman's correlations between top 20 OS-related HPCs and immune landscape features. (c) Standard-of-care group GSEA between the top OS-related HPCs and major cancer hallmark pathways. HPCs (columns of the matrix) were colored according to the beta984 coefficients estimated from the optimized regularized Cox regression, with red indicating HPCs related to worse survival and green indicating HPCs related to better survival. The color bar at the upper left corner indicates the value of the correlation coefficients with red denoting enrichment and blue denoting underrepresentation in a gene pathway. (d) AVANT-experimental treated group GSEA for the top 20 OS-related HPCs. The immune landscape analysis (N=355 patients) and GSEA analysis (N=265 patients) were performed using data available from TCGA. AVANT, Bevacizumab-Avastin® adjuvant trial. GSEA, gene set enrichment analysis. HPC, histomorphological phenotype cluster. OS, overall survival.

In the AVANT-experimental group (Figure 6d), we observed a strong correlation between hierarchical clustering based on oncogenic enrichment scores and survival-related HPCs. Several HPCs exhibited enrichment in pathways that potentially be specific to bevacizumab through its target of *VEGFa* expression and *KRAS* signaling-up pathway. HPC 5, characterized by necrosis, was linked to elevated *VEGFa* expression ($\rho=0.163$, two-sided $p=0.005$), indicating that patients with higher pre-treatment necrosis levels may have benefited from the bevacizumab treatment. HPC 27, 18, and 31 were associated with enriched hypoxia (HPCs 31, 27) and angiogenesis (HPCs 18, 31), and unfolded protein (HPCs 31, 1, 5) pathways which involve the *VEGFa* gene. In addition, other survival-favorable HPCs may not be specific to bevacizumab but related to the standard cytotoxic chemotherapy of XELOX and FOLFOX-4. Survival-unfavorable HPCs (12, 38, 14, 24, 20) were linked to the *KRAS* signaling-up pathway, which may have a negative impact on the prognosis of patients treated with FOLFOX. Certain survival-favorable HPCs exhibited enrichment in pathways related to cell cycle regulation, signaling, DNA repair, and growth, including G2/M checkpoint (34, 31, 5), E2F targets (5, 31, 34, 26, 18), Myc targets (5, 31, 34, 26, 18), and mTORC1 signaling (5, 31, 34, 26, 18), DNA repair (18, 31). In contrast, survival-unfavorable HPCs demonstrated depletion in these pathways.

Discussion

In this study, we derived and independently validated a total of 47 distinct HPCs that were extracted from colon cancer H&E WSIs using a self-supervised algorithm. These HPCs possess distinctive histomorphologic features carefully identified and assessed by the pathologists and were also linked several immune features and oncogenic pathways. The HPCs showed state-of-the-art performance on OS prediction. Furthermore, the HPC-based risk classifier was an independent prognostic factor after adjusting for important clinical and demographic variables, suggesting additional insight beyond the current clinical prognostication. The unique AVANT trial enabled us to endeavor in identifying possible mechanisms of response to bevacizumab and standard chemotherapy using the HPCs.

Moreover, we hereby emphasize the importance of the tumor microenvironment, or tumor stroma, and its effect on survival. Tumor stroma is composed of extracellular matrix, vasculature, immune cells and cancer-associated fibroblasts, forming a complex and close interaction with tumor epithelial cells [36, 37]. Subject to increasing research the past decades, this dynamic entity has been found to modulate tumor behavior through its cross-talk, and ultimately influence patient-related outcomes. Specifically, regarding the amount of stroma and stromal architecture, i.e. alignment or categorization of the desmoplastic reaction, and immune infiltrate in the stroma, our results corroborate previous literature [31,38–47]. However, such assessments have not been implemented in standard pathology diagnostics.

We identified common histopathological patterns associated with survival as observed in both the standard and bevacizumab-treated group. In line with our results, immune cell infiltration (e.g. HPCs 13, 31) [40,46] and aligned, stroma-low tumors (e.g. HPCs 2, 27) [43,47] were associated with better survival, while poor-to-undifferentiated (high-grade) tumor epithelium (e.g. HPC 45) or mucinous tumor epithelium (e.g. HPCs 12, 14) [1,3,48,49], and disorganized stroma-high tumors (e.g. HPCs 40, 0) [43,47] were linked to worse survival. This pathological phenotype correlated with the corresponding genetic immune profile (e.g. increased leukocyte fractions correlated with HPCs 13 and 31; and a stroma-high category was seen in HPCs 0, 21, 17, 11) and enrichment in oncogenic pathways (e.g. epithelial-to-mesenchymal transition pathway, contributing to tumor stroma amounts [36], correlated with HPCs 17 and 11). We also observed an association, though imperfect, between mucinous tumor, poor-to-undifferentiated tumor epithelium, and survival. One explanation may arise from the absence of contextual information in small images. The differentiation between well-differentiated and undifferentiated, or adenocarcinoma and more mucinous tumor types, is established based on whether each tissue type constitutes more than 50% of the total tumor, underscoring the importance of considering the overall context [1,3,48,49].

Another interesting discovery emerges from the survival-favoring HPCs 26 (AVANT-experimental group) and 46 (standard-of-care group). These HPCs contain primarily poor-to-undifferentiated tumor epithelium but with a high influx of immune cells. Such a histopathological pattern is frequently observed in MSI tumors [50]. MSI-high tumors have been linked to a favorable prognosis [2,51], characterized by lower differentiation grade, increased T-cell infiltration, and reduced susceptibility to invasiveness and *KRAS* mutation [51], however, are commonly identified through separate MSI analysis and/or additional immunohistochemical staining for mismatch repair enzymes in pathology diagnostics [2,3,51].

Furthermore, we also noted HPC 39 containing predominately healthy colon tissue associated with better survival. Interpretation hereof lies in the nature of the multivariable analysis, where the 47 HPCs were modeled simultaneously. One can interpret this result as, while holding the other 46 HPCs constant, patients with relatively more healthy colon tissue, showed an improved survival. The higher proportion of healthy colon tissue may thus indicate relatively smaller or less aggressive tumors. Indeed, within routine TNM assessments, lower pathological T-stage is known to lead to an improved survival [1,3]. Additionally, the favorable OS could potentially also be linked to an absence of the cancer field effect, leaving this healthy colon tissue unaffected from the cancerous lesion [52].

In the GSEA analysis, we noted a remarkable concurrence in AVANT-experimental group between clustering based on cancer hallmark pathways and outcome-related HPCs, while such alignment was much less pronounced in the standard-of-care group. One possible explanation is the heterogeneity of patients in the TCGA-COAD dataset. In contrast to the well-defined treatment protocol in AVANT, TCGA-COAD patients encompass diverse disease and demographic profiles. Consequently, this diversity led to a wide spectrum of treatments, including surgical, neoadjuvant, and adjuvant therapies. The survival-related HPCs discovered in the TCGA-COAD could therefore be, if at all, related to multiple distinctive biological pathways. A general alignment observed in the AVANT-experimental group was hence not anticipated in the standard-of-care group.

In the AVANT-experimental group however, several survival-favoring HPCs either directly correlated with *VEGFA* expression or were associated to enrichment in oncogenic pathways involving *VEGFA* gene, indicating a favorable responses to the contentious bevacizumab treatment. In particular, HPC 5, primarily characterized by necrosis, emerged as a significant contributor to enhanced survival, displaying a positive correlation with *VEGFA* expression, which is the target of bevacizumab. Necrosis

promotes the expression of *VEGF α* , as dying tumor cells release signals that stimulate the growth of new blood vessels [33,53]. In addition, survival-favoring HPCs 18, 27 and 31 were associated with enriched hypoxia and angiogenesis pathways, which also involve the *VEGF α* gene. We hypothesize that patients exhibiting a higher abundance of HPC 5, 18, 27, and 31 may correspondingly express elevated levels of bevacizumab's target *VEGF α* , which in turn results in a more favorable response to this treatment.

Oxaliplatin plus 5-fluorouracil-based regimens of XELOX and FOLFOX-4 are standard chemotherapy for colorectal cancer [2]. Nonetheless, the treatment response is still modest with an estimated rate of approximately 50% [1,54], and the prediction of which patients will respond to this adjuvant chemotherapy remains challenging. We observed enrichment in oncogenic pathways that may be within the context of XELOX and FOLFOX-4 treatments in the AVANT trial. Survival-unfavorable HPCs (12, 38, 14, 24, 20) were linked to the *KRAS* signaling-up pathway. In line with previous literature, several *KRAS* mutations activate downstream signaling pathways and the *KRAS G12D* mutation was predictive of an inferior response to FOLFOX [55]. HPCs 12, 14, and 38 contained mucinous tumors tissue, which has also been linked to *KRAS* mutational burden [50,56]. Survival-favoring HPCs (5, 31, 34, 26, 18, 27) were associated with enrichment in pathways of cell cycle, signaling, DNA repair, and growth (i.e. G2/M checkpoint, E2F targets, Myc targets, and mTORC1 signaling), while survival-unfavorable HPCs (12, 41, 38, 14, 24, 40, 20) linked to a depletion of those pathways. Interestingly, patients with altered DNA repair capacity showed greater benefits from treatment with oxaliplatin [54,57]. Moreover, carriers of *MNAT1* gene, which is one of the leading genes in the G2/M checkpoint pathway [58], were linked to better treatment outcome of FOLFOX [54]. A plausible mechanism is that patients exhibiting activated oncogenic activities within these pathways might harbor a greater abundance of targets suitable for oxaliplatin-based cytotoxic chemotherapy.

Another finding was our HPC-based prediction on OS outperforming the clinical baseline model. This HPC-based risk classifier remained an independent prognostic factor (HR=2.50, 95% CI=1.18-5.31) after adjusting for crucial clinical and demographic variables including TSR and tumor stage. This finding aligns with recent findings reported by Jiang et al. [13]. Of interest, although both HPCs and the TSR scores were derived from H&E slides, they appeared to encapsulate distinct non-overlapping information. A potential explanation could be that the HPC-based classifier captures intricate details, while TSR assessment requires a broader contextual understanding, not fully attainable with small image patches [37,38,59]. Comparing our OS prediction directly with previous studies poses further challenges due to differences in cancer types (e.g., colorectal [12,13,60] instead of colon cancer only), varied outcomes (e.g., 5-year disease-free survival [DFS] [12]), diverse statistical measures (e.g., hazard

ratios [60]), and absence of independent test sets [15] in prior studies. Nevertheless, a recently study reported a test set c-index of 0.65 for OS prediction in colorectal cancer [13] aligning with our reported c-index in the context of colon cancer.

Building upon the aforementioned findings, this study showcases the prospective clinical utility of AI-generated HPCs (Figure 7). Cancer WSIs were preprocessed into image patches and subsequently used to train SSL encoders and to form HPCs. These HPCs serve as condensed representations of the original WSIs, ready to be inspected by pathologists and enabling flexible linkage to multimodal omics data. These HPCs hold promise in classifying various tumor characteristics, potentially predicting patient prognosis and discerning distinct sensitivity groups to various therapies. Although this study was already trained on multicenter TCGA data and validated in an external, clinical multicenter cohort, we plan to conduct additional validations in other population-based external cohorts to strengthen its clinical applicability. In a pathology lab, this could be implemented as followed: Alongside the routine pathology report, an AI-generated report would provide personalized prognostic risk quantification (e.g., based on the patient's HPC composition and SHAP values). The report would also include tissue composition descriptions for each HPC, granting pathologists a complete overview per patient. Subsequently, with these reports used in the multidisciplinary team meetings, a colon cancer patient can be granted an optimal personalized treatment strategy.

Given the significant potential implications of these HPCs on OS and therapy response, implementation of this AI-based analysis may be advocated to international guideline organizations, such as the TNM evaluation committee of the Union for International Cancer Control (UICC). Our HPC-based analysis not only summarizes the TNM classification, but also correlates to parameters assessed in standard pathology diagnostics, such as the International Collaboration on Cancer Reporting (ICCR) [61]. Of note, ICCR parameters pertain to a global assessment on whole slide or patient level, while HPCs can provide a local assessment of histopathological characteristics on the tile level. However, these tile-level histopathological characteristics can still be linked to ICCR parameters. For example, a core element of the ICCR, 'Histological tumor grade', can be correlated to HPCs containing tumor epithelium with different differentiation statuses. HPCs 2 and 3 are for instance characterized by well-moderately differentiated (or 'low-grade') tumor epithelium, while HPC 45 is formed by poorly-to-undifferentiated (or 'high-grade') tumor epithelium-containing tiles. Moreover, 'Histological tumor type' in the ICCR can be linked to HPC 12, which contains predominantly mucinous tumors. Lastly, the TNM classification, and particularly tumor size, in the ICCR reporting guidelines may be linked to HPCs containing healthy colon tissue (e.g. HPC 39), since WSIs with more of the healthy HPC 39 may correspond to a fractionally smaller tumor size (i.e. lower T-stage).

Despite the interesting findings, the study also has several limitations. The identification of HPCs was based on small image tiles as is imperial to model training, while information regarding the larger context is likely lost, as also stated above. For example, it is often challenging to distinguish aligned and organized tumor stromal strands from muscle tissue through traditional microscopic assessment [59]. Therefore, pathologists typically make this distinction based on contextual cues, color variances, or use additional immunohistochemical stainings, all of which is not available within the small image patches. Another limitation pertains to using TCGA as the training set. Although the TCGA is a large open-sourced database, it depends on the availability of registered clinicopathological data. This introduces potential bias and variability in data quality across participating institutions. Moreover, due to data availability, we were only able to focus on predicting OS rather than DFS which may better reflect tumor behavior and biology [29,62]. Lastly, due to the AVANT treatment regimen design, separate analysis regarding bevacizumab or oxaliplatin-based chemotherapy was not attainable. Nonetheless, we made efforts to differentiate between their distinct mechanisms by conducting histopathological inspections and correlating the findings with immune landscape and oncogenic pathways.

In conclusion, our study employed a self-supervised approach to identify and validate histopathological features in colon cancer that are recognizable by human eyes and relevant to prognosis. These features were interpreted through a pathology-focused perspective. Our results highlighted the clinical significance of tumor tissue type, stromal amount and architecture, and the involvement of immune cells. Integration of histopathological features with genetic and gene expression data unveiled potential insights into oncogenic pathways and their relation to patient survival. Utilizing data from the clinical AVANT trial, we proposed mechanisms influencing patient sensitivity to diverse treatments. Future research should focus on refining prediction accuracy and validating the proposed mechanisms regarding the therapeutic strategies in colon cancer.

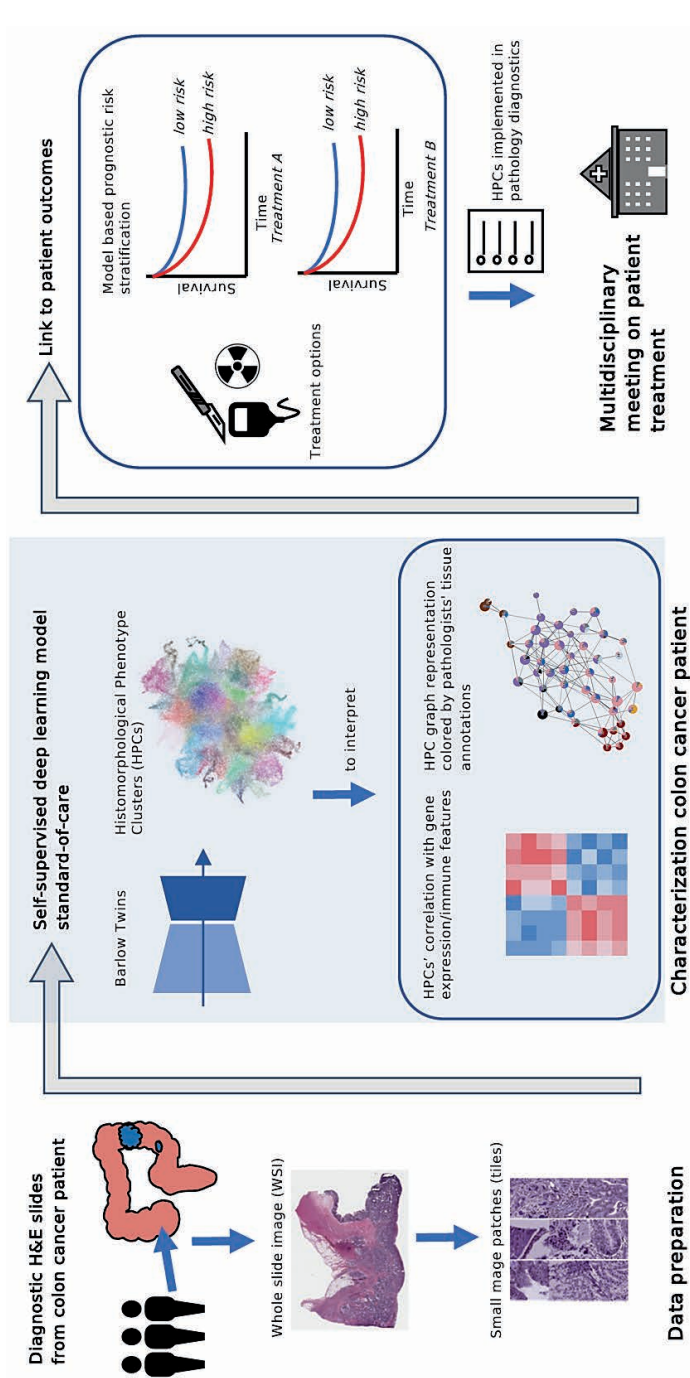


Figure 7. Clinical application of AI-derived HPCs in prediction of patient outcomes. The clinical algorithm consists of three key stages: data preparation, cancer patient characterization, and AI-supported multidisciplinary treatment meetings. Data preparation involves collecting histopathology WSIs, segmenting them into small image tiles. Patient characterization encompasses SSL model training, yielding HPCs via clustering. HPCs are easily interpretable by pathologists, linkable to omic data. Most importantly, HPCs are valuable for predicting diagnosis, patient outcomes, and treatment responses. In treatment-related outcomes, AI-predicted high/low risk groups aid multidisciplinary meetings, enabling personalized treatment plans by oncologists, pathologists, and other physicians. AI, artificial intelligence. HPC, histomorphological phenotype cluster. SSL, self-supervised learning, WSI, whole slide image.

Acknowledgements

The work is supported by the Swedish Research Council (BL, 2019-06360), NCI/NIH Cancer Center Support Grant (AT, P30CA016087), the Stichting Fonds Oncologie Holland and the Bollenstreekfonds, Hillegom, Netherlands (WEM, no grant numbers). We thank the team of NYU Langone High Performance Computing (HPC) Core's resources supporting us to perform the analysis. We would like to thank the NYU Applied Bioinformatics Laboratories (ABL) for providing bioinformatics support and helping with analysis of the data. ABL is a shared resource partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center. We would also like to thank Leslie Solorzano for her input in data preprocessing. The AVANT-trial (BO17920) was originally financed by Genentech Inc., Roche, Switzerland. The authors thank Nikolas Jan Rakebrandt as liaison from Roche.

References

- James D Brierley, Mary K Gospodarowicz, and Christian Wittekind. TNM classification of malignant tumours. John Wiley & Sons, 2017.
- G. Argilés, J. Tabernero, R. Labianca, D. Hochhauser, R. Salazar, T. Iveson, P. Laurent-Puig, P. Quirke, T. Yoshino, J. Taieb, E. Martinelli, and D. Arnold. Localised colon cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up†. *Annals of Oncology*, 31(10):1291–1305, 2020.
- Martin R Weiser. Ajcc 8th edition: colorectal cancer. *Annals of surgical oncology*, 25:1454–1455, 2018.
- A Cervantes, R Adam, S Roselló, D Arnold, N Normanno, J Taieb, J Seligmann, T De Baere, P Osterlund, T Yoshino, et al. Metastatic colorectal cancer: Esmo clinical practice guideline for diagnosis, treatment and follow-up. *Annals of Oncology*, 34(1):10–32, 2023.
- Eileen Morgan, Melina Arnold, A Gini, V Lorenzoni, CJ Cabaasag, Mathieu Laversanne, Jerome Vignat, Jacques Ferlay, Neil Murphy, and Freddie Bray. Global burden of colorectal cancer in 2020 and 2040: Incidence and mortality estimates from globocan. *Gut*, 72(2):338–344, 2023.
- Aoife Maguire and Kieran Sheahan. Controversies in the pathological assessment of colorectal cancer. *World journal of gastroenterology: WJG*, 20(29):9850, 2014.
- Kaustav Bera, Kurt A Schalper, David L Rimm, Vamsidhar Velcheti, and Anant Madabhushi. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nature reviews Clinical oncology*, 16(11):703–715, 2019.
- Hamid R Tizhoosh, Phedias Diamandis, Clinton JV Campbell, Amir Safarpour, Shivam Kalra, Danial Maleki, Abtin Riasatian, and Morteza Babaie. Searching images for consensus: can ai remove observer variability in pathology? *The American journal of pathology*, 191(10):1702–1708, 2021.
- Wouter Bulten, Hans Pinckaers, Hester van Boven, Robert Vink, Thomas de Bel, Bram van Ginneken, Jeroen van der Laak, Christina Hulsbergen-van de Kaa, and Geert Litjens. Automated deep-learning system for gleason grading of prostate cancer using biopsies: a diagnostic study. *The Lancet Oncology*, 21(2):233–241, 2020.
- Mohsin Bilal, Shan E Ahmed Raza, Ayesha Azam, Simon Graham, Mohammad Ilyas, Ian A Cree, David Snead, Fayyaz Minhas, and Nasir M Rajpoot. Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: a retrospective study. *The Lancet Digital Health*, 3(12):e763–e772, 2021.
- Veenu Rani, Syed Tufael Nabi, Munish Kumar, Ajay Mittal, and Krishan Kumar. Self-supervised learning: A succinct review. *Archives of Computational Methods in Engineering*, pages 1–15, 2023.
- Ellery Wulczyn, David F Steiner, Melissa Moran, Markus Plass, Robert Reihs, Fraser Tan, Isabelle Flament-Auvigne, Trissia Brown, Peter Regitnig, Po-Hsuan Cameron Chen, et al. Interpretable survival prediction for colorectal cancer using deep learning. *NPJ digital medicine*, 4(1):71, 2021.
- Xiaofeng Jiang, Michael Hoffmeister, Hermann Brenner, Hannah Sophie Muti, Tanwei Yuan, Sebastian Foersch, Nicholas P West, Alexander Brobeil, Jitendra Jonnagaddala, Nicholas Hawkins, et al. End-to-end prognostication in colorectal cancer by deep learning: a retrospective, multicentre study. *The Lancet Digital Health*, 6(1):e33–e43, 2024.
- Han Xiao, Zongpeng Weng, Kaiyu Sun, Jingxian Shen, Jie Lin, Shuling Chen, Bin Li, Yiyu Shi Ming Kuang, Xinming Song, et al. Predicting 5-year recurrence risk in colorectal cancer: development and validation of a histology-based deep learning approach. *British Journal of Cancer*, pages 1–10, 2024.
- Richard J Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Jana Lipkova, Zahra Noor, Muhammad Shaban, Maha Shady, Mane Williams, Bumjin Joo, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell*, 40(8):865–878, 2022.
- Omar SM El Nahhas, Chiara ML Loeffler, Zunamys I Carrero, Marko van Treeck, Fiona R Kolbinger, Katherine J Hewitt, Hannah S Muti, Mara Graziani, Qinghe Zeng, Julien Calderaro, et al. Regression-based deep-learning predicts molecular biomarkers from pathology slides. *Nature communications*, 15(1):1253, 2024.
- Nicolas Coudray and Aristotelis Tsirigos. Deep learning links histology, molecular signatures and prognosis in cancer. *Nature Cancer*, 1(8):755–757, 2020.
- Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 30(3):850–862, 2024.
- Eugene Vorontsov, Alican Bozkurt, Adam Casson, George Shaikovski, Michal Zelechowski, Kristen Severson, Eric Zimmermann, James Hall, Neil Tenenholtz, Nicolo Fusi, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nature Medicine*, pages 1–12, 2024.
- Mingu Kang, Heon Song, Seonwook Park, Donggeun Yoo, and Sérgio Pereira. Benchmarking self-supervised learning on diverse pathology datasets. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3344–3354, 2023.
- Richard J Chen and Rahul G Krishnan. Self-supervised vision transformers learn visual concepts in histopathology. *arXiv preprint arXiv:2203.00585*, 2022.
- Sophia J Wagner, Daniel Reisenbüchler, Nicholas P West, Jan Moritz Niehues, Jiefu Zhu, Sebastian Foersch, Gregory Patrick Veldhuizen, Philip Quirke, Heike I Grabsch, Piet A van den Brandt, et al. Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study. *Cancer Cell*, 41(9):1650–1661, 2023.

23. Adalberto Claudio Quiros, Nicolas Coudray, Anna Yeaton, Xinyu Yang, Bojing Liu, Hortense Le, Luis Chiriboga, Afreen Karimkhan, Navneet Narula, David A Moore, et al. Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unannotated pathology slides. *Nature Communications*, 15(1):4596, 2024
24. Francesco Cisternino, Sara Ometto, Soumick Chatterjee, Edoardo Giacomuzzi, Adam P Levine, and Craig A Glastonbury. Self-supervised learning for characterising histomorphological diversity and spatial rna expression prediction across 23 human tissue types. *Nature Communications*, 15(1):5906, 2024.
25. Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised809 learning via redundancy reduction. In *International Conference on Machine Learning*, pages 12310–810 12320. PMLR, 2021
26. Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2127–2136. PMLR, 10–15 Jul 2018
27. Chengkuan Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Andrew J Schaumberg, and Faisal Mahmood. Fast and scalable search of whole-slide images via self-supervised deep learning. *Nature Biomedical Engineering*, 6(12):1420–1434, 2022.
28. Corentin Gueréndel, Phil Arnold, and Ben Torben-Nielsen. Creating small but meaningful representations of digital pathology images. In Manfredo Atzori, Nikolay Burlutskiy, Francesco Ciompi, Zhang Li, Fayyaz Minhas, Henning Müller, Tingying Peng, Nasir Rajpoot, Ben Torben-Nielsen, Jeroen van der Laak, Mitko Veta, Yinyin Yuan, and Inti Zlobec, editors, *Proceedings of the MICCAI Workshop on Computational Pathology*, volume 156 of *Proceedings of Machine Learning Research*, pages 206–215. PMLR, 27 Sep 2021
29. The cancer genome atlas program (tcga). <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>. Accessed: 2023-05-09.
30. Aimery de Gramont, Eric Van Cutsem, Hans-Joachim Schmoll, Josep Tabernero, Stephen Clarke, Malcolm J Moore, David Cunningham, Thomas H Cartwright, J Randolph Hecht, Fernando Rivera, et al. Bevacizumab plus oxaliplatin-based chemotherapy as adjuvant treatment for colon cancer (avant): a phase 3 randomised controlled trial. *The lancet oncology*, 13(12):1225–1233, 2012
31. Stéphanie M Zunder, Gabi W van Pelt, Hans J Gelderblom, Christoph Mancao, Hein Putter, Rob A Tollenaar, and Wilma E Mesker. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage ii and stage iii colon cancer. *British Journal of Cancer*, 119(2):164–169, 2018.
32. F Alexander Wolf, Fiona K Hamey, Mireya Plass, Jordi Solana, Joakim S Dahlin, Berthold Göttgens, Nikolaus Rajewsky, Lukas Simon, and Fabian J Theis. Paga: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome biology*, 20:1–9, 2019.
33. Herbert Hurwitz, Louis Fehrenbacher, William Novotny, Thomas Cartwright, John Hainsworth, William Heim, Jordan Berlin, Ari Baron, Susan Griffing, Eric Holmgren, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England journal of medicine*, 350(23):2335–2342, 2004.
34. Leonard B Saltz, Stephen Clarke, Eduardo Díaz-Rubio, Werner Scheithauer, Arie Figer, Ralph Wong, Sheryl Koski, Mikhail Lichinitser, Tsai-Shen Yang, Fernando Rivera, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase iii study. *Journal of clinical oncology*, 26(12):2013–2019, 2008.
35. Fairouz F Kabbinavar, Julie Hambleton, Robert D Mass, Herbert I Hurwitz, Emily Bergsland, and Somnath Sarkar. Combined analysis of efficacy: the addition of bevacizumab to fluorouracil/leucovorin improves survival for patients with metastatic colorectal cancer. *Journal of clinical oncology*, 23(16):3706–3712, 2005
36. Tessa P Sandberg, Maaïke PME Stuart, Jan Oosting, Rob AEM Tollenaar, Cornelis FM Sier, and Wilma E Mesker. Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumor-stroma ratio in colorectal cancer. *BMC cancer*, 19(1):1–9, 2019
37. Gabi W van Pelt, Tessa P Sandberg, Hans Morreau, Hans Gelderblom, J Han JM van Krieken, Rob AEM Tollenaar, and Wilma E Mesker. The tumour–stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 73(2):197–206, 2018.
38. Wilma E Mesker, Jan Junggebur, Karoly Szuhai, Pieter de Heer, Hans Morreau, Hans J Tanke, and Rob AEM Tollenaar. The carcinoma–stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Analytical Cellular Pathology*, 29(5):387–398, 2007.
39. Peter Friedl and Stephanie Alexander. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*, 147(5):992–1009, 2011.
40. Alexander Bagaev, Nikita Kotlov, Krystle Nomie, Viktor Svekolkina, Azamat Gafurov, Olga Isaeva, Nikita Osokin, Ivan Kozlov, Felix Frenkel, Olga Gancharova, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer cell*, 39(6):845–865, 2021.
41. TJA Dekker, A Charehbili, VTHBM Smit, Peter ten Dijke, E Meershoek-Klein Kranenbarg, CJH van de Velde, JWR Nortier, RAEM Tollenaar, WE Mesker, and JR Kroep. Disorganised stroma determined on pre-treatment breast cancer biopsies is associated with poor response to neoadjuvant chemotherapy: Results from the neozotac trial. *Molecular oncology*, 9(6):1120–1128, 2015

42. MTA Strous, TKE Faes, ALHM Gubbels, RLA van der Linden, WE Mesker, K Bosscha, CM Bronkhorst, MLG Janssen-Heijnen, FJ Vogelaa, and AP de Bruijne. A high tumour-stroma ratio (tsr) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clinical and Translational Oncology*, 24(6):1047–1058, 2022.
43. Stéphanie Zunder, Priscilla Van der Wilk, Hans Gelderblom, Tim Dekker, Christoph Mancao, Anna Kiialainen, Hein Putter, Rob Tollenaar, and Wilma Mesker. Stromal organization as predictive biomarker for the treatment of colon cancer with adjuvant bevacizumab; a post-hoc analysis of the avant trial. *Cellular Oncology*, 42:717–725, 2019
44. Hideki Ueno, Yukihide Kanemitsu, Shigeki Sekine, Megumi Ishiguro, Eisaku Ito, Yojiro Hashiguchi, Fukuo Kondo, Hideyuki Shimazaki, Yoshiki Kajiwar, Koichi Okamoto, et al. A multicenter study of the prognostic value of desmoplastic reaction categorization in stage ii colorectal cancer. *The American journal of surgical pathology*, 43(8):1015–1022, 2019
45. Hideki Ueno, Yoshiki Kajiwar, Yoich Ajioka, Tamotsu Sugai, Shigeki Sekine, Megumi Ishiguro, Atsuo Takashima, and Yukihide Kanemitsu. Histopathological atlas of desmoplastic reaction characterization in colorectal cancer. *Japanese Journal of Clinical Oncology*, 51(6):1004–1012, 2021
46. Franck Pagès, Bernhard Mlecnik, Florence Marliot, Gabriela Bindea, Fang-Shu Ou, Carlo Bifulco, Alessandro Lugli, Inti Zlobec, Tilman T Rau, Martin D Berger, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 391(10135):2128–2139, 2018.
47. M Polack, MA Smit, GW van Pelt, AGH Roodvoets, E Meershoek-Klein Kranenbarg, H Putter, H Gelderblom, ASLP Crobach, V Terpstra, G Petrushevska, et al. Results from the united study: a multicenter study validating the prognostic effect of the tumor–stroma ratio in colon cancer. *ESMO open*, 9(4):102988, 2024.
48. Hideki Ueno, Yoshiki Kajiwar, Hideyuki Shimazaki, Eiji Shinto, Yojiro Hashiguchi, Kuniaki Nakanishi, Kazunari Maekawa, Yuka Katsurada, Takahiro Nakamura, Hidetaka Mochizuki, et al. New criteria for histologic grading of colorectal cancer. *The American journal of surgical pathology*, 36(2):193–201, 2012.
49. ID Nagtegaal, Mark Arends, and Robert Odze. Tumours of the colon and rectum: Who classification of tumours of the colon and rectum, tnm staging of carcinomas of the colon and rectum and the introduction. In *World Health Organization Classification of Tumours of the Digestive System*, pages 157–162. IARC press, 2019.
50. Jinru Shia, Nikolaus Schultz, Deborah Kuk, Efsevia Vakiani, Sumit Middha, Neil H Segal, Jaclyn F Hechtman, Michael F Berger, Zsofia K Stadler, Martin R Weiser, et al. Morphological characterization of colorectal cancers in the cancer genome atlas reveals distinct morphology–molecular associations: clinical and biological implications. *Modern pathology*, 30(4):599–609, 2017.
51. C Richard Boland and Ajay Goel. Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6):2073–2087, 2010
52. Paul Lochhead, Andrew T Chan, Reiko Nishihara, Charles S Fuchs, Andrew H Beck, Edward Giovannucci, and Shuji Ogino. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Modern Pathology*, 28(1):14–29, 2015.
53. Napoleone Ferrara. Vegf and the quest for tumour angiogenesis factors. *Nature Reviews Cancer*, 2(10):795–803, 2002
54. Elisabeth J Kap, Petra Seibold, Swantje Richter, Dominique Scherer, Nina Habermann, Yesilda Balavarca, Lina Jansen, Natalia Becker, Katrin Pfütze, O Popanda, et al. Genetic variants in dna repair genes as potential predictive markers for oxaliplatin chemotherapy in colorectal cancer. *The pharmacogenomics journal*, 15(6):505–512, 2015
55. Gongmin Zhu, Lijiao Pei, Hongwei Xia, Qiulin Tang, and Feng Bi. Role of oncogenic kras in the prognosis, diagnosis and treatment of colorectal cancer. *Molecular cancer*, 20(1):1–17, 2021.
56. Shih-Ching Chang, Pei-Ching Lin, Jen-Kou Lin, Chien-Hsing Lin, Shung-Haur Yang, Wen-Yi Liang, Wei-Shone Chen, and Jeng-Kai Jiang. Mutation spectra of common cancer-associated genes in different phenotypes of colorectal carcinoma without distant metastasis. *Annals of surgical oncology*, 23:849–855, 2016.
57. Pavel Vodicka, Sona Vodenkova, Tomas Buchler, and Ludmila Vodickova. Dna repair capacity and response to treatment of colon cancer. *Pharmacogenomics*, 20(17):1225–1233, 2019
58. Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P Mesirov, and Pablo Tamayo. The molecular signatures database hallmark gene set collection. *Cell systems*, 1(6):417–425, 2015.
59. Gabi W van Pelt, Sanne Kjær-Frifeldt, J Han JM van Krieken, Raed Al Dieri, Hans Morreau, Rob AEM Tollenaar, Flemming B Sørensen, and Wilma E Mesker. Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Archiv*, 473:405–412, 2018
60. Jakob Nikolas Kather, Johannes Krisam, Pornpimol Charoentong, Tom Luedde, Esther Herpel, Cleo-Aron Weis, Timo Gaiser, Alexander Marx, Nektarios A Valous, Dyke Ferber, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS medicine*, 16(1):e1002730, 2019.
61. MB Loughrey, M Arends, I Brown, LJ Burgart, C Cunningham, JF Flejou, S Kakar, R Kirsch, M Kojima, A Lugli, et al. Colorectal cancer histopathology reporting guide. International Collaboration on Cancer Reporting. Sydney, Australia, 2020
62. Jianfang Liu, Tara Lichtenberg, Katherine A Hoadley, Laila M Poisson, Alexander J Lazar, Andrew D Cherniack, Albert J Kovatich, Christopher C Benz, Douglas A Levine, Adrian V Lee, et al. An integrated tcga pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*, 173(2):400–416, 2018
63. Carmen J Allegra, Greg Yothers, Michael J O’Connell, Saima Sharif, Nicholas J Petrelli, Linda H Colangelo, James N Atkins, Thomas E Seay, Louis Fehrenbacher, Richard M Goldberg, et al. Phase iii trial assessing bevacizumab in stages ii and iii carcinoma of the colon: results of nsabp protocol c-08. *Journal of Clinical Oncology*, 29(1):11, 2011
64. Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyo, Andre L. Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24, 2018.

65. Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, 21, 2001
66. Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1):5233, 2019
67. Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017
68. Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29), 2018.
69. Aravind Subramanian, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, Scott L Pomeroy, Todd R Golub, Eric S Lander, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005
70. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E. The immune landscape of cancer. *Immunity*. Apr 17;48(4):812-30. 2018

Methods

Ethics statement

Our research complies with all relevant ethical regulations. The study protocol was approved by Applied Bioinformatics Laboratories of New York University Grossman School of Medicine and the Department of Surgery of Leiden University Medical Center. The analyses were performed using anonymized archival material, not necessitating additional informed consents. Data from TCGA-COAD was open-accessed, ensuring patient anonymity without risk of patient identification. All institutions contributing annotated biospecimens provided documentation to the TCGA, and have obtained ethical approvals to use the sample and data according to the human subjects protection and data access policies in TCGA program. Archival material derived from the AVANT-trial (BO17920) was performed in accordance with the declaration of Helsinki [30,31]. Protocol approval was obtained from the local medical ethics review committees or institutional review boards at participating sites.

Study population

The TCGA-COAD dataset was used for training and extracting features and histologic patterns using SSL. This dataset consisted of 451 WSIs from 444 unique patients [29] with matched genetic and transcriptomic information. We excluded duplications and WSIs with erroneous resolution that were not suitable for the analyses (i.e. only several kilobytes in size). The final TCGA training set included 435 WSIs from 428 patients with a diagnosed pathological TNM-stage I-IV colon carcinoma (333 alive, 94 dead, and 1 missing vital status). We referred to the source population which TCGA-COAD representing as the "standard-of-care" group, contrasting it with the clinical trial external test data described below.

As external dataset we leveraged a study comprising 1213 colon cancer patients with available diagnostic H&E WSIs (one WSI per patient) as part of the clinical Bevacizumab-Avastin® adjuvant (AVANT) trial [30,31]. Bevacizumab, a VEGF monoclonal antibody, had initially been shown to improve the OS in patients with metastatic colon cancer when jointly used with the standard chemotherapy [33]. In the phase III AVANT trial, with an intent-to-treat population of 3451 patients, an open-label design was used [30]. Patients were randomly assigned in a 1:1:1 ratio to three different treatment regimens: FOLFOX-4 (intravenous 5-fluorouracil/folinic acid plus oxaliplatin), bevacizumab-FOLFOX-4, and bevacizumab-

XELOX (oral capecitabine plus intravenous oxaliplatin) [30]. The study aimed to investigate whether adding bevacizumab to the standard oxaliplatin-based adjuvant chemotherapy could improve DFS among patients with stage II-high risk and III colon cancer [30]. However, the trial was prematurely terminated due to the serious adverse effect on the patient's OS in the bevacizumab-treated group [30]. The AVANT trial was chosen also due to the previously studied potential correlation of VEGF, the stromal compartment (e.g. TSR), and patient prognosis [31,38,43]. For a detailed overview of the trial and patient characteristics, see Zunder et al [31].

Given the unique treatment regimen of bevacizumab and its adverse effect in non-metastatic colon cancer, as also proven by a predecesing clinical trial, the NSABP protocol C-08 trial [63], we decided to primarily validate OS prediction, trained in the TCGA-COAD, within the control group who only received FOLFOX-4 (without bevacizumab) and refer to it as the "AVANT control group". Subsequently, as several phase III trials had demonstrated that FOLFOX and XELOX are comparable in the context of metastatic colorectal cancer [33–35], we thus combined the bevacizumab+FOLFOX-4 and bevacizumab+XELOX groups into a unified "AVANT-experimental group" and conducted a separate analysis to predict OS within the bevacizumab-treated patients.

Data pre-processing

Tissue segmentation and image tiling

We used the preprocessing methods described in our previous study [64]. Tissue areas in WSIs were segmented against background at 10x magnification level (pixel size approximate 1.0 μ m). WSIs were divided into non-overlapping image tiles of size 224x224 pixels. The selection of a 10x magnification was based on two key considerations. Firstly, it aligns with the standard magnification utilized by pathologists during microscopic assessments in clinical practice. Secondly, we conducted visual inspections of tiles at 20x, 10x, and 5x magnifications, and found that tiles at 10x magnification provided an optimal balance of capturing sufficient detail while also offering a reasonably sized overview of the morphological structure. In addition, to overcome the variability of color stains from different scan facilities in the TCGA and AVANT cohorts, we further applied the color normalization [65]. In total, we obtained 1,117,796 tiles in the TCGA training set (i.e. TCGA-COAD), and 4,827,055 tiles in the AVANT external test set (AVANT-COAD, consisting of the standard and experimental treatment groups).

Extracting image features using Barlow Twins

We trained the SSL Barlow Twins feature extractor based on 250,000 image tiles randomly selected from the TCGA-COAD dataset. The Barlow Twins extracted unique latent vectors (128 dimension) from the preprocessed image tiles in TCGA. The model is based on ResNet-like architecture consisted of several ResNet layers and one self-attention layer [23]. In essence, the Barlow Twins calculates the cross-correlation matrix between the embedding outputs of two identical twin networks, both fed with distorted versions of the same image tile [25]. It is optimized to make the correlation matrix close to the identity matrix [25]. We used the batch size of 64 trained on a single NVIDIA® Tesla V100 GPU for 60 epochs. The Barlow Twins feature extractor was frozen after the training and used to project image tiles into latent representation in the entire TCGA training set. To facilitate the downstream analysis, we also applied five-fold CV partition in the TCGA-COAD training set on patient level balanced the on AJCC TNM stage, OS outcomes (i.e. death or censor), and binned survival time categories.

Leiden community detection algorithm

The Leiden clustering algorithm is a graph-based clustering algorithm that aims to identify distinct communities or clusters within graph data [66]. In brief, it optimizes the modularity function (Equation 1), in such way to maximize the difference between the actual number of edges and the expected number of edges in a community [66]. This modularity function also includes a resolution parameter γ , with higher values leading to more clusters and lower values leading to fewer clusters. Leiden clustering is initiated by assigning each node in the graph to its own individual cluster, treating them as separate communities, then iteratively optimizes the modularity function by moving a node from its current cluster to a neighboring cluster or by merging clusters until the algorithm converged and revealing distinct communities or clusters. We employed the Leiden clustering in a particular workflow. We began by constructing a neighborhood graph using the $K=250$ nearest neighbors from a pool of 200,000 randomly selected latent image vectors from a training set. Subsequently, we applied the Leiden algorithm to identify clusters within this neighborhood graph. These cluster labels were then propagated to each individual image tile across the entire dataset, once more utilizing the K-nearest neighbors approach.

$$\mathcal{H} = \frac{1}{2m} \sum_c \left(e_c - \gamma \frac{K_c^2}{2m} \right)$$

Equation 1. Modularity measures the difference between the actual number of edges in a community and the expected number of such edges. e_c denotes the actual number of edges in community c and the expected number of edges is expressed as $\frac{K_c^2}{2m}$, where K_c is the sum of the edges of the nodes in community c and m is the total number of edges in the network. γ is the resolution parameter, with higher values leading to more clusters and lower values leads to fewer clusters.

Identification of HPCs

We identified HPCs as clusters obtained from the Leiden community detection algorithm operated on 128-dimensional image features extracted through the Barlow Twins encoder. The Leiden method was also used for quality control to eliminate artifacts and underfocused image tiles. Quality control was carried out after training the Barlow Twins and extracting image features. In the overall process, we initially generated a substantial number of Leiden clusters. Next, we visually examined sample image tiles from each cluster and removed clusters exhibiting artifacts such as air bubbles, foreign objects, etc., or those containing under-focused images. In particular, within the training set of a randomly selected CV fold (fold 0), we randomly selected 200,000 latent image vectors to generate Leiden clusters. We obtained 125 clusters at a high resolution of $\gamma=6$. The Leiden labels were propagated into the entire TCGA set using again the KNN methods. Next, we inspected randomly selected sample tiles (N=32) from each cluster and identified 12 clusters containing predominately artifacts or underfocused images. Image tiles labeled by these 12 clusters were subsequently removed from the further analyses. The HPCs were newly derived in this cleaned dataset by re-running the Leiden clustering. Optimization of the HPCs was conducted using primarily unsupervised methods and secondarily confirmed using supervised methods. Importantly, both the two approaches converged on the same optimal Leiden resolution, as elaborated below.

Optimization of HPCs using unsupervised methods

Leiden resolutions (i.e. $\gamma=0.4, 0.7, 1.0, 1.5, 2.5$, and 3.0) were optimized using three unsupervised statistical tests: the Disruption score, Silhouette score, and Daves-Boundin index. Due to the potentially high variance in data from the different institutions in the TCGA, all three scores were weighted by the mean percentage of the institution presence in each cluster. We consistently identified the optimal Leiden resolution as 1.5 through the three aforementioned statistical tests (Supplementary Figure 1a).

Optimization of HPCs and the prediction of OS using Cox regressions with L2 regularization

To identify optimal HPC configurations and their associations with patient OS, we trained L2 regularized Cox regressions for OS prediction using 5-fold CV. The Cox regressions were trained separately among standard-of-care colon cancer patients (i.e. TCGA-COAD) and among the AVANT-experimental group.

OS prediction in the standard-of-care group

Prediction of OS from HPCs among the standard-of-care group was trained within TCGA-COAD using 5-fold CV and tested in the independent AVANT control group. Specifically, for each CV fold within TCGA-COAD, we began by generating a range 587 of Leiden clustering configurations at various resolutions, including gamma values of $0.4, 0.7, 1.0, 1.5, 2.5$, and 3.0 , using the method described earlier. Next, at each Leiden resolution, we calculated the compositional representation of HPCs for each WSI (see Main Figure 1c), followed by a center-log-ratio transformation (Equation 2). This transformation was designed to mitigate inter-dependencies among HPCs, ensuring that the independence assumptions required for subsequent Cox regression analysis were met. The L2 regularized Cox regressions were then trained at the patient level, with one WSI per patient considered. At each Leiden resolution, we performed a multivariable L2 regularized Cox regression, incorporating all center-log-ratio-transformed HPCs specific to that resolution. We finetuned L2 regularizer (alpha) through an iterative process involving 50 steps, spanning the alpha range from 10^{-4} to 10^4 . This sequence of steps was repeated across all five cross-validation folds. The optimal Leiden resolution and L2 regularizer was selected based the CV C-index.

Through this optimization process, the optimal Leiden resolution was determined to be 1.5 , and the L2 regularizer alpha was fixed at 0.1842 . Of the note, this optimal Leiden resolution of 1.5 concurred with the

result from the unsupervised approaches. An HPC-based classifier was determined by the median predicted hazard obtained in the TCGA-COAD. Once the Cox model is optimized, we evaluated the final model performance in the external test set consisting of the AVANT-standard care group. First, the 47 HPCs were integrated into the AVANT-standard care group by employing the K-nearest neighbors method (K=250), where each AVANT tile's HPC label was determined based on the majority votes from its nearest neighbors in the TCGA training set. Next, the trained Cox model, with optimized regularization and parameter estimates for the HPCs, was then applied to the AVANT-standard care group to test the prediction of OS (Supplementary Figure 3b).

Furthermore, employing the same CV method, we trained a clinical baseline model on OS in TCGA-COAD using L2 regularized Cox regression incorporating age, sex, TNM staging, and TSR as predictors. We observed a c-index of 0.58 (bootstrap 95% CI=0.49-0.67) in the independent AVANT-standard care group and the model-based risk classifier did not reach the statistical significance level (Supplementary Figure 3c). This baseline model illustrates a simulation of decision-making in clinical practice as control, using the most readily available and relevant clinical and demographic variables. Our HPC-based model outperformed this clinical baseline model. In addition, we explored whether the HPC-based classifier add additional prognostic value to the existing important clinical predictors. We fitted an ordinary multivariable Cox regression within the external AVANT-standard care group, including HPC-based classifier, sex, age, TSR, and AJCC TNM stage (Main Figure 4a).

$$\text{clr}(x_i) = \ln \left(\frac{x_i}{g} \right), x_i = \frac{\text{counting of tiles in } C_i}{\text{total number of tiles in each WSI}}, \quad g = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$

Equation 2. The center-log-ratio transformation (clr) calculates the natural logarithm of the ratio (x_i) of compositional data for a specific cluster (C_i) to the geometric mean (g) of the compositional data across all clusters. The compositional data (x_i) is obtained by dividing the counting of tiles in cluster C_i by the total number of tiles in each WSI. The geometric mean (g) is computed by multiplying the compositional data values (x_i) for each cluster (C_i), ranging from $i = 1$ to n (where n is the total number of clusters), and raising the resulting product to the power of $\frac{1}{n}$.

OS prediction in the AVANT-experimental group

Bevacizumab, a unique intervention investigated in the AVANT trial, was unlikely accessed by patients from the TCGA-COAD cohort. Considering the significant poor prognostic outcome associated with bevacizumab in non-metastatic colorectal cancer, we postulated that the relationship between HPCs on OS might be influenced by this intervention. Furthermore, we hypothesized that estimates of HPCs on OS trained from the general COAD population, such as TCGA, might not be applicable to bevacizumab-treated patients.

To address these hypotheses, we conducted a separate 5-fold CV estimating the relationship between HPCs and OS within the bevacizumab-treated patients (AVANT-experimental group). Similarly, we generated 5-fold train-validation split stratified by TNM stage and survival time. We used the same sets of HPCs obtained at the previously optimized Leiden 1.5 resolution. Similarly, we modelled the center-log-transformed compositional data of HPCs on OS using Cox regressions with L2 regularization. We fine-tuned the L2 regularizer specifically for AVANT-experimental group. The model performance of HPCs on OS prediction was evaluated using c-index in the 5-fold CV validation sets (Supplementary Figure 3d).

To understand the importance of each HPC on OS, we employed SHAP values [67]. SHAP values measure the marginal contribution of a HPC towards the predicted OS, considering all possible combinations of features. We highlighted top 20 important HPCs favoring and hindering survival for both standard-of-care and AVANT-experimental treated COAD patients.

Interpretation of HPCs*Plotting HPCs using UMAP and PAGA plot*

We applied UMAP dimensionality reduction [68] to TCGA-COAD tile vector representations (128 dimensional vectors), color-coded by 47 HPC IDs using optimized Leiden clustering configuration from the prior step. Next, we generated a PAGA plot where each HPC is a node connected by lines based on vector similarity. Pie charts within each HPC node showed annotated tissue type percentages annotated independently by three experts (ASLPC, JHJMvK, and MP) (see below). We analyzed this plot to identify and describe significant interconnected clusters (Main Figure 2b). Based on the PAGA plot, we also defined

the "super-clusters" (Main Figure 2c) according to the interconnections among HPCs 646 and tissue composition.

Pathologist assessment of HPCs

Histopathological assessment and characterization of HPCs

The histomorphological features of 47 HPCs derived in TCGA training set were comprehensively and independently analyzed by two expert pathologists (ASLPC and JHJMvK) and one medical researcher (MP). The analysis was exclusively conducted using image tiles and the assessors were kept blinded to results from other analyses. Specifically, we randomly selected a set of 32 tiles from each HPC within TCGA and each individual tile was examined with specific focus on tumor epithelium, tumor stroma, and immune cells. Attention was also paid to general tumor differentiation grade, tumor stromal amount, stromal classification, i.e. aligned or disorganized, stromal neovascularization or other notable patterns (e.g. dysplastic tissue, fatty tissue, muscle tissue fibers, blood vessels, erythrocytes, etc.) [3,43,59]. Each assessor evaluated the tissue composition based on 32 randomly selected tiles per HPC, providing an average tissue composition for each HPC. In cases of conflicting tissue annotations, the assessors carried out discussions to achieve consensus. In addition, a short label was given to each HPC based on tissue annotation either according to the first and second predominant tissue types and patterns (e.g. "aligned stroma – well differentiated tumor epithelium"), or as the single most dominant tissue type covering $\geq 70\%$ area of an average tile (e.g. "necrotic"). The tissue description, composition, and short labels of all 47 HPCs were displayed in Supplementary Table 1.

Assessment of HPC consistency within and across TCGA and AVANT cohorts

We performed various qualitative and quantitative analyses to evaluate the within-cluster and between-cluster heterogeneity of HPCs, as well as their transferability from TCGA to AVANT cohorts. Initially, qualitative visual assessments were independently conducted by three experts (ASLPC, JHJMvK, and MP) to evaluate the concordance within and discordance between HPCs. This evaluation followed the established protocol for tissue type analysis, utilizing the previously randomly selected 32 tiles per HPC. The assessors noted consistent histomorphological patterns within HPCs and diverse patterns across various HPCs.

In addition, we conducted quantitative objective tests within TCGA and AVANT tiles separately to determine if the learned morphological features of the HPCs could be replicated by human eyes. We displayed three rows, each containing five tiles, which is referred to as a "question". Among these rows, two belonged to the same HPC, while the third row, also referred to as the "odd HPC", were tiles randomly selected from a different HPC. All example tiles were randomly selected within each HPC. The researcher (MP) was required to identify the "odd HPC", and upon doing so, the next question was presented (example shown in Supplementary Figure 2). In order to calculate the success rate, we ran 50 questions per HPC and we repeated the experiment for all 47 HPCs. The test was conducted on a closed website accessible through login and was conducted separately for tiles randomly selected from TCGA and AVANT. The conducted experiment exhibited an average completion time of under 10 seconds per question. The success rate was determined for each HPC by dividing the number of incorrect answers by the total number of questions (50 questions per HPC). We hypothesized that most HPCs would be distinguishable. However, some HPCs may be challenging to differentiate due to tissue similarities, either within the same super-cluster or through similar tissue morphology (e.g., muscle tissue and tumor stroma). The results showed that HPCs in close proximity to each other in the PAGA or belonging to the same super-cluster were indeed more prone to being mistaken.

To validate whether HPCs derived from the TCGA training set generalized well in the external AVANT test set, three assessors independently carried out visual examination of tissue characteristics in 32 randomly selected tiles from each HPC in both the AVANT trial and the TCGA set, comparing their histomorphological features regarding tissue types and composition established in the previous step. Assessors confirmed that the consistent and meaningful patterns, as learned by the Barlow Twins, could be replicated across other clinical cohorts like the AVANT trial. To validate our observation, we compared the objective test results between TCGA and AVANT, focusing on the overlap in both misclassified and correctly classified HPCs across the datasets.

Linkage between survival-related HPCs, immune landscape, and gene expression data

We calculated Spearman correlations between HPCs and RNASeq-derived immune features from TCGA-COAD data, correcting for false discovery rate (0.05) using Benjamini-Hochberg correction. We created

bi-hierarchical clusters for immune features and top 20 significant HPCs linked to OS in both standard-of-care and AVANT-experimental groups. Clustering was based on correlation coefficients using Euclidean distance and Ward's aggregation.

In addition, we performed GSEA to explore the potential associations between HPCs and the major cancer hallmark pathways, as previously described [69]. The analysis was conducted within 282 TCGA-COAD patients with available gene expression data for 20,530 genes. For each HPC, we calculated the Spearman correlation coefficients with the expression data of these 20,530 genes and sorted the resulting correlation coefficients to rank the genes in a descending order. Subsequently, enrichment scores for 50 major cancer hallmark pathways, as defined by the "MSigDB_Hallmark_2020" gene set [58], were computed for each HPC based on the generated ranked gene list. A positive enrichment score indicated higher composition of a HPC was associated with gene enrichment in a pathway and a negative value

suggested higher composition of a HPC was associated with underrepresentation of a pathway. We plotted the cancer hallmark enrichment scores for the top 20 important HPCs related to OS, for both the general and experimental treatment groups. We set a significance level of 0.01 of the false discovery q value. The analysis was conducted using the gseapy (1.0.4) Python package. Two-way hierarchical clustering was based on the Euclidean distance of the enrichment scores and Ward's aggregation method.

Lastly, we created a Supplementary Table 2 to grant a detailed overview of the significant HPCs associated with OS, including the direction of the association (worse or better OS) for both treatment groups (TCGA and AVANT standard, or AVANT-experimental) and in-depth histopathological analyses potentially explaining these associations with appropriate references.

Data and Code Availability Statement

The publicly available TCGA-COAD can be accessed at Genomic Data Commons portal (<https://gdc.cancer.gov/>). The immune landscape signatures for TCGA-COAD is available in Thorsson et al. [70]. The AVANT data that support the findings of this study are available from Genentech Inc., Roche. However, access to these data is restricted as they were used under license for the current study and are not for commercial use to protect patient privacy. Data may however be available from the authors upon request

for research purposes with permission from Genentech Inc., Roche. Data analyses are based upon publicly available Python software packages and codes are available from our previous publication [23] (<https://github.com/AdalbertoCq/Histomorphological-Phenotype-Learning>). The remaining data are available within the Article, Supplementary Information or Source Data file. Source code and Python software packages are available from our previous publication [23] (<https://github.com/AdalbertoCq/Histomorphological-Phenotype-Learning>).

- Chapter 6

Supplementary Material

Supplementary Table 1. Tissue composition for all 47 HPCs

HPC Super-cluster	Label	General description	Tumor epithelium	Tumor stroma	Immune cells	Other	HPC tissue type composition
0	Tumor stroma	Mostly disorganized stroma, sometimes some aligned strands and sometimes some muscle tissue, immune cells in between, with a few isolated tumor cells	A couple isolated tumor cells	Primarily tumor stroma on tiles. Stroma-high tiles, more disorganized than aligned stroma	Some immune cells in hotspots, immune-high overall	Some muscle tissue	Stroma 70%, Muscle tissue 20%, Immune cells 10%
1	Tumor stroma	Loose, edematous and disorganized stroma with background, sometimes vessels and an artefact. Most likely not tumor-induced stroma but healthy serosa as there is also some fatty tissue	A couple isolated tumor cells	Primarily loose, edematous and disorganized stroma-high tiles with background, sometimes vessels with some erythrocytes. Most likely not tumor-induced stroma but healthy serosa looking at pattern, and there is also some fatty tissue	A couple immune cells, immune-low	Some fatty tissue	Stroma 80%, Fatty tissue 10%, Background 10%
2	Tumor stroma	Aligned stroma Mostly up until 50% stroma in aligned, straight strands. Well differentiated tumor in between, with a little dirty necrosis	Mostly well differentiated tumor epithelial cells	Mostly aligned, straight strands of tumor stroma on the stroma-low tiles	A couple immune cells, and some (dirty) necrosis, immune-low	N/A	Stroma 40%, Tumor epithelial cells 60%
3	Tumor epithelium	Moderately differentiated tumor - high-grade dysplasia, and little stroma, (dirty) necrosis and background	Much well-moderately differentiated tumor epithelium	Little tumor stroma, stroma-low tiles	A couple immune cells, and some (dirty) necrosis, immune-low couple vessels with some erythrocytes	Some high-grade dysplasia, some background. A couple vessels with some erythrocytes	Tumor epithelial cells 70%, Stroma 10%, Necrosis 10%, Background 10%
4	Tumor epithelium	Moderately differentiated tumor - aligned in between strands of stroma, some immune cells and a few spots of dirty necrosis	Well-moderately differentiated tumor	Some strands of aligned tumor stroma, stroma-low tiles	Some immune cells, and a few spots of dirty necrosis, immune-high overall	N/A	Tumor epithelial cells 60%, Stroma 30%, Immune cells 10%
5	Necrosis	Necrotic Mostly (dirty) necrosis, some tumor epithelial cells and some vessels, with background	A couple isolated tumor cells	Some aligned tumor stroma strands, stroma-containing low tiles	Primarily tiles stroma-containing (dirty) necrosis	Some vessels with dirty necrosis and some background or	Necrosis 80%, Tumor epithelial cells 10%, Background 10%

				similarly fatty tissue or mucin	
6	Healthy and dysplastic colon tissue	Inflamed healthy colon tissue	Healthy and some dysplastic colon tissue, with many immune cells in proliferative and often vascularized stroma, aspect of inflammation	No tumor epithelium inflamed and vascularized 'healthy' stroma Many immune cells	Healthy or dysplastic colon tissue 70%, Immune cells 30%
7	Necrosis	Avial tumor with necrosis	Tiles are approx. half background, some stroma and necrosis, but mostly moderately differentiated tumor	Mostly moderately differentiated tumor epithelium Some strands of aligned tumor stroma, stroma-low tiles Some immune cells and necrosis, immune-high	Tumor epithelial cells 40%, Background 30%, Stroma 10%, Necrosis 20%
8	Tumor epithelium	Mixed tumor epithelial - infiltrated stroma	Some variation, mostly tumor and some stroma in different levels of differentiation and organization.	Mostly moderately differentiated tumor epithelium Some stroma on mostly stroma-low tiles, some aligned strands but more disorganized in between tumor epithelium Part immune cells, immune-high	Tumor epithelial cells 40%, Stroma 20%, Immune cells 20%, Background 10%, Erythrocytes 10%
9	Tumor epithelium	Moderately differentiated tumor - aligned stroma	Some background, but mostly moderately differentiated tumor and some stroma in between, a few spots of mucin	Mostly well-moderately differentiated tumor epithelium Some aligned strands of stroma, tiles are stroma-low A few immune cells, immune-low	Tumor epithelial cells 70%, Stroma 20%, Background 10%
10	Muscle tissue	Muscle tissue (longitudinal fibers)	Muscle tissue fiber strands as well as similar looking aligned strands of stroma. Some vessels with erythrocytes, and immune cells in between	No tumor epithelium Some tiles contain strands of aligned stroma, then the tile is stroma-high Some immune cells, immune-low in stroma tiles Mostly muscle tissue, stroma longitudinal fibers and some vessels	Muscle 60%, Stroma 30%, Immune cells 10%
11	Tumor stroma	Vascularized and infiltrated stroma	Mostly much tumor stroma with neovascularization, and with many immune cells in between, potentially parts of lymphoid tissue. Few isolated tumor epithelial cells	Few isolated tumor epithelial cells Mostly much and disorganized tumor stroma with neovascularization, stroma-high High amount of immune cells, potentially part of lymphoid tissue. Immune-high	Stroma 70%, Immune cells 30%

12	Mucinous	Mucinous tumor	Some background in otherwise mucinous tumor containing tiles. Notably high contrast in colors, little to no stroma	Predominantly mucinous tumor epithelium	Some stroma, disorganized spots. Stroma-low	Some immune cells, immune-low	Mostly mucus, some background and one tile with fatty tissue	Tumor epithelial cells 40%, Mucin 30%, Background 20%, Stroma 10%
13	Immune cells	Infiltrated stroma	The mostly aligned tumor stroma is highly infiltrated with immune cells, with some well-moderately differentiated tumor epithelial tissue and some background	Well differentiated tumor epithelium, high-grade dysplasia	Mostly aligned, straight strands of tumor stroma on the stroma-low tiles	High amount of immune cells in stroma, potentially part of lymphoid tissue. Also some dirty necrosis. Immune-high	N/A	Immune cells 30%, Stroma 30%, Tumor epithelial cells 30%, Necrosis 10%
14	Mucinous	Mucinous tumor	Tumor epithelial cells with some stroma, on tiles consisting approx. half of mucin, some background	Predominantly mucinous tumor epithelium	A part disorganized tumor stroma, some aligned strands. Officially stroma-high tiles, excluding the mucus	A few immune cells, immune-low	Mostly mucus, some background	Mucin 40%, Tumor epithelial cells 30%, Stroma 20%, Background 10%
15	Fatty tissue	Fatty tissue	Much fatty tissue, some vessels (also round shape), strands of mostly aligned stroma, some immune cells	No tumor epithelium	Strands of aligned stroma, potentially not only tumor-induced but some healthy as well	A part immune cells, some lymphoid tissue, in stroma then immune-high overall	Predominantly fatty tissue on the tiles, and some other round shapes like vessels, some with erythrocytes, and a spot of mucin	Fatty tissue 60%, Stroma 30%, Immune cells 10%
16	Tumor epithelium	Well-differentiated tumor	Mostly well-moderately differentiated tumor, some stroma in between, and dirty necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Some aligned strands of stroma, tiles are stroma-low	Some immune cells, some dirty necrosis, immune-low	Some background	Tumor epithelial cells 60%, Stroma 20%, Necrosis 10%, Background 10%
17	Tumor stroma	Infiltrated aligned stroma-high	Some immune cells and tumor epithelial cells in the much most prevalent aligned strands of stroma (stroma-high). Sometimes a neural bundle and a few spots more similar to muscle fibers	Some parts tumor epithelium, well-differentiated	Predominantly tumor stroma, mostly aligned strands on stroma-high tiles	Part immune cells infiltrated in tumor stroma, immune-high overall	Some spots background, a couple tiles with some muscle tissue. A sporadic neural bundle	Stroma 60%, Immune cells 20%, Tumor epithelial cells 10%, Muscle tissue 10%
18	Immune cells	Infiltrated stroma	Cluster abundant in cells. Much stroma, mostly disorganized, with many immune cells and some isolated tumor cells	Part isolated tumor epithelial cells	Predominantly tumor stroma, mostly disorganized stroma-high tiles	Many immune cells infiltrated in tumor stroma, immune-high	N/A	Stroma 60%, Immune cells 30%, Tumor epithelial cells 10%

19	Tumor stroma	Edematous disorganized stroma almost no immune cells but with tumor epithelium, some necrosis	Some parts tumor epithelium, well-differentiated	Much edematous tumor stroma, mostly disorganized stroma-high tiles	Few immune cells, some (dirty) necrosis, immune-low	A few spots mucin	Stroma 70%, Tumor epithelial cells 30%
20	Tumor epithelium	Disorganized and vascularized stroma-low epithelial cells in disorganized stroma, some immune cells	Well-moderately differentiated tumor	Mostly disorganized stroma, stroma-low tiles, some vascularization	A part immune cells, some (dirty) necrosis, immune-low mostly	Some background mucin, a few spots signet cell component	Tumor epithelial cells 50%, Stroma 30%, Immune cells 10%, Background 10%
21	Tumor stroma	Vascularized stroma	A couple isolated tumor cells	Predominantly aligned tumor stroma strands, stroma-high, vascularized (or other round shapes)	Part immune cells	Some background, a few spots of fatty tissue. A couple tiles muscle tissue in stead of stroma	Stroma 70%, Immune cells 10%, Tumor epithelial cells 10%, Muscle tissue 10%
22	Immune cells	Infiltrated stroma	Some well differentiated tumor epithelium, high-grade dysplasia	Healthy and tumor stroma, which is stroma-high	High number of immune cells, some necrosis, immune-high	Some healthy and dysplastic tissue	Immune cells 30%, Stroma 30%, Tumor epithelial cells 20%, Background 10%, Healthy and dysplastic colon tissue 10%
23	Healthy and dysplastic colon tissue	Dysplastic colon tissue	A couple isolated tumor cells	Dysplastic stroma	Many immune cells in dysplastic stroma	Predominantly dysplastic colon tissue, some background and mucin	Healthy and dysplastic colon tissue 40%, Immune cells 20%, Mucin 20%, Background 20%
24	Muscle tissue	Vessel-like stroma/muscle	A couple isolated tumor cells	Stroma in the shape of vessels, these are mostly aligned tumor stroma and stroma-high	A few immune cells, some dirty necrosis, immune-low	Muscle tissue, mostly the walls from vessels and background, a spot of erythrocytes within these vessels, some mucin	Muscle tissue 30%, Stroma 30%, Background 30%, Mucin 10%

25	Necrosis	Moderately differentiated tumor - necrotic	Moderately differentiated tumor epithelial with (dirty) necrosis, and some stroma in between. Some background and mucin as well	Mostly well differentiated tumor epithelial cells	Some strands of aligned tumor stroma, stroma-low tiles	Part immune cells in (dirty) necrosis, otherwise immune-low	In general busy and disorganized tiles, round shapes, some background, some mucin	Tumor epithelial cells 50%, Necrosis 20%, Stroma 10%, Mucin 10%, Background 10%
26	Tumor epithelium	Infiltrated poor-undifferentiated tumor	Cluster rich in cells. Tiles differentiated with solid growing poor-undifferentiated tumor, others immune cells and some stroma in between. Potentially microsatellite instability	Tiles with solid growing poor-undifferentiated epithelium	A couple strands of aligned tumor stroma-low tiles	Many immune cells, some tiles only containing immune cells, immune-high	Cluster rich in cells, one tile a slight signet cell component, a spot background	Tumor epithelial cells 80%, Immune cells 10%, Stroma 10%
27	Tumor epithelium	Moderately differentiated tumor - aligned stroma	Some background or mucin. Well-moderately differentiated tumor - aligned epithelial cells in some aligned strands of stroma and some immune cells	Well differentiated tumor epithelium, high-grade dysplasia	Some strands of aligned tumor stroma-low tiles	Some immune cells, also in (dirty) necrosis, overall immune-high	Some background	Tumor epithelial cells 50%, Stroma 30%, Immune cells 10%, Background 10%
28	Muscle tissue	Muscle (longitudinal fibers)	Mostly muscle tissue, longitudinal sections, sometimes aligned stroma strands and some background. A few tumor epithelial cells and neural bundles	A couple isolated tumor cells	A couple tiles contain strands of aligned stroma, these are stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly longitudinal fibers, some background or a spot of fatty tissue. A sporadic neural bundle	Muscle tissue 70%, Stroma 20%, Background 10%
29	Muscle tissue	Muscle (longitudinal fibers)	Mostly longitudinal sectioned muscle tissue and/or stroma strands, a bit edematous, some vessels, a neural bundle, not many immune cells	A couple isolated tumor cells	A couple tiles contain strands of aligned stroma, then the tile is stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly longitudinal fibers, some background or a spot of fatty tissue. A sporadic neural bundle. Small vessels often in between	Muscle tissue 80%, Stroma 20%
30	Immune cells	Mixed, immune-high	Aligned stroma strands and well-moderately differentiated tumor epithelial cells, immune cells in between, some dirty necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Aligned stroma strands on mostly stroma-low tiles	Many immune cells infiltrated in tumor stroma, some spots of (dirty) necrosis	N/A	Tumor epithelial cells 30%, Stroma 30%, Immune cells 30%, Necrosis 10%

31	Immune cells Mixed, infiltrated stroma	Often aligned stroma with many immune cells and moderately differentiated tumor epithelium. Often some background, some necrosis	Well differentiated tumor epithelium, high-grade dysplasia	Aligned, loose stroma strands on mostly stroma-low tiles	Part immune cells often infiltrated in tumor stroma, immune-high	Some background, a couple spots of mucin or fatty tissue	Stroma 40%, Immune cells 20%, Tumor epithelial cells 20%, Background 20%
32	Muscle tissueMuscle (axial section)	Axial sections of muscle tissue, some disorganized strands of stroma and some background. Also, some tiles that contain erythrocytes, similar looking to the muscle fiber sections	No tumor epithelium	No tumor stroma	A few immune cells	Tiles contain mostly muscle tissue, with the fibers in axial sections. Some are part of a vessel wall, with sometimes also erythrocytes in vessels. Some background	Muscle tissue 70%, Erythrocytes 10%, Stroma 10%, Background 10%
33	Muscle tissueInvasive tumor buds in muscle, fatty tissue	Muscle fibers in buds in muscle, longitudinal sections with fatty tissue, some vessels and only few immune cells. Some tumor budding present	Some tumor buds	A couple tiles contain strands of aligned stroma, similar to muscle tissue fibers, then the tile is stroma-high	A couple immune cells, immune-low in stroma tiles	Fatty tissue, in between loose bundles of longitudinal muscle fibers, muscle tissue as part of vessels	Muscle tissue 50%, Stroma 20%, Fatty tissue 20%, Background 10%
34	Muscle tissueMuscle (axial section)	Mostly muscle tissue in axial plane, some tiles disorganized tumor stroma, with some vessels; overall few nuclei or immune cells	A couple isolated tumor cells	A couple tiles contain strands of disorganized tumor stroma, these are tiles stroma-high	A few immune cells, immune-low in stroma tiles	Predominantly axial sections of muscle tissue, a few spots of background or mucus	Muscle tissue 70%, Stroma 30%
35	Immune cells Immune cells	Mostly immune cells in stroma or lymphoid tissue, with some tumor epithelial cells or necrosis	A couple isolated tumor cells	Tumor- and healthy stroma, stroma-high tiles	Highly infiltrated (tumor) stroma or necrosis, sometimes part of lymphoid tissue, immune-high	Some background, a few spots of mucin or fatty tissue, a few spots of dysplastic colon tissue	Immune cells 60%, Stroma 20%, Tumor epithelial cells 10%, Background 10%
36	Necrosis Avital tumor with necrosis	Loose and pieces of tissue, mostly avital tumor epithelium or necrosis, in tiles with much background, some strands of stroma	Loose and pieces of tumor epithelium, well-moderately differentiated	Some spots of stroma, mostly disorganized, stroma-low tiles	Immune cells in much present necrosis	Much background, a few spots of mucin	Necrosis 30%, Tumor epithelial cells 30%, Background 30%, Stroma 10%

37	Healthy and dysplastic colon tissue	High-grade dysplasia - adenoma	Adenoma or well differentiated tumor, with much background, some mucin, and few spots of dirty necrosis	Adenoma, high-grade dysplasia or well differentiated tumor epithelium	A few strands of aligned (tumor) stroma, stroma-low	Some immune cells in dysplastic and tumor stroma, a spot of dirty necrosis	Some background and/or mucin	Healthy or dysplastic colon tissue 60%, Stroma 10%, Immune cells 10%, Background 10%, Mucin 10%
38	Mucinous	Mucinous tumor stroma	Mucin, mostly from tumor stroma, some background	A couple isolated tumor cells	Some disorganized strands of tumor stroma in between the stroma, tiles are stroma-high	A few immune cells, immune-low	Predominantly mucin	Mucin 80%, Stroma 20%
39	Healthy and dysplastic colon tissue	Healthy colon tissue	Mostly clear, healthy colon tissue with mucin in goblet cells and the villi, a small component dysplasia, some background and immune cells in between	No tumor epithelium	No tumor stroma	Some immune cells in healthy stroma, normal healthy colon tissue, some background and/or mucin	Predominantly some background and/or mucin	Healthy or dysplastic colon tissue 80%, Background 10%, Mucin 10%
40	Tumor stroma	Disorganized stroma, growing infiltratively	Polymorphic and poorly differentiated tumors in much prevalent edematous and disorganized stroma, growing infiltratively, a few immune cells	Part polymorphic and poorly differentiated tumor epithelium, growing infiltratively	Much prevalent edematous and disorganized tumor stroma, often vascularized, stroma-high	Part immune cells, immune-low	N/A	Stroma 60%, Tumor epithelial cells 30%, Immune cells 10%
41	Tumor stroma	Disorganized and loose stroma	Heterogeneous cluster. Some background and/or (loose) stromal tissue, not all tumor-associated, with some vessels and immune cells, a few spots of mucin	A couple isolated tumor cells	Predominantly loose aligned stroma fibers, some tumor-induced, then stroma-high. Vascularized	Part immune cells, immune-low	Some tiles contain muscle tissue from vessels, some background, a few spots of mucin	Stroma 70%, Muscle tissue 10%, Background 10%, Immune cells 10%
42	Tumor epithelium	Moderately differentiated tumor	Mostly moderately differentiated tumor epithelial with a few aligned strands of stroma in between. Some background and a few artefacts	Mostly well-moderately differentiated tumor epithelium	Some aligned strands of stroma, tiles are stroma-low	Few immune cells, some spots (dirty) necrosis, immune-low	Some background	Tumor epithelial cells 70%, Stroma 20%, Background 10%
43	Necrosis	Necrotic	Necrosis with anucleic cells, or fields with mostly immune cells. Much background and some tumor epithelial cells	Some moderately tumor epithelium, some isolated tumor epithelial cells	A few strands of aligned (tumor) stroma	Immune cells in much. Some background present necrosis	Some background	Necrosis 50%, Immune cells 20%, Background 20%, Tumor epithelial cells 10%

44	Tumor stroma	Disorganized, immune-low stroma (desert-type)	Mostly aligned stroma, desert type with few nuclei and immune cells, potentially some not tumor-induced with a few isolated tumor epithelial cells	A couple isolated tumor cells	Much prevalent edematous and disorganized tumor stroma, often vascularized, stroma-high	A few immune cells, immune-low	A couple tiles contain longitudinal muscle fibers	Stroma 90%, Immune cells 10%
45	Tumor epithelium	Poor-undifferentiated tumor	Poor-undifferentiated tumor epithelium, with some disorganized stroma, some background	Much poor-undifferentiated tumor epithelium	Some disorganized tumor stroma, stroma-low	A few immune cells, immune-low	Some background	Tumor epithelial cells 70%, Stroma 20%, Background 10%
46	Tumor epithelium	Infiltrated undifferentiated tumor	Cluster rich in cells. Solid growing, undifferentiated tumor, with some immune cells. Potentially micro-satellite instability. Component signet cell carcinoma and squamous cells	Much poor-undifferentiated tumor epithelium	A few strands of aligned (tumor) stroma, stroma-low	Many immune cells, also a couple tiles with only immune cells, immune-high	A tile with slight component signet cell, one with necrosis	Tumor epithelial cells 70%, Immune cells 20%, Stroma 10%

The table shows all HPCs with tissue compositions, including labels, super-clusters, and detailed description of aspect and percentage of tissue types. HPC, histomorphological phenotype cluster. N/A, not applicable.

Supplementary Table 2. Summary of the associations between HPCs and overall survival

HPC	Association	Treatment group	Potential associations with OS	References
0	Worse survival	TCGA and AVANT-standard	Disorganized stroma indicates active remodeling and an intense, yet failing, host response to aggressive invasive tumor cells. The tiles are also stroma-high, consistent with results from the immune landscape analysis, also leading to a worse survival.	Mesker et al (2007), Friedl et al (2011), Zunder et al (2018), Zunder et al (2020), Ueno et al (2021), Strous et al (2022), Polack et al (2024)
2	Better survival	TCGA and AVANT-standard	Aligned stroma indicates a successful defense of the host against less invasive tumor cells. Moreover, the tiles are mostly stroma-low with well-differentiated tumor epithelium leading to a better survival.	Mesker et al (2007), Friedl et al (2011), Weiser et al (2011), Ueno et al (2012), Zunder et al (2018), Zunder et al (2020), Ueno et al (2021), Strous et al (2022), Polack et al (2024)
3	Better survival	TCGA and AVANT-standard	Well-to-moderately differentiated (low-grade) tumor epithelium, along with some high-grade dysplasia and low levels of stromal contents, are well-established indicators of a better survival.	Mesker et al (2007), Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Weiser et al (2018), Zunder et al (2018), Strous et al (2022), Polack et al (2024)
4	Better survival	AVANT-experimental	This cluster consists mostly of well-to-moderately differentiated (low-grade) tumor epithelium, with stroma-low tiles and an influx of immune cells, which is associated with a favorable survival.	Mesker et al (2007), Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Pages et al (2018), Weiser et al (2018), Zunder et al (2018), Bagaev et al (2021), Strous et al (2022), Polack et al (2024)
5	Better survival	AVANT-experimental	Much necrosis, promoting the expression of VEGFa as target of bevacizumab, is present in this cluster, leading to a better response to bevacizumab. Enrichment in pathways of cell cycle and DNA repair may also indicate more targets to the cytotoxic chemotherapy and promote survival.	Ferrara et al (2002), Hurwitz et al (2004), Kap et al (2015), Roos et al (2016), Vodicka et al (2019)
6	Worse survival	TCGA and AVANT-standard	This cluster contains a significant amount of inflamed and partially dysplastic colon tissue, indicating a large amount of remodeling. The presence of potential chronic inflammation predisposes the development of cancer as well as gives rise to a possible cancer field effect, hence leading to a worse survival.	Ullman et al (2011), Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018)
7	Better survival	TCGA and AVANT-standard	Avital tumor and necrosis in this cluster are signs of poor neovascularization and thus potentially indicating an less aggressive tumor. Additionally, downregulation in epithelial-to-mesenchymal transition pathways could lead to favorable outcomes.	Ferrara et al (2002), Brierley et al (2016), Weiser et al (2018), Sandberg et al (2019)
9	Better survival	AVANT-experimental	Besides relatively favorable histopathology aspects similar to some previous HPCs (e.g. HPC 4, low-grade tumor epithelium), downregulation of pathways related to epithelial-to-mesenchymal transition and angiogenesis, as well as lower tumor necrosis factor-alpha could lead to an improved survival.	Ferrara et al (2002), Zins et al (2007), Weiser et al (2011), Ueno et al (2012), Weiser et al (2018), Sandberg et al (2019)

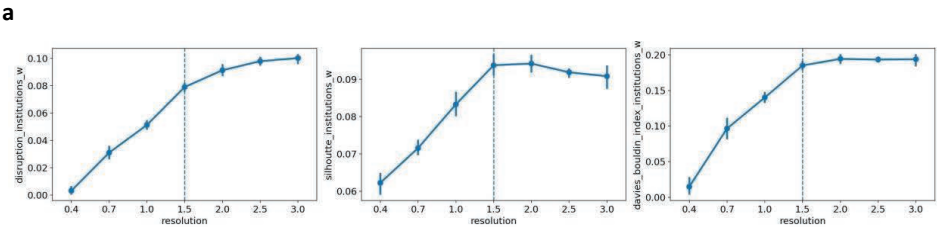
11	Worse survival	TCGA and AVANT-standard	Neovascularization in stroma indicates active host response with wound healing features, known to have tumor promoting properties. Tiles are stroma-high and linked to an epithelial-to-mesenchymal transition in the gene set enrichment analysis, associated to worse outcomes.	Ferrara et al (2002), Mesker et al (2007), Zunder et al (2018), Sandberg et al (2019), Strouss et al (2022), Polack et al (2024)
12	Worse survival	AVANT-experimental	Mucinous tumors are well-known to lead to worse prognosis and can be linked to a higher mutational burden. <i>KRAS</i> signaling-up pathways and depletion in cell cycle regulatory pathways such as DNA repair, overall lead to worse survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Shia et al (2017), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Zhu et al (2021)
13	Better survival	TCGA and AVANT-standard	The immune cell influx has a positive correlation with survival, indicating a good host response; an increased leukocyte fraction was seen in the immune landscape analysis as well.	Pages et al (2018), Bagaev et al (2021)
14	Worse survival	AVANT-experimental	Similar to HPC 12, mucinous tumors are well-known to lead to a worse prognosis and can be linked to a higher mutational burden. <i>KRAS</i> signaling-up pathway and depletion in cell cycle regulatory pathways such as DNA repair, overall lead to worse survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Shia et al (2017), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Zhu et al (2021)
17	Worse survival (TCGA and AVANT-standard) AND Better survival (AVANT-experimental)	Both treatment groups	Mixed associations to survival are seen here. This could be due to high amount of stroma and enrichment in epithelial-to-mesenchymal transition gene pathways, which are poor prognostic factors and lead to a worse survival among the standard-treated patients. However, both the aligned stroma with immune infiltration and the enrichment in the hypoxia pathway (involving the expression of VEGFa gene as the target of bevacizumab), lead to a better survival in experimentally-treated patients.	Ferrara et al (2002), Hurwitz et al (2004), Mesker et al (2007), Friedl et al (2011), Pages et al (2018), Zunder et al (2018), Sandberg et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021), Strouss et al (2022), Polack et al (2024)
18	Better survival	AVANT-experimental	The immune cell component may indicate an effective response from the host and thus lead to a favorable survival. Enrichment in the angiogenesis pathway (involving VEGFa as the target of bevacizumab) may indicate an improved survival in patients receiving bevacizumab.	Hurwitz et al (2004), Kap et al (2015), Pages et al (2018), Bagaev et al (2021)
19	Worse survival	TCGA and AVANT-standard	Stroma-high tiles categorized by disorganized stroma, including an absence of immune cells, indicate an active remodeling with a failing host response, which collectively may lead to a poor survival outcome.	Friedl et al (2011), Kap et al (2015), Pages et al (2018), Zunder et al (2018), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021), Strouss et al (2022), Polack et al (2024)
20	Worse survival	AVANT-experimental	Disorganized stroma can be a sign of insufficient host response and matrix remodeling. Enrichment in pathways of <i>KRAS</i> signaling-up and hypoxia, along with depletion in cell cycle regulation pathways, may contribute to an unfavorable survival through several plausible mechanisms.	Friedl et al (2011), Kap et al (2015), Chang et al (2016), Roos et al (2016), Vodicka et al (2019), Zunder et al (2020), Ueno et al (2021), Zhu et al (2021)

21	Worse survival	TCGA and AVANT-standard	Neovascularization in stroma-high tiles, as indicator of the failure of host response, and enrichment in oncogenic pathways may cause an unfavorable survival in the standard-treated patients.	Ferrera et al (2002), Kap et al (2015), Pages et al (2018), Bagaev et al (2021)
23	Worse survival	TCGA and AVANT-standard	Tiles are characterized by precancerous dysplastic colon tissue and chronic inflammation, both of which predispose to cancer development and suggest a cancer field effect, potentially leading a worse survival.	Ullman et al (2011), Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018)
24	Better survival (AVANT-standard) AND Worse survival (AVANT-experimental)	Both treatment groups	The underlying mechanisms of the mixed results between the standard and experimental treatment groups remains unclear. However, we hypothesize that this may be due to the complex interactions between mixed histopathological patterns and gene expression patterns. On one hand, the worse survival may be attributed to the stroma-high histopathological features of stromal strands in the muscle fibers. This is further confirmed by the upregulation of epithelial-to-mesenchymal transition pathway that has been linked to a poor survival. Enrichment in <i>KRAS</i> signaling-up and depletion cell cycle regulatory pathways may be associated with a worse response to chemotherapy. On the other hand, enrichment in interferon-gamma response and inflammatory response pathways may indicate a strong host immune response which may lead to a better survival.	Ferrera et al (2002), Hurwitz et al (2004), Zins et al (2007), Friedl et al (2011), Kap et al (2015), Chang et al (2016), Roos et al (2016), Pages et al (2018), Sandberg et al (2019), Vodicka et al (2019), Jorgovanovic et al (2020), Zunder et al (2020), Bagaev et al (2021), Zhu et al (2021), Polack et al (2024)
25	Worse survival	Both treatment groups	Necrotic and mucinous components may be associated with a worse survival in general. We also observed depletion in expression of genes in several immune related pathways (e.g. interferon-gamma response). In addition, no enrichment was observed in the pathways involving VEGFa, potentially indicating a low number of targets for bevacizumab.	Ferrera et al (2002), Brierley et al (2016), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Jorgovanovic et al (2020), Bagaev et al (2021)
26	Better survival	AVANT-experimental	Tiles are high in immune cells despite the presence of poorly-to-undifferentiated (high-grade) tumor epithelium, which has been linked to tumors with microsatellite instability. This is known to lead to a better survival. Enrichment in cell cycle regulatory pathways may indicate more targets for cytotoxic chemotherapy, potentially enhancing the survival.	Boland et al (2010), Brierley et al (2016), Shia et al (2017), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Bagaev et al (2021)
27	Better survival	AVANT-experimental	Aligned stroma and influx of immune cells are both predictors for a good response to cytotoxic chemotherapy. Enrichment in the hypoxia pathway, which involves the VEGFa gene, potentially indicate more targets and a better response to bevacizumab. Together, these factors may result in an improved outcome in the experimentally treated patients.	Hurwitz et al (2004), Pages et al (2018), Kap et al (2015), Bagaev et al (2021), Ueno et al (2021)
28	Better survival	Both treatment groups	Higher fractions of this muscle tissue-abundance cluster may indicate more healthy tissue, which may lead a better prognosis. No tumor buds was observed (in contrast to HPC 33), which may indicate a less aggressive tumor.	Brierley et al (2016), Lugli et al (2017), Weiser et al (2018), Nagtegaal et al (2019)

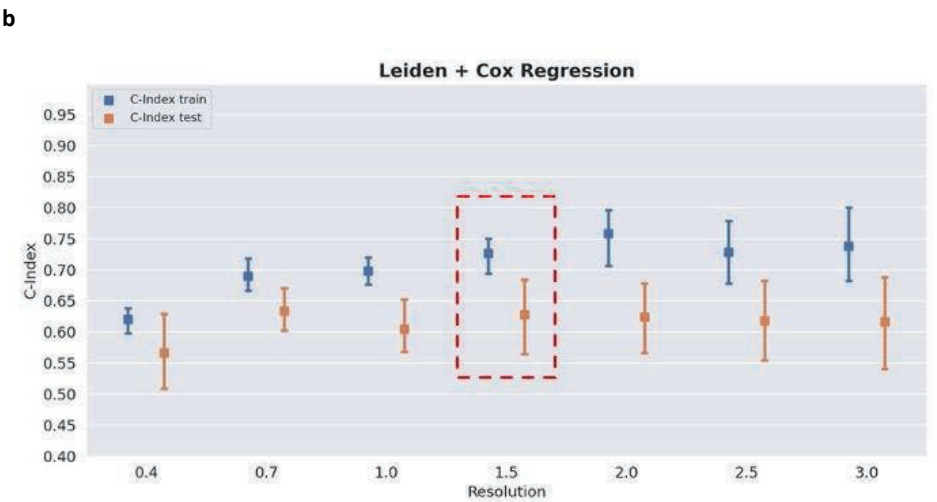
31	Better survival	Both treatment groups	Aligned stroma and an immune component are both favorable factors regarding response to cytotoxic therapy. This is consistent with the high leukocyte fraction observed in the immune landscape analysis. Moreover, enrichment in pathways of hypoxia and angiogenesis, which involves VEGFa, may indicate more targets for bevacizumab. Enrichments in cell cycle regulation (e.g. G3/M checkpoint, mTORC1) may suggest more targets for the chemotherapy.	Ferrara et al (2002), Hurwitz et al (2004), Roos et al (2016), Pages et al (2018), Vodicka et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021)
33	Worse survival	AVANT-experimental	The worse survival may be related to tumor buds invading the serosa and fatty tissue, a known indicator for aggressive tumors. Fatty tissue is also known as a cancer inducer, providing energy for cancer cells. The relative absence of immune cells observed in the tiles suggests a weak host response, which is further reflected in the depletion in inflammatory and angiogenesis pathways. Together, these factors may contribute to poor survivals.	Ferrara et al (2015), Lugli et al (2017), Pages et al (2018), Bagaev et al (2021), Brown et al (2023)
34	Better survival	AVANT-experimental	Muscle tissue abundance may suggest fractionally more healthy tissue, which in turn suggests a smaller tumor size and thus a better prognosis. Moreover, we observed neovascularization in the tiles, which in the experimentally-treated group could enhance survival. Enrichment in pathways related to cell cycle regulations may indicate more targets to the cytotoxic chemotherapy.	Ferrera et al (2002), Kap et al (2015), Brierley et al (2016), Roos et al (2016), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019)
38	Better survival (AVANT-standard) AND Worse survival (AVANT-experimental)	Both treatment groups	Mechanisms underlying these different associations remain unclear. On one hand, mucinous tumors are known to lead to a worse prognosis. <i>KRAS</i> signaling-up pathways and depletion in cell cycle regulatory pathways may suggest a poor response to chemotherapy. On the other hand, enrichment in several inflammatory response pathways may suggest a strong host immune response that results in a better survival.	Kap et al (2015), Brierley et al (2016), Chang et al (2016), Roos et al (2016), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Vodicka et al (2019), Jorgovanovic et al (2020), Bagaev et al (2021), Zhu et al (2021)
39	Better survival	Both treatment groups	Tiles contain healthy colon tissue. Higher fraction of healthy colon tissue tiles may indicate relatively smaller tumor size and potentially less cancer field effect in a larger tissue area (i.e. little disturbances in the healthy tissue) and leads to a better survival.	Lockhead et al (2015), Brierley et al (2016), Weiser et al (2018), Nagtegaal et al (2019)
40	Worse survival	Both treatment groups	Disorganized stroma, indicating an insufficient host response, and poor-to-undifferentiated (high-grade) tumor epithelium could both lead to an unfavorable survival. A cancer field effect negatively influences surrounding tissue, adding to the poor association. Depletion of cell cycle pathways may suggest few targets for cytotoxic chemotherapy and thus worse outcomes.	Friedl et al (2011), Weiser et al (2011), Ueno et al (2012), Kap et al (2015), Lockhead et al (2015), Roos et al (2016), Nagtegaal et al (2019), Vodicka et al (2019), Zunder et al (2020), Ueno et al (2021)
41	Worse survival	AVANT-experimental	The combination of disorganized and loose stroma as well as absence of immune components indicates active remodeling and failure of the immune response. Depletion in cell cycle pathways may suggest fewer targets for cytotoxic chemotherapy and result in a worse survival.	Friedl et al (2011), Kap et al (2015), Roos et al (2016), Vodicka et al (2019), Zunder et al (2020), Bagaev et al (2021), Ueno et al (2021)

45	Worse survival	TCGA and AVANT-standard	Poor-to-undifferentiated (high-grade) tumor epithelium is a well-known indicator of poor survival.	Weiser et al (2011), Ueno et al (2012), Brierley et al (2016), Weiser et al (2018), Nagtegaal et al (2019)
46	Better survival	TCGA and AVANT-standard	Abundance in immune cell components is positively correlated with a better survival. Poorly-to-undifferentiated (high-grade) tumor epithelium together with enrichment in immune-related gene pathways may potentially indicate microsatellite instability which is known to indicate a better survival.	Boland et al (2010), Brierley et al (2016), Shia et al (2017), Pages et al (2018), Weiser et al (2018), Nagtegaal et al (2019), Bagaev et al (2021)

The table summarizes HPCs that were significantly associated with OS per treatment group, including references. Color legend: Orange, HPCs associated with worse survival in one group. Red, HPCs associated with worse survival in both groups. Blue, HPCs associated with improved survival in one group. Green, HPCs associated with improved survival in both groups. Grey, HPCs with mixed associations with OS in the two treatment groups. AVANT, Bevacizumab-Avastin® adjuVANT trial. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. VEGF, vascular endothelial growth factor.



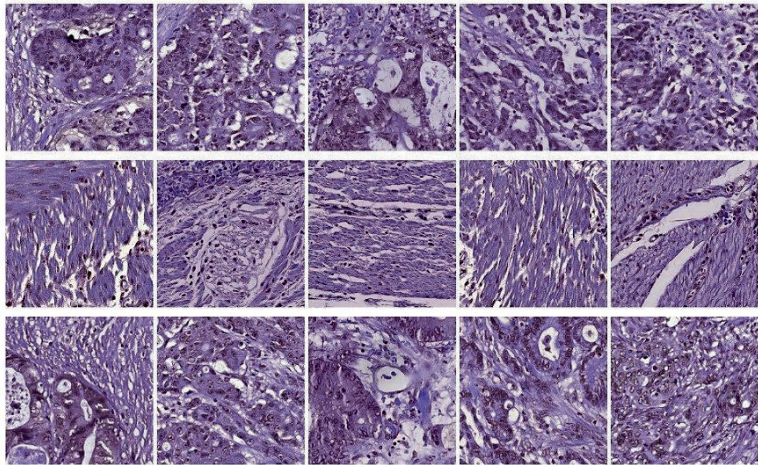
Leiden resolution optimization using unsupervised methods in TCGA



Leiden resolution optimization on OS prediction from regularized Cox regression in TCGA

Supplementary Figure 1: Optimization of Leiden resolution (a) Disruption score, Silhouette score, and Daves-Boundin index weighted by mean percentage of the institution presence in each HPC reached a consensus on optimal Leiden resolution at 1.5 in TCGA (n=405 patients). (b) We trained regularized Cox regressions including all 47 HPCs for each Leiden resolution using 5-fold CV in TCGA (n=405 patients). The Leiden resolution 0.7 and 1.5 showed the highest validation c-index. We chose Leiden 1.5 as the optimal resolution because it led to more explorable HPCs. CV, cross-validation. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. n, number of patients included in the analyses after those with missing clinical data. Source data are provided as a Source Data file.

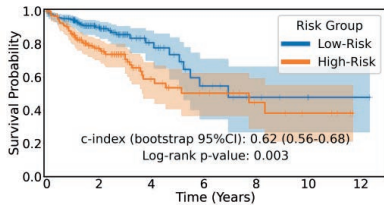
Cluster: 19 / 47, Test: 42 / 50



Supplementary Figure 2. Objective test for all HPCs. Screenshot of the online objective test. Three rows of tiles were presented at each test, with 2 rows were from the same HPC and the other row from a randomly selected different HPC.

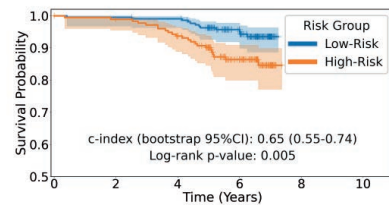
a HPC-based classifier in five-fold CV

in TCGA



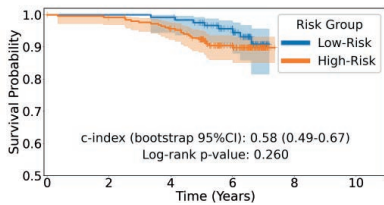
b HPC-based classifier in the external

AVANT control group



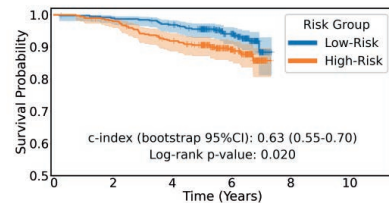
c Clinical baseline classifier in the ex-

ternal AVANT control group



d HPC-based classifier in five-fold CV in

AVANT-experimental group



Supplementary Figure 3. Regularized Cox regression for OS prediction. (a) Kaplan-Meier plot showing the HPC-based classifier stratifying OS risk in the five-fold CV in TCGA (two-sided log-rank $p=0.003$, $n=405$ patients). (b) The HPC-based classifier stratifying OS risk in the external AVANT control test set (two-sided log-rank $p=0.005$, $n=391$ patients). (c) The clinical baseline model (predictors including age, sex, tumor-stroma ratio, AJCC TNM staging) stratifying OS risk in the independent AVANT control test set (two-sided log-rank $p=0.260$, $n=378$ patients). (d) The HPC-based classifier stratified OS risk AVANT-experimental group (two-sided log-rank $p=0.020$, $n=780$ patients). AVANT, Bevacizumab-Avastin® adjuvant trial. CV, cross-validation. HPC, histomorphological phenotype cluster. OS, overall survival. TCGA, The Cancer Genome Atlas. n , number of patients included in the analyses after excluding those with missing clinical data. Source data are provided as a Source Data file.

References

- Alexander Bagaev, Nikita Kotlov, Krystle Nornie, Viktor Svekolkina, Azamat Gafurov, Olga Isaeva, Nikita Osokin, Ivan Kozlov, Felix Frenkel, Olga Gancharova, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer cell*, 39(6):845–865, 2021.
- C Richard Boland and Ajay Goel. Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6):2073–2087, 2010.
- James D Brierley, Mary K Gospodarowicz, and Christian Wittekind. *TNM classification of malignant tumours*. John Wiley & Sons, 2017.
- Kristy A Brown and Philipp E Scherer. Update on adipose tissue and cancer. *Endocrine reviews*, 44(6):961–974, 2023.
- Shih-Ching Chang, Pei-Ching Lin, Jen-Kou Lin, Chien-Hsing Lin, Shung-Haur Yang, Wen-Yi Liang, Wei-Shone Chen, and Jeng-Kai Jiang. Mutation spectra of common cancer-associated genes in different phenotypes of colorectal carcinoma without distant metastasis. *Annals of surgical oncology*, 23:849–855, 2016.
- Napoleone Ferrara. Vegf and the quest for tumour angiogenesis factors. *Nature Reviews Cancer*, 2(10):795–803, 2002.
- Peter Friedl and Stephanie Alexander. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell*, 147(5):992–1009, 2011.
- Herbert Hurwitz, Louis Fehrenbacher, William Novotny, Thomas Cartwright, John Hainsworth, William Heim, Jordan Berlin, Ari Baron, Susan Griffing, Eric Holmgren, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England journal of medicine*, 350(23):2335–2342, 2004.
- Dragica Jorgovanovic, Mengjia Song, Liping Wang, and Yi Zhang. Roles of ifn- γ in tumor progression and regression: a review. *Biomarker research*, 8:1–16, 2020.
- Elisabeth J Kap, Petra Seibold, Swantje Richter, Dominique Scherer, Nina Habermann, Yesilda Balavarcia, Lina Jansen, Natalia Becker, Katrin Pfütz, O Popanda, et al. Genetic variants in dna repair genes as potential predictive markers for oxaliplatin chemotherapy in colorectal cancer. *The pharmacogenomics journal*, 15(6):505–512, 2015.
- Paul Lochhead, Andrew T Chan, Reiko Nishihara, Charles S Fuchs, Andrew H Beck, Edward Giovannucci, and Shuji Ogino. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Modern Pathology*, 28(1):14–29, 2015.
- Alessandro Lugli, Richard Kirsch, Yoichi Ajioka, Fred Bosman, Gieri Cathomas, Heather Dawson, Hala El Zimaity, Jean-François Fléjou, Tine Plato Hansen, Arndt Hartmann, et al. Recommendations for reporting tumor budding in colorectal cancer based on the international tumor budding consensus conference (itbcc) 2016. *Modern pathology*, 30(9):1299–1311, 2017.
- Wilma E Mesker, Jan Junggeburst, Karoly Szuhai, Pieter de Heer, Hans Morreau, Hans J Tanke, and Rob AEM Tollenaar. The carcinoma–stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Analytical Cellular Pathology*, 29(5):387–398, 2007.
- ID Nagtegaal, Mark Arends, and Robert Odze. Tumours of the colon and rectum: Who classification of tumours of the colon and rectum, tm staging of carcinomas of the colon and rectum and the introduction. In *World Health Organization Classification of Tumours of the Digestive System*, pages 157–162. IARC press, 2019.
- Franck Pages, Bernhard Mlecnik, Florence Marliot, Gabriela Bindea, Fang-Shu Ou, Carlo Bifulco, Alessandro Lugli, Inti Zlobec, Tilman T Rau, Martin D Berger, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet*, 391(10135):2128–2139, 2018.
- M Polack, MA Smit, GW van Pelt, AGH Roodvoets, E Meershoek-Klein Kranenbarg, H Putter, H Gelderblom, ASLP Crobach, V Terpsstra, G Petrushevska, et al. Results from the united study: a multicenter study validating the prognostic effect of the tumor-stroma ratio in colon cancer. *ESMO open*, 9(4):102988, 2024.
- Wynand P Roos, Adam D Thomas, and Bernd Kaina. Dna damage and the balance between survival and death in cancer biology. *Nature Reviews Cancer*, 16(1):20–33, 2016.
- Tessa P Sandberg, Maaïke PME Stuart, Jan Oosting, Rob AEM Tollenaar, Cornelis FM Sier, and Wilma E Mesker. Increased expression of cancer-associated fibroblast markers at the invasive front and its association with tumor-stroma ratio in colorectal cancer. *BMC cancer*, 19(1):1–9, 2019.
- Jinru Shia, Nikolaus Schultz, Deborah Kuk, Efsevia Vakiani, Sumit Middha, Neil H Segal, Jaclyn F Hechtman, Michael F Berger, Zsafia K Stadler, Martin R Weiser, et al. Morphological characterization of colorectal cancers in the cancer genome atlas reveals distinct morphology–molecular associations: clinical and biological implications. *Modern pathology*, 30(4):599–609, 2017.
- MTA Strous, TKE Faes, ALHM Gubbels, RLA van der Linden, WE Mesker, K Bosscha, CM Bronkhorst, MLG Janssen-Heijnen, FJ Vogelaar, and AP de Bruijne. A high tumour-stroma ratio (tsr) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clinical and Translational Oncology*, 24(6):1047–1058, 2022.
- Hideki Ueno, Yoshiki Kajiwara, Yoichi Ajioka, Tamotsu Sugai, Shigeki Sekine, Megumi Ishiguro, Atsuo Takashima, and Yukihide Kanemitsu. Histopathological atlas of desmoplastic reaction characterization in colorectal cancer. *Japanese Journal of Clinical Oncology*, 51(6):1004–1012, 2021.
- Hideki Ueno, Yoshiki Kajiwara, Hideyuki Shimazaki, Eiji Shinto, Yojiro Hashiguchi, Kuniaki Nakanishi, Kazunari Maekawa, Yuka Katsurada, Takahiro Nakamura, Hidetaka Mochizuki, et al. New criteria for histologic grading of colorectal cancer. *The American journal of surgical pathology*, 36(2):193–201, 2012.
- Thomas A Ullman and Steven H Itzkowitz. Intestinal inflammation and cancer. *Gastroenterology*, 140(6):1807–1816, 2011.
- Pavel Vodicka, Sona Vodenkova, Tomas Buchler, and Ludmila Vodickova. Dna repair capacity and response to treatment of colon cancer. *Pharmacogenomics*, 20(17):1225–1233, 2019.

25. Martin R Weiser. Ajcc 8th edition: colorectal cancer. *Annals of surgical oncology*, 25:1454–1455, 2018.
26. Martin R Weiser, Mithat Gönen, Joanne F Chou, Michael W Kattan, and Deborah Schrag. Predicting survival after curative colectomy for cancer: individualizing colon cancer staging. *Journal of Clinical Oncology*, 29(36):4796, 2011.
27. Gongmin Zhu, Lijiao Pei, Hongwei Xia, Qiulin Tang, and Feng Bi. Role of oncogenic kras in the prognosis, diagnosis and treatment of colorectal cancer. *Molecular cancer*, 20(1):1–17, 2021.
28. Karin Zins, Dietmar Abraham, Mouldy Sioud, and Seyedhossein Aharinejad. Colon cancer cell–derived tumor necrosis factor- α mediates the tumor growth–promoting response in macrophages by up-regulating the colony-stimulating factor-1 pathway. *Cancer research*, 67(3):1038–1045, 2007.
29. Stéphanie M Zunder, Hans Gelderblom, Rob A Tollenaar, and Wilma E Mesker. The significance of stromal collagen organization in cancer tissue: An in-depth discussion of literature. *Critical Reviews in Oncology/Hematology*, 151:102907, 2020.
30. Stéphanie M Zunder, Gabi W van Pelt, Hans J Gelderblom, Christoph Mancao, Hein Putter, Rob A Tollenaar, and Wilma E Mesker. Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage ii and stage iii colon cancer. *British Journal of Cancer*, 119(2):164–169, 2018.



Chapter 7

Summary, general discussion
and future perspectives

Summary, general discussion and future perspectives

The tumour microenvironment, with all its facets, is no longer deemed an innocuous compartment of a cancerous growth. As much research these past decades has proven, this thesis subsequently showcases the undeniable influence of the tumour stroma specifically on tumour behaviour and patient-related outcomes in colorectal cancer. The tumour-stroma ratio (TSR) is herein par excellence a biomarker capturing that effect, and is the subject of this current work. Commencing with the results of the by-international-instances-recommended, prospective study validating the TSR as prognosticator in colon carcinoma patients, subsequently proving the value of the TSR in predicting response to neoadjuvant therapy in rectal carcinoma patients, and researching this tumour stroma more in depth, the overview given here is concluded by a self-supervised and unbiased artificial intelligence-algorithm confirming the importance of tumour stroma. This thesis forms a theoretical framework, a step-by-step plan and concrete checklist ultimately aiming implementation of the TSR in national and international guidelines.

Previously, the TSR has iteratively shown to be a reproducible, solid and accurate biomarker [1-4]. However, upon the initial presentation of this evidence to the Union of International Cancer Control (UICC) and College of American Pathologists (CAP), advocating implementation of the TSR in routine pathology diagnostics, first, a prospective study was recommended [5]. Finalising the work of predecessors, this current work therefore starts off with the long-awaited results of that recommendation: the international multicentred UNITED study (**Chapter 2**). Over 1,500 stage II and III colon carcinoma patients were included in total by 27 centres from 12 countries worldwide. The TSR was scored by certified pathologists, participating after successfully finishing the official TSR E-learning [6], on the primary resected tumour material through standardized microscopic assessment [7]. Categorised with a stroma-high ($>50\%$) or stroma-low ($\leq 50\%$) colon carcinoma, included patients were in follow-up for a median of three years to evaluate the influence of the TSR on disease-free survival, and secondarily, overall survival and assessment of benefit from adjuvant chemotherapy.

Indeed, disease-free survival was significantly shorter in the patients with a stroma-high colon carcinoma in comparison to their stroma-low counterparts, illustrated by the respective 3-year disease-free survival rates of 70% and 83%. Correcting for potential bias of other risk factors, the TSR stayed an independent prognosticator for disease-free survival in multivariate analysis as well, with a hazard ratio of 1.49. Although the study was not specifically powered for the secondary outcomes, a strong trend towards worse overall survival was visible in stroma-high patients. Moreover, in the stage II and stage III stroma-high patients, a worse disease-free survival was observed despite their received adjuvant treatment, indicating a potential therapy resistance by the tumour stroma.

Hence, the UNITED study hereby unequivocally validates the TSR as an independent prognostic biomarker in predicting worse disease-free survival in colon cancer patients [8]. These results have been once again presented to the UICC, as well as to the Dutch national guideline committee [9], strongly urging implementation of the TSR in the tumour-node-metastasis (TNM) classification [10] as additional prognosticator. In response, first, the UICC congratulated us on the UNITED study, asking to share our work with their collaborators the American Joint Committee on Cancer (AJCC), and promised to add the TSR to the list of important prognostic factors in the upcoming next version of the TNM classification for colon and rectal cancer. Moreover, as the UICC suggested to work with the International Collaboration on Cancer Reporting (ICCR) committee to incorporate the TSR in the diagnostic pathology synoptic reports, we reached out to the ICCR, and in response, they agreed that it needs consideration for the next ICCR review [11]. Although the Dutch national guidelines committee also acknowledged the prognostic value and promised to submit the implementation request in their next meeting, the committee emphasized that a clinical need for ultimate validation of the predictive value of the TSR remained. An initial exploration of that value, predicting therapeutic effect or possible benefit from this therapy, is already endeavoured in this thesis in **Chapter 3**.

Here, the influence of the TSR score on biopsy material and response to neoadjuvant therapy in patients with rectal carcinoma is studied. As neoadjuvant therapy alters the tumour microenvironment, resulting in e.g. fibrosis [12], and a total neoadjuvant therapy regime, including chemoradiation variations, is currently standard protocol in rectal carcinomas [13], a biopsy-based and upfront selection of potential responders is imperative. Aiming to confirm the findings of the UNITED study as well as other existing, though insufficient and scarce, literature [14, 15], we hypothesised that a biopsy-scored stroma-high rectal carcinoma patient would less often reach a major response to their neoadjuvant therapy. Ultimately, these potential non-responders could be for instance counselled in a shared decision-making setting to opt for a surgery-based approach, sparing them the burdensome and inefficient neoadjuvant therapy.

Harnessing data from two clinical trials, i.e. RAPIDO[16] and PROCTOR-SCRIPT[17], and integrating our LUMC patient database, a large multicentre cohort was therefore established. After first validating the TSR scoring method in biopsy material in our study as per literature [14, 15], we could indeed observe significantly less major response rates in a total neoadjuvant therapy approach in the biopsy-scored stroma-high rectal carcinoma patients compared to stroma-low patients (hazard ratio 0.63). Although the TSR had no singular prognostic effect on disease-free nor overall survival, this was however the case for response, to which the TSR was a major contributor. Lastly, emphasizing their inherent aggressiveness, locally advanced rectal carcinomas were more often stroma-high. Hence, where the TSR was proven to be a prognostic biomarker in chapter 2, here in chapter 3, the TSR is also a proven predictive biomarker, albeit in a retrospective analysis and in rectal carcinoma. Currently,

novel established collaborations are utilised to accurately determine the predictive potential of the TSR in colon carcinomas as well, aiming ultimately unavoidable implementation in (inter)national guidelines: the UNITED-II. The clinical consequences of implementing this important parameter are far-reaching, identifying those stroma-high patients with predicted worse prognosis and less expected benefit from therapy. In frail elderly, this could potentially lead to a more supportive approach with a focus in shared-decision making on maintaining quality of life, while for the younger and fitter patients, more research is necessary for experimental and personalised therapies.

One of those potential experimental personalised therapies comprises the use of theragnostics [18]. Combining imaging and therapy by a single radioactive biologically-targeted substance, theragnostics can also pertain to pathways or cells in the tumour stroma compartment [19]. Specifically, the initially quiescent fibroblast, activated by an epithelial-to-mesenchymal transition and giving rise to the so-called cancer-associated fibroblast (CAF), is subject to an exponential amount of research [20]. CAFs, actively remodelling the tumour microenvironment and being important modulators of tumour aggressiveness, form therefore an attractive target [21]. Despite their large heterogeneity, a general CAF marker expressed and used in a majority of research is the fibroblast activation protein (FAP) [22]. Radioactively labelled and chemically modified, the FAP inhibitor (FAPI) is currently increasingly studied in improving tumour diagnostics and staging, specifically in identifying tumour-positive lymph nodes, laying groundwork for future theragnostics [23]. Hence, this thesis subsequently dives more into the tumour stroma in lymph node metastases, followed by an assessment of FAP-expressing CAFs in colorectal carcinomas, aiming to elucidate this phenomenon and forming a theoretical framework.

In **Chapter 4**, we first analyse tumour-positive lymph nodes and specifically, the stromal metastases therein. Since detailed literature on this was scarce, we aimed to determine whether stromal-based detection methods like the FAPI tracer would indeed theoretically improve nodal staging [24]. Protocols in imaging state that lymph nodes with a diameter in size >5-10 millimetres, along with other aspects like roundness of shape, are suspect for malignancy, but this is still prone to erroneous interpretations [25, 26]. An overview is provided in this work containing all features, sizes and tumour-stromal content of lymph node metastases in colon carcinoma material. We observed that lymph node size had positive correlations to the presence and the size of metastases, however, also small lymph nodes <5 millimetres in diameter could contain stromal metastases [27]. A theoretical basis was thus formed for the indeed often clinically observed improvement of nodal staging with the upcoming targeted FAPI tracer used in nuclear imaging compared to the common fluorodeoxyglucose tracer, but also the potential limitations thereof [28].

Subsequently, in **Chapter 5**, we look even more in depth at the FAP marker by using immunohistochemical staining in colorectal carcinoma tissue. Aiming to fill the current gap in

knowledge and to form an extensive histological reference for future correlation to FAPI imaging studies [29], we described patterns of FAP expression in primary resected tumours, lymph nodes and biopsy material. Moreover, we assessed the correlation between the TSR and FAP expression, which later could constitute FAPI uptake: if positively correlated, on future imaging high FAPI uptake would thus identify stroma-high colorectal carcinomas, potentially enabling an even more improved upfront therapy selection. While the majority of colorectal carcinoma tissue did stain positively for FAP, this was mostly heterogeneous, and not only between tumours, but also within tumours. Also healthy colorectal and lymphoid tissue could stain for FAP. Furthermore, the TSR, mostly scored in the tumour centre, did not correlate to FAP, which was generally expressed at the invasive front. Therefore, care has to be taken directly translating FAP expression and inherently future FAPI uptake to disease activity and extent. This chapter emphasizes the importance of a multidisciplinary approach in studies, and the need for further correlation studies, some of which are currently ongoing in collaborative aspects, before patient treatment can be endeavoured, i.e. the FAPI-CRC1 (NCT05209750) and FoCus (NCT06191120) trials.

Historically, from the sixteenth century onward, the ‘anatomists’ increasingly gained knowledge of pathology. With the invention of the microscope, improved organisation and the distribution of that knowledge of the many pathological observations and drawings through publishing, this is often coined as the start of modern pathology [30]. However, after centuries, a new era has been slowly dawning: that of the digital pathology [31]. The traditional glass slides are currently generally scanned, giving rise to digitalized whole slide images. Using exponential computing power, artificial intelligence has emerged as an important tool to aid pathologists [32]. Initially a supportive feature, creating annotations or detailed assessments, saving and sharing analyses, nowadays the new era pertains to the development of more advanced algorithms gaining more and novel information from slides than with mere microscopy, and performing increasingly complicated tasks, for instance with deep learning algorithms. Deep convolutional networks, a large network of connected ‘neurons’ creating a computerized brain-like model, can for instance be trained on pathologists’ annotations in order to predict tissue types on unannotated whole slide images [33]. This supervised deep learning is increasingly becoming more accurate, and enables prediction of patient-related outcomes as well as implementation of other biomarkers, such as the TSR, which our research group analysed previously [34].

However, to adequately train supervised learning models, time-consuming and extensive annotations by pathologists are required. Moreover, the results are limited to the prespecified outcomes, such as tissue types. Contrastingly, self-supervised learning merely necessitates unannotated whole slide images as input [35]. In the last scientific chapter of this thesis, **Chapter 6**, a self-supervised learning algorithm, specifically the Barlow Twins encoder, is trained to automatically extract histological patterns from group small image patches (tiles) of the whole slide images from colon carcinoma patients [36]. These

tiles are grouped based on the similarity of the extracted features in so-called histomorphological phenotype clusters (HPCs). An independent clinical trial cohort was then used to test their reproducibility. Ultimately, a total of 47 unique HPCs were identified. Each HPC was examined in detail, histopathologically as well as with an immune landscape and gene set enrichment analysis, and associated to overall survival and response to treatment therein. Of note, traits pertaining to tissue type, tissue quantity and tissue architecture were independently highlighted by the model as clinically significant. Moreover, survival predictions with this HPC-based risk grouping were state-of-art. The resulting atlas even uncovered some HPCs identifying stage II/III colon carcinoma patients who could potentially have benefited from the experimentally tested and controversially detrimental deemed bevacizumab [37]. Lastly, this unbiased self-supervised learning method specifically emphasised the importance of the tumour stromal compartment. Future HPC-assisted analyses could aid decision-making in multidisciplinary team meetings for an improved personalized colon carcinoma patient treatment.

Summarizing, this thesis confirms, from bench to bedside, the clinical importance of the TSR as parameter. Throughout these chapters, unequivocal evidence is collected, in line with previous literature. While this evidence in our current, modern, evidence-based health care is crucial, implementation of a biomarker in guidelines is not as straightforward as merely presenting this evidence. Part of an evidence-based approach is analysing this evidence and ranking or scoring this in a predefined hierarchical order [38]. However, this order is based on interventional medicine studies, mostly promoting randomised controlled trials, and pathology parameters are often inapplicable and poorly representable therein [39]. Hence, aiming to find gaps in the current World Health Organisation Classification of Tumours (WCT) [40], which is deemed the absolute fundament of all evidence and containing all carcinoma classifications, recently, a large Delphi study was performed amongst pathologists and other associated experts worldwide [41]. Collecting expertise, opinions, and discussions, this resulted in a consensus for a modified hierarchy of these levels of evidence for pathological parameters pertaining tumours, ranking evidence in level P1 (highest) to P5 (lowest) and specified for the various WCT topics. For the TSR, two are of relevance: those with a research focus on prognostics, and those pertaining to predictive biomarkers.

First, the prognostic biomarker assessment. Following the example research question, ‘What features influence the five-year survival?’, a list of study types is given in the sequence of weighed robustness, e.g. ‘greatest confidence’, to answer this question. Level P1 merely states that a systematic review with predominantly prospective cohort studies or studies derived from randomised-controlled trials, i.e. level P2, would suffice as evidence. For predictive biomarkers, answering a generic ‘can this marker predict treatment?’ research question, also randomised-controlled trials themselves, as an additional level P2 evidence, could be included in the similarly only deemed sufficient P1 systematic review.

Simultaneously, the work group admits that still pathological research is not always includable in these levels of evidence and that establishing this hierarchy was challenging and requires some nuances, also due to some unique considerations per topic. Although the large, international, multicentre UNITED study, would technically be deemed as level P2 evidence, combining all results with many (systematic) reviews and meta-analyses, previously performed by our research group [2] as well as worldwide [42-44], the TSR as prognostic biomarker can sufficiently with great confidence answer the research question, as the stroma-high feature indeed significantly negatively influences the five-year disease-free survival. Thus, implementation of the TSR as prognosticator additional to the TNM classification, as currently is being implored to the UICC, is undeniably pivotal. The TSR as predictive biomarker is still subject to ongoing, high-level research, of which the results will subsequently be presented in an implementation proposal as well.

Interestingly, no mention of a potential existence of expert consensus on the specific prognostic or predictive biomarker, the cost-effectiveness, educational nor distributive aspect of said biomarker is given in the novel modified hierarchy. However, in comparison to the relatively recently implemented tumour budding biomarker [45] or the Immunoscore [46], the TSR surpasses these parameters. Not only is the TSR endorsed by the European Society of Pathology with a consensus on how to score this parameter, with their support assuring the qualitative aspects, an official E-learning was developed to distribute knowledge on this consensus. Previously, our research group had shown that this E-learning could relatively easily instruct pathologists worldwide, experienced and those in training, the TSR scoring method, also on a long-term basis with consistently high interobserver agreement Cohen's kappa's >0.70 [6]. Moreover, the TSR requires no additional, time-consuming nor expensive analyses or staining, but is scored on the same haematoxylin-and-eosin stained routine diagnostic pathology slide used in staging of the local extent of the tumour, i.e. T-stage [7]. From 2007 onward, our research group was the first to describe this parameter, but nowadays, overwhelming evidence from exponential amounts of studies published worldwide has been collected [1].

In conclusion, the TSR is relatively easy, cost-effective, reproducible and valid pathological prognosticator. Stroma-high tumours are more aggressive and less likely to respond to treatments, based on biopsy data, the primary tumour itself, and lymph nodes. The TSR can support patient selection for specific treatment strategies. For example, frail and/or older patients with stroma-high tumours may be offered less or no (neo-)adjuvant therapy, while fitter and/or younger patients may be offered more intensive or experimental treatments. New techniques, such as FAPI PET-CT scans and artificial intelligence, may help map the TSR and other tumour characteristics more quickly and accurately in the future. Moreover, the TSR qualifies for the as abovementioned stated 'robustness' with highest levels of evidence, and although the TSR as predictive biomarker will be proven in the foreseeable high future, the TSR as prognostic biomarker should urgently be implemented in (inter)national guidelines.

References

1. Mesker, W.E., et al., The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Cell Oncol*, 2007. 29(5): p. 387-98.
2. van Pelt, G.W., et al., The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 2018. 73(2): p. 197-206.
3. Huijbers, A., et al., The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*, 2013. 24(1): p. 179-85.
4. Zunder, S.M., et al., Predictive potential of tumour-stroma ratio on benefit from adjuvant bevacizumab in high-risk stage II and stage III colon cancer. *Br J Cancer*, 2018. 119(2): p. 164-169.
5. Smit, M., et al., Uniform Noting for International Application of the Tumor-Stroma Ratio as an Easy Diagnostic Tool: Protocol for a Multicenter Prospective Cohort Study. *JMIR Res Protoc*, 2019. 8(6): p. e13464.
6. Smit, M.A., et al., e-Learning for Instruction and to Improve Reproducibility of Scoring Tumor-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study. *JMIR Form Res*, 2021. 5(3): p. e19408.
7. van Pelt, G.W., et al., Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*, 2018. 473(4): p. 405-412.
8. Polack, M., et al., Results from the UNITED study: a multicenter study validating the prognostic effect of the tumor-stroma ratio in colon cancer. *ESMO Open*, 2024. 9(4): p. 102988.
9. Richtlijndatabase. Colorectaal carcinoom (CRC). 2024 08-04-2024 28-11-2024]; Available from: https://richtlijndatabase.nl/richtlijn/colorectaal_carcinoom_crc/startpagina_-_colorectaal_carcinoom.html.
10. O'Sullivan, B., et al., UICC manual of clinical oncology. Vol. 267. 2015: Wiley Online Library.
11. Loughrey, M.B., et al., Dataset for Pathology Reporting of Colorectal Cancer: Recommendations From the International Collaboration on Cancer Reporting (ICCR). *Ann Surg*, 2022. 275(3): p. e549-e561.
12. Nagtegaal, I., et al., Morphological changes in tumour type after radiotherapy are accompanied by changes in gene expression profile but not in clinical behaviour. *J Pathol*, 2004. 204(2): p. 183-92.
13. Glynne-Jones, R., et al., Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2018. 29(Suppl 4): p. iv263.
14. Strous, M.T.A., et al., Tumour-stroma ratio to predict pathological response to neo-adjuvant treatment in rectal cancer. *Surg Oncol*, 2022. 45: p. 101862.
15. Park, J.H., et al., Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. *J Pathol Clin Res*, 2020. 6(1): p. 30-39.
16. Bahadoer, R.R., et al., Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *The Lancet Oncology*, 2021. 22(1): p. 29-42.
17. Breugom, A.J., et al., Adjuvant chemotherapy for rectal cancer patients treated with preoperative (chemo)radiotherapy and total mesorectal excision: a Dutch Colorectal Cancer Group (DCCG) randomized phase III trial. *Ann Oncol*, 2015. 26(4): p. 696-701.
18. Langbein, T., W.A. Weber, and M. Eiber, Future of Theranostics: An Outlook on Precision Oncology in Nuclear Medicine. *J Nucl Med*, 2019. 60(Suppl 2): p. 13s-19s.
19. Mori, Y., et al., Fibroblast Activation Protein Inhibitor Theranostics: Early Clinical Translation. *PET Clin*, 2023. 18(3): p. 419-428.
20. Joshi, R.S., et al., The Role of Cancer-Associated Fibroblasts in Tumor Progression. *Cancers (Basel)*, 2021. 13(6).
21. Koustoulidou, S., et al., Cancer-Associated Fibroblasts as Players in Cancer Development and Progression and Their Role in Targeted Radionuclide Imaging and Therapy. *Cancers (Basel)*, 2021. 13(5).
22. Loktev, A., et al., Development of Fibroblast Activation Protein-Targeted Radiotracers with Improved Tumor Retention. *J Nucl Med*, 2019. 60(10): p. 1421-1429.
23. Koerber, S.A., et al., The Role of (68)Ga-FAPI PET/CT for Patients with Malignancies of the Lower Gastrointestinal Tract: First Clinical Experience. *J Nucl Med*, 2020. 61(9): p. 1331-1336.
24. Giesel, F.L., et al., (68)Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with Various Cancers. *J Nucl Med*, 2019. 60(3): p. 386-392.
25. Tudyka, V., et al., EURECCA consensus conference highlights about colon & rectal cancer multidisciplinary management: the radiology experts review. *Eur J Surg Oncol*, 2014. 40(4): p. 469-75.
26. Brouwer, N.P.M., et al., Clinical lymph node staging in colorectal cancer; a flip of the coin? *Eur J Surg Oncol*, 2018. 44(8): p. 1241-1246.
27. Polack, M., et al., Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT. *Clin Transl Oncol*, 2022.
28. Chen, H., et al., Usefulness of [(68)Ga]Ga-DOTA-FAPI-04 PET/CT in patients presenting with inconclusive [(18)F]FDG PET/CT findings. *Eur J Nucl Med Mol Imaging*, 2021. 48(1): p. 73-86.
29. Mona, C.E., et al., Correlation of (68)Ga-FAPI-46 PET Biodistribution with FAP Expression by Immunohistochemistry in Patients with Solid Cancers: Interim Analysis of a Prospective Translational Exploratory Study. *J Nucl Med*, 2022. 63(7): p. 1021-1026.
30. van den Tweel, J.G. and C.R. Taylor, A brief history of pathology: Preface to a forthcoming series that highlights milestones in the evolution of pathology as a discipline. *Virchows Arch*, 2010. 457(1): p. 3-10.
31. Pallua, J.D., et al., The future of pathology is digital. *Pathol Res Pract*, 2020. 216(9): p. 153040.
32. Bera, K., et al., Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, 2019. 16(11): p. 703-715.

33. Coudray, N., et al., Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*, 2018. 24(10): p. 1559-1567.
34. Smit, M.A., et al., Deep learning based tumor-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform*, 2023. 14: p. 100191.
35. Claudio Quiros, A., et al., Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unannotated pathology slides. *Nat Commun*, 2024. 15(1): p. 4596.
36. Zbontar, J., Barlow Twins: Self-Supervised Learning via Redundancy Reduction. 2021.
37. de Gramont, A., et al., Bevacizumab plus oxaliplatin-based chemotherapy as adjuvant treatment for colon cancer (AVANT): a phase 3 randomised controlled trial. *Lancet Oncol*, 2012. 13(12): p. 1225-33.
38. Sackett, D.L., et al., Evidence based medicine: what it is and what it isn't. 1996, British Medical Journal Publishing Group. p. 71-72.
39. Indave, B.I., et al., Evidence-levels in pathology for informing the WHO classification of tumours. *Histopathology*, 2022. 81(4).
40. del Aguila Mejia, J., et al., Understanding the use of evidence in the WHO Classification of Tumours: a protocol for an evidence gap map of the classification of tumours of the lung. *BMJ open*, 2022. 12(10): p. e061240.
41. Colling, R., et al., A New Hierarchy of Research Evidence for Tumor Pathology: A Delphi Study to Define Levels of Evidence in Tumor Pathology. *Modern Pathology*, 2024. 37(1): p. 100357.
42. Sullivan, L., et al., Tumor Stroma Ratio and Its Significance in Locally Advanced Colorectal Cancer. *Current Oncology*, 2022. 29(5): p. 3232-3241.
43. Wu, J., et al., Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis. *Oncotarget*, 2016. 7(42): p. 68954-68965.
44. Gao, J., et al., Impact of tumor-stroma ratio on the prognosis of colorectal cancer: a systematic review. *Frontiers in oncology*, 2021. 11: p. 738080.
45. Lugli, A., et al., Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Modern Pathology*, 2017. 30(9): p. 1299-1311.
46. Pagès, F., et al., International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*, 2018. 391(10135): p. 2128-2139.



Chapter 8

Nederlandse samenvatting

Nederlandse samenvatting

De hoeksteen van de huidige behandeling van kanker is gebaseerd op de uitgebreidheid van de ziekte volgens internationale richtlijnen. Voor deze uitgebreidheid wordt wereldwijd de TNM-classificatie gehanteerd, gedefinieerd door lokale ingroei en grootte van de tumor (*Tumour*), of het zich heeft verspreid naar nabijgelegen lymfeklieren (*Node*), en of het eventueel zelfs naar organen op afstand is uitgezaaid (*Metastasis*) [1, 2]. Het voor de patiënt resulterend TNM-stadium wordt meegewogen met andere karakteristieken die invloed hebben op potentiële overleving en succes van behandeling [3-6]. Deze classificatie is echter niet optimaal. Hierdoor wordt er immer onderzoek verricht om de classificatie te verbeteren, zo ook in kanker van de dikke darm en endeldarm (hier samengevoegd in de term ‘darmkanker’) [7]. Darmkanker is de nummer drie meest voorkomende vorm van kanker ter wereld en eist nog steeds veel mensenlevens [8-10]. Bij ongeveer 50% van de patiënten met darmkanker in eenzelfde TNM-stadium zien we terugkeer van de ziekte ondanks de behandeling [3, 5, 11, 12]. Aanvullende tumorkenmerken die het voorspellen van de behandelresultaten voor de individuele patiënt met darmkanker kunnen verbeteren, en zodoende in de toekomst overbehandeling of onderbehandeling verminderen, zijn daarom essentieel.

Biomarkers in tumor stroma

Biologische parameters (biomarkers) van tumoren die het ziektebeloop bij de patiënt kunnen voorspellen zijn inmiddels in de richtlijnen opgenomen [13, 14], zoals bepaalde genetische mutaties, of ingroeïende kleine groepen tumor epitheelcellen (tumor *buds* of tumor *deposits*) [3, 15-17]. Deze biomarkers hebben echter slechts betrekking op de tumor epitheelcellen zelf. Initieel werd gedacht dat een tumor een simpele klomp is met cellen die zich ongeremd hebben kunnen delen. De laatste jaren is in tegenstelling daarvan duidelijk geworden dat de omgeving waar de tumor zich in bevindt, de zogenaamde ‘tumor micro-omgeving’, minstens net zo belangrijk is voor de tumor en ook het gedrag van de tumor beïnvloedt [18-21]. De tumor micro-omgeving, ook wel het tumor stroma genoemd, bestaat uit onder andere immuun cellen, kleine bloedvaten, bindweefsel zoals collageen, en fibroblasten [22-24]. Een fibroblast vormt normaliter bindweefsel en houdt het bijeen (denk aan een genezende wond), maar kan door invloed van een tumor zich ontwikkelen tot een kanker-gerelateerde fibroblast (*cancer-associated fibroblast*, CAF), die een stimulerende werking heeft op de tumor [25, 26]. Het tumor stroma vormt een dynamisch geheel, en is, indien in grote hoeveelheden aanwezig in een tumor, uiteindelijk verantwoordelijk voor een agressievere, ingroeïende tumor met resistentie voor behandelingen zoals radiotherapie en chemotherapie [23, 27-30]. Een biomarker die dat agressieve effect goed weergeeft, is de tumor-stroma ratio (TSR). Hiermee wordt volgens een vast protocol de hoeveelheid van het tumor stroma gescoord, telkens in stappen per tien procent, en daarna gecategoriseerd in stroma-laag ($\leq 50\%$) en stroma-hoog ($> 50\%$) [31, 32]. De TSR, specifiek in darmkanker, is het onderwerp van dit proefschrift.

Al in 2007 werd deze biomarker ontdekt door onze onderzoeksgroep op hematoxyline-en-eosine (HE) gekleurde objectglaasjes (coupes) met darmkankerweefsel, die normaliter routinematig door pathologen worden onderzocht [31]. Na deze eerste publicatie volgden nog vele andere onderzoeken, niet alleen in darmkanker [33] maar ook borstkanker [34], longkanker [35] en slokdarmkanker [36]. Al deze onderzoeken toonden aan dat veel stroma in een tumor (stroma-hoog) tot een kortere overleving leidde en minder succes van behandeling met chemotherapie en/of radiotherapie, dan als er weinig stroma in een tumor was (stroma-laag). Door alle onderzoeken die de TSR bewezen, niet alleen van onze eigen onderzoeksgroep maar inmiddels ook wereldwijd [27, 37, 38], werd destijds een verzoek tot implementeren van de TSR in de richtlijnen gestuurd. Onder andere de TNM-evaluatie commissie (*Union for International Cancer Control*, UICC) en de *College of American Pathologists* (CAP) ontvingen een oproep. Ondanks dat de commissies de waarde van de TSR erkenden, achtten zij eerst nog een grote studie nodig, specifiek één die patiënten in de toekomst vervolgt (prospectief), met een kwaliteitscontrole, vooraleer implementatie overwogen kon worden. Zodoende zette onze onderzoeksgroep in 2017 de UNITED-studie (*Uniform Noting for International application of the Tumour-stroma ratio as Easy Diagnostic tool*) op [39].

TSR in dikke darmkanker

Eerst, ook om het werk van voorgangers af te sluiten, beginnen we in dit proefschrift zodoende met de resultaten van deze langverwachte UNITED-studie (**hoofdstuk 2**). Meer dan 1500 patiënten met TNM-stadium II (grote, lokale tumoren) en III (tumoren die naar de lymfeklieren zijn uitgezaaid) dikke darmkanker werden geïncludeerd. Van deze patiënten werden karakteristieken verzameld van henzelf (zoals leeftijd en geslacht) en de tumor (zoals TNM-stadium en de gescoorde TSR), en vervolgens werden zij gemiddeld 3 jaar gevolgd om te kijken of er een terugkeer van ziekte was (ziektevrije overleving). De TSR werd gescoord door pathologen die waren getraind met een eerder door onze onderzoeksgroep opgezette E-learning [40]. In totaal deden 27 centra uit 12 verschillende landen over de hele wereld mee aan de studie. Naast de ziektevrije overleving werd ook gekeken naar de invloed van de TSR op het slagen van de chemotherapie die sommige patiënten na een operatie krijgen om overgebleven tumorcellen te behandelen (adjuvante chemotherapie, van het Latijn *adjuvare* = helpen) [3]. In patiënten met stroma-hoge dikke darmkanker zagen we een significant kortere ziektevrije overleving ten opzichte van stroma-lage tumoren. Een patiënt met een stroma-hoge tumor had dus een grotere kans op terugkeer van ziekte. Zelfs nadat er werd gecorrigeerd voor andere risicofactoren, zoals leeftijd of uitgebreidheid van de ziekte, bleef de TSR een onafhankelijke voorspellende parameter voor een slechte uitkomst. Tot slot zagen we bij de groep stroma-hoge patiënten een kortere ziektevrije overleving ondanks dat die adjuvante chemotherapie hadden kregen. Mogelijk wijst dit op therapieresistentie van het tumor stroma.

TSR in endeldarmkanker

In hoofdstuk 3 is vervolgens een eerste stap gezet om te kijken naar deze resistentie door de predictieve waarde van de TSR, het voorspellende effect op het wel of niet slagen van behandeling, verder te bewijzen. Dit hebben we onderzocht in uitsluitend endeldarmkanker. Ondanks dat tumoren van de dikke darm en de endeldarm vaak samen worden genoemd, zijn het eigenlijk biologisch gezien verschillende ziektes. Daardoor, alsmede door de andere locatie in het lichaam, ondergaan ze verschillende behandelingen [6, 41]. Endeldarmkanker wordt in opzet orgaan-sparend behandeld en hier zijn afgelopen jaren grote stappen in gezet [6, 42]. Neoadjuvante therapie (van het Latijn *neo* = nieuw; *adjuvare* = helpen) [43] is tegenwoordig de gouden standaard in endeldarmkanker en wordt gegeven aan patiënten vóór een operatie, omdat het, samen met een verbetering in operatietechniek (de totale mesorectale excisie) de kans op terugkeer van de tumor verkleint [44, 45]. Er werd zelfs door combinaties van chemotherapie en radiotherapie steeds vaker gezien dat de tumor soms helemaal verdween, waarbij een zware operatie mogelijk niet meer nodig was: de zogenaamde ‘complete respons’ [42, 46, 47]. De response op neoadjuvante therapie is alleen vaak heterogeen. Bovendien kwam de respons zoals werd gezien op beeldvorming (‘klinische complete respons’), niet altijd overeen met de respons in het weefsel (‘pathologische complete respons’) [48-51]. Daarom is het belangrijk de selectie te verbeteren van zowel patiënten die waarschijnlijk baat hebben bij neoadjuvante therapie, als patiënten die er waarschijnlijk geen baat bij hebben.

Eerder werd in studies gezien dat tumoren met veel stroma niet alleen een kortere overleving hebben, maar ook minder reageren op behandelingen [36, 52, 53]. Neoadjuvante behandeling tast echter de tumor en tumor micro-omgeving aan, waarna bijvoorbeeld verlittekening (fibrosering: een toename van fibroblasten) optreedt [54-56]. Het scoren van de TSR is op behandeld weefsel daarom niet betrouwbaar meer. Biopten vormen hierbij een uitkomst. Dit zijn kleine ‘hapjes’ weefsel die zijn afgenomen tijdens een coloscopie op het moment dat er een verdacht gezwel in de darm zit [7]. De biopten zijn van dan nog onbehandeld weefsel en dus geschikt voor een TSR bepaling. Eerder werd al gezien dat de TSR van deze biopten overeenkomt met de TSR van de daadwerkelijke tumor zelf (de primaire tumor). Bovendien bleek in de studies ook dat stroma-hoge tumoren minder reageerden op behandeling, en dus dat de TSR ook een predictieve biomarker kan zijn [57, 58]. Dit waren alleen vaak kleine studies die in slechts een enkel ziekenhuis was uitgevoerd [59-61]. Om een grote studie met data van verschillende ziekenhuizen (multicenter) te verkrijgen, en daarmee een hogere bewijskracht voor de TSR, hebben we in hoofdstuk 3 twee grote klinische studies (RAPIDO [48] en PROCTOR-SCRIPT [62]) met onze eigen LUMC cohort gecombineerd.

De hypothese van onze grote multicenter onderzoek was dat neoadjuvante behandeling bij patiënten met een stroma-hoog biop van endeldarmkanker minder vaak zou slagen. We scoorden eerst de TSR

op de bipten [57, 60]. Vervolgens werd de respons op celniveau (pathologische respons) na de neoadjuvante behandeling [63] van stroma-hoge tumoren vergeleken met die van stroma-lage tumoren. Niet alleen werd er inderdaad minder vaak een grote respons gezien in stroma-hoge tumoren, dit effect was ook onafhankelijk van andere risicofactoren. TSR was alleen niet direct één-op-één voorspellend voor de ziektevrje en algehele overleving. Er bleken daarvoor te veel andere risicofactoren te zijn die de overleving mede beïnvloedden, zoals tumorgrootte en stadium van de ziekte. Daarnaast bleken de tumoren die groter waren of in de omringende weefsels ingroeien ('lokaal geavanceerd') ook vaker stroma-hoog, wat weer illustratief is voor de mate van agressiviteit van tumor stroma. Dit onderzoek bewijst dus de predictieve waarde van de TSR, al is dit in een retrospectieve manier met achteraf geïnccludeerde patiënten en alleen in specifiek endeldarmkanker onderzocht [64]. Inmiddels loopt er door nieuwe samenwerkingen een grote prospectieve vervolgstudie om de predictieve waarde van de TSR ook in patiënten met dikke darmkanker te bewijzen: de UNITED-II.

Tumor stroma op celniveau en nieuwe doelgerichte beeldvorming

De eerdergenoemde CAFs zijn een groot onderdeel van tumor stroma. CAFs passen hun omgeving actief aan (remodelleren) en zijn daarmee een belangrijke invloed op tumor agressiviteit, ook op het uitzaaingsvermogen naar lymfeklieren. Aanwezigheid van tumorcellen in de lymfeklieren geeft aan hoe uitgebreid de ziekte is (N-stadium) en of aanvullende behandeling nodig is. Met huidige beeldvorming (met CT of MRI-scans) blijft het echter lastig om accuraat te voorspellen welke lymfeklieren verdacht zijn voor het bevatten van een uitzaaing. De PET-CT scan is een redelijk nieuwe soort beeldvorming waarbij door een radioactief gelabelde stof (een tracer: meestal fluorodeoxyglucose, FDG) de activiteit en locatie van cellen in het lichaam wordt weergegeven. De ongeremde celdeling in kanker leidt tot een hoge celactiviteit en laat dus in principe een sterkere opname zien op de scan, ook in lymfeklieren [3, 65, 66]. Door hun tumor-specificiteit en betrokkenheid bij uitzaaïen zijn CAFs tegenwoordig vaak het onderwerp van studies, ook voor PET-CT scans [25, 26, 67]. De door hen het meest tot expressie gebracht eiwit is de *fibroblast activation protein* (fibroblast activatie eiwit, FAP) [68-70]. Een veelbelovende PET-CT tracer die recent hiervoor ontwikkeld is, is de *FAP-inhibitor* (FAP-remmer, FAPI). Hiermee wordt FAP niet alleen gebonden en op de scan weergegeven, maar zou men FAP, en daarmee CAFs en tumor activiteit, mogelijk zelfs kunnen remmen. De FAPI-tracer kan in theorie dus helpen met de diagnostiek, door de uitgebreidheid van ziekte, de tumorgrootte en ook met name de verspreiding naar lymfeklieren met meer accuratesse te bepalen [71-75]. Hierdoor zou in de toekomst een betere selectie gemaakt kunnen worden van patiënten voor bepaalde behandelstrategieën. Ondanks dat er veel nieuwe FAPI-studies in patiënten zijn, is er nog weinig onderzoek gedaan naar de theoretische grondslag op celniveau hiervan [76, 77], daarom hebben wij dat hier in dit proefschrift getracht.

Eerst hebben wij in **hoofdstuk 4** gekeken naar de lymfeklieren en de uitzaaiingen daarin, en specifiek hoeveel tumor stroma die bevatten. Onze onderzoeksgroep had al eerder gezien dat de TSR in deze lymfeklier uitzaaiingen ook de overleving van patiënten met dikke darmkanker beïnvloedt [78], en in dit hoofdstuk geven we er een verder algemeen overzicht over. Met huidige beeldvorming wordt een verdachte lymfeklier bepaald op basis van diameter (>5 -10mm doorsnede) en bijvoorbeeld rondheid van de lymfeklier, maar dit is nog vaak aan fouten onderhevig [79, 80]. Niet alleen zagen we dat een grotere lymfeklier inderdaad een grotere kans had op het bevatten van een (grotere) uitzaaiing, maar belangrijker, dat ook juist kleine lymfeklieren <5 mm in doorsnede ook een stromale uitzaaiing kon bevatten [81]. Zo bleek dat de FAPI-tracer in theorie potentie had om beter lymfeklier uitzaaiingen te kunnen opsporen [82]. In **hoofdstuk 5** analyseerden we daarom vervolgens het weefsel van de primaire tumor, bipten en lymfeklieren in darmkanker. We vergeleken coupes gekleurd met de standaard HE-kleuring, met coupes waarbij specifiek voor dat FAP-eiwit gekleurd is (immunohistochemische kleuring).

Door de patronen van FAP-kleuringen op deze coupes te beschrijven, konden we inzicht krijgen of er inderdaad een biologische basis is voor de FAPI-tracer [76]. Daarnaast keken we naar de correlatie tussen de TSR en de sterkte van de aankleuring van het FAP-eiwit. Ondanks dat in ons onderzoek darmkanker meestal aankleurde voor FAP, bleek het vaak heterogeen. Er waren niet alleen verschillen in sterkte en patroon tussen tumoren, maar ook binnen dezelfde tumor van eenzelfde patiënt. Daarnaast zagen we dat niet tumor-bevattend weefsel soms kon aankleuren voor FAP, zoals weefsel dat is behandeld door neoadjuvante therapie en een grote respons heeft doorgemaakt. Mogelijk kleurde dat aan door de fibrosering. Ook gezonde actieve lymfeklieren kleurden sporadisch aan voor FAP. Bovendien zagen we dat stroma-hoge tumoren niet altijd een sterkere aankleuring voor FAP hadden, waarschijnlijk door verschil in locatie van het scoren van de twee parameters in de tumor. Concluderend dient men voorzichtig te zijn met direct vertalen van aankleuring van FAP (en in de toekomst dus FAPI opname) naar activiteit en uitgebreidheid van ziekte, en dient er meer onderzoek verricht te worden [76].

TSR en kunstmatige intelligentie

Al sinds halverwege de negentiende eeuw is de microscoop de vertrouwde techniek en gouden standaard voor pathologen [83, 84]. Na al deze jaren is er echter een nieuw tijdperk op komst: dat van de digitale pathologie [85]. Traditionele objectglaasjes kunnen tegenwoordig worden ingescand waardoor een digitaal plaatje van de gehele coupe wordt gevormd (*whole slide image*, WSI). Deze WSIs kunnen dan niet alleen bekeken worden, maar inmiddels ook door steeds krachtigere computers met kunstmatige intelligentie worden geanalyseerd [86-88]. Waar het begon om de patholoog te ondersteunen, van automatisch simpele aantekeningen op coupes (annotaties) laten zetten, is door zogeheten *deep learning* het mogelijk geworden om patronen op de WSIs te leren herkennen en

voorspellen, en zelfs de samenhang met patiënt-gerelateerde uitkomsten daarvan te analyseren [89, 90]. Deep learning maakt gebruik van neurale netwerken, als een soort brein, waardoor het niet altijd duidelijk is hoe de computer tot een conclusie komt, als een soort ‘*black box*’. In **hoofdstuk 6** pogen we dit fenomeen te verhelderen.

Eerder heeft onze onderzoeksgroep met deep learning getracht de TSR automatisch te bepalen op WSIs [91]. Het trainen vergde echter veel tijd en werk, met name door de benodigde extensieve annotaties van pathologen. In ons onderzoek nu hanteerden we een revolutionaire aanpak, waardoor we WSIs konden gebruiken zonder dat er annotaties op stonden. Deze zelfgestuurde vorm van deep learning heet ook wel *self-supervised learning*, en is compleet onafhankelijk van kennis en input van pathologen [92]. We verdeelden eerst de grote WSIs in kleine beeldstukjes (‘tegels’), die door ons self-supervised deep learning model (specifiek het Barlow-Twins netwerkmodel [93]) werd gegroepeerd (geclusterd) op basis van kenmerken op celniveau (histologie). Dit resulteerde in 47 unieke groepen, *histomorphological phenotype clusters* (HPCs) genoemd. Elk HPC werd in detail onderzocht, op zowel tumorcel eigenschappen, als immuunsysteem en genetische eigenschappen. Vervolgens werd de verhouding en samenhang onderzocht tussen de verschillende HPCs, en werd dit gecorreleerd met algehele overleving en respons op therapie. Opmerkelijk was dat het geheel onafhankelijke deep learning model zelfstandig weefsel types (tumor, spier, slijm) en hoeveelheden (veel/weinig) van elkaar onderscheidde en per soort groepeerde. Nog opmerkelijker was dat tumor stroma ook als aparte entiteit werd gezien, zonder dat wij daar invloed op hadden gehad. Daarnaast bleek dat hogere hoeveelheden tumor stroma ook leidden tot een slechtere uitkomst. Op basis van de door ons gevormde HPCs konden wij de overleving van patiënten voorspellen, accurater dan wat er wereldwijd tot dan toe gepubliceerd was. Tevens konden we voorspellen welke patiënten beter op bepaalde gegeven behandelingen zouden reageren. In de toekomst zouden deze HPC-analyses het stellen van de diagnose middels pathologie kunnen verbeteren, en eventueel zelfs kunnen helpen met het selecteren van een behandelstrategie [94].

Toekomstperspectieven en TSR implementeren

Dit proefschrift bevestigt het belang van tumor stroma en specifiek de TSR, van het laboratorium tot aan het bed van patiënten. Door de hoofdstukken bladerend kan een ruim overzicht aan bewijsmateriaal gevonden worden, aansluitend op bestaande literatuur, en specifiek in onderzoeken en proefschriften van voorgangers. Ondanks dat dit overzicht belangrijk is, is het implementeren van een nieuwe biomarker in richtlijnen niet zo eenvoudig als het louter presenteren daarvan. In de discussie van dit proefschrift wordt hierbij stilgestaan. De *evidence-based* (op onderzoeksbewijs-gebaseerde) werkwijze, die tegenwoordig wordt gehanteerd in de geneeskunde om richtlijnen op te stellen, is gebaseerd op het systematisch analyseren van al het bewijsmateriaal en dit rangschikken in een vooraf gedefinieerde, hiërarchische volgorde [95]. Onderzoeken naar tumor kenmerken in het veld van de pathologie lenen zich hier echter niet goed voor [96]. In een poging dit te verbeteren werd recent een grote Delphi-studie

verricht onder pathologen en andere deskundigen. Dit is een studie, die door het verzamelen van kennis, meningen en ervaringen, en na het verrichten van grote discussies, resulteert in een overeenstemming (consensus). Deze Delphi-studie leidde tot een gewijzigde hiërarchische volgorde voor pathologische parameters van tumoren [97]. Ondanks dat de UNITED-studie officieel volgens deze nieuwe hiërarchie wellicht niet het hoogste niveau van bewijs geeft, werd door de werkgroep ook al aangegeven dat deze nuance vereist, ook vanwege enkele unieke aspecten per onderwerp en biomarker.

De UNITED-studie bewijst namelijk met hoogste betrouwbaarheid het prognostische effect van de TSR: het is een onafhankelijke factor die voorspelt dat patiënten met stroma-hoge dikke darmtumoren een kortere ziektevrije overleving hebben. Ook resultaten van eerdere onderzoeken, onder andere van (systematische) reviews en andere meta-analysen, van zowel onze onderzoeksgroep [98] als wereldwijd [99-101], worden meegenomen in de UNITED-studie. Zodoende hebben wij weer een herhaalde oproep verstuurd om de TSR te implementeren aan de TNM-evaluatiecommissie van de UICC. De commissie feliciteerde ons met de behaalde resultaten en zegden toe de TSR daadwerkelijk in de volgende aankomende TNM-classificatie in de lijst te zetten van belangrijke tumorkenmerken die gescoord moeten worden in het pathologisch diagnostisch proces. De UICC suggereerde ook om de resultaten te delen met de *International Collaboration on Cancer Reporting* (ICCR) om de TSR voortaan standaard in pathologie verslagen te laten implementeren [84]. De ICCR gaf ook toe dat de TSR de overweging verdiende voor de volgende versie. De Nederlandse richtlijnencommissie [102], waar de resultaten tot slot naar verstuurd werden, erkende de prognostische waarde van deze biomarker en beloofde de TSR als zodanig toe te voegen in de lijst met optioneel te bepalen tumorkenmerken door de patholoog. Voor de predictieve waarde was er volgens hen eerst alleen nog een specifieke grote studie nodig die dit zou bewijzen. Inmiddels is daarom de UNITED-II begonnen.

Conclusie

De TSR is een belangrijke prognostische factor in darmkanker. Stroma-hoge tumoren zijn agressiever en reageren minder vaak op behandelingen, zowel gescoord op bipten, de tumor zelf, als de lymfeklieren. De TSR kan daardoor de selectie van patiënten voor bepaalde behandelstrategieën ondersteunen. Zo kan bij kwetsbare en/of oudere stroma-hoge patiënten gekozen worden om minder of geen (neo-)adjuvante therapie te geven, en bij fittere en/of jongere patiënten juist intensieve of experimentele behandelingen. Nieuwe technieken, zoals FAPI PET-CT scans en kunstmatige intelligentie, kunnen in de toekomst tot slot helpen om de TSR en andere tumorkenmerken sneller en nauwkeuriger in kaart te brengen. De TSR is niet alleen door onze onderzoeksgroep bewezen, maar inmiddels door overweldigende hoeveelheden gepubliceerde onderzoeken wereldwijd [31]. Dit proefschrift roept daarom op tot het implementeren van de TSR in internationale richtlijnen, waar de eerste stappen voor zijn gezet.

Referenties

1. Brierley, J.D., Gospodarowicz, M.K., Wittekind, C., *The TNM Classification of Malignant Tumours*. 8 ed. 2016: Wiley Blackwell.
2. Weiser, M.R., *AJCC 8th Edition: Colorectal Cancer*. *Ann Surg Oncol*, 2018. 25(6): p. 1454–1455.
3. Argiles, G., et al., Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2020. 31(10): p. 1291–1305.
4. Cervantes, A., et al., Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*, 2023. 34(1): p. 10–32.
5. Baxter, N.N., et al., Adjuvant Therapy for Stage II Colon Cancer: ASCO Guideline Update. *J Clin Oncol*, 2022. 40(8): p. 892–910.
6. Glynne-Jones, R., et al., Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*, 2018. 29(Suppl 4): p. iv263.
7. Weitz, J., et al., Colorectal cancer. *Lancet*, 2005. 365(9454): p. 153–65.
8. Sung, H., et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2021.
9. Siegel, R.L., et al., Cancer statistics, 2022. *CA Cancer J Clin*, 2022. 72(1): p. 7–33.
10. Bray, F., et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024. 74(3): p. 229–263.
11. Kim, H.S., et al., Clinicopathological and biomolecular characteristics of stage IIB/IIC and stage IIIA colon cancer: Insight into the survival paradox. *J Surg Oncol*, 2019. 120(3): p. 423–430.
12. Li, H., et al., Re-Evaluation of the Survival Paradox Between Stage IIB/IIC and Stage IIIA Colon Cancer. *Front Oncol*, 2020. 10: p. 595107.
13. Chuang, J.P., et al., Comprehensive Review of Biomarkers for the Treatment of Locally Advanced Colon Cancer. *Cells*, 2022. 11(23).
14. Chen, K., et al., Pathological Features and Prognostication in Colorectal Cancer. *Curr Oncol*, 2021. 28(6): p. 5356–5383.
15. Lugli, A., et al., Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod Pathol*, 2017. 30(9): p. 1299–1311.
16. Roth, A.D., et al., Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *J Natl Cancer Inst*, 2012. 104(21): p. 1635–46.
17. Ueno, H., et al., Optimal colorectal cancer staging criteria in TNM classification. *J Clin Oncol*, 2012. 30(13): p. 1519–26.
18. Fidler, I.J., Seed and soil revisited: contribution of the organ microenvironment to cancer metastasis. *Surg Oncol Clin N Am*, 2001. 10(2): p. 257–69, vii–viii.
19. Ouahoud, S., et al., Bidirectional tumor/stroma crosstalk promotes metastasis in mesenchymal colorectal cancer. *Oncogene*, 2020. 39(12): p. 2453–2466.
20. Sudhakar, M., H. Vignesh, and K.N. Natarajan, Crosstalk between tumor and microenvironment: Insights from spatial transcriptomics. *Adv Cancer Res*, 2024. 163: p. 187–222.
21. Werb, Z. and P. Lu, The Role of Stroma in Tumor Development. *Cancer J*, 2015. 21(4): p. 250–3.
22. Balkwill, F.R., M. Capasso, and T. Hagemann, The tumor microenvironment at a glance. *J Cell Sci*, 2012. 125(Pt 23): p. 5591–6.
23. Bremnes, R.M., et al., The role of tumor stroma in cancer progression and prognosis: emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *J Thorac Oncol*, 2011. 6(1): p. 209–17.
24. Sugai, T., et al., Microenvironmental markers are correlated with lymph node metastasis in invasive submucosal colorectal cancer. *Histopathology*, 2021. 79(4): p. 584–598.
25. Chen, Y., K.M. McAndrews, and R. Kalluri, Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*, 2021. 18(12): p. 792–804.
26. Kobayashi, H., et al., Cancer-associated fibroblasts in gastrointestinal cancer. *Nat Rev Gastroenterol Hepatol*, 2019. 16(5): p. 282–295.
27. Park, J.H., et al., The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. *Ann Oncol*, 2014. 25(3): p. 644–651.
28. Tsujino, T., et al., Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clin Cancer Res*, 2007. 13(7): p. 2082–90.
29. Zhuyan, J., et al., Critical steps to tumor metastasis: alterations of tumor microenvironment and extracellular matrix in the formation of pre-metastatic and metastatic niche. *Cell Biosci*, 2020. 10: p. 89.
30. Strous, M.T.A., et al., A high tumour-stroma ratio (TSR) in colon tumours and its metastatic lymph nodes predicts poor cancer-free survival and chemo resistance. *Clin Transl Oncol*, 2022. 24(6): p. 1047–1058.
31. Mesker, W.E., et al., The carcinoma-stromal ratio of colon carcinoma is an independent factor for survival compared to lymph node status and tumor stage. *Cell Oncol*, 2007. 29(5): p. 387–98.
32. van Pelt, G.W., et al., Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch*, 2018. 473(4): p. 405–412.
33. Huijbers, A., et al., The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol*, 2013. 24(1): p. 179–85.
34. Vangangel, K.M.H., et al., The prognostic value of the tumor-stroma ratio is most discriminative in patients with grade III or triple-negative breast cancer. *Int J Cancer*, 2020. 146(8): p. 2296–2304.

35. Smit, M.A., et al., The prognostic value of the tumor-stroma ratio in squamous cell lung cancer, a cohort study. *Cancer Treat Res Commun*, 2020. 25: p. 100247.
36. van Pelt, G.W., et al., The value of tumor-stroma ratio as predictor of pathologic response after neoadjuvant chemoradiotherapy in esophageal cancer. *Clin Transl Radiat Oncol*, 2020. 20: p. 39–44.
37. Sullivan, L., et al., Tumor Stroma Ratio and Its Significance in Locally Advanced Colorectal Cancer. *Curr Oncol*, 2022. 29(5): p. 3232–3241.
38. Souza da Silva, R.M., et al., Standardized assessment of the tumor-stroma ratio in colorectal cancer: interobserver validation and reproducibility of a potential prognostic factor. *Clinical pathology*, 2021. 14: p. 2632010X21989686.
39. Smit, M., et al., Uniform Noting for International Application of the Tumor-Stroma Ratio as an Easy Diagnostic Tool: Protocol for a Multicenter Prospective Cohort Study. *JMIR Res Protoc*, 2019. 8(6): p. e13464.
40. Smit, M.A., et al., e-Learning for Instruction and to Improve Reproducibility of Scoring Tumor-Stroma Ratio in Colon Carcinoma: Performance and Reproducibility Assessment in the UNITED Study. *JMIR Form Res*, 2021. 5(3): p. e19408.
41. Wei, E.K., et al., Comparison of risk factors for colon and rectal cancer. *International journal of cancer*, 2004. 108(3): p. 433–442.
42. Beets, G.L., et al., A new paradigm for rectal cancer: Organ preservation: Introducing the International Watch & Wait Database (IWWD). *European Journal of Surgical Oncology (EJSO)*, 2015. 41(12): p. 1562–1564.
43. Wood, W.C., Neoadjuvant chemotherapy, in *Adjuvant Therapy of Breast Cancer*, I.C. Henderson, Editor. 1992, Springer US: Boston, MA. p. 279–291.
44. Peeters, K.C., et al., The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma. *Ann Surg*, 2007. 246(5): p. 693–701.
45. Kapiteijn, E., et al., Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *New England Journal of Medicine*, 2001. 345(9): p. 638–646.
46. Garcia-Aguilar, J., et al., Organ Preservation in Patients With Rectal Adenocarcinoma Treated With Total Neoadjuvant Therapy. *J Clin Oncol*, 2022. 40(23): p. 2546–2556.
47. Custers, P.A., et al., Long-term Quality of Life and Functional Outcome of Patients With Rectal Cancer Following a Watch-and-Wait Approach. *JAMA Surgery*, 2023. 158(5): p. e230146–e230146.
48. Bahadoer, R.R., et al., Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *The Lancet Oncology*, 2021. 22(1): p. 29–42.
49. Lambregts, D.M.J., T.N. Boellaard, and R.G.H. Beets-Tan, Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights Imaging*, 2019. 10(1): p. 15.
50. Temmink, S.J., et al., Complete response rates in rectal cancer: Temporal changes over a decade in a population-based nationwide cohort. *European Journal of Surgical Oncology*, 2023. 49(11): p. 106991.
51. Nahas, S.C., et al., Diagnostic performance of magnetic resonance to assess treatment response after neoadjuvant therapy in patients with locally advanced rectal cancer. *Abdom Radiol (NY)*, 2019. 44(11): p. 3632–3640.
52. Hagenaaers, S.C., et al., Tumor-stroma ratio is associated with Miller-Payne score and pathological response to neoadjuvant chemotherapy in HER2-negative early breast cancer. *Int J Cancer*, 2021.
53. Hansen, T.F., et al., Tumor-stroma ratio predicts recurrence in patients with colon cancer treated with neoadjuvant chemotherapy. *Acta Oncol*, 2018. 57(4): p. 528–533.
54. Nagtegaal, I., et al., Morphological changes in tumour type after radiotherapy are accompanied by changes in gene expression profile but not in clinical behaviour. *J Pathol*, 2004. 204(2): p. 183–92.
55. Stone, H.B., et al., Effects of radiation on normal tissue: consequences and mechanisms. *Lancet Oncol*, 2003. 4(9): p. 529–36.
56. Miller, K.E., J.L. Ostrowski, and C.M. Quinn, Effects of chemotherapy on breast cancer tissue. *Histopathology*, 1997. 30(4): p. 397–8.
57. Park, J.H., et al., Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. *J Pathol Clin Res*, 2020. 6(1): p. 30–39.
58. Courrech Staal, E.F., et al., Reproducibility and validation of tumour stroma ratio scoring on oesophageal adenocarcinoma biopsies. *Eur J Cancer*, 2011. 47(3): p. 375–82.
59. Scheer, R., et al., Tumor-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: A retrospective cohort study. *World J Gastrointest Oncol*, 2017. 9(12): p. 466–474.
60. Strous, M.T.A., et al., Tumour-stroma ratio to predict pathological response to neo-adjuvant treatment in rectal cancer. *Surg Oncol*, 2022. 45: p. 101862.
61. Zhu, Y., et al., Prognostic Value of Tumor-Stroma Ratio in Rectal Cancer: A Systematic Review and Meta-analysis. *Front Oncol*, 2021. 11: p. 685570.
62. Breugom, A.J., et al., Adjuvant chemotherapy for rectal cancer patients treated with preoperative (chemo)radiotherapy and total mesorectal excision: a Dutch Colorectal Cancer Group (DCCG) randomized phase III trial. *Ann Oncol*, 2015. 26(4): p. 696–701.
63. Mandard, A.M., et al., Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. *Clinicopathologic correlations*. *Cancer*, 1994. 73(11): p. 2680–6.
64. Polack, M., et al., The tumour–stroma ratio as predictive aid towards a biopsy-based treatment strategy in rectal carcinoma. *Histopathology*, 2025.
65. Godefroy, J., et al., Perceptual omission errors in Positron emission tomography/Computed tomography reporting. *Q J Nucl Med Mol Imaging*, 2021.
66. Lu, Y.Y., et al., A systematic review and meta-analysis of pretherapeutic lymph node staging of colorectal cancer by 18F-FDG PET or PET/CT. *Nucl Med Commun*, 2012. 33(11): p. 1127–33.

67. Joshi, R.S., et al., The Role of Cancer-Associated Fibroblasts in Tumor Progression. *Cancers (Basel)*, 2021. 13(6).
68. Henrich, L.M., et al., The Impact of Cancer-Associated Fibroblasts on the Biology and Progression of Colorectal Carcinomas. *Genes (Basel)*, 2024. 15(2).
69. Liu, F., et al., Fibroblast activation protein overexpression and clinical implications in solid tumors: a meta-analysis. *PLoS One*, 2015. 10(3): p. e0116683.
70. Janani, M., et al., Association of future cancer metastases with fibroblast activation protein- α : a systematic review and meta-analysis. *Front Oncol*, 2024. 14: p. 1339050.
71. Koerber, S.A., et al., The Role of (68)Ga-FAPI PET/CT for Patients with Malignancies of the Lower Gastrointestinal Tract: First Clinical Experience. *J Nucl Med*, 2020. 61(9): p. 1331–1336.
72. Giesel, F.L., et al., (68)Ga-FAPI PET/CT: Biodistribution and Preliminary Dosimetry Estimate of 2 DOTA-Containing FAP-Targeting Agents in Patients with Various Cancers. *J Nucl Med*, 2019. 60(3): p. 386–392.
73. Kratochwil, C., et al., (68)Ga-FAPI PET/CT: Tracer Uptake in 28 Different Kinds of Cancer. *J Nucl Med*, 2019. 60(6): p. 801–805.
74. Mori, Y., et al., FAPI PET: Fibroblast Activation Protein Inhibitor Use in Oncologic and Nononcologic Disease. *Radiology*, 2023. 306(2): p. e220749.
75. Dendl, K., et al., The Role of Fibroblast Activation Protein Ligands in Oncologic PET Imaging. *PET Clin*, 2021. 16(3): p. 341–351.
76. Mona, C.E., et al., Correlation of (68)Ga-FAPI-46 PET Biodistribution with FAP Expression by Immunohistochemistry in Patients with Solid Cancers: Interim Analysis of a Prospective Translational Exploratory Study. *J Nucl Med*, 2022. 63(7): p. 1021–1026.
77. Strating, E., et al., Fibroblast activation protein identifies Consensus Molecular Subtype 4 in colorectal cancer and allows its detection by (68)Ga-FAPI-PET imaging. *Br J Cancer*, 2022. 127(1): p. 145–155.
78. van Pelt, G.W., et al., Stroma-high lymph node involvement predicts poor survival more accurately for patients with stage III colon cancer. *J Med Surg Pathol*, 2016. 1(02).
79. Tudyka, V., et al., EURECCA consensus conference highlights about colon & rectal cancer multidisciplinary management: the radiology experts review. *Eur J Surg Oncol*, 2014. 40(4): p. 469–75.
80. Brouwer, N.P.M., et al., Clinical lymph node staging in colorectal cancer; a flip of the coin? *Eur J Surg Oncol*, 2018. 44(8): p. 1241–1246.
81. Polack, M., et al., Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT. *Clin Transl Oncol*, 2022.
82. Chen, H., et al., Usefulness of [(68)Ga]Ga-DOTA-FAPI-04 PET/CT in patients presenting with inconclusive [(18)F]FDG PET/CT findings. *Eur J Nucl Med Mol Imaging*, 2021. 48(1): p. 73–86.
83. van den Tweel, J.G. and C.R. Taylor, A brief history of pathology: Preface to a forthcoming series that highlights milestones in the evolution of pathology as a discipline. *Virchows Arch*, 2010. 457(1): p. 3–10.
84. Loughrey, M.B., et al., Dataset for Pathology Reporting of Colorectal Cancer: Recommendations From the International Collaboration on Cancer Reporting (ICCR). *Ann Surg*, 2022. 275(3): p. e549–e561.
85. Pallua, J.D., et al., The future of pathology is digital. *Pathol Res Pract*, 2020. 216(9): p. 153040.
86. Hinton, G.E., 20 - CONNECTIONIST LEARNING PROCEDURES11This chapter appeared in Volume 40 of *Artificial Intelligence in 1989*, reprinted with permission of North-Holland Publishing. It is a revised version of Technical Report CMU-CS-87-115, which has the same title and was prepared in June 1987 while the author was at Carnegie Mellon University. The research was supported by contract N00014-86-K-00167 from the Office of Naval Research and by grant IST-8520359 from the National Science Foundation, in *Machine Learning*, Y. Kodratoff and R.S. Michalski, Editors. 1990, Morgan Kaufmann: San Francisco (CA). p. 555–610.
87. Hamet, P. and J. Tremblay, Artificial intelligence in medicine. *Metabolism*, 2017. 69s: p. S36–s40.
88. Chen, R.J., et al., Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 2024. 30(3): p. 850–862.
89. Bera, K., et al., Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology*, 2019. 16(11): p. 703–715.
90. Echle, A., et al., Deep learning in cancer pathology: a new generation of clinical biomarkers. *British Journal of Cancer*, 2021. 124(4): p. 686–696.
91. Smit, M.A., et al., Deep learning based tumor-stroma ratio scoring in colon cancer correlates with microscopic assessment. *J Pathol Inform*, 2023. 14: p. 100191.
92. Claudio Quiros, A., et al., Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unannotated pathology slides. *Nat Commun*, 2024. 15(1): p. 4596.
93. Zbontar, J., Barlow Twins: Self-Supervised Learning via Redundancy Reduction. 2021.
94. Liu, B., et al., Self-Supervised Learning Reveals Clinically Relevant Histomorphological Patterns for Therapeutic Strategies in Colon Cancer. *bioRxiv*, 2024: p. 2024.02.26.582106.
95. Sackett, D.L., et al., Evidence based medicine: what it is and what it isn't. 1996, British Medical Journal Publishing Group. p. 71–72.
96. Indave, B.I., et al., Evidence-levels in pathology for informing the WHO classification of tumours. *Histopathology*, 2022. 81(4).
97. Colling, R., et al., A New Hierarchy of Research Evidence for Tumor Pathology: A Delphi Study to Define Levels of Evidence in Tumor Pathology. *Modern Pathology*, 2024. 37(1): p. 100357.
98. van Pelt, G.W., et al., The tumour-stroma ratio in colon cancer: the biological role and its prognostic impact. *Histopathology*, 2018. 73(2): p. 197–206.
99. Sullivan, L., et al., Tumor Stroma Ratio and Its Significance in Locally Advanced Colorectal Cancer. *Current Oncology*, 2022. 29(5): p. 3232–3241.

- Chapter 8

100. Wu, J., et al., Association between tumor-stroma ratio and prognosis in solid tumor patients: a systematic review and meta-analysis. *Oncotarget*, 2016. 7(42): p. 68954–68965.
101. Gao, J., et al., Impact of tumor–stroma ratio on the prognosis of colorectal cancer: a systematic review. *Frontiers in oncology*, 2021. 11: p. 738080.
102. Richtlijndatabase. Colorectal carcinoom (CRC). 2024 08–04–2024 28-11-2024]; Available from: https://richtlijndatabase.nl/richtlijn/colorectaal_carcinoom_crc/startpagina_-_colorectaal_carcinoom.html.





A

Appendices

List of publications

.

Curriculum Vitae

.

Dankwoord

List of publications

Polack M, van Pelt GW, Crobach ASLP, de Geus-Oei LF, Tollenaar RAEM, van Krieken JHJM, Mesker WE. 'Fibroblast activation protein-expression in colorectal carcinomas and implications for clinical application.' *Cancer Treatment and Research Communications*. 2025 [Jul 14;44:100964](#)

Polack M, van Pelt GW, van den Heuvel DH, Klein-Kranenburg EM, Roodvoets AGH, Putter H, Crobach ASLP, Nagtegaal ID, Peeters KCMJ, Tollenaar RAEM, van Krieken JHJM, Mesker WE. 'The tumour-stroma ratio as predictive aid towards a biopsy-based treatment strategy in rectal carcinoma.' *Histopathology*. 2025 [Jul ;87\(1\):44-57](#)

Polack M, Liu B, Coudray N, Claudio Quiros A, Sakellariopoulos T, Le H, Karimkhan A, Crobach ASLP, van Krieken JHJM, Yuan K, Tollenaar RAEM, Mesker WE, Tsirogas A. 'Self-supervised learning reveals clinically relevant histomorphological patterns for therapeutic strategies in colon cancer.' *Nature Communications*. 2025 [Mar 8;16\(1\):2328](#).

Polack M, Smit MA, van Pelt GW, Roodvoets AGH, Meershoek-Klein Kranenburg E, Putter H, Gelderblom H, Crobach ASLP, Terpstra V, Petrushevska G, Gašljević G, Kjær-Frifeldt S, de Cuba EMV, Bulkman NWJ, Vink GR, Al Dieri R, Tollenaar RAEM, van Krieken JHJM, Mesker WE and the UNITED Collaboration. 'Results from the UNITED study: a multicenter study validating the prognostic effect of the tumour-stroma ratio in colon cancer.' *ESMO Open*. 2024 Apr;9(4):102988. [Epub 2024 Apr 12](#).

Kastelein AM, **Polack M**, Vervuurt M. 'PhD regulations in the Netherlands.' *Report by the PhD candidates Network Netherlands*. [2022 Nov 9](#)

Polack M, Hagenaars SC, Couwenberg A, Kool W, Tollenaar RAEM, Vogel WV, Snaebjornsson P, Mesker WE. 'Characteristics of tumour stroma in regional lymph node metastases in colorectal cancer patients: a theoretical framework for future diagnostic imaging with FAPI PET/CT.' *Clinical Translational Oncology*. [2022 Apr 28](#).

Ravensbergen CJ, Kuruc M, **Polack M**, Crobach S, Putter H, Gelderblom H, Roy D, Tollenaar RAEM, Mesker WE. 'The Stroma Liquid Biopsy Panel Contains a Stromal-Epithelial Gene Signature Ratio That Is Associated with the Histologic Tumour-Stroma Ratio and Predicts Survival in Colon Cancer.' *Cancers*. 2022; [14\(1\):163](#)

Anholt R, **Polack M** (*Promovendi Netwerk Nederland*). 'De stille crisis in de wetenschap.' *ScienceGuide*. [Sept 2021](#).

Ravensbergen CJ, **Polack M**, Roelands J, Crobach S, Putter H, Gelderblom H, Tollenaar RAEM, Mesker WE. 'Combined Assessment of the Tumour–Stroma Ratio and Tumour Immune Cell Infiltrate for Immune Checkpoint Inhibitor Therapy Response Prediction in Colon Cancer.' *Cells*. 2021; [10\(11\):2935](#)

Smits AJ, **Polack M**, Deunk J, Bloemers FW. 'Combined anteroposterior fixation using a titanium cage versus solely posterior fixation for traumatic thoracolumbar fractures: A systematic review and meta-analysis.' *Journal of Craniovertebral Junction Spine*. 2017 [Jul-Sep;8\(3\):168-178](#)

Curriculum vitae

Meaghan Polack was born on September 7th, 1994 in Leiden. During her time at the St. Ignatius gymnasium (Amsterdam), she rowed at Rowing Ignatius College (RIC). After graduating in 2012, she enrolled at the Faculty of Medicine of the Vrije Universiteit (VU) Amsterdam. She continued rowing on a high level at the R.S.V.U. 'Okeanos', including Freshmen Women as well as later the Lightweight Women. She furthermore tutored anatomy lessons and during her last year of the Bachelor, she was in the Year Representatives' Board.

She started her Master's in 2015 with her thesis research at the dept. Cardiothoracic Surgery, after which she had chosen to take a lapse year to continue her representatives' work into the Students Council. She took extra courses during that year, i.e. Health Economics and Health Law and Ethics, after which she continued with her Master of Medicine, while doing extra research at the dept. Trauma Surgery and as student member of the Scientific Committee of the Faculty of Medicine of the Vrije Universiteit. She successfully finished her Master as semi doctor at the dept. Surgery in Noordwest Ziekenhuisgroep Alkmaar, and an internship of choice at the Neonatal Intensive Care Unit of VUmc, after which she decided to see if the field of paediatrics was a fit.

She worked as resident not in training (ANIOS) at Deventer Hospital for 10 months, during which she was parttime scheduled to do extra research. However, in search of long-desired PhD research and wanting to work in the field of surgery again, she applied at the lab of dr. Wilma Mesker and prof. Rob Tollenaar at the Leiden University Medical Center and was accepted to finish the UNITED study, for which prof. Han van Krieken was also her promotor. Despite COVID-19, she started her PhD at the University of Leiden in May 2020. Her rowing on a high level had made place for organization and coaching meanwhile, for which she was Head Coach of the club Eights.

She also continued her representatives work at the dept. Surgery, where she was the contact person for all PhD candidates of the department during her whole PhD trajectory ('oudste promovendus'). In the PhD Candidates Network Netherlands (PNN) during the first part of her PhD, she was elected as vice chair/secretary for a year, and the subsequent year as chair, representing all PhDs from the whole of the Netherlands, e.g. for the Ministry of Education, Culture and Science. After her time at PNN, she continued board work as member of the national board for communication and external relations of Medical Business while doing her PhD research.

She presented the UNITED study results at the European Society of Pathology congress in 2023 together with her work with prof. Aristotelis Tsirigos, which both resulted in a top-5 nomination of best oral presentations. The collaboration with the lab of prof. Aristotelis Tsirigos of New York University Grossman School of Medicine pertained to artificial intelligence on pathology images of colorectal cancer. This resulted in an internship in 2024 at his lab in New York City, United States of America. Upon return, she applied for clinical work again at the dept. Surgery of the Alrijne Hospital, at which she started as ANIOS in December 2024, and became a member of the PR committee of The Young Specialist.

Dankwoord

De totstandkoming van dit proefschrift was alleen mogelijk met hulp en steun van velen. Graag wil ik daar iedereen kort maar krachtig voor bedanken, een aantal hier in het bijzonder:

Ten eerste mijn promotoren, Prof. Dr. Tollenaar, em.Prof. Dr. van Krieken, en recent ook Dr. Mesker: Beste Rob, jouw kennis over de praktijk (en wijn) waren essentieel voor het translationele aspect van mijn proefschrift: tijdens onderzoekersoverleggen hield je voor ons de klinische relevantie voor ogen. Wanneer je er was, maakten we met waardevolle meetings altijd een verdiepingsslag, waarvoor veel dank!

Beste Han, met jouw helikopterview kon je mij bijsturen naar (politiek) correcte antwoorden, en waar wij eerst gedurende mijn PNN-tijd naast mijn PhD konden sparren over bestuurlijke/organisatorische/politieke ontwikkelingen, hebben wij later ook samen nog gewoon gezellig coupes gekeken. Ik voelde mij vereerd om genoemd te worden in je afscheidsoratie, dank voor jouw immer meelevende nuchterheid.

Beste Wilma, mijn (eo-)promotor en dagelijks begeleider, waar je initieel misschien moest wennen aan mijn (over-)gestructureerdheid, groeide onze samenwerking al snel, en was het een eer om bij te kunnen dragen aan de naderende implementatie van de TSR in (inter)nationale richtlijnen! Dank voor de vele snelle reacties voor deadlines, begeleiding en betrokkenheid, maar bovenal, dank voor jouw altijd positieve uitkijk ondanks tegenslagen, waardoor je mij weer wist te motiveren.

Prof. Tsigos, dear Aris, thank you so much for the amazing opportunity to collaborate and for my exciting internship at NYU. Hope to see you in person again soon, also to celebrate and toast with ouzo to our finally published article! Also, my co-author buddy Bojing, without you, I could not have made that manuscript nor the ESP presentation, thank you so much! Of course the whole PMed-gang too, such a warm welcome during my internship; keep PMed-Thursdays going strong!

Support van het Clinical Research Center, Annet en Elma, was onmiskenbaar voor het succes van de studies in dit boek; hartelijk dank voor jullie werk en meedenken! Ook patholoog Stijn Crobach bedankt voor de goede samenwerking!

Also, a major thank-you goes out to all coauthors of my articles, UNITED collaboratives, the PLCRC work force, and the RAPIDO steering group. Without all your input, inclusions, work and support, it would have been impossible to finish the studies in this thesis, for which I'm incredibly thankful. I will cherish the collaborations and contacts made throughout the years. Daarnaast ook dank aan mijn voorgangers, dankzij wie ik de uiteindelijk door hen opgezette UNITED kon afmaken.

Mijn Guidance Committee. Eerst, Dr. Koen Peeters. Waar het begon als klinische begeleiding van mijn onderzoek, ik met je kon sparren en een link kreeg naar klinische studies zoals de RAPIDO, heb ik je ook als behandelaar mogen meemaken. Bewonderenswaardig was dit, dank voor de goede zorgen, respect en lach of overleg tussendoor. Dr. Frank Speetjens wil ik ook bedanken, niet alleen voor initiële guidance committee-werk, maar tevens voor jouw begeleiding.

Labwerk was in dit boek niet mogelijk zonder hulp van Geeske en Ronald. Dank voor het begeleiden van mij en bovenal ook mijn verschillende studenten. Dank gaat ook uit naar deze waardevolle studenten Lois, Diana, Shriya, Cindy en Davita.

Dank aan alle collega's, Heelkunde en ook Onco, door heel LUMC verspreid. Bij de THREMA of lunch was het altijd gezellig, maar vooral de promovendiborrels en de Chirurgendagen/cup zullen mij (grotendeels) bijblijven! Een speciaal plekje hebben de C7 (Ruben, Labrinus, Victor, Floris, Sophie en Annelieke) en de P5 peeps. De huidige Wilma's Angels (Cor en Layla) en mijn opvolger Floor wil ik

ook bedanken en nog veel succes en plezier wensen! Daarnaast ook veel dank aan inmiddels mijn niet-meer zo nieuwe collega's van het Alrijne Ziekenhuis voor het warme bad, de langverwachte terugkeer naar de kliniek verliep (hopelijk) soepel mede dankzij jullie support!

Aan al mijn vrienden, teveel om nu op te noemen, van de basisschool, middelbare school, studie tot werk, extracurriculair van commissies en besturen, van roeien en wonen: dankjulliewel voor alles. Samen zijn we door diepe dalen gegaan maar hebben ook hoge pieken verkend. Stoom afblazen en/of sporten hielp mij er doorheen! Extra veel dank gaat ook uit naar mijn fijne paranimfen, Laurens en Anneke, de gevallenen.

Dear family, from near (NL) and from far away (UK/NZ/CA), thank you so much for your support, hospitality, fun, cozy and crazy times. Missing those who aren't here anymore, Oma Oma, grandma Heather, Bev and dr. grandNed. Alison, belly, vroeger altijd vechtend zijn we inmiddels naar elkaar toegegroeid, ondanks de overzeese afstand nu, en elkaars steun geworden in de afgelopen zware periodes; Lieve mama en papa. Dank voor het immer luisterend oor. Aan jullie draag ik dit proefschrift op.

Tot slot: H(aa)i mijn Ridder, duikbootman, dank voor je onvoorwaardelijke steun en toeverlaat. Jouw liefde, een onverwachts cadeautje, maakt van mij een gelukkige forens. Ik kijk uit naar nog vele mooie momenten en reizen samen.



M. Polack
M.S.

