# Microenvironment, systemic inflammatory response and tumor markers considering consensus molecular subtypes of colorectal cancer

Thesis

**Anna Jakab, MD**

Doctoral School of Semmelweis University
Division of Pathology and Oncology

Supervisor:

Tamás Micsik, MD, Ph.D.

Official reviewer: Judit Halász, MD, Ph.D.

Tamás Szamosi, MD, Ph.D.

Head of the Complex Examination Committee:

Professor Kulka Janina, MD, D.Sc.

Members of the Complex Examination Committee:

Erika Tóth, MD, Ph.D.

Professor Béla Molnár, MD, D.Sc.

Budapest

2024

## 1. Introduction

Besides the TNM classification, many other aspects of malignancies affect the outcome, which is the case in colorectal cancer (CRC) as well. These predictors could be certain histological features, molecular profile, and even the host response.

Host response can be represented both at the level of tumor microenvironment and systemic inflammatory response. Despite the complex nature of these mechanisms, some of these descriptors yield robust prognostic power while being easily assessable such as simple morphometric analysis or interpretation of routine laboratory tests.

Within the tumor microenvironment, certain cell populations aid an effective host response and the destruction of tumor cells, while other populations facilitate invasion, chemoresistance and metastatic properties. Some of these can be evaluated on HE-stained slides: the Klintrup-Makinen score descirbes the antitumoral inflammatory response, while the Tumor-stroma ratio quantifies the desmoplastic reaction in colorectal cancer, which reflects the activity of cancer associated fibroblasts.

The systemic host response can be examined using a variety of markers, such as CRP, albumin, absolute neutrofil, lymphocyte and platelet count as well as combined, SIR-based scores.

Routinely used tumor markers are also associated with outcome to some extent.

Colorectal cancer is also known to have heterogenous genetic background with some characteristic alterations and driver mutations that influence biological therapy, such as microsatellite instability, KRAS and BRAF mutations.

These molecular features and other, aforementioned characteristics were integrated to predict the biological behaviour of colorectal cancer. Based on the comprehensive gene expression profile analysis of over 5000 CRC cases, the consensus molecular subtypes were created. Altogether, four distinct Consensus Molecular Subtypes (CMS) had been identified.

CMS1 aka. the immune subtype. These tumors are often regarded to as „hypermutated" which is due to mismatch repair deficiency (dMMR) in the vast majority of this group. Cases are often present with prominent antitumoral inflammatory response which is explained by the overexpression of tumoral neoantigens. This group has the best prognosis in early stages, and responds well to immune checkpoint inhibitors.

CMS2 aka the canonic subtype. Both by appearance and molecular pathogenesis (Chromosomal instability – Vogelstein-Fearon pathway), this is the most frequent and typical subtype of CRC with intermediate outcome.

CMS3 aka the metabolic subtype is characterized by enrichment of KRAS mutations, and metabolic de-regulation. Regarding

morphology and outcome, it shares many similarities with CMS2. Some authors don't differentiate between the two subtypes, therefor the two CMSs are often referred to as „epithelial subtypes".

CMS4, aka. the mesenchymal subtype is present with prominent desmoplastic reaction and angiogenesis. These tumors have the worst outcome and response to both targeted and standard chemotherapy. It is believed that the TME has an immunosuppressive signature as well.

The host response, especially the TME is greatly important in two subgroups – CMS1 and 4, in the best and the worst subgroup of CRC.

## 2. Objectives

First, the objective of the authors was to consider a method that aids the precise quantification of TSR with the help of a commercially available, ML-based digital image analysis software. A crucial point was testing the accuracy of the software-aided TSR assessment and investigating the agreement with visual TSR evaluation. Of further focus was to compare the two methods in matters of prognostic power and relationship to clinicopathological variables.

Second, our research was focused on TME markers, and systemic inflammation, due to their good reproducibility and availability in routine histopathological and clinical evaluation and also the CMS. The aim of the authors was to analyze the relationship between the aforementioned factors, as well as exploring and comparing their prognostic significance while assessing their influence on the biological behaviour of CRC.

Third, in order to identify patients with particularly poor outcome, we attempted to create a novel scoring system that integrates host response both on the microenvironmental and on the systemic level as well.

## 3. Methods

Altogether, 185 stage I-IV CRC patients were selected in a retrospective cohort. All patients underwent surgical resection and were diagnosed in the Institue of Pathology and Experimental Cancer Research (Budapest, Semmelweis University, Hungary) between 2009 and 2017. All relevant slides and formaline fixed paraffin embedded (FFPE) blocks were retrieved from the archives of the Institute. Slides containing the deepest site of invasion were scanned using Pannoramic 1000 scanner (3DHistech, Budapest, Hungary). Exclusion criteria included neoadjuvant treatment, death within 30 days of surgery, diagnosis of other malignancy in medical history. Anamnestic data and laboratory results were obtained via the internal medical database of Semmelweis University (MedSolution, T-Systems, Budapest, Hungary). Preoperative serum CA19-9 and CEA levels were measured routinely using Abbott Architect CA 19-9 XR immunoassay (Chicago, IL, United States of America) and Abbott Architect CEA immunoassay (Chicago, IL, United States of America). Staging and evaluating surgically resected CRC specimen was performed according to the UICC TNM classification, 8$^{th}$ edition. The studies were approved by the Hungarian Scientific Council National Ethics Committee for Scientific Research (no. 216/2020).

### *Visual TME assessment*

For TSR assessment, the area of interest was selected at the invasive front using a 10x objective. This “hotspot” area had to contain the highest percentage of stromal compartment. Furthermore, tumor cells had to be present at all four poles of the examined area.The stromal content was estimated per 5% increments. Tissue compartments consisting of necrotic debris, mucus, smooth muscle, nerves, and large vessels were excluded from the stromal compartment as recommended. In case the TSR was ≥50%, the case was classified as TSR-high, otherwise, it was classified as TSR-low.

The KM grade evaluates the inflammatory reaction in a semi-qunatitative manner. When there was no or insignificant increase in inflammatory cells at the invasive front, host reaction was graded as KM-low. In case of abundant, band- or cup-like infiltrate at the margin with destructed tumor cells, the inflammation was classified as KM-high. The combination of KM grade and TSR gives a slightly more sophisticated notion on the local host response. The presence of extensive inflammation (KM-high) always results in GMS0 score, even with pronounced stromal infiltration, though, such cases are uncommon. When there isn’t any prevalent inflammation (KM-low) nor stromal infiltration (TSR-low), the case is graded as GMS1. Abundant stromal content (TSR-high) in combination with weak inflammatory reaction yields a GMS2 score.

All of the aforementioned parameters were graded by two independent observers (AJ and TM) in a blinded manner. The final

results in cases of discordant scoring were determined after discussion between the two raters.

***TME assessment using ML-based digital pathology software***

The "hotspot" area selected for visual TSR assessment was evaluated using the ML-assisted software as well, which was annotated as a circular area of 3.77 $mm^2$ representing a 10x objective.

To establish "gold standards" for validation of software performance, 52 cases were randomly selected. Different tissue types (tumor epitheli, smooth muscle, necrosis, background/lumina, stroma) were manually annotated within the hotspot area, which resulted in accurate quantification of each tissue compartment. These "gold standard" annotations were the basis of comparing the pathologist's and the software's performance.

The software of choice was the PatternQuant module of QuantCenter, which is the extension of SlideViewer (3DHistech, Budapest, Hungary). PatternQuant is a trainable, machine learning-based pattern recognition algorithm, which identifies certain tissue types based on texture patterns and hue intensity via wavelet transformation. An individually trained algorithm was created for all 185 slides, that could differentiate between five distinct patterns representing particular tissue compartments (tumor epithelium, stroma, smooth muscle, necrosis/debris, and background). All cases were assessed by both software ($TSR_{software}$) and visual analysis ($TSR_{visual}$).

***Microarray construction and immunohistochemistry (IHC)***

Tissue microarracy (TMA) blocks were created from surgically derived FFPE blocks of 167 patients using TMA Master1000. At least two representative cores were selected per case from the tumor centre. The following IHC reactions were carried out on 4 um thick sections of the TMAs: anti-MLH1, anti-PMS2, anti-MSH2 and anti MSH6 for mismatch repair status assessment, and anti-pancytokeratin, anti-CDX2, anti-FRMD6 and anti-ZEB1 stains for CMS classification. All IHC reactions were assessed by two observers (AJ and TM).

***CMS classification***

Mismatch repair deficient cases were classified as CMS1. The remaining cases were classified using an online, TMA-based and validated online classifier (https://crcclassifier.shinyapps.io/appTesting/), based on staining intensity of ZEB1, HTR2B, FRMD6 and CDX2. This method classifies cases as either epithelial (CMS2 and CMS3) or mesenchymal (CMS4). Typically, low FRMD6 and HTR2B staining intensities, lack of nuclear ZEB1 expression and strong CDX2 stain correlates with epithelial subtypes (CMS2/3); while strong positive FRMD6 and HTR2B, loss of CDX2 and positive nuclear ZEB1 reaction is expected in mesenchymal CRCs (CMS4). Where probability of estimated CMS was between 0.5 and 0.6, the case was

automatically excluded from our analysis. In total, 12 cases were excluded due to uncertain subtyping.

***Creating the stroma-tumor marker score***

To reflect the biological behaviour of certain cancer subtypes classified by TME and systemic response, a novel scoring system was established by combining CA19-9 and TSR, the most robust prognosticators independent of stage, into stroma-tumor marker (STM) score. In case of TSR-low and CA 19-9 low cases, STM 0 score was given. If either CA19-9 or TSR was classified as high, but the other marker as low, an STM 1 score was given. When both markers were classified as „high", STM 2 score was given.

## 4. *Results*

Out of the 185 patients who were included in our cohort, 155 patients had available CEA, and 135 patients had available CA19-9 results. CMS classification was carried out in 155 patients.

Accuracy for tumor epithelium recognition was 80%, 72.4% for stromal recognition, respectively. Cohen's κ for TSRvisual between observers was 0.778, and between TSRGoldStandard and TSRvisual it was 0.711. κ values for between TSRsoftware and TSRvisual were 0.472.

TSR-high tumors were associated with higher pT and pN stages and advanced disease and with lymphatic and perineural invasion as well. Similarly, KM-low cases correlated with more advanced pT and pN stage and distant metastasis as well. As expected, GMS was also associated similarly with clinicopathological features. KM, TSR or GMS were not associated with any SIR markers.

The elevation of serum CRP was assocaited with increasing age, pT, distant metastasis, higher histological grade, vascular invasion and there was a tendency towards perineural invasion. Elevation of absolute platelet count (APC) was elevated in males, and was associated with right-sidedness, CMS1, and lymphatic invasion, and there was a trend towards distant metastasis and vascular invasion as well. The mGPS showed significant association with higher grade and a tendency towards elevated pT. CEA was significantly lower in left sided tumors. Elevated CEA levels were associated with stage,

pT and distant metastasis and also showed a tendency towards higher pN. CA19-9 was also associated with stage, pT, distant metastasis, lymphatic invasion, and GMS. There was a tendency towards vascular and perineural invasion.

CMS1 was significantly associated with right colonic localization and higher histological grades. CMS4 was associated with higher stage, lymphatic and perineural invasion, pN and M descriptors, and there was a tendency towards TSR-high just failing to be significant. We did not find any significant correlation between the examined tumor markers (CEA and CA19-9) and CMS.

For TSRvisual, the 5-year OS for patients with TSR-high versus TSR-low was 49% versus 74%, and TSRsoftware yielded similar results (TSR-high versus TSR-low: 50% vs 73%). On multivariate analysis both visual and software TSR were found to be associated with poorer OS. Patients with high GMS, high ANC, low albumin, elevated CRP, elevated CEA and CA19-9, as well as higher mGPS and mesenchymal subtype (CMS4) had poorer overall survival.

Apart from TSR, in the univariate Cox regression analysis GMS, mGPS, ANC, CRP, Albumin, CEA and CA19-9 were significantly associated with OS; CMS presented a tendency (towards increased risk of death in CMS4 patients. In the multivariate analysis mGPS, Albumin, CRP, CA19-9, and STM-score were significant predictors of OS.

In our research the strongest independent TME-based marker was TSRvisual. Also, CA19-9, a tumor marker often, though not routinely used in colorectal cancer follow up, came out as a predictor of overall survival in our analysis. Incorporating these two, STM was created. Cases classified as STM2 were associated with younger age, higher pN and M, as well as higher TNM stage, and presence of lymphatic and perineural invasion. The mesenchymal subtype of CMS was also more prevalent in STM1 and STM2 groups. There was a tendency towards higher pT, also, preoperative serum CRP and CEA levels correlated with STM1 and STM2.

The STM score significantly stratified 5-year overall survival (86% versus 54% versus 42%) with Kaplan-Meier analysis. In the multivariate Cox-regression analysis STM was found to be an independent prognosticator of OS (independent of sex, grade, stage and vascular invasion.

## 5. Conclusions

Using a ML-based digital image analysis platform, the PatternQuant, $TSR_{software}$ delivered similar prognostic power as $TSR_{visual}$ while presenting acceptable accuracy and inter rater agreement. TSRvisual helped identifiyfing a subset of CRC patients with rather aggressive phenotype and poor outcome, and was also significantly associated with elevated levels of CEA and CA19-9 tumor markers and with CMS4, which was concordant with previous studies.

The KM grade was also associated with more advanced stage, however, it didn't yield significant prognostic power. The combination of KM and TSR, the GMS score, was similarly associated with adverse clinicopathological features and poorer survival.

Amongst SIR markers, elevated serum CRP outstandingly identified cases with adverse histopathological features (stage, lymphatic and vascular invasion) while also being an independent predictor of OS, however, it wasn't associated with any of the TME markers, nor CMS. The mGPS, a combined SIR marker comprising of CRP and albumin also reflected similar associations. Out of the SIR markers, only

elevated APC was associated with CMS1, while also showing tendency towards advanced disease.

Strikingly, the authors couldn't find significant connection between the TME and SIR, which had been expected based on preceding results.

The elevation of CEA and CA19-9 tumor markers also indicated advanced disease and poor outcome. Utilising the combination of the most robust TME marker, the TSR and CA19-9, resulted in the STM-score, which stratified the outcome of CRC patients significantly and came out as an independent predictor of OS in the multivariate analysis, while also identifying a group of patients with adverse clinicopathological charateristics.

For CMS classification a simplified and validated IHC-based method was applied. The CMS4 a.k.a. mesenchymal subtype represented an immensely adverse phenotype of CRC while showing weak correlation with TSR. CMS1 tumors were also associated with right-sidedness and higher histological grade. Remarkably, none of the CMSs were significantly associated with SIR or tumor markers. CMS4 showed a tendency towards poorer outcome, however, CMS didn't yield considerable prognostic value in our study.

Taken together, our study could not reveal a significant connection between CMS and host repsonse, still, both characteristics can help anticipating disease outcome and clinical decision making by distinguishing individuals who require closer follow up or vigorous treatment.

## 6. Bibliography of the candidate's publications

***Publications related to the thesis***

**Jakab, Anna;** Patai, Árpád V ; Micsik, Tamás

Digital image analysis provides robust tissue microenvironment-based prognosticators in stage I-IV colorectal cancer patients

HUMAN PATHOLOGY 128 pp. 141-151. , 11 p. (2022)

IF: 3,3

**Jakab, Anna;** Patai, Árpád V. ; Darvas, Mónika ; Tormássi-Bély, Karolina ; Micsik, Tamás

Microenvironment, systemic inflammatory response and tumor markers considering consensus molecular subtypes of colorectal cancer

PATHOLOGY AND ONCOLOGY RESEARCH 30 Paper: 1611574, 20 p. (2024)

IF: 2,8

***Publications not related to the thesis***

**Jakab, Anna ;** Patai, Árpád ; Micsik, Tamás

Colorectalis carcinoma: molekuláris biomarkerek, precíziós terápia és új perspektívák

ORVOSKÉPZÉS 96 : 3 pp. 387-402. , 16 p. (2021)

IF: -

Micsik, Tamás ; **Jakab, Anna**

A tápcsatorna daganatainak molekuláris genetikája

MAGYAR BELORVOSI ARCHIVUM 75 : 2-3 pp. 72-92. , 21 p. (2022)

IF: -

Micsik, Tamás ; **Jakab, Anna** ; Lehoczki, Csaba ; Patai, Árpád V.

Traditional Serrated Adenoma of the Gallbladder, a Case Report

PATHOLOGY AND ONCOLOGY RESEARCH 28 : 4 Paper: 1610133 , 5 p. (2022)

IF: 2,8

**Jakab, A ;** Micsik, T

Molekuláris biológia/genetika az emésztőrendszer daganataiban

In: Tulassay, Zsolt (szerk.) Gasztroenterológia

Budapest, Magyarország : Medicina Könyvkiadó (2023) 1,180 p. pp. 1081-1107. , 27 p.

IF: -