Microsatellite instability at tetranucleotides (EMAST) in colorectal cancer: clinical relevance, mechanisms and immune markers

Martin M Watson

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2020

UNIVERSITY OF BERGEN

Microsatellite instability at tetranucleotides (EMAST) in colorectal cancer: clinical relevance, mechanisms and immune markers

Martin M Watson

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 18.06.2020

Year: 2020

Title: Microsatellite instability at tetranucleotides (EMAST) in colorectal cancer: clinical relevance, mechanisms and immune markers

Name: Martin M Watson

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

The research of this PhD thesis was undertaken at the Stavanger University Hospital as part of the Gastrointestinal Translational Research Group and the Research Group for Gastrointestinal Surgery. The group is affiliated with the Department of Clinical Medicine (Klinisk Institutt 1, K1) at the University of Bergen (UiB), Norway.

Part of this work was carried out in collaboration with the University of Stavanger (UiS), Norway.

Main financial support was provided by grants from Folke Hermansen Fond, Stavanger, Norway.

Additional intramural grants were received from SUH, UiB and UiS.

Acknowledgements

I am extremely grateful and indebted to my main supervisor, Prof. Kjetil Søreide. Above all mentors and teachers, you put the *translation* in translational research for me. I've made you work holidays, nights and between shifts, with most of our correspondence happening after 11 PM (95% C.I. 10:32-12:02). Under your supervision I learned how to be pragmatic, navigating the jungle of possible leads, dead ends and maybes, without losing the passion for research (which by definition is made of possible leads, dead ends and maybes).
I am a better researcher and - not to be underestimated - a better writer thanks to you.

To my co-supervisor, ass. Prof. Hanne R Hagland: all of the above, plus the psychological support required to cope with all of the above. You've been a colleague, a friend, a teacher and a sister through this journey, thank you for your guidance and everlasting positive attitude. If there isn't a silver lining in a storm you'll go up and paint it.

Thanks to Ramesh Khajavi, for being an exceptional resource with all the administrative work behind the ACROBATICC cohort biobanking process and including nearly every single patient in the study. All the surgeons of Gastrokir. Department and all the patients that participated in this study: without you this study could literally not have happened.

I am indebted to my colleagues at the Hillevåg lab, too, for the positive environment and encouragement and for tolerating my absence from most daily lunches. I can now say that the intermittent fasting diet did not produce a statistically relevant decrease in my body weight ($p>>>0.05$).
Special thanks to Kvantitativ Patologi's finest: Emma Rewcastle and Melinda Lillesand! Not only you helped from the start with the collection of the project's samples, you were always willing to go the extra mile for me and my experiments. I truly value your friendship and professionalism.

To all the pathologists at SUS, thanks for helping me collecting samples when Melinda was out of office, and apologies for countless lunch breaks interrupted. Thanks, in particular, to Einar Gudlaugsson and Dordi Lea for your high-quality contributions to my analyses despite having little time at your disposal. Your no-nonsense approach and research ethics, Dordi, kept me with my feet on the ground in more than one occasion.

If I was allowed to roam free between the pathology department and the Hillevåg lab, encouraged to use resources and to pick people's brains at all times is only thanks to prof. Emiel Janssen. The same is true for Ivar Skaland, immunohistochemistry *grand master* and guru, willing to eat his lunch sitting at the Visiopharm PC to teach me the ropes of digital image analysis and share his endless knowledge on the topic of IHC.

Just like in WWI trenches, those who struggle together form strong bonds. Thanks to my fellow PhD candidates Dordi Lea, Nina Egeland and Tia Tidwell, for the reciprocal support and stimulating conversations.
Thanks to my friends, who still consider me one despite seeing me once every solar eclipse and usually on a half-hour notice.

Mamma e Papà, thank you for always believing in me, always giving me the means to follow my dreams and become whatever I wanted to become in life. Even if it meant giving up a lot, including the right to be close to your grandchildren, you've never stopped trusting me. The respect, admiration and love I have for you cannot be described. I only wish Papà had seen all this.
Thanks to my acquired parents too; where I come from "in-laws" is often used as a not-so-flattering figure of speech, while I challenge anyone to find a more encouraging and loving extended family.

Finally, all my love to my two sons Matteo and Simon and my rock Solveig. You took our household entirely on your shoulders in the past few weeks, never complaining and always understanding while working three jobs, all the while being the best mum a mother can be. You and the boys are my everything.

Thank you very much.

Martin M. Watson

Do not fear to be eccentric in opinion,
for every opinion now accepted was once eccentric.
Bertrand Russel

Abbreviations

EMAST	Elevated microsatellite instability at selected tetranucleotides
CD3/8	Cluster of differentiation 3/8
CEA	Carcinoembryonic antigen
CMS	Consensus Molecular Subtype
CRC	Colorectal cancer
FFPE	Formalin-fixed, Paraffin-embedded
HNPCC	Hereditary non-polyposis colorectal cancer
IHC	Immunohistochemistry
IS	Immunoscore
MMR	Mismatch repair
MSH2/3/6	MutS homolog 2/3/6
MLH1	MutL homolog 1
MSI	Microsatellite instability
PCR	Polymerase chain reaction
PD-1	Programmed death receptor 1
PD-L1	Programmed death Ligand 1
PMS2	Post-meiotic segregation 2
TNM	Tumour-Node-Metastasis

Abstract

Colorectal cancer (CRC) is the second-leading cause of cancer-related death and a global burden in terms of incidence, person-years of life lost and public health costs. About 10-20% of sporadic CRCs develop through a pattern of DNA mismatch repair (MMR) insufficiency that leads to hypermutation, microsatellite instability (MSI) and a strong immunogenicity, exemplified by the higher number of immune cells infiltrating the tumours.

Over the course of the past two decades, MSI was investigated and defined in the context of CRC and other solid tumours. MSI can be measured by length variation at selected mononucleotide repeats across the normal/tumour genomes, with specific panels and thresholds, and/or by immunohistochemical (IHC) analysis of MMR members such as MSH2 and MLH1. MSI CRCs are often located in the proximal colon, have better prognosis, and are to some degree refractory to a common CRC chemotherapy (5-fluorouracil). Further, MSI CRCs are better responders to anti-PD-1/PD-L1 immunotherapy, which is an exciting development in modern oncology. A variation of MSI, investigated exclusively at tetranucleotides, is elevated microsatellite alterations at selected tetranucleotides (EMAST). Studies on EMAST in CRC are scarce, no consensus on a panel of markers and their cut-off exists, and neither its clinicopathological implications nor its relationship with MSI is defined at present.

In this work, EMAST was investigated in a highly selected archival cohort of non-metastatic CRC and associated with MSI, proximal colon location and a lower number of lymph nodes harvested at surgery (**paper I**). Loss of MSH3 has been proposed in the literature as a possible mechanism leading to instability at tetranucleotides, and therefore implicated with EMAST in CRC. To test this hypothesis, automated digital image analysis of IHC staining for MSH3 was performed in a consecutive cohort of CRC included prospectively in an ongoing biobank established in 2013 (**paper II**). Negligible degrees of MSH3 loss were found in this cohort, and were not associated with EMAST, thus suggesting alternative

mechanisms of instability at tetranucleotides. In the same second cohort, EMAST was found again to overlap with MSI, at higher incidences in the proximal colon, and with an improved recurrence-free survival (**paper III**). Notably, however, EMAST patients were shown to be older and display a higher incidence of pre-operatively recorded markers of frailty, such as anaemia, hypoalbuminemia and greater loss of weight.

To establish whether EMAST, as MSI, correlates with a higher immunogenicity and could serve as a predictive marker for PD1 blockade immunotherapy, IHC analyses of T-lymphocyte markers CD3 and CD8, as well as PD-L1 were carried out. EMAST tumours were found to have higher degrees of T-cells, including CD8 cytotoxic T-cells, and to express PD-L1, suggestive of an active immune evasion mechanism (**paper IV**). Prognostically, PD-L1 expression in tumour cells was not indicative, although its expression on peritumoral immune cells correlated with an improved recurrence-free survival, independent of EMAST status. Higher densities of T-lymphocytes were also highly prognostic, both in terms of recurrence-free and overall survival.

The data hereby presented suggests a role for EMAST in CRC clinicopathological analysis, although further subgroup analyses are required to establish whether MSI and EMAST are distinct mechanisms of instability. EMAST seems to be of comparable immunogenicity to MSI and shows expression of PD-L1 possibly as an antagonising mechanism to it. Finally, results here presented suggest a prognostic value for PD-L1 expression in non-tumour cells, pointing at the importance of the cellular distribution of the PD1/PD-L1 axis.

List of Publications

I. Watson MM, Lea D, Rewcastle E, Hagland HH, Søreide K. Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal cancers with and without high-frequency microsatellite instability: same, same but different?
Cancer Medicine (2016) *Vol. 5, 7:1580-1587*

II. Watson MM, Lea D, Hagland HH, Søreide K. Elevated microsatellite alterations at selected tetranucleotides (EMAST) is not attributed to MSH3 loss in stage I-III colon cancer: an automated, digitalized assessment by immunohistochemistry of whole slides and hot spots.
Translational Oncology (2019) *Vol. 12, 12:1583-1588*

III. Watson MM, Kanani A, Lea D, Khajavi R, Søreide JA, Kørner H, Hagland HH, Søreide K. Elevated microsatellite alterations at selected tetranucleotides in colorectal cancer is associated with an elderly, frail phenotype and improved recurrence-free survival.
Annals of Surgical Oncology (2020) Vol 27, 4:1058-1067

IV. Watson MM, Lea D, Skaland I, Gudlaugsson E, Hagland HR, Søreide K. PD-L1 expression is associated with EMAST, density of peritumoral T-cells and recurrence-free survival in colorectal cancers.
Cancer Immunology, Immunotherapy (2020) Epub ahead of print).

Contents

INTRODUCTION

1.1 COLORECTAL CANCER

1.1.1 Epidemiology

Colorectal cancer (CRC) is the fourth most diagnosed cancer world-wide, with an estimated 1.8M new cases and >800.000 deaths in 2018 [1]. In Norway, CRC is the second most diagnosed cancer in both men and women, at near-equal ratio, with 4332 new cases in 2017 [2]. Median age at diagnosis is 73 for colon and 69 for rectal cancer, and it is thus mainly considered as diseases of the elderly[2]. Early-onset (diagnosis at <50 years of age) of CRC is however an emerging trend, especially in westernised countries [3]. Norwegian age-standardised incidence (**Figure 1**) is increasing steadily, and was higher than the rest of northern/western Europe and the fourth-highest world-wide in 2018 (42.9 per person-years) [1,4,5]. When considered independently, however, rectal cancer's incidence remained stationary for over 3 decades, as opposed to an increase in colon cancer in both men (+2.3%) and women (+6.6%) in the previous 5-year period only [2].

Globally, CRC is a "disease of the western world", the risk of which is affected by lifestyle and environmental factors [6,7]. Western-style dietary habits such as red and processed meat as well as elevated alcohol consumption are consistently associated with increased risk, while long-term, low-dose NSAIDs intake such as aspirin, and a high-fibre diet are generally considered as protective for CRC [6,8-12]. The most prominent risk factor is age, as the incidence rate for precursor lesion increases steeply and steadily over the age of 50 [13]. Other risk factors include family history of CRC, inflammatory bowel disease, obesity, male gender and ethnicity [14] (**Figure 2**).

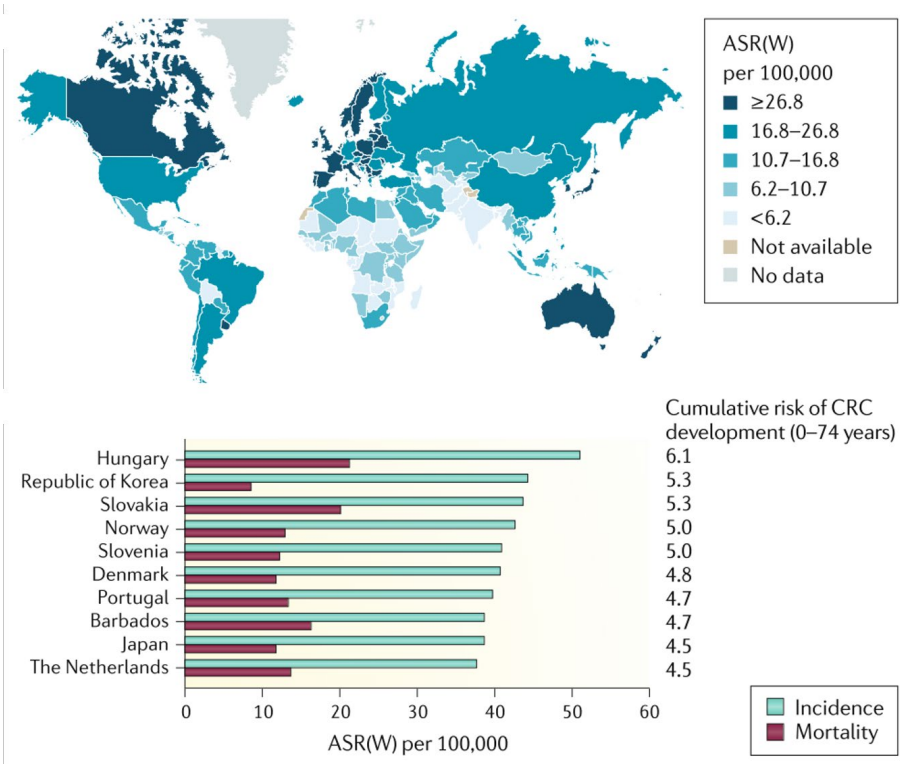


Figure 1. Age-standardised incidence rate of CRC world-wide. Reproduced with permission from [3], Copyright © 2019 Springer Nature.

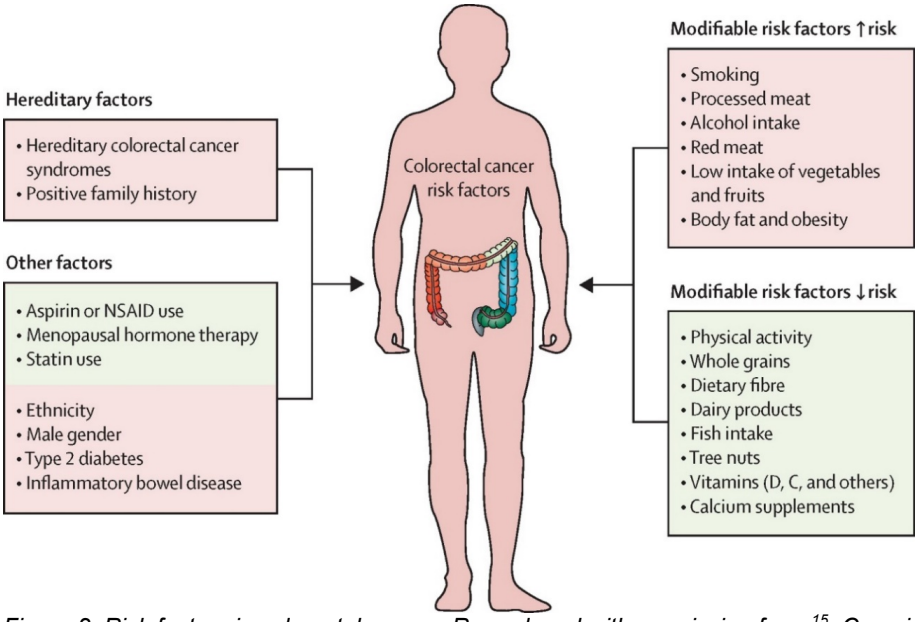


Figure 2. Risk factors in colorectal cancer. Reproduced with permission from [15], Copyright © 2019 Elsevier Ltd

Most CRCs are sporadic, but up to 30% are estimated to be of some inherited nature, although only ca. 5% are due to established congenital mutations. These include familial adenomatous polyposis (FAP), hereditary non-polyposis colorectal cancer (HNPCC, now known as Lynch syndrome) or juvenile polyposis (JPS) [16-18].

1.1.2 Diagnosis and screening

Colonoscopy is the primary method of CRC diagnosis, which is confirmed histologically on a biopsy taken during the procedure. Other, explorative techniques include CT colonography, flexible sigmoidoscopy and barium enema, each with its own advantages (slightly reduced invasiveness) and limitations (sensitivity) and however still considered as inferior to colonoscopy [13,19,20]. Given the uncomfortable nature of such invasive tests, compliance to undertake investigation is considered low. Non-invasive test alternatives have been developed in the past decades, such as stool-based tests like faecal immunochemical (FIT) and faecal occult blood (FOBT), to date the only screening test that has shown to decrease mortality in randomised controlled trials [21-23]. Stool DNA tests to date lack standardisation, in some cases sensitivity, and/or are too expensive for integration into public health systems [13,24]. A recent report from the international agency for research on cancer (IARC) reviewed the available literature on harm and benefits of various endoscopic and stool-based screening test, concluding that both FOBT and FIT, together with endoscopic techniques, reduced the risk of death from CRC [25]. During the initiation of this thesis, no formal screening programme was in place in Norway. Hence, patients are largely diagnosed based on symptoms (e.g. blood in stool, anaemia, change in bowel habits, abdominal pain or weight loss) or incidentally during work-up for other conditions.

1.1.3 Treatment

Early lesions such as benign polyps and a small proportion of stalked polyps with non-infiltrating adenocarcinomas can be resected endoscopically, provided no evidence of poor cellular differentiation or infiltration in blood and lymphatic vessels

is present. Surgery is however the only curative intervention for CRC, and adjuvant and neoadjuvant treatment is offered on the basis of Norwegian Health Council's guidelines [19].

A proportion of rectal cancers are administered neo-adjuvant radiotherapy in hope of tumour remission and improvement of circumferential resection margins, considered proportionally related to survival [26,27]. In primary colon cancer, a range of adjuvant chemotherapy regimens are offered to stage III (see later - staging) and "high-risk" stage II cancers. Regimens for stage III cancer are based on fluoropyrimidine drugs such as 5-fluorouracile (5FU, given in combination with leucovorin) or capecitabine. These are given in combination with oxaliplatin (FOLFOX, FLOX or XELOX/CAPOX), to patients <70 yo, based on degree of tumour invasion in the submucosa, extent of node involvement and presence of microsatellite instability (MSI). Depending on clinicians' patient-to-patient individual assessment (age, general condition, comorbidity), monotherapies of either one fluoropyrimidine are considered for patients older than 70 years.

Stage II patients are defined as "high-risk" when perforations are present near the tumour area, tumour growing through the serosa and low number (≤ 12) of assessed lymph nodes. In such cases patients are tested for microsatellite instability and offered a fluoropyrimidine monotherapy if microsatellite stable (MSS).

1.1.4 Prognosis

The relative 5-year survival for the 2014-2018 period in Norway is estimated at 65.4% and 69.8% for colon and rectum, respectively [2]. In a recent multi-register, global surveillance study (CONCORD-3, referring to the period 2012-2014) Norway was placed in the top 15% of 68 reporting countries for 5-year net survival of CRC [28]. The increasing survival and declining mortality from rectal cancer in both sexes are likely due to the improvement in patient care, as for example the introduction of total mesorectal excision and of preoperative radiotherapy in selected patients. However, circa 1600 patients a year still die from CRC in Norway [2,29].

Staging

Staging of colorectal cancer, traditionally based on anatomical features, has served pathologists as a prognostic tool since the 1920s. In 1932, Cuthbert E. Dukes, a pathologist at St. Mark's Hospital in London, described an association between survival of patients who underwent surgery for rectal cancer and the extent of growth of the tumour in the surrounding tissue [30]. Dukes' classification is still often reported in pathology alongside the newer staging system, adopted in 1954 and maintained up to date by the American joint Committee on cancer (AJCC), known as tumour-node-metastases (TNM). The 8th version of the AJCC cancer staging manual [31] was recently adopted, whilst the 7th edition (shown in **figure 3)**was in use at the time the present study was carried out.

The TNM staging system aims at stratifying patients in prognostic stage groups according to anatomical extent of their disease, at the same time guiding clinicians to choose the appropriate treatment strategy. Staging also provides a tool for comparison of results and stratify patients in trials. Patients are assigned a stage I-IV, with tumour (T) invasion through the muscularis propria (T3) as the limiting step for stage I to II switch. Presence of tumour deposits and/or positive lymph nodes (nodal status N), in turn, is what determines the stage to advance from II to III, independent of T status. Any distant metastasis (M) at the time of diagnosis automatically classifies the disease as stage IV CRC.

Definitions

Primary Tumor (T)

- **TX** Primary tumor cannot be assessed
- **T0** No evidence of primary tumor
- **Tis** Carcinoma in situ: intraepithelial or invasion of lamina propria[1]
- **T1** Tumor invades submucosa
- **T2** Tumor invades muscularis propria
- **T3** Tumor invades through the muscularis propria into pericolorectal tissues
- **T4a** Tumor penetrates to the surface of the visceral peritoneum[2]
- **T4b** Tumor directly invades or is adherent to other organs or structures[2,3]

Regional Lymph Nodes (N)[4]

- **NX** Regional lymph nodes cannot be assessed
- **N0** No regional lymph node metastasis
- **N1** Metastasis in 1–3 regional lymph nodes
- **N1a** Metastasis in one regional lymph node
- **N1b** Metastasis in 2–3 regional lymph nodes
- **N1c** Tumor deposit(s) in the subserosa, mesentery, or nonperitonealized pericolic or perirectal tissues without regional nodal metastasis
- **N2** Metastasis in 4 or more regional lymph nodes
- **N2a** Metastasis in 4–6 regional lymph nodes
- **N2b** Metastasis in 7 or more regional lymph nodes

Distant Metastasis (M)

- **M0** No distant metastasis
- **M1** Distant metastasis
- **M1a** Metastasis confined to one organ or site (for example, liver, lung, ovary, nonregional node)
- **M1b** Metastases in more than one organ/site or the peritoneum

ANATOMIC STAGE/PROGNOSTIC GROUPS

Stage	T	N	M	Dukes*	MAC*
0	Tis	N0	M0	–	–
I	T1	N0	M0	A	A
	T2	N0	M0	A	B1
IIA	T3	N0	M0	B	B2
IIB	T4a	N0	M0	B	B2
IIC	T4b	N0	M0	B	B3
IIIA	T1–T2	N1/N1c	M0	C	C1
	T1	N2a	M0	C	C1
IIIB	T3–T4a	N1/N1c	M0	C	C2
	T2–T3	N2a	M0	C	C1/C2
	T1–T2	N2b	M0	C	C1
IIIC	T4a	N2a	M0	C	C2
	T3–T4a	N2b	M0	C	C2
	T4b	N1–N2	M0	C	C3
IVA	Any T	Any N	M1a	–	–
IVB	Any T	Any N	M1b	–	–

NOTE: cTNM is the clinical classification, pTNM is the pathologic classification. The y prefix is used for those cancers that are classified after neoadjuvant pretreatment (for example, ypTNM). Patients who have a complete pathologic response are ypT0N0cM0 that may be similar to Stage Group 0 or I. The r prefix is to be used for those cancers that have recurred after a disease-free interval (rTNM).
* Dukes B is a composite of better (T3 N0 M0) and worse (T4 N0 M0) prognostic groups, as is Dukes C (any TN1 M0 and Any T N2 M0). MAC is the modified Astler-Coller classification.

Figure 3. The tumour-node-metastases (TNM) staging system as reported from the 7th edition of the AJCC cancer staging manual. Reprinted with permission, Copyright © 2017 Springer.

The recording of further prognostic factors is encouraged by the latest (8th) version of the AJCC manual, and these include:

- Serum CEA (both pre-operatively and at regular intervals during surveillance)
- Tumour regression score (in case of neoadjuvant chemo/radiotherapy)

- Circumferential resection margin (CRM)
- Perineural invasion (PNI)
- Microsatellite instability - MSI
- Mutations in the RAS/RAF pathway (KRAS, BRAF, NRAS)

Issues with current staging

While the TNM system remains the strongest recognised prognosticator in CRC, it received much criticism during the past years. This is especially true since the stratification of stages II and III introduced from the 5th to the 6th edition of the AJCC staging manual. Amendments introduced in subsequent versions of the TNM system have led to "stage migration" in some groups, which attracted criticism [32-34]. Among the main critiques was introduction of stage IIb (T4N0), for example, which showed poorer survival over the more advanced stage IIIa (T1-2N1) [35,36]. This could be explained by the fact that, following the guidelines, stage III patients received adjuvant chemotherapy and stage II did not. Further, stage II patients could be under-staged if an insufficient number of lymph nodes are resected at surgery or microscopically investigated, thus possibly missing N+ observations that could assign a stage III diagnosis. The number of lymph nodes resected during CRC surgery has been shown to be of prognostic relevance, especially in stage II-III patients [37,38]. Despite suggestions upon ideal minimum number of lymph nodes to harvest at surgery, this is subject to a number of factors such as location of the tumour (proximal vs. distal), length of the resected specimen and age of the patient, as well as biological characteristics such as MSI status or mutations in the RAS pathway [39-43].

All revisions of the TNM strongly rely on nodal status (pN) to assess prognosis and recommend adjuvant chemotherapy, however up to 25% of stage II and 40% of stage III CRC present with relapse, regardless of treatment offered. The main shortcoming of the TNM system lies in its inability to account for differences in tumour biology, such as genetic and epigenetic heterogeneity, molecular composition and interaction with surrounding tissue.

1.2 CRC tumorigenesis

1.2.1 Large intestine architecture and homeostatic dynamics

Throughout the gastrointestinal (GI) tract, architectural conformation of the epithelium lining the lumen evolved to serve specific purposes. The proximal (from the small intestine to the first two thirds of the transverse colon) and distal (last third of transverse to the upper anal canal) portions of the intestine generate embryonically from the midgut and the hindgut, respectively [44]. The large intestine, the terminal tract of the GI tract, is composed right-to-left (or proximal-to-distal) by caecum, colon (ascendens, transversum and descendens and sigmoideum) and rectum. While the right-side colon is still partly dedicated to the remainder of the digestive function through absorption, the distal portion is mainly involved in compacting and lubricating the stool and allowing it to pass through by peristalsis.

To increase overall surface and effective absorption, and to protect the stem cell compartment vital to tissue renewal and homeostasis, the lumen of the intestine folds into pits named crypts of Lieberkühn [45]. At the bottom of every crypt, in the so-called stem cell niche, resides a group of undifferentiated cells, which have been extensively studied with the aid of mouse models, and were shown to express stem cell markers such as LGR5 and EPHB2 [46-50]. Their "stemness" is believed to be maintained by pericryptal myofibroblasts and – in the colon - by CD24 and KIT-expressing Goblet cells via WNT, NOTCH, EGF and BMP pathways signalling [51-53]. By means of asymmetric division, the 6-14 stem cells in each crypt's niche are alone able to recapitulate the entire diversity of intestinal cell populations [54,55].

Daughter cells originating from a stem cell's asymmetric division initiate a series of rapid division cycles (ca. 2 per day), whilst migrating upwards through the crypt's mid-section (named transient-amplifying compartment), towards the lumen.

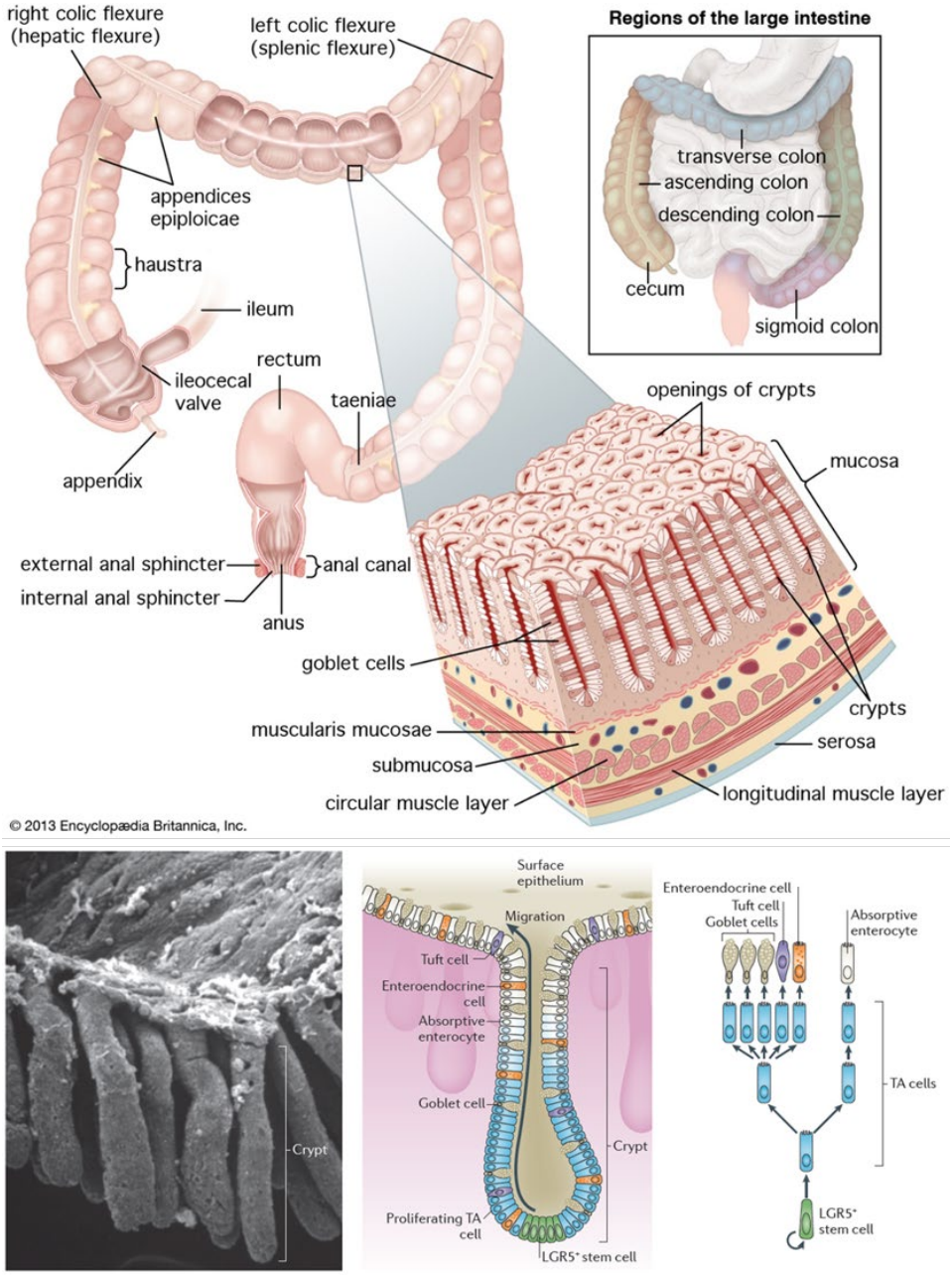


*Figure 4. (**Top**) Macro- and microscopic structure of the human large intestine. Reproduced with permission (Copyright © 2013 Encyclopaedia Britannica). (**Bottom**) the structural and cellular organisation of the large intestine. Reproduced with permission from [48], Copyright © 2014 Nature publishing group.*

Differentiation takes place during this division-migration process, and transient amplifying cells become increasingly committed to one of the handful of terminally differentiated cells composing the intestinal epithelium (**Figure 4 - bottom**). These include mucin-secreting goblet cells, tuft cells, M cells, absorptive enterocytes and hormone-secreting enteroendocrine cells. Once reaching the surface, terminally differentiated cells undergo apoptosis and shed into the lumen, renewing the entire lining of the intestinal lumen within 6-8 days. This cell proliferation/death cycle is tightly regulated and kept at equilibrium under normal circumstances. A shift towards proliferation rather than cell death disrupts cellular homeostasis, causing the formation of neoplasms.

1.2.2 Carcinogenesis

In the current model of carcinogenesis, neoplasms undergo several rounds of random mutations that may ultimately lead to their malignant transformation into cancer. Pertinent to the stochastic model of mutation accumulation, no two cancer are genetically the same. Ultimately, however, the wealth of genetic and epigenetic alterations in each tumour serves the cancer's independence from normal physiological signals in a set of cellular functions.

The range of capabilities that neoplasms acquire during carcinogenesis were described as the 'Hallmarks of cancer' (**Figure 5**) [56,57]. Self-sufficient propagation is achieved by the hijacking of proliferative signalling, and insensitivity to its anti-proliferative counterpart. Evasion from programmed cell death is enabled by loss of intrinsic apoptotic pathways and/or inhibition of extrinsic ones. Gain of telomerase activity provides a "life extension" to cancer cells via circumvention of cellular senescence, which is normally induced by telomeres shortening at each replication cycle. Angiogenesis allows vascularization of the tumour, permitting access to the oxygen and nutrients needed to the growing lesion. Finally, modifying the cell-cell and cell-extracellular matrix adhesion mechanisms allows a portion of tumour cells to detach from the primary frame and use the newly built vascular connections to travel to distant organs. This is what ultimately leads to metastases, which cause most

cancer mortalities.

Depending upon whether their amplification or inactivation aids the carcinogenic process, genes found to be involved in cancer are generally defined as oncogenes or tumour-suppressing genes, respectively.

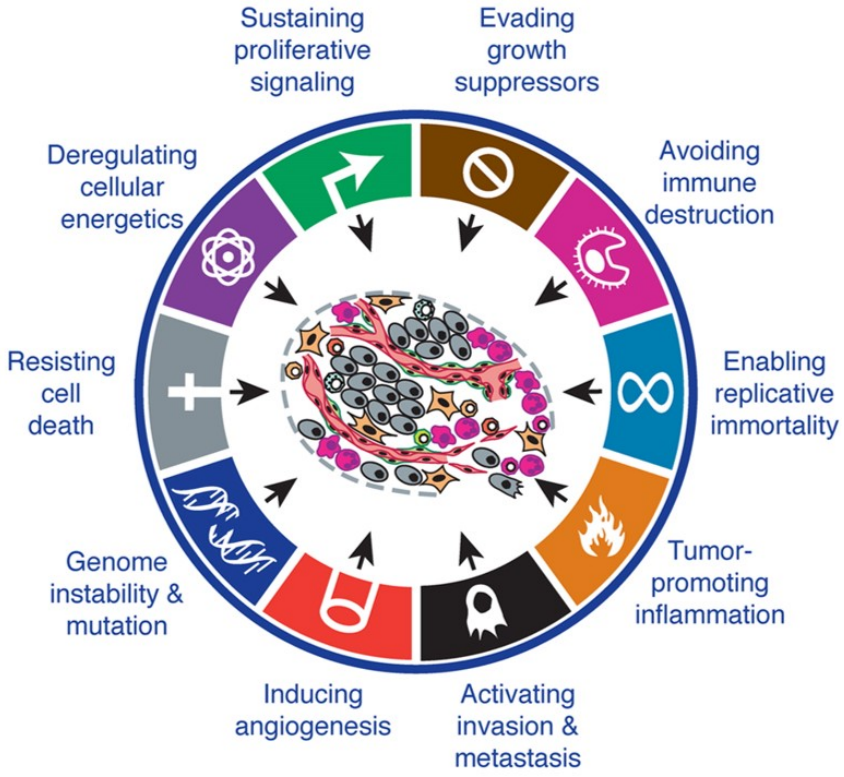


Figure 5. Hallmarks of cancer. Adapted with permission from [57], Copyright © 2011 Elsevier.

Carcinogenic transformation is initiated by factors that are caused by the environment and/or are inherited, therefore producing somatic and/or inherited mutations. These two factors, however, do not fully explain the great variability in organ-specific risk of developing tumours. Such variability is for example evident in the GI tract, where cancers of the upper GI (i.e. stomach, small intestine, etc.) are less frequent than

those of the colorectum [2,58]. The correlation of cancer incidence with increasing age and with tissue-specific rate of stem-cell division is well documented [58]. Additionally, it is known that only a portion of the genetic mutations found in any cancer are identified as *driver mutations* (those affecting the hallmarks of cancer, i.e. APC, KRAS, TP53) (**Figure 6**), while the rest are *passenger mutations*, or silent. Thus, individual cancer risk should be considered as the complex interaction of environmentally driven mutations, inherited changes and the chance of acquiring hallmarks of cancer-hitting mutations [59].

1.2.3 Adenoma-carcinoma sequence and pathways of colorectal carcinogenesis

Chromosomal and microsatellite instability are the two main forms of genetic disorders that drive tumorigenesis of sporadic colorectal cancer (**Figure 6)**. Based on discoveries on clonality and mutations underlying CRC, during the 1980s the 'adenoma-carcinoma sequence' was developed as a model of carcinogenesis in CRC [60-62]. This is a stepwise timeline of genetic and epigenetic mutations underlining the transformation of normal epithelium to adenomas, to carcinomas and metastases. (**Figure 6**).

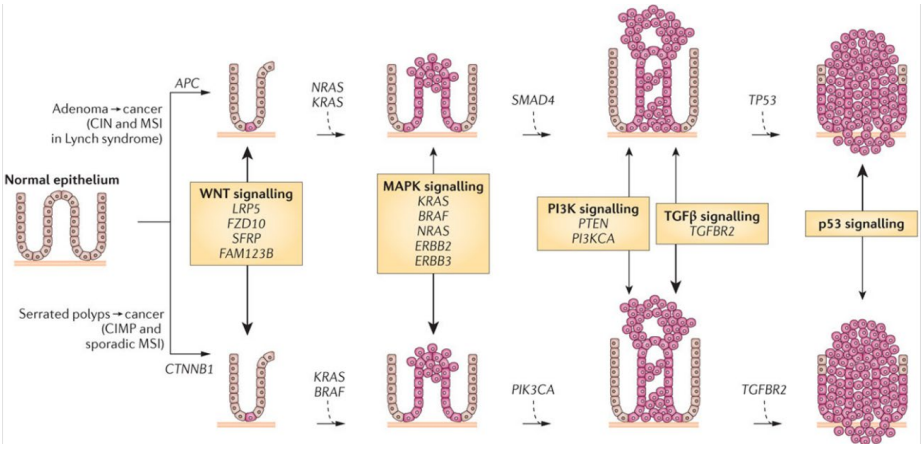


Figure 6 Adenoma-carcinoma sequence showing molecular and cellular changes. Reproduced with permission from [63], Copyright © 2015 Springer Nature.

The adenoma-carcinoma sequence is driven by chromosomal instability (CIN) and is understood today that most (up to 85-90%) but not all sporadic CRCs develop through this pathway [14,15]. CIN involves a fault during the chromosomal segregation at mitosis, and cells following this pathway are characterised by aneuploid karyotype, copy number variations and further chromosomal aberrations [64]. The earliest change of the normal intestinal epithelium is the formation of aberrant crypt foci (ACF), hyper- and sometimes dysplastic crypts that are not easily detected during normal endoscopic investigations [65]. The proliferative advantage of ACF can stem from alterations of both MAPK (mainly *KRAS* gene) and WNT (either *APC or beta-catenin*) signalling pathways. Only a small portion of ACFs develop into adenomas. Adenomatous polyps, or adenomas, are pre-cancerous lesions that are widely present throughout the population (up to 53% in individuals older than 50 years the US) and are classified according to their size and histological features [13]. More than 99% of early adenomas (<10mm in size) are benign. A small share of adenomas eventually undergoes malignant transformation into full adenocarcinomas, by losing tissue architecture and cell polarization and growing to invade the underlying *muscularis mucosae*. Adenoma-to-carcinoma switch is usually initiated by mutations and loss of heterozygosity in members of the TGF-β signalling pathway such as *SMAD4*, as well as in the tumour suppressor gene *TP53* [63].

A subset (10-30%) of CRC develops from a different type of polyps called sessile serrated polyps, which are commonly found in the proximal side of the colon and were until recently considered as unable to progress into adenocarcinomas [66,67]. Defects in DNA mismatch repair (MMR) are more commonly associated with the serrated polyps pathway and are the cause of MSI, which is reported at rates between 15 and 30% [68]. MSI is the cell's failure to repair single or multiple mismatched nucleotides during DNA replication. It can be experimentally evaluated by comparing the length of repetitive DNA sequences (microsatellites) obtained from normal and tumour tissue.

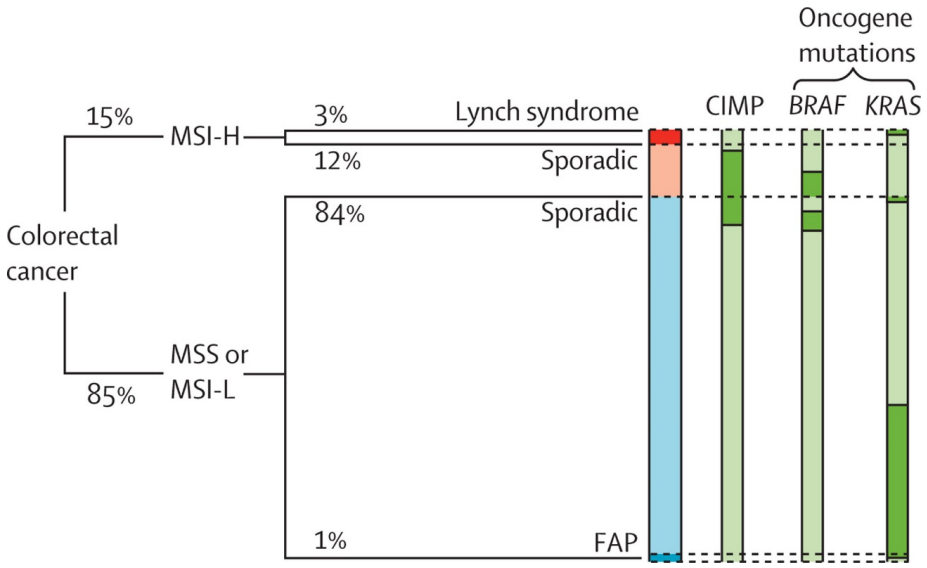


Figure 7. Pathways of colorectal neoplasia. Reproduced with permission from [14], Copyright © 2014 Elsevier.

MSI cancers are predominantly located in the proximal colon, are mucinous and undifferentiated, with a more prominent immune infiltration and generally less aggressive. Molecularly, these cancers are associated with less frequent alterations of *APC*, *KRAS* and *TP53*, but rather with activating mutations in *BRAF* oncogene and the CpG island methylator phenotype (CIMP) [69,70] (**Figure 7**). CIMP is a type of epigenetic aberration that shows extensive methylation at CpG islands within promoter regions of tumour suppressor genes, thereby inhibiting their transcription and affecting their expression [63,64,71].

Moreover, MSI is also present in >95% of *hereditary non-polyposis colorectal cancer* (HNPCC or Lynch syndrome [72]). Lynch is an autosomal syndrome that greatly increases risk to develop a series of cancers, of which CRC and endometrium cancer are the most prominent [73].

1.2.4 Consensus Molecular Subtypes (CMS) of CRC

Cancer is a rather heterogeneous disease and therefore challenging to characterize functionally. Tumour behaviour and cellular diversity are hardly derived by mutation profiles alone but are intrinsically linked by the way each singular gene is expressed and contributes to the overall picture. The advent of next generation sequencing (NGS) techniques and data sharing platforms in the past decade produced a surge in data available from gene expression studies. Multiple, large cohort studies based on gene expression profiling were produced in the past decade that provided an eye-opening insight into tumour behaviour. For the first time, CRC was being classified according to its *transcriptome* (the degree of expression of each gene), rather than its *genome*, with implications for the development of the concept of personalised medicine. These efforts produced several gene expression signatures capable of grouping most CRCs into distinct molecular subtypes, albeit with limited inter-study reproducibility [74-79]. In 2015, an international collaboration (the "CRC subtyping consortium") developed a network-based approach applying each of the algorithms to 18 pooled datasets, comprising >4000 CRCs. This effort lead to the identification of 4 common molecular subtypes (**Figure 8**), defined Consensus Molecular Subtypes (CMS1-4) of CRC [80].

The proposed CMS model identified four (1-immune, 2-canonical, 3-metabolic and 4-mesenchymal) subtypes which not only recapitulated the genetic model of CRC carcinogenesis, but considerably expanded its understanding. The CMS1-immune subtype correlates well with the MSI carcinogenic pathway, with higher mutational burden, gene methylation and immune infiltrate. The other three CMSs further subdivided CIN-driven carcinogenesis according to genomic, epigenomic and transcriptomic pathways, type and degree of immune infiltration and its interaction with surrounding stroma, and clinical features of the individual tumours (**Figure 9**) [80,81].

CMS1 MSI Immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
BRAF mutations		*KRAS* mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGFβ activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

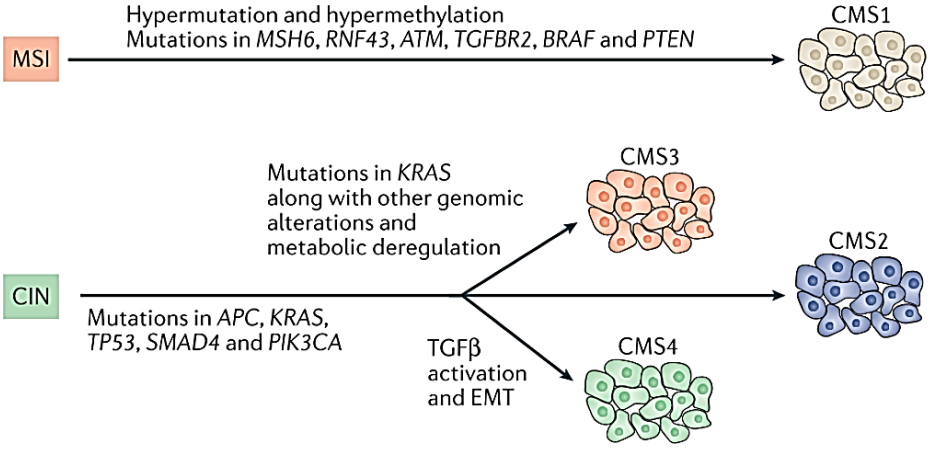


Figure 8. Consensus molecular subtypes (CMSs) and their molecular/clinical associations. Reproduced with permission from [80] (top, Copyright © 2015 Springer Nature) and [81] (bottom, Copyright © 2017 Springer Nature)

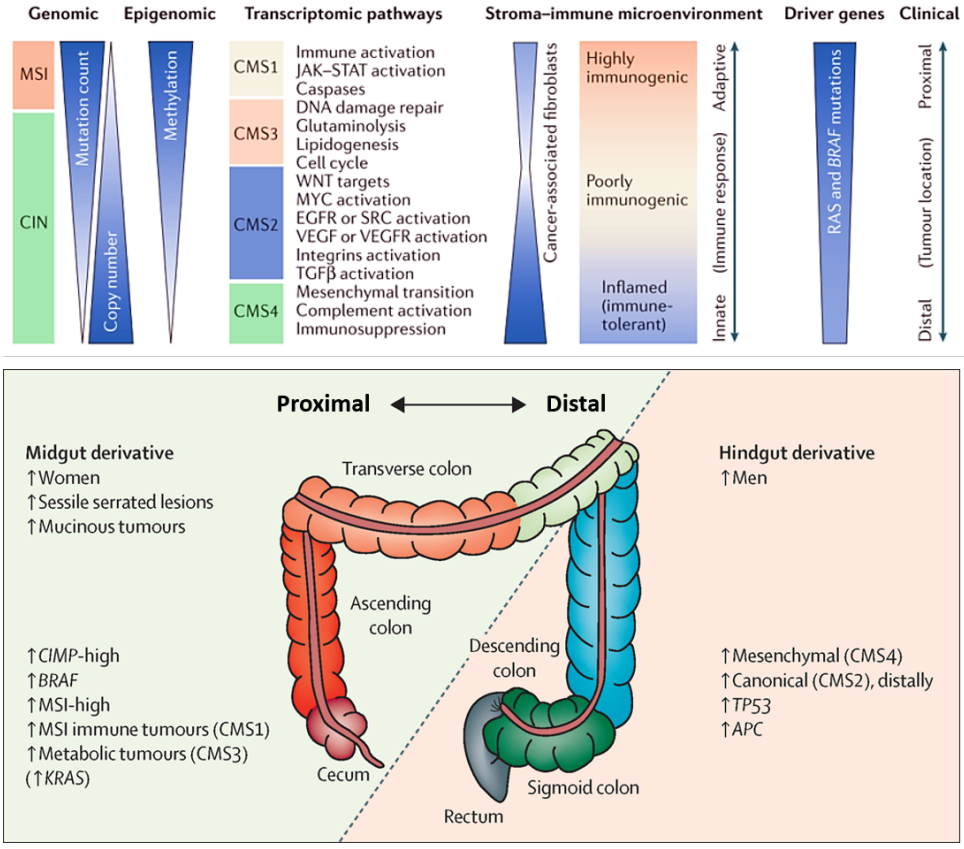


Figure 9. (Top) Schematic representation of CRC subtypes and their molecular characterisation and (bottom) their association with clinical traits. Reprinted with permission from [81] (top, Copyright © 2017 Springer Nature) and [15] (bottom, Copyright © 2019 Elsevier)

1.3 Microsatellite Instability in colorectal cancer

Microsatellites are stretches of DNA containing repetitive blocks of 1-6 base pairs (e.g. A_n, AG_n, AAG_n, etc.), abundant and ubiquitous throughout eukaryotic and prokaryotic genomes [82]. Also known as short tandem repeats (STRs) or simple-sequence repeats (SSRs), they constitute about 3% of the whole human genome, with mononucleotide repeats (A_n, C_n, etc.) being the most abundant [83]. Microsatellites in humans are mostly polymorphic by nature but conserved and of even length within each cell of the same individual, therefore useful in forensic, paternity and gene mapping applications [84,85].

In 1993, using a technique known as arbitrarily-primed polymerase chain reaction (PCR), it was discovered that tumour DNA of a small subset of CRC displayed small deletions, or band shifts, when compared to normal DNA from the same patient [86]. This was particularly true in repetitive elements of DNA, and the authors described this subset of CRC as more prevalent in the proximal colon, in younger and non-metastatic patients [86].

The same year, Thibodeau and Schaid reported increased survival and proximal location of tumours in patients demonstrating a similar pattern of instability in repetitive regions of DNA [87], which they termed microsatellite instability or MIN. Microsatellite instability was at the same time found in Lynch syndrome patients, referred to as RER (replication error) [88,89], and linked to a genetic locus subsequently identified as one of the mismatch repair (MMR) genes.

After these initial findings, several other laboratories described the same phenomenon and a variety of methods to investigate it. Workshops held by the National Cancer Institute (NCI) in 1996-7 in Bethesda (MD, USA), eventually agreed upon and endorsed guidelines [90,91]. Since the Bethesda workshop and over the course of the past two decades, a great deal of research was centred on MSI and its detection, molecular mechanisms and its clinical implications in terms of prognosis and response to therapy.

In Norway today, a CRC patient may be considered for MSI testing if diagnosed under the age of 60 or generally considered at risk for Lynch syndrome, as well as those belonging to high-risk group in stage II (and therefore a candidate for adjuvant chemotherapy) [19]. Moreover, guidelines suggest collecting MSI status of stage IV patients, as it could prove important should immunotherapy with pembrolizumab and nivolumab be approved in the near future by the health council [19].

1.3.1 Genetics of MSI and MMR deficiency

DNA replication by polymerases during S phase of eukaryotic cells is the essential mechanism that ensures that daughter cells receive the same genetic information during cell division. DNA polymerases are to some degree error-prone [92,93], and occasionally introduce nucleotide mismatches during DNA replication. Such mishaps happen at a higher rate in microsatellite sequences, owing to their increased self-complementarity and easier formation of a range of tertiary structures [85,94,95].
The mismatch repair system (MMR) is the cell's own proofreading mechanisms ensuring the fidelity of DNA replication. Single nucleotide mismatches and insertion-deletion loops (IDLs, indels) occurring due to strand slippage are rapidly recognised and corrected by this system using a variety of enzymes [96,97]. If left uncorrected, IDLs produce an equivalent loss or gain of nucleotides in the daughter cell, according to whether the IDL happened in the template or daughter strand, respectively (**Figure 10)**. Uncorrected gain/losses of nucleotides at or near protein coding sequences cause frameshift mutations, where the reading frame of a protein is altered due to insertion or deletion of base pairs, affecting protein functionality [68,70]. Several genes involved in regulation of cell proliferation, mismatch repair and apoptosis contain microsatellite sequences that are affected in MSI [64,98], potentially contributing to MSI-driven CRC development.

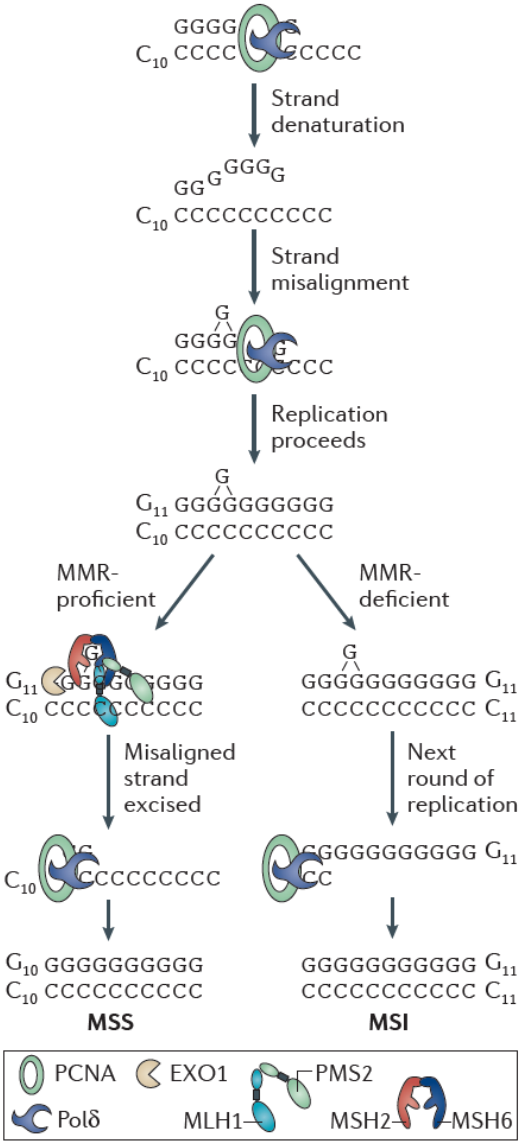


Figure 10. Mechanisms of microsatellite instability. During DNA replication, «strand slippage» (strand denaturation and mis-realignment) can occur, particularly at high-complementary sequences such as microsatellites. In an MMR-proficient cell the error is quickly corrected (see also figure 11). In an MMR-deficient environment, replication continues, producing a loss or gain in nucleotides in the newly synthesised DNA. Reprinted with permission from [73], Copyright © 2015 Springer Nature.

In sporadic CRC, members of the MMR system, particularly *MLH1* and *PMS2*, are epigenetically silenced by promoter methylation, enhancing the association of MSI with CIMP phenotype [68,99,100]. On the other hand, MSI in hereditary CRC such as Lynch syndrome is caused by germline defects in MMR genes, autosomal dominant mutations (and epimutations) that are inherited in kindreds [68,101]. Owing in part to the discoveries by 2015 chemistry Nobel laureates Modrich and Lahue on mismatch repair function in bacteria [102,103], initial cloning of the human *MSH2* gene were achieved in 1993 [104]. The authors mapped its location to chr. 2p, close to a locus that in the same year was associated with MSI in Lynch syndrome, making *MSH2* the first MMR gene associated with MSI [89,104,105]. Yeast and mammals' MMR homologue genes were cloned from bacterial *mutS* and *mutL*, thus acquiring the names *MSH* and *MLH* (*mut S/L* homologue), respectively. In humans, five *MSH* (*MSH2*, *MSH3*, *MSH4*, *MSH5*, *MSH6)* and four *MLH* (*MLH1*, *MLH3*, *PMS1*, *PMS2*) homologues were additionally cloned.

MSH2 form heterodimers with either MSH6 (to form the MutSα complex) or MSH3 (MutSβ) to scan and identify single mismatches and IDLs [64,97]. MutSα has a higher affinity for single base mismatches [106], and recruits the MutLα complex to the mismatch site. The MutSα-MutLα complex is in turn able to slide across the newly synthesised DNA until reaching the DNA polymerase. The whole complex then uses proliferating cell nuclear antigen (PCNA) and an exonuclease enzyme to backtrack and excise all the newly bound nucleotides until the mismatch is reached and corrected (F**igure 11**).

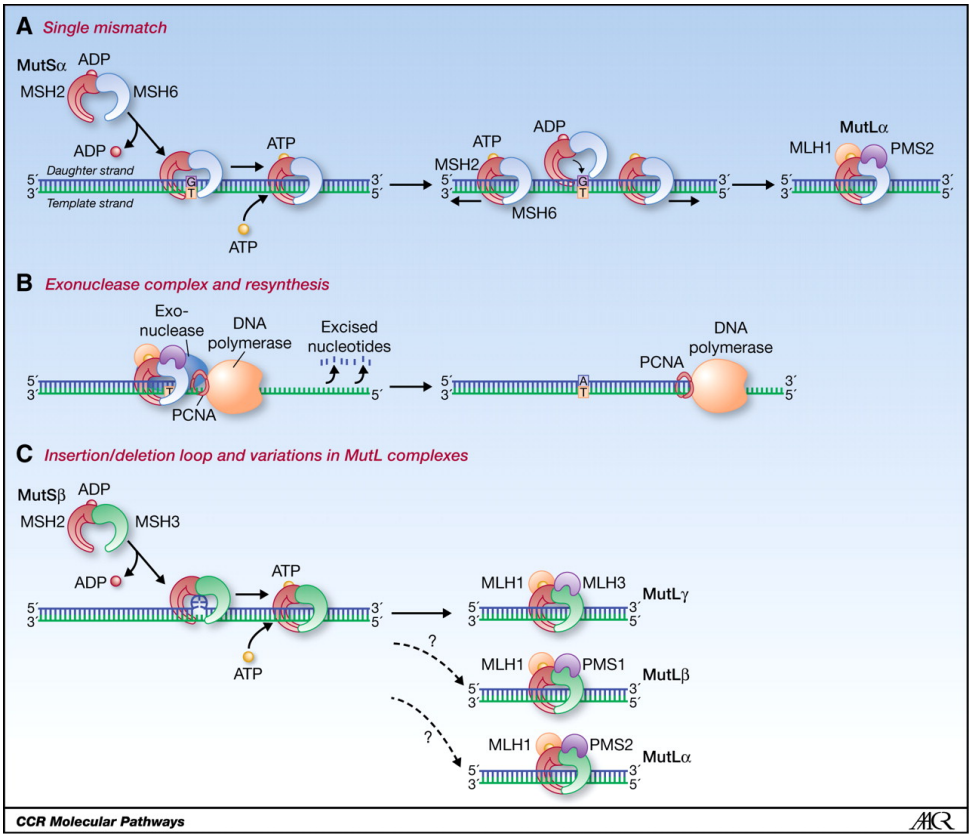


Figure 11. Different MMR heterodimers interactions and functions. (A) Single mismatch recognition MutSα and recruitment of MutLα to the site. (B) exonuclease/PCNA-mediated base excision and subsequent resynthesis by DNA polymerase. (C) Possible MMR heterodimers variations upon insertion/deletion loop formation.
Reprinted with permission from [107], Copyright © 2008 Elsevier.

1.3.2 Testing for MSI in CRC

The Bethesda guidelines defined microsatellite instability as "*a change of any length due to either insertion or deletion of repeating units, in a microsatellite within a tumour when compared to normal tissue*" [90], suggesting a panel of five microsatellite loci to assess for instability (**Table 1**). Three classes of CRC were also defined: MSI-H (high) showing MSI at ≥2/5 loci, MSI-L (low) were those CRCs exhibiting instability at 1/5 loci, and MSS (microsatellite stable) or MSI-L CRCs, where no

unstable marker was detected out of the suggested five [90].
The biological and clinical relevance of a MSI-L class is debated, as CRCs demonstrating instability at <30% of loci examined did not differ clinically from MSS ones [90]. The significance of MSI-L is still widely contested at present [108-111], often quoting the intrinsic instability of CRCs and the possibility to find all CRCs as MSI-L should enough markers be analysed [112,113]. It is thus not unusual to find MSI-L and MSS cancers grouped together in CRC studies.

The panel of mono- and dinucleotide markers suggested by the Bethesda group has been the subject of debate and revision, with a panel of five mononucleotides-only, of quasimonomorphic distribution, being considered as more specific for MSI-H (**Table 1**) [112,114,115]. The new panel was also put forward to address concerns that dinucleotide repeats may be more specific to the MSI-L group than MSI-H, therefore once again highlighting discordance upon the subject. Notably, in the revised Bethesda guidelines [112] it was suggested to shift the threshold of MSI-H detection in the new mononucleotide panel from 2/5 (40%) to 3/5 (60%) markers. The revised quasimonomorphic panel is widely used today, albeit with the 40% threshold

Table 1. Original (Bethesda) and revised (Quasimonomorphic) MSI panels

Panel	Marker	Locus	GenBank accession number	Gene	Type of repeat	Repeat motif
Bethesda original						
	BAT-25	4q12	U62834	c-kit	mono	(A)25
	BAT-26	2p21	U41210	MSH2	mono	(A)26
	D5S346	5q21-22	NM_005669	APC	di	(CA)26
	D2S123	2p16	Z16551	MSH2	di	(CA)21
	D17S250	17q11.2-12	X54562	NF1	di	(CA)24
Quasimonomorphic						
	BAT-25	4q12	L04143	c-kit	mono	(A)25
	BAT-26	2p21	U41210	MSH2	mono	(A)26
	NR-21	14q11	XM_033393	SLC7A8	mono	(T)21
	NR-24	2q11	X60152	ZNF-2	mono	(T)24
	NR-27	11q22	AF070674	BIRC2	mono	(A)27

Technically, both PCR and immunohistochemistry (IHC) for members of the MMR systems have been used, alone or in conjunction, to test for MSI. PCR is performed

using primers annealing at regions spanning either end of each microsatellite region in the five markers panels, and the resulting lengths of the amplified regions are compared between normal and tumour DNA. Immunohistochemistry is a colorimetric technique that can be used to show expression (or lack thereof) of proteins such as MLH1 and MSH2 in tumour tissue. Lack of MLH1 expression coincides in most cases with a defective MMR and consequently with MSI.
Recently, the European society of molecular oncology (ESMO) systematically reviewed the current knowledge on definitions and methodologies of MSI testing [116]. Among the study conclusions, the authors recommended the use of IHC for MLH1, MSH2, MSH6 and PMS2 as the primary method of MSI testing. In the case of difficult interpretation of IHC results, the more sensitive PCR method was recommended, and particularly with the use of the quasimonomorphic mononucleotides panel.

1.3.3 MSI as a prognostic and predictive biomarker

The notion that MSI represents a marker of less aggressive disease was noted immediately [87], and the panel held some years later in Bethesda suggested further exploration in the matter [90]. In 2005, a systematic review of 32 studies and >7500 patients with CRC concluded that patients with MSI had a significantly better overall survival (OS) than MSS ones [117]. The exact mechanisms by which the immune system modulates MSI tumour growth and aggressiveness are yet to be fully understood. It is, however, widely accepted that the higher immune infiltration in MSI CRCs is the biological basis for their prognostic advantage over MSS cancers. Colorectal cancer arising through a defective MMR system harbours a notably higher number of frameshift and other mutations that produce truncated, dysfunctional proteins [111,118,119]. Upon degradation, fragments of such proteins (termed neoantigens), are presented on the surface of tumour cells by major histocompatibility complex (MHC) molecules (**Figure 12**). By co-binding to MHCs on tumour or antigen-presenting cells, immune T-cells initiate maturation and become activated, with the whole process eliciting a specific immune reaction [118,120,121]. This is for

example the case of *TGFβRII*, a gene containing an A_8 microsatellite targeted by MSI within its coding sequence. The truncated peptides arising from its frame-shifted translation have been shown to interact with the population of immune cells infiltrating the tumour [122,123]. At the same time, MSI-associated mutation in *TGFβRII* was also associated with improved survival in stage III CRC [124].

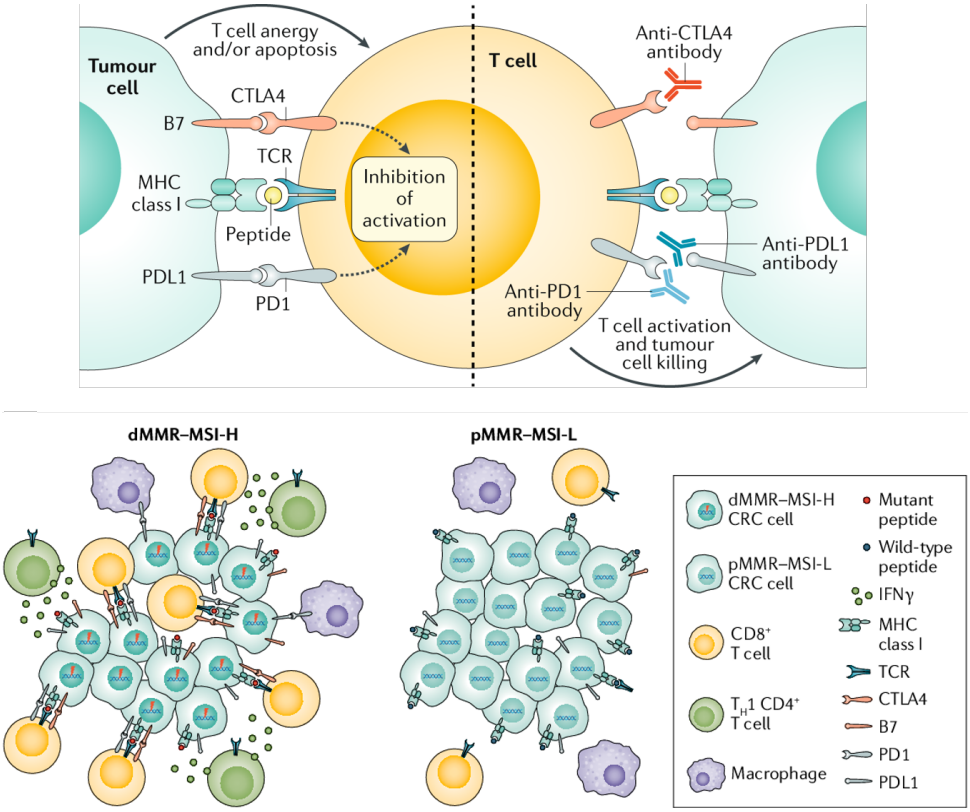


Figure 12. Top: Mechanism of MHC-mediated T-cell activation/inhibition of activation by immune checkpoints (PD1:PD-L1, CTLA4:B7) and immunotherapeutic checkpoint inhibition. Bottom: differences in tumour-immune microenvironment between deficient- and proficient-mismatch repair CRCs (dMMR and pMMR, respectively). Adapted from [125], copyright © 2019 Springer Nature.

Other than overall survival, MSI in CRC is thought to represent a prognostic biomarker of increased recurrence-free survival (RFS), too, as highlighted in reports

and meta-analyses [126,127]. This observation holds true even when stratifying for stage II [128-130], which is important to define high-risk patients that may benefit from adjuvant treatment. The beneficial effect of MSI on RFS has also been reported as conditional to clinicopathological features such as proximal location [131,132], suggesting that right and left-sided MSI cancer behave differently.
Some reports highlighted how the small proportion of MSI CRCs that develop recurrence do so preferentially at loco-regional (colorectum, peritoneum) rather than distant (lung, liver) sites. OS is usually worse in these patients [130,133]. This is likely due to the higher rate of perforating tumours, eligibility issues for metastatic curative resection and the intrinsic resistance to common 5FU-based chemotherapies.

The role of MSI as a predictive biomarker is multifaceted. When evaluated against traditional regimens of adjuvant chemotherapy, it is generally accepted that MSI patients do not benefit from 5FU adjuvant therapy, as opposed to MSS patients [107,117,134-136]. Resistance of MSI CRCs to 5FU-based regimens has been attributed to the overexpression of enzymes such as thymidylate synthase, a precursor of DNA synthesis specifically targeted by the drug [137,138]. Treatment of stage III CRC is therefore mostly based on oxaliplatin. A study on stage III patients showed an improved RFS in MSI patients when irinotecan was supplemented to 5FU and leucovorin [132].
Preclinical studies showed improved sensitivity of MSI tumour cells to irinotecan and mitomycin C, but the results need validation in clinical studies [139]. Notably, the lack of precise data on MSI as a predictive marker of response to all the available chemotherapy warrants further investigation especially in stage II CRCs, where adjuvant therapeutic intervention is decided on a risk basis [140].

The same augmented immune response that is believed to improve prognosis of MSI CRC has recently proven of predictive value, too. At stage II, MSI CRCs have a better prognosis, however those that do progress to metastasise show poorer survival [141], as well as an increased expression of immune checkpoints regulators such as programmed death receptor 1 (PD1), its ligand PD-L1 and cytotoxic T lymphocyte antigen 4 (CTLA4) (**Figure 12)** [125,142]. This observation prompted the

notion that MSI cancers would respond better to immunotherapy, as recently confirmed in the case of PD-1 and PD-L1 blockade immunotherapy [143], not limited to CRC [144]. At present, the Food and Drug Administration (FDA) in the US approved the use of anti-PD1 immune checkpoint inhibitors nivolumab and pembrolizumab in MSI CRCs, whilst the European medicine agency (EMA) has yet to evaluate the results of phase III studies [125].

1.4 MSI in tetranucleotides (EMAST)

After the initial discovery of microsatellite instability in 1993, researchers looked at a variety of microsatellites (mono-, di-, tri-, tetra-, penta-, and hexanucleotides) across a range of cancers. During the Bethesda meeting on MSI held in 1997, a variant of MSI concerning exclusively tetranucleotides was reported and defined as Elevated Microsatellite Alterations at Selected Tetranucleotides, or EMAST [90]. Unstable tetranucleotides were by then found in lung, liver, head and neck, bladder, cervical, prostate, breast, colorectal and other cancers [145-151]. EMAST, however, was not described in its molecular mechanisms and relationship to the better-defined MSI and was only briefly mentioned in the Bethesda report.

Tetrameric repeats are approximately half as abundant as monomeric ones, and their density is higher in intronic and intergenic than exonic regions (**Figure 13**) [83]. Exonic tetranucleotides densities are highest at chromosomes 7 and 22, and the most represented tetramers throughout the genome are AAAT, AAAG and AAAC [83].

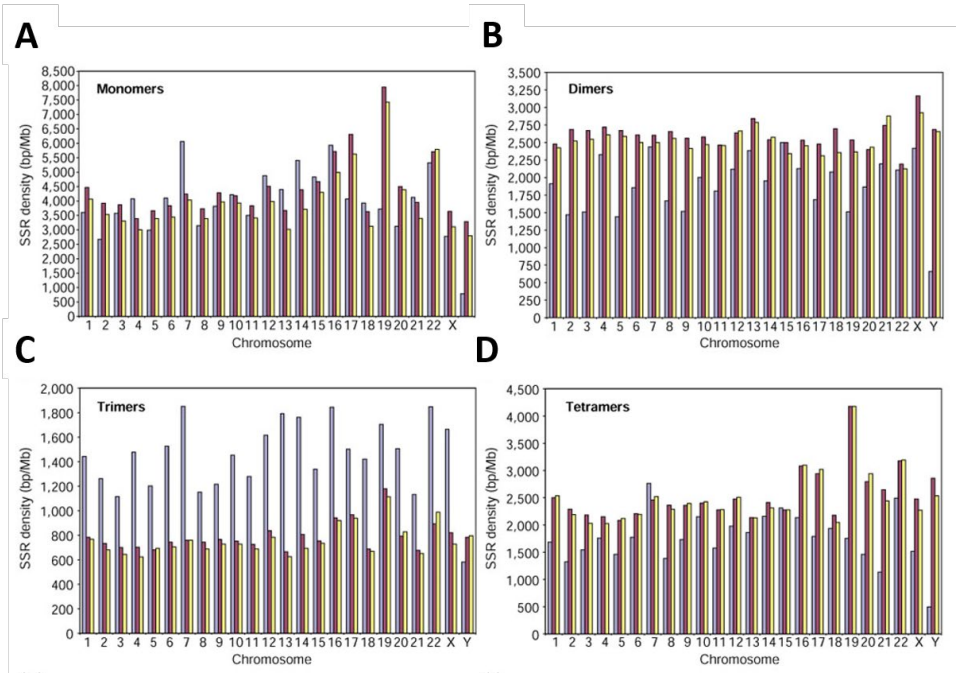


Figure 13. Microsatellites density in exonic, intronic and intergenic regions on individual human chromosomes. (A) Monomers (B) dimers; (C) trimers; (D) tetramers. Blue bars, exons; red bars, introns; yellow bars, intergenic regions. Adapted under creative commons (CC BY 4.0) from [83].

Over the course of the past two decades, an increasing number of reports pointed to the presence of EMAST in a wide range of solid tumours, with diagnostic and pathological implications (**Table 2**, expanded from [152], **appendix 1**).

Unlike MSI and despite the increasing number of reports linking EMAST to prognosis, no consensus exists to date upon which tetranucleotide markers and what cut-offs to use for its determination. Most of the EMAST markers used are found in non-coding regions, and it has been shown that tetranucleotides found in exonic/coding regions are not commonly unstable [153]. A group of five markers, namely MYCL1, D20S82, D20S85, D9S242, D8S321, together with the two of five (40%) unstable markers as a cut-off is mostly used.

1.4.1 EMAST in non-colonic cancers

In non-small cell lung cancer (NSCLC), instability at tetranucleotides was initially investigated in the early 2000s. EMAST was found at frequencies ranging from 35% to 65% of cases, associated with TP53 mutations and lacking phenotypical similarities with MSI [154,155]. Further, in two studies reporting clinicopathological associations, EMAST was shown to correlate significantly with nodal metastases and squamous differentiation in one [156], and worse 5-year OS with no effect on DFS in the other [157]. In three of the above-mentioned NSCLC studies EMAST was assessed as instability in at least one of 10-13 markers , whilst the last study grouped microsatellite assessment of di- and tetranucleotides together [156].

EMAST was found also in tumours of the reproductive system, such as ovarian (13%) [158], endometrial (39%) [159] and prostate (3-25%) [145,160-162] cancer, usually with little or no overlap with MSI.

Urinary tract cancers were also reported to display EMAST at incidences slightly higher in the bladder than in the kidney. In bladder cancer, association of EMAST with TP53 mutation is discordant [163,164], and association with clinical descriptives is yet to be fully described.

Other non-colonic solid tumours displaying EMAST are head, neck, skin and pancreatic ductal adenocarcinoma [155,163,165,166], and a wider overview is given in **appendix 1** [152].

Table 2. Prevalence, study size and cut-offs of EMAST in published studies.

Site	Author, year	Cohort size (n)	markers cut-off	Prevalence
Bladder	Xu et al. 2001 [155]	38	≥1/12	21%
	Danaee et al. 2002 [163]	57	≥1/7	44%
	Catto et al. 2003 [167]	89	≥1/8	45%
	Burger et al. 2006 [164]	117	≥1/10	9%
NSCLC	Ahrendt et al. 2000 [154]	88	≥1/13	35%
	Xu et al. 2001[155]	47	≥1/12	51%
	Arai et al. 2013[157]	65	≥1/10	65%
Prostate	Perinchery et al. 2000 [162]	40	≥1/4	25%
	Burger et al. 2006 [160]	81	≥1/10	5%
	Azzouzi et al. 2007 [161]	50	≥2/4	4%
Renal	Xu et al. 2001 [155]	25	≥1/12	12%
	Catto et al. 2003 [167]	71	≥1/8	23%
Head and neck	Xu et al. 2001 [155]	18	≥1/12	56%
	Temam et al. 2004 [165]	54	≥1/5	48%
NM skin	Danaee et al. 2002 [163]	61	≥1/7	75%
Ovarian	Singer et al. 2004[158]	53	≥1/6	13%
Endometrial	Choi et al. 2008 [159]	39	≥1/6	39%
PDAC	Mori et al. 2018 [166]	40	≥2/5*	45%
CRC	Haugen et al. 2008 [168]	117	≥1/7*	60%
only rectum	Devaraj et al. 2010 [169]	147	≥2/5*	33%
	Yamada et al. 2010 [170]	88	≥1/7	65%
	Lee et al. 2010/2012 [171,172]	108	≥2/5*	50%
	Hamaya et al. 2015 [173]	230	≥2/5	44%
	Venderbosch et al. 2015 [174]	183	≥2/5*	46%
	Watson et al. 2016 [175]	151	≥2/5*	23%
	Koi et al. 2016 [176]	88	≥1/7* and/OR ≤2/7 of MSI panel	62%
	Lee et al. 2016 [177]	100	≥2/5*	22%
	Chen et al. 2019 [178]	1505	≥2/5*	11%
	Torshizi et al. 2019 [179]	159	≥2/5	42%
	Kuan et al. 2019 [180]	509	≥2/5*	13%
	Mohammadpour et al. 2019 [181]	157	≥2/5*	40%

Abbreviations: NSLC non-small cell lung cancer; NM skin non -melanoma skin; PDAC pancreatic duct adenocarcinoma; CRC colorectal cancer; MSI microsatellite instability. Studies marked with * utilized the same panel of five EMAST markers as in the present study, with or without additional markers.

1.4.2 EMAST in CRC

Prevalence

EMAST, as MSI, remains most studied in CRC, where much of the available literature is concentrated. The frequency of EMAST in CRC is generally higher (up to 60%, **Table 2**) than the frequency of classical MSI (usually 15–20%) [152,170,171]. Geographically, its prevalence was investigated prevalently in Asian cohorts, such as in Japan [168,170,176,182,183] (same core cohort, prevalence ca. 60%), Korea [171,172,176,177] (22-50%) and Taiwan [178,180] (10-13%). In European [173,174] and Iranian [179,181] cohorts the recorded EMAST rate is at 40-46%. The only study in a USA-based cohort was based on rectal cancers separately, with a uniquely reported high rate of 33% for EMAST [169].

The disparity of panels and cut-offs used, especially during the early discoveries on EMAST (Japanese cohorts), lead to a great variation in terms of frequencies reported. Of note, this issue is also supplemented by variation in cohort ethnicity background which, known to affect rates of canonical MSI, might have an influence in EMAST prevalence, too.

Clinicopathological features and survival

Presence of EMAST in adenomas and early-stage CRC is not well documented, with reports ranging from 0% [172] to 33% [171]. In selected rectal cancers, EMAST was associated with advanced (stage III-IV) CRC [169]. In unselected CRC, results are discordant [178,181], with most reports lacking significant association with TNM stage. Sidedness is also not unanimous, albeit a recent, large cohort study associated EMAST with cancers of the proximal colon [178].

Histologically, patients with EMAST are often associated with poorly differentiated or mucinous (high grade) tumours [171,178,181,184], although some reports found no association with tumour grade [170,174,177]

EMAST CRCs were shown to harbour a greater infiltration of CD8+ T-cells in both the tumour and surrounding stroma [169,172], a feature usually associated with

better prognosis in MSI CRC.

The relationship of EMAST with patients' prognosis and survival is however not straightforward. On one hand, EMAST is most often reported as having no significant influence on survival [170,174,175,177,180,185]. On the other hand, some reports associated EMAST with a worse prognosis, a shorter time to recurrence and development of distant metastasis in stage II and III CRC [179,182]. These studies, however, grouped EMAST cancers with the MSI-L group, without addressing individual differences. Only one study, which is the largest to date (n >1500), reported improved disease-specific survival in EMAST-positive patients [178].

Biology of EMAST mechanisms

The biological causes of EMAST are yet to be fully understood. Whether it is simply an exacerbation of the MSI mechanism, or a distinct pathway of instability is still unclear. MLH1 and/or MSH2 are the most common genetically (Lynch) or epigenetically (sporadic) silenced proteins in mononucleotides-confined MSI. MutSα (MSH2-MSH6) outcompetes MutSβ (MSH2-MSH3) at single mismatches repair, while repair IDLs longer than 1 bp (2-8) is shared between the two complexes in the current MMR model [106,186-190].

On these bases, in 2008 Haugen and co-workers used the CRC cell line HCT116, harbouring biallelic inactivation of both *MLH1* and *MSH3* to examine stability at a series of mono and tetranucleotide repeats [168]. The authors demonstrated an increase in the stability of microsatellites analysed when restoring either MLH1 or MSH3. Restoration of the MMR gene was achieved by whole chromosome transfer (chr. 3 for MLH1 and chr. 5 for MSH3), and a complete stabilisation of microsatellite repeats was observed upon concurrent 3 + 5 transfer. Although the effect on tetranucleotide loci was similar upon restoration of either member of the MMR system, the investigators concluded that MSH3 was alone responsible for instability at tetranucleotides (EMAST). The *in vitro* observation of limited mismatch repair activity at di- and tetranucleotides in MSH3-deficient CRC cell lines led to the suggestion that MSH3 deficiency could represent a shared mechanism for both

EMAST and MSI-L.

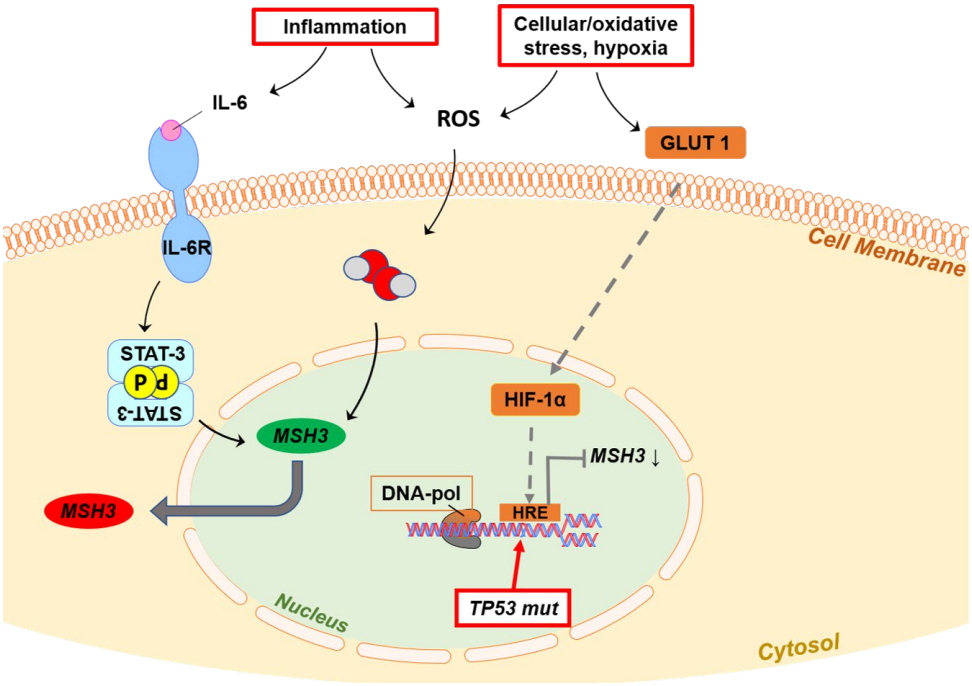


Figure 14. Putative culprits of MSH3 inactivation in EMAST, developed from data in [183,191,192]. IL-6: Interleukin-6. HIF-1α: Hypoxia-induce factor 1 alpha. ROS: reactive oxygen species.

A proposed mechanism of MSH3 inactivation involves its translocation outside of the nucleus, and consequent loss of function, mediated by a chronic inflammation state and oxidative stress [185]. Both Interleukin-6 and reactive oxygen species (ROS, mimicked by H_2O_2 supplementation) were shown to dislocate MSH3 from the nucleus in CRC cell lines [191,192]. Moreover, *in vitro* IL-6 treatment induced EMAST in otherwise microsatellite-stable CRC cells, in a process likely mediated by its downstream co-signalling molecule, STAT3 [192] (an overview of the proposed mechanisms is given in **Figure 14**).

Transcriptional downregulation of MSH3 expression was also proposed as a mechanism for EMAST. GLUT1 (hypoxic marker) was found to be overexpressed on

EMAST CRCs, which were also linked to high-rate TP53 loss of heterozygosity [183]. These observations led to the discovery of two putative hypoxia response element (HRE) in the promoter region of the *MSH3* gene. It was thus hypothesised that two isoforms of hypoxia-induced factor 1 alpha (HIF-1α) could bind to such HREs under hypoxic condition, thus driving MSH3 protein downregulation [183] (**Figure 14**).

DNA slippage in the poly(A) microsatellite contained in the coding region of the MSH3 gene, as well as single nucleotide variations, are relatively common (>30%) events in MSI CRCs [111,178]. When isolating EMAST cases (EMAST+/MSI-), no specific MSH3 mutation correlation is seen at the genetic level [178]. Epigenetic and post-translational modifications are, however, mechanism that can affect MSH3 protein functionality without showing on sequencing experiments.
Although mounting, the evidence on MSH3 involvement in EMAST is confined to *in vitro* models, as MSH3 expression in EMAST CRCs has been investigated by means of immunohistochemistry with discordant results [168,171,174,177,181,183].

1.5 The immune system and cancer

The cells and function of the human immune system are grossly subdivided into two compartments, innate and adaptive (**Figure 15**). Innate immunity uses toll-like receptors to quickly recognise and tackle a broad spectrum of infecting microbes via either cytolysis or phagocytosis. Adaptive immunity, on the other hand, relies on T-cell receptors' exposure to antigens for a slower but highly specific immune response [193,194]. Other than cells, a series of soluble molecules such as chemo- and cytokines modulate both the immune reaction and the crosstalk between innate and adaptive immunity.

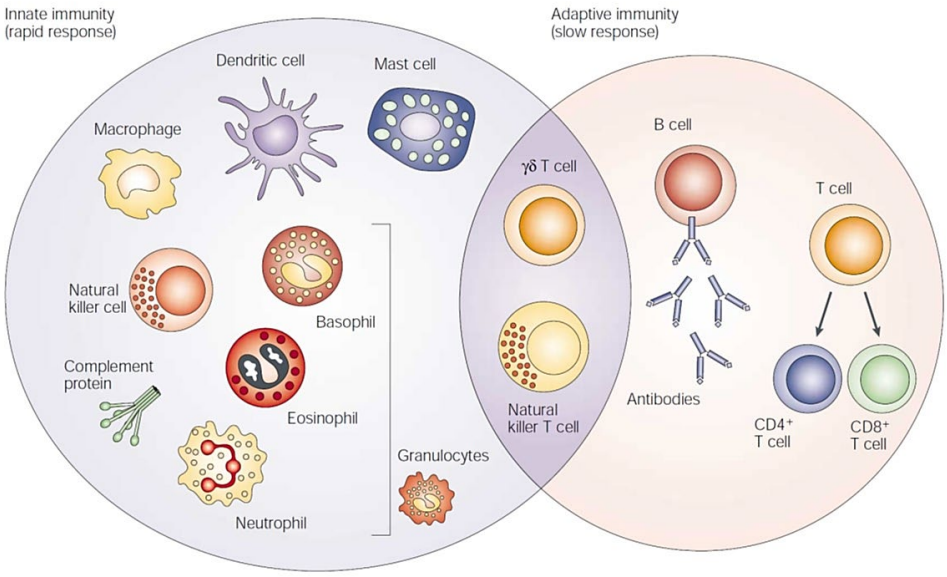


Figure 15. Components of the innate and adaptive immunity. Reproduced from [194], copyright © 2004 Springer Nature

The host immune system presents with two opposing roles towards cancer development and progression. On the one hand, infectious pathogens, prolonged exposure to irritants or failure of self-regulating immune mechanisms all induce a state of chronic inflammation that can lead to inflammation-related cancers. Examples are *H. pylori* in the stomach, cigarette smoke and asbestos in the lungs, ulcerative colitis in the colon. Chronic inflammation resulting from exposure to such stimuli increases the local infiltration by immune effector cells belonging to the innate immune system, and may lead to pre-cancerous lesions [120].

On the other hand, adaptive immunity can detect cancer-derived fragments of truncated, inactive proteins resulting from mutated and epi-mutated genes, called antigens [195,196]. Tumour antigens are presented on the surface of immune cells engulfing dying tumour cells (or by the tumour cells themselves) via MHC molecules, effectively "educating" the adaptive immune system to react towards cancer, mounting an efficient immunosurveillance[118,120]. Main effectors of the adaptive immune system are B- and T-lymphocytes, of which multiple subclasses

exist. T-cells expressing the Cluster of Differentiation 8 (CD8) on their surface are referred to as *cytotoxic T-cells*, as they are able to induce apoptosis in those cells displaying antigens that activate an immune response [197].

T-lymphocyte-mediated immune responses are initiated by the binding of their T-cell receptors to antigen-MHC complexes on antigen-presenting cells. This initial activation is complemented by a range of further receptor-ligand interactions, which can be co-stimulatory or co-inhibitory, thereby potentiating or suppressing immune response [196,198]. Owing to their higher tumour mutational burden, and thus often referred to as hypermutated, MSI CRCs are associated with an increase output of tumour antigens. Their consequent higher immunogenicity is widely accepted [199] and found confirmation in the recent consensus molecular subtyping of CRC. As mentioned earlier, these cancers were classified on the basis of their transcriptomes within the "CMS1 MSI-immune" subtype [80].

To survive, however, a tumour needs to escape and adapt to the host immune system and its multifaceted role[57]. In recent years, more light was shed on the acquired abilities of tumour cells to exploit some of the cellular mechanism fundamental to immune response to dampen or totally evade immunosurveillance (**Figure 5**). The reality of human anti-tumour immunity is extremely complex and the balance between immunity and tolerance is affected by a multitude of factors such as type of cells, molecules, metabolism, bacterial flora, as well as host and tumour's genomic make-ups [196].

The current knowledge in cancer-immune dynamics is only the "tip of the iceberg", and already offers therapeutic targets and biomarkers that are currently under evaluation.

1.5.1 PD-L1

An example of co-inhibitory mechanism parallel to antigen-TCR receptor binding are immune checkpoints, a complex system of intra- and intercellular signalling used by the adaptive immune response to decrease its magnitude. The discovery of co-inhibitory signalling of CTLA-4 [200] and programmed death receptor 1 (PD-1) [201], together with their suitability as immunotherapy targets in cancer, was awarded with the 2018 Nobel Prize in Physiology or Medicine to JP Allison and T Honjo [202].
PD-L1 (also known as B7-H1 or CD274) is a single-pass transmembrane protein constitutively expressed on antigen-presenting cells and other non-hemopoietic cells [203,204]. PD-L1 expression can also be induced by cytokines and molecules such as interferons, tumour necrosis factor α (TNF-α), and vascular endothelial growth factor (VEGF) secreted by cells of the immune system in a paracrine fashion [203].
Binding of T-cells' PD-1 to its ligands PD-L1 and PD-L2 affects signalling of a range of molecular pathways (such as RAS, mTOR and p38 pathways) leading to diminished secretion of cytokines, as well as impaired proliferation and survival [203-205]. While fundamental to immune tolerance in healthy tissue, most cancers (including CRC) have been found to express PD-L1 and thus actively suppress the host T-cell-mediated tumour toxicity (**Figure 16**) [203,205-208]. PD-L1 expression in cancer is believed to be the result of either constitutive endogenous mechanisms, or induced as an adaptive response to the strong infiltration of PD-1- and interferon-γ (IFNγ)-expressing lymphocytes [205,209].

Targeting of the PD-1/PD-L1 axis became thus of clinical relevance to overcome the resistance of a number of cancers to the surveillance of the immune system, with initial promising results in melanoma, lung and other cancers [205,207,210]. Originally a poor therapeutic target in CRC [210], PD-1 blockade therapy regained momentum after the seminal finding of the increased expression and benefits of anti-PD-1 immunotherapy in mismatch repair-deficient (dMMR)/MSI cancers [143]. These effects were seen in terms of both objective response rate and progression-free survival. As a result, multiple phase II and III clinical trials are to date ongoing to assess the benefits of anti-PD1/PD-L1 therapy in this subgroup of CRC [125,211] (**Figure 16**). As mentioned above, the FDA approved immunotherapeutic drugs

pembrolizumab and nivolumab in 2017 for second-line treatment of MSI CRC [125]. The proportion of metastatic CRC with MSI is however low (2-5%), and reliable predictive biomarkers of anti PD-L1 immunotherapies are needed to expand the CRC population benefitting from it [211].

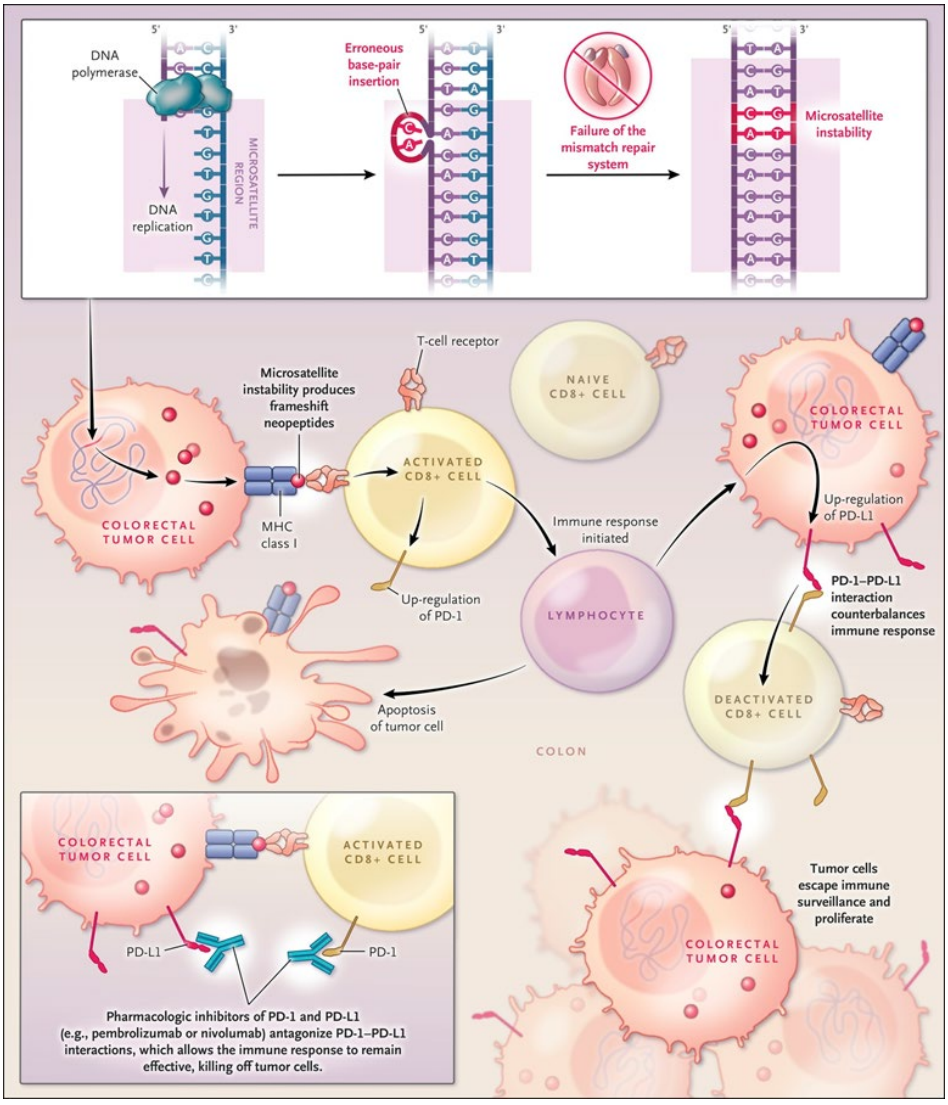


Figure 16. Tumour immune response and its regulation. MSI produces high numbers of neoantigens. Neoantigen presentation by tumour cells to TCRs on naïve T-lymphocytes. In presence of co-stimulatory binding (not shown) determines the maturation of T-cells into effector lineages, thereby inducing immunity. Tumour cells able to express PD-L1 can

induce immunosuppression of effector T-cells. Reproduced with permission from[18] , Copyright © 2018 Massachusetts Medical Society.

This is especially true in light of the development of a class of immune checkpoint inhibitors (such as atezolizumab or durvalumab, currently in phase I-III combination trials [125]) that targets PD-L1.

1.5.2 Immunoscore

Although not a new concept, the improved prognosis of CRC patients exhibiting a higher density of tumour-infiltrating T-lymphocytes has instigated further investigation in tumour immunology research in the past decade [120,212-214].

One such example is the Immunoscore® (IS), which quantifies certain immune cells infiltrating the tumour area and correlates it to prognosis in CRC [215,216]. In the initial development of the IS, expression of markers related to two lymphocyte populations (CD8, cytotoxic, and CD45RO, memory T-cells) was assessed in tumour centre (TC) and invasive margin (IM) of CRCs [217]. Aided by digital image analysis, counting positive cells in each field produces a score between I0 and I4, representing the degree of immune reaction strength (**Figure 16**) [218]. Patients with high densities of CD8 and CD45RO cells have better prognosis, with recurrence rates of I4 patients (high IS) as low as ca. 5%, and a 5-year overall survival rate of 86%. On the other hand, relapse rate of patients with low densities of CD8 and CD45RO-positive cells (low IS) was as high as 75%, with a 5-year overall survival of ca. 28% [217].

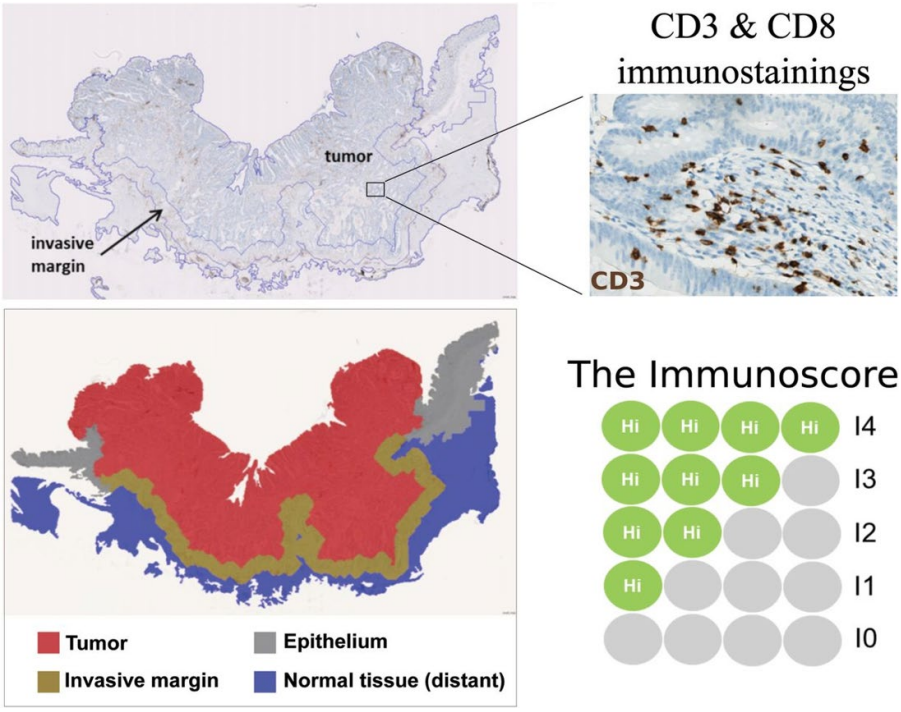


Figure 17. Immunoscore classification. The densities of CD3+ and CD8+ (immunohistochemistry, top panels) are categorized into Hi (high) or Lo (low) in each tumour region (TC and IM, lower left panel), giving four individual scores. Patients are then stratified accordingly into I0 to I4 groups, depending on the total number of high densities observed (lower right panel). Adapted from [219] Copyright © 2016 Oxford University Press.

The general T-lymphocyte marker CD3 later substituted CD45RO to obviate recurrent background staining problems related to the memory T-cell marker [218]. To further increase prognostic ability, both CD3 and CD8 are assessed separately in the tumour centre and invasive margins of each tumour [216,218,220].

Both individually and – more strongly – combined, CD3 and CD8 abundance has been shown to be consistently lower in relapsing patients and superior to the TNM staging system in terms of prognostic ability [219,221,222]. The prognostic ability of IS was also individually confirmed in rectal cancer [223].

In a recent multicentre study, IS was validated by an international consortium as highly prognostic in both uni- and multivariate analyses independent of patient age, gender, T and N stage, MSI, and other prognostic factors [220].

The scoring system was "refined" in the last report to use three (low-intermediate-high) rather than five (I0 through I4) categories. Methods, including cut-offs and mathematical operations, were also disclosed to aid reproducibility of the study [220]. The consortium advocates for the implementation of the IS as an integrative component of cancer staging, designated TNM-immune (TNM-I).

1.6 Biomarkers

1.6.1 Tumour biomarkers

The bridge between the notion of no two tumours being alike (nor can be treated alike) and the recent excitement towards "personalised medicine" are cancer biomarkers. The National Institute of Health (NIH) defines biomarkers as

A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease [224]

A broader way to define biomarkers would include any measurable change imputable to a disease that would classify patients into subgroups relevant to a clinical endpoint. According to which endpoint it serves, in the clinical settings a biomarker is often described as:

Diagnostic: as able to identify the presence/absence of a disease, therefore guiding towards a given diagnosis of a disease.

Prognostic: as able to foresee the outcome of a diagnosed disease, for example survival.

Predictive: as able to forecast the response to a therapy, as in the case of mutations in the RAS-pathway that predict no response to anti-EGFR therapies, as

the KRAS mutation is downstream of the EGFR receptor and anti-EGFR therapy is thus not able to block this pathway.

Biomarkers can also serve as surrogate endpoints in very specific and validated cases, where reaching the clinical endpoint of an ideal trial is impractical or unethical [225]. An example is using progression-free survival as a surrogate endpoint of overall survival for fluoropyrimidine use in advanced CRC [226].

Consideration of patients', tumours' and tumour biology's heterogeneity have today replaced the outdated concept of "one drug fits all" [227]. Fluoropyrimidines such as 5FU, however, are still the pillar of most adjuvant combination treatments in CRC, despite the progress in the understanding of cancer biology made since they were first proposed [228].
Thanks to the rapid advancement of technology in the field of biomedical research (e.g. high-throughput techniques such as next-generation sequencing or microarrays), large numbers of promising putative biomarkers have emerged. Nevertheless, the gap between the number of proposed biomarkers and those entering clinical practice is vast [229,230], as are the resources lost in the process. The need for reliable biomarkers reaching clinical studies is therefore increasing and is addressed by the qualitative improvement of pre-clinical ones.

1.6.2 Good practice in biomarker research

A good biomarker candidate should satisfy a series of parameters assessing its analytical and clinical validity, as well as ethical and financial implications [229,231].
Is the biomarker's test reproducible and accurate? Are sensitivity and specificity of the test satisfactory? Does detection of the biomarker directly or indirectly improve health care? Do the benefits derived from the test outweigh the financial impact on public health? These all are issues that can halt the clinical implementation of a novel biomarker along the development process and need be addressed with rigorous, sound and peer-reviewed analysis.

To reach analytical validity, arguably the greatest issue in biomarker research is reproducibility of results from pre-clinical studies. Far too often the promising results generated in academic environment fail to reproduce in more advanced stages of biomarker discovery, as for example in the case of drug targets reaching phase II stage in industry [232,233].

In 2011, a study from Bayer Health Care assessed reproducibility of 67 projects by comparing in-house obtained data with that originally generated in academic, pre-clinical studies [234]. The rate of concordance was as low as 20-25%. Similarly, only 11% of pre-clinical studies considered could be reproduced at the biotechnology company Amgen, in results published the following year [235].

Reasons for such a high rate of failure encompass everything from poorly designed trials to undetailed published methods, lack of standardisation and all the way down to inter-laboratory environmental and technical conditions [236].

Sample acquisition

In both prospective and retrospective cohort-based studies such as the present work, patient-derived substrates such as tissue and fluids constitute the primary source of study material. Sample acquisition practices, from collection to storage and further processing, can have a profound effect on the quality of nucleic acids, proteins and other molecules derived from such materials [237]. In a typical example of formalin-fixed, paraffin-embedded (FFPE) tissue samples, exposure to formalin for extended periods of time is known to induce fragmentation, cross-linking, strand breaks and denaturation of nucleic acids [238]. Factors such as time from collection to fixation and time spent in fixative (formalin) solution can even affect protein expression, and therefore dramatically affect downstream analyses, such as alignment of sequencing reads or detection of protein by IHC [239,240].

Standardisation of methods

Selection bias is a common threat in cohort and case-control studies, and is defined as "*a non-random imbalance among treatment groups of the distribution of factors capable of influencing the end points*" [241]. Even the highest standards of sample handling can be nullified *a priori* by improper or (both voluntarily and involuntarily) biased subject selection. Efforts to avoid sources of bias and to openly address them upon their presentation is thus imperative for transparency and reproducibility of scientific research.

A further issue, particularly common in explorative studies, is the standardisation of the methodology used. Gene/marker panels, cut-offs, antibodies, detection systems, laboratory consumables are but a few examples of the tools for which an ample selection is available both on the market and the scientific literature. Each can in turn add to the variability to pre-clinical experiments [236]. In the example of EMAST, over the course of the past decade several different panels and individual markers have been used to assess its incidence, with obvious implication on results (**Table 1**).

Insufficient description of the methods, or improper use of statistical methods have also become issues adding to the difficulties in reproducing results [236,242,243].

A further layer of complexity can be found in analyses that are typically evaluated by subjective scoring, such in the case of IHC. Other than the fixation/storage processes and the type of antibody/staining protocol employed, inter-operator subjectivity can affect the final results of these type of analyses [244,245].

The increasing use of automation, such as robot pipettors, autostainer instruments and digital image analysis software can however reduce technical variability and are being increasingly adopted in wet-lab practice [246,247].

Calls for quality improvement in pre-clinical studies

Transparency and accuracy of study reports are indispensable to assess their strengths, weaknesses, and impact. Scientists to clinicians to public healthcare legislators use the existing reported evidence in order to plan experiments, treatment strategies and healthcare guidelines. In other words, successful integration or

substitution of current practices depends on evidence, and an improperly reported study loses value even when it is flawlessly designed.
The EQUATOR (Enhancing the QUAlity and Transparency Of health Research) network [248] is an international initiative aimed at the implementation of robust guidelines on study reporting, to strengthen the quality and usefulness of published health research.
The initiative works as a repository for reporting guidelines tailored to specific study types, from case reports to randomised trials (CONSORT), systematic reviews (PRISMA), observational cohort studies (STROBE) and animal pre-clinical studies. Moreover, specific extensions to practices reporting such as biospecimen handling (BRISQ) and tumour biomarker prognostic studies (REMARK) are available to improve quality of reports. Guidelines commonly consist of n-item checklists for authors to fill and submit together with their manuscripts, of which we report three examples.

The REMARK (REporting recommendations for tumour MARKer prognostic studies) [249] was developed jointly by the US National Cancer Institute (NCI) and European Organisation for Research and Treatment of Cancer (EORTC). The guidelines were issued to address the disproportionally low output of clinically useful biomarkers stemming from an ever-increasing number of research efforts and reports. The REMARK checklist is aimed at encouraging authors to consider and report on a number of common issues regarding design, method and analysis of tumour biomarker and prognostic studies [250,251]. The study in **paper IV** (Watson et al., *submitted*) was submitted with the REMARK checklist.

The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist refers to the standardisation and improvement of observational study reports [252,253]. It aims at correct presentation of planning, methods and findings, and was made available in multiple languages, both as a whole (combined) checklist and as sub-checklists specific to cohort, case-control and cross-sectional studies. The study in **paper II** (Watson et al, *Ann Sur Oncol*, 2019) and **paper IV** (Watson et al. *submitted*) were submitted alongside the STROBE checklist.

The BRISQ (Biospecimen Reporting for Improved Study Quality) checklist [254] was developed to address the gap existing in EQUATOR-endorsed guidelines on handling of human biospecimens. It is thus a checklist intended to be used alongside reporting guidelines for all studies involving human biospecimens. It is of particular importance especially when considering the experimental variability deriving from non-uniform handling of patient material before downstream analysis. The Methods section of this dissertation (**3.2 Material and data collected**) are written in adherence with the BRISQ guidelines.

2. AIMS OF THE STUDY

The aim of this PhD work is to characterize a defined cohort of patients with CRC in terms of its microsatellite instability at both mono and tetra-nucleotides, as well as the type of immune cells infiltrating the tumours. The purpose is to explore potentially useful biomarkers that will aid in classification of patients into prognostic groups beyond the current TNM system. The identification of clinical and molecular traits that have an influence on patient prognosis serves to improve accuracy of cohort stratification and may allow for advanced management and surveillance.

Specific aims of this project are:

1. To investigate EMAST to MSI distribution in CRC, and the relationship with clinicopathological characteristics of patients <75 years undergoing systematic surveillance after surgery (**paper I**).

2. To investigate MSH3 expression in CRC and assess whether loss of MSH3 is related to the EMAST phenotype (**paper II**).

3. To further evaluate the clinical relevance of EMAST in a prospectively collected cohort of stage I-III CRC, including age- and comorbidity-related aspects (**paper III**).

4. To investigate the immunological context in EMAST by quantitative profiling of CD3+ immune cells and CD8+ cytotoxic T-lymphocytes and PD-L1-expressing cells in stage I-III CRC (**paper IV**).

3. METHODS

3.1 Study populations

Two different cohorts were used in the present work. In both cases, patients were recruited at Stavanger University Hospital (SUS), Stavanger (Norway). The hospital covers today a primary catchment area of approximately 370,000 inhabitants of no extreme socio-economic disparities, under the Norwegian universal health coverage system. There are no private practices or competing clinical services in the region. The study cohorts can therefore be regarded as population-representative.

The first study population (**paper I**) consists of a sub-cohort (N = 196) derived from larger (N = 314) study population of consecutive patients with stage I through III CRC who underwent surgery with curative intent with negative resection margins R_0 between 1996 and 1999. Of the 314 patients who underwent surgical resection, 196 were enrolled on a systematic surveillance program in an intent-to-treat basis and according to the guidelines of the Norwegian Gastrointestinal Cancer Group (NGICG) at the time. This excluded patients >75 years and stage III (pN+) not fit for adjuvant chemotherapy. Further exclusion from the sub-cohort was applied in the case of distant recurrence, where the patients were deemed not fit for a second surgery.
Clinicopathological information was recorded, and follow- up was updated as of July 23, 2011, thus providing up to 15 years follow-up after surgery.
The study was approved as a quality assurance project by the Regional Ethics Committee (REK) of the Health Trust of Western Norway (Helse Vest), document #2010/3414.

The second cohort (**papers II-IV**) is a sub-group of the first patients recruited consecutively in the Assessment of Clinically Related Outcomes and Biomarker

Analysis for Translational Integration in Colorectal Cancer (ACROBATICC) project (Described in [255], **appendix 2**). ACROBATICC is a prospective and consecutive population-based biobank of primary and metastatic colorectal cancer, registered with the identifier NCT01762813 (www.clinicaltrials.gov).
Patient recruitment started as of January 2013 and is to date ongoing (N>1100). Upon scheduling of curative surgery (either primary, metastatic or both), all consecutive patients aged ≥18 years of age amenable to curative-intent surgery for CRC and who could provide written informed consent were eligible for inclusion.
The study was approved by the regional ethics committee (REK, #2012/742).

3.2 Material and data collected

3.2.1 Biobanking and database building

Personal identifying information such a social security number (fødselsnummer) and names/surnames of the patients were censored via the use of a unique patient identifier stored in a password-protected file. This information was therefore not available to the analysts during data processing.
Under the ongoing ACROBATICC study, each patient is assigned an ACRO-number upon inclusion in the study via signed informed consent. Each sample of patient-derived material collected is then marked with unique material (m####) identifier numbers. Further, each processed material such as DNA extracted from tissue is also assigned a unique identifier (p####) (**Figure 18**). The vessels containing three aliquots of DNA extracted from a single patient-derived specimen would therefore each bear a label with a unique p-number, the same m-number linking to the original material from which it was derived, and an ACRO-number belonging to an individual patient.
Fresh-frozen and FFPE tissue are obtained at each round of surgery (i.e. new collection upon eventual metastatic surgery) under the ACROBATICC study. Blood, plasma and serum are collected both pre-surgery (at inclusion) and again at the first follow up visit (scheduled according to health council guidelines [19]).

For data analysis, patients' clinicopathological information was retrieved from the electronic patient record (EPR). All variables considered were appropriately classified and re-anonymised, connecting them to each patient's ACRO-number. Coding of categorical variables was performed where appropriate and all the information plotted in a password-secured SPSS file (v. 25, IBM software). Statistical methods used were described in each individual paper.

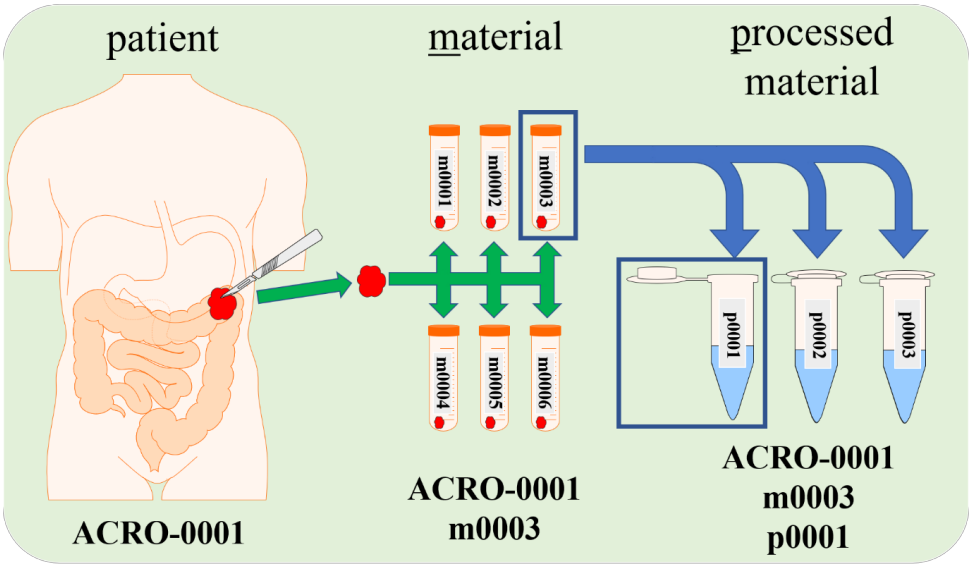


Figure 18. Example of biobanking and unique identifier assignation

3.2.2 Formalin-fixed, paraffin embedded tissue (FFPE):

Patient-derived tissue is fixated in a 10% neutral-buffered formalin solution, for a period that ranges between 8 hrs (e.g. biopsies, small specimens) to 96 hrs (e.g. hemicolectomies), according to sample size and to ensure full penetration of the fixative. Extended macroscopic analysis and excision of the tumour and normal sections are then performed by an experienced pathologist. Sections resulting from

the resection margins, deepest infiltration, tumour centre, eventual lymph nodes, eventual omentum, and normal epithelium are placed in colour-coded cassettes and embedded in paraffin. Formalin-fixed, paraffin-embedded blocks are then sectioned on a microtome to a thickness of 3-5µm and mounted on Superfrost glass slides for haematoxylin and eosin (H&E) staining, as well as eventual immunohistochemical colouring if requested by a pathologist. H&E slides are then used to confirm or eventually update the clinical staging of the tumour and provide additional microscopical information such as tumour cell content, degree of tumour infiltration, necrosis and mucinous content. Paraffin blocks are then stored in an archival room kept at 15 degrees Celsius (unstained, glass-mounted slides kept at 4 degrees) until further sectioning is required.

3.3 Techniques

3.3.1 Multiplex PCR and Fragment analysis

PCR is a technique that relies upon modified polymerases, specifically designed primers and unbound nucleotide supplementation to amplify desired stretches of DNA from a template of choice. Primers are designed to be complementary to regions flanking the segment of interest. An initial step at high temperature allows the template DNA to denature, allowing the two complimentary strands to dissociate. Thereafter, temperature is cycled constantly at specific levels to allow annealing of the primers and the ligation of the correct series of nucleotides to replicate the template DNA. The constantly cycling denaturation-annealing-extension steps and the stoichiometric abundance of both primers and unbound nucleotides allows for an exponential number of copies of DNA to be created (2^n copies at 100% efficiency, where n is the number of cycles). Length and G/C content of the primers greatly affect their melting temperature (Tm), at which at least 50% of the primers and their

complementary sequences on the template DNA are bound together. To allow for specific annealing of the primers, their careful design and relative reaction temperature conditions need to be achieved, as the energy required for each of the four nucleotides to bind to its complementary is specific.
Multiplex PCR reactions can be carried out where more than one stretch of DNA is amplified, via the use of multiple pairs of primers. The high specificity of primers' annealing temperatures therefore proves these experiments challenging in terms of optimisation. All the primers in the pool used should be designed to achieve:

- A narrow Tm range (within 3-5ºC) and a balanced G/C content to allow for simultaneous PCR annealing
- High degree of specificity to the targeted region to avoid biased amplification of one amplicon over another caused by competition of primers in the same reaction
- Lack of inter-primer complementarity, which would cause primer dimerization and consequent stoichiometric imbalances

The above-mentioned parameters are only a few of those affecting the outcome of a PCR reaction, and even established protocols need to undergo lab to lab optimisation. In the present work, both MSI and EMAST were assessed through two independent multiplex reactions, with five primer pairs each.

3.3.2 Immunohistochemistry

IHC is a technique used to spatially visualise specific antigens in a tissue (-*histo*) sample, which relies on specificity of antigen-antibody (*immuno*-) binding. To aid visualisation of antibody-bound antigens at the cell level under light microscopy, secondary antibody-linked molecules are used that precipitate upon incubation with a substrate (**Figure 19**).
After sectioning to a thickness of 2-5µm, the FFPE tissue slice is mounted on a glass slide, which cycles through a series of steps. These are:

- rehydration of the sample
- retrieval of eventually masked antigens by either heat-induced or enzymatic methods
- blocking of endogenous enzymes that can interfere with the detection system
- primary antibody binding and detection system (**Figure 19 A-E**)
- substrate addition and chromogen precipitation (**Figure 19 F-G**)
- counterstaining
- dehydration of the sample
- application of a glass cover

Sample collection, pre-treatment (fixation, embedding) and storage can however affect performance of IHC, requiring optimisation of each staining protocol.

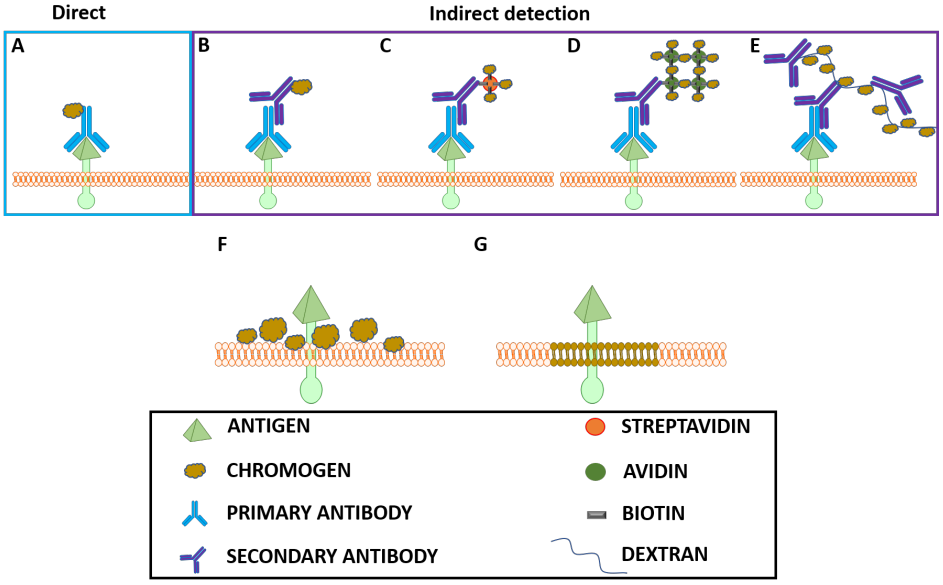


*Figure 19. Principles of immunohistochemistry. (**A**) direct antigen detection. (**B-E**) indirect antigen detection with secondary antibody conjugation methods. (**B**) secondary antibody conjugation. (**C**) Labelled streptavidin-biotin (LSAB). (**D**) Avidin-biotin complex (ABC) (**E**) Dextran backbone polymer conjugation. (**F**) Upon addition of a substrate, the chromogen precipitates. (**G**) The original antigen location is marked by the chromogen colour.*

Choice of appropriate positive (e.g. a tissue known to contain the antigen, genetic knock-in) and negative controls (e.g. isotype controls, genetic knock-outs, tissue known not to express the antigen), especially in the clinical setting, is also important [256]. Selection of antibody, pH and chemical formulation of buffers, blocking of endogenous enzymatic activity and antigen retrieval methods need often fine-tuning before a staining can be considered true [244].

The widespread use of IHC for diagnostic purposes brought the need for standardisation of methods, to avoid the multiple source of experimental variation this technique is subject to. Most protocols can therefore be largely automated today, with benefits towards reproducibility.

3.3.3 Digital Image Analysis

Traditionally used in pathology as a diagnostic tool, IHC is today increasingly employed in prognostic and predictive tests. As guidelines for treatment often rely on IHC analyses for disease biomarkers, interpretation of results should be as unequivocal as possible.

Other than the cells of interest (those where the biomarker assessed has validity), histological samples contain numerous other structures, such as different cell types, blood vessels and lymph nodes. Moreover, rarely expressed antigens and artefacts induced by for example tissue folding or poor dispersion of reagents are all disturbances which might further complicate scoring of immunostaining. Manual (visual) interpretation of IHC is therefore time-consuming and requires highly skilled and experienced pathologists. One major pitfall is however the variation and involuntary bias intrinsic in the subjectivity of the scoring process, which often affects inter- and intra- operator reproducibility [257-259].

Various semi-quantitative scoring methodologies have been proposed and are currently in place across laboratories. These are commonly obtained by multiplying a categorised score for staining intensity (e.g. absent, weak, strong) by the pure or classified percentage of cells stained. Examples are the histology-score (H-score, range typically 0-300) [260,261], staining index (SI, range 0-9) score or the Allred score,

where intensity and extent scores are added, not multiplied (range 0-8) [262]. Although helping towards standardisation of methods, these score systems still largely rely on observer accuracy and establishment of cut-offs, thus open to interpretation variability.

The development of automated and semi-automated digital image analysis software have the potential to reduce discrepancies in interpretation, and increase output [263]. Digital image analysis means using computer-based algorithms to extract quantitative and qualitative information from digitalised images. The process of digitalisation converts an image into discrete quantitative variables (e.g. pixel depth, pixel connectivity, colour vectoring such as RGB or HSV). Once converted into numerical values, the images can be pre-processed and manipulated to achieve a series of enhancements, such as background noise reduction or sharpening of colour contrasts (**Figure 20**). The enhanced image can then be segmented using a range of supervised or unsupervised methods. Image segmentation refers to the process of effectively separating objects in defined classes based on any feature, such as texture in the case of tumour-stroma separation, or colour in positive-negative IHC stain. Classification algorithms can be applied in-line (e.g. one after the other) and developed using different methods. One such methods relies on Bayes' probabilistic theorem [264], which simply stated makes decisions based on probabilities derived from the results of pre-existing data. Bayesian classification is a supervised method of segmentation, as it relies upon the establishment of a training set.
Once appropriately classified, the image can be further manipulated to aid extraction of data, by for example colour-coding the created classes. Visual aids can be implemented, by stacking additional algorithms, to select areas in an unbiased way. An example is the inspection of rarely or over-expressed antigens, where manual picking of areas of max/min positivity can prove challenging. Creating visual hotspots based on classified positive/negative labels can help selecting the right region of interest. The hotspot approach has been implemented in the Norwegian, Danish and Swedish pathology guidelines [265-267] for its documented utility in the evaluation of biomarkers of difficult interpretation, such as Ki67 in breast cancer [268].

The use of digital platforms to capture, store, share, analyse and report pathological examinations on glass slides is collectively known as Digital Pathology [269].

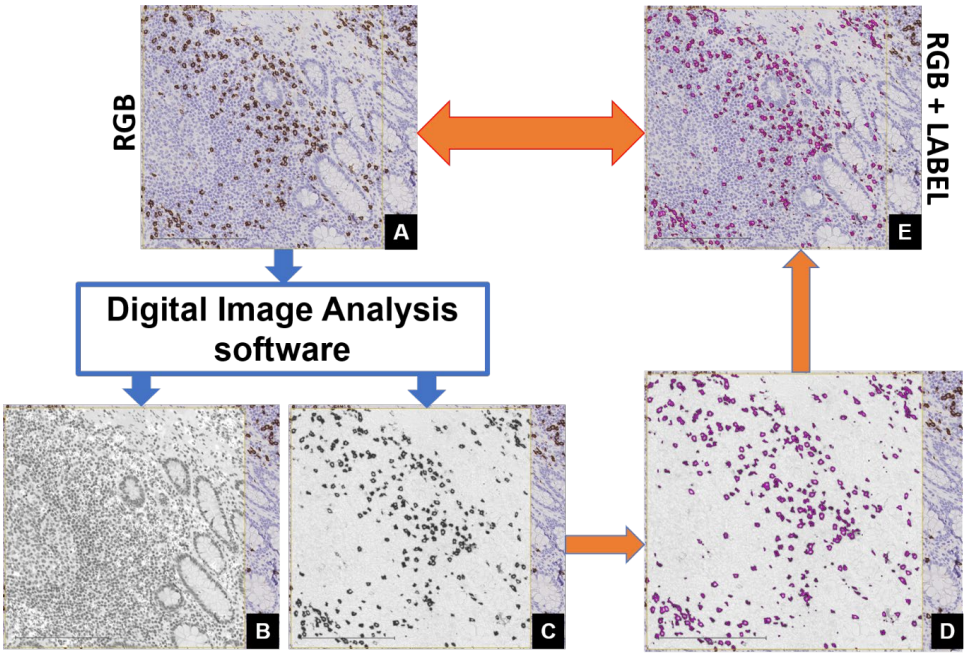


*Figure 20. Colour deconvolution of a digital image, in this case CD8 staining (DAB staining, brown) in colon tissue (**A**). Through digital image analysis, the red, green and blue (RGB) components of the image can be isolated and manipulated, to highlight colour contrasts. In this case the blue haematoxylin (**B**), or the brown DAB staining (**C**) counterstain can be highlighted to provide the training parameters to a classification algorithm. The algorithm can then overlay a custom-coloured label to the component wanted (in this case membrane-bound DAB, labelled in pink) (**D**). The result can be viewed in the original RGB format, with labelled components (**E**) and mathematical operations can be performed based on label characteristics such as area, or absolute number.*

4. RESULTS OF THE STUDY

In paper I (Watson et al., *Cancer Medicine*, 2016) we evaluated EMAST prevalence and association with MSI and other clinical and tumour-specific characteristics in a consecutive, population-based series of stage I–III colorectal cancers <75 years who were eligible for systematic surveillance according to national guidelines at the time (1996-1999). MSI and EMAST were assessed using multiplex PCR with primers amplifying two sets of five microsatellite markers. Of 151 patients included, 33 (21.8%) had MSI and 35 (23.2%) were EMAST+, with an overlap of 77% for positivity, (odds ratio [OR] 61; $p < 0.001$), and 95% for both markers being negative. EMAST was more prevalent in colon than rectum (86% vs. 14%, $p = 0.004$). EMAST+ cancers were significantly more frequent in proximal colon (77 vs. 23%, $p = 0.004$), had advanced T-stage (T3-4 vs. T1-2 in 94% vs. 6%, respectively; $p = 0.008$), were larger (≥5 cm vs. <5 cm in 63% and 37%, respectively; $p = 0.022$) and had poorly differentiated tumour grade (71 vs. 29%, $p < 0.001$). Furthermore, EMAST+ tumours had a higher median number of harvested lymph nodes than EMAST− (11 vs. 9 nodes; $p = 0.03$). No significant association was found between EMAST status and age, gender, presence of distant metastases or metastatic lymph nodes, and overall survival. Overall, survival was not influenced by the presence of EMAST, although a non-significant difference toward worse survival in node-negative colon cancers was noted.

In paper II (Watson et al., *Translational Oncology*, 2019), we sought to investigate and potentially validate whether EMAST was attributed to loss of MSH3 protein expression by IHC. Results from fragment analysis of multiplex PCR used to assess MSI and EMAST were cross-examined with MSH3 protein expression, using digital heatmap-derived hot spot image analysis (digital pathology). Of 152 patients, EMAST was found in 50 (33%) and exclusively in the colon. Most EMAST-positive cancers had instability at all five markers, and EMAST overlapped with MSI-H in

42/50 cases (84%). The most frequently altered tetranucleotide markers were D8S321 (38.2% of tumours) and D20S82 (34.4%). Subjective evaluation of MSH3 expression by IHC in tumour cells found <10% negative cells in all samples, most being <5% negative. Digital analysis improved the detection but showed a similar spread of MSH3 loss (range 0.1-15.7%, mean 2.2%). Hotspot MSH3 negativity ranged between 0.1 to 95.0%, (mean 8.6%) with significant correlation with the whole slide analysis (Spearman's rho = 0.677 $p < 0.001$). Loss of MSH3 expression did not correlate with EMAST in our study, thus rejecting the causal mechanism hypothesis.

In paper III (Watson et al., *Annals of Surgical Oncology*, 2019) we analysed a population-based, consecutive sub-cohort of surgically treated stage I–III CRC patients, derived from the first 200 included in the ACROBATICC study. MSI and EMAST, together with clinicopathological characteristics and associations thereof were reported as OR and survival was presented as hazard ratios (HR) with 95% CIs. Of 161 patients included, 25% were aged >79 years EMAST (31.7%) and MSI-H (27.3%). We found a large overlap in the prevalence of the two types of microsatellite instability, where 82.4% of EMAST were also MSI. EMAST had the highest prevalence in the proximal colon (OR 15.9, 95% CI 5.6–45.1; $p < 0.001$) and in women (OR 4.1, 95% CI 1.9–8.6; $p < 0.001$), and were poorly differentiated (OR 5.0, 95% CI 2.3–10.7; $p < 0.001$). Compared with EMAST-negative patients, EMAST-positive patients were older (median age 77 vs. 69 years; $p < 0.001$), leaner (median weight 67.5 vs. 77 kg; $p = 0.001$), had significantly higher rates of hypoalbuminemia (24% vs. 6%; OR 2.3, 95% CI 1.5–3.6; $p = 0.002$) and anaemia (45% vs. 20%; OR 3.3, 95% CI 1.6–6.8; $p = 0.001$), and had elevated preoperative C-reactive protein (CRP) levels (51% vs. 34%; OR 1.9, 95% CI 1.0–3.9; $p = 0.046$). Improved recurrence-free survival was found in both MSI and EMAST subtypes. In multivariable analysis, node status (pN+), together with elevated CRP and MSI-positive, were the strongest prognostic factors for recurrence-free survival. We thus found EMAST in CRC to be associated with an older, leaner, and frailer phenotype with a lower risk of recurrence.

In paper IV (Watson et al., *submitted*) the relationships between EMAST and PD-L1, CD3 and CD8 expression in the invasive margin or tumour centre (Immunoscore) were investigated in relation to risk of disease recurrence. A total of 149 stage I-III CRCs patients, with a median follow up of 60.1 months were included, with recurrence-free survival (RFS) as the main endpoint. Difference in survival between groups were assessed by log rank test with univariate Cox regression for hazard ratios (HR) and 95% confidence intervals. Patients with PD-L1+ tumours (7%) were older (median 79 vs. 71 years, p = 0.045) and more likely to have EMAST+ cancers (OR 10.7, 95% CI 2.2-51.4, p = 0.001). Recurrence-free survival was better in cancers with PD-L1+ immune cells (HR 0.35, 95%CI 0.16-0.76, p = 0.008, independent of EMAST) and high Immunoscore (HR 0.10, 95%CI 0.01-0.72, p=0.022). Patients with PD-L1+ immune cells showed also better disease-specific survival (HR 0.28, 95%CI 0.10-0.77, p = 0.014).

Only tumour-centre (but not peritumoral-immune) PD-L1 expression correlated with EMAST. Lymphocytic infiltrate and peritumoral rather than intratumoral PD-L1 expression had prognostic value in CRC.

5. DISCUSSION

Despite the constant improvements in health care and disease management, CRC registers almost 900,000 deaths every year, its rates are steadily increasing, and it is projected to reach 2.5 million new cases in 2035. Improving accuracy of early diagnostic tools and patient stratification through reliable biomarkers is therefore of high importance. The prognostic and predictive value of biomarkers such as microsatellite instability and the concert of tumour-immune interaction, as they get better elucidated, show promise in this advancement. Although applying to relatively small subsets of the CRC population, ca 20% MSI cases out of 2.5 million in 2035 still account for 500,000 patients, among which may be those who benefit from targeted therapies. Should EMAST, occurring at higher rates of both incidence and instability, prove of equal clinical value in the future, the effect may be even greater. It is thus important to keep on generating data that help resolve patient heterogeneity.

5.1 The patient with EMAST in CRC

We found EMAST incidences of 22% and 32% in our two cohorts. While the latter is more in line with other reports on EMAST in CRC, the lower incidence in the archival (**paper I**) cohort may be attributable to both the younger age (<75 years) and the higher prevalence of rectal cancers. In both cohorts we indeed associated EMAST with colon rather than rectum, a finding concurrent with other studies [174,177,178,184]. The reported incidence of EMAST is subject to the range of markers analysed and cut-offs used in individual analyses. Initial reports described EMAST as being present at high rates in CRC (>60%) [168,170]. The authors, however, used a low cut-off

(1/7 markers) and therefore identified high numbers of EMAST-positive patients [168]. EMAST was also initially associated with the MSI-L subgroup of CRCs [168], however in the CRC population analysed by the authors both MSI-H and MSI-L overlapped completely with EMAST. It is thus unclear how the same association was not drawn between EMAST and MSI-H. Moreover, as mentioned earlier, there is debate upon whether MSI-L truly represents a subclass of CRC, and rather a consensus that it does not [111]. Notwithstanding the more frequently reported overlap with MSI-H phenotype [174,175,177,178], the MSI-L/EMAST group is still sometimes reported in studies [176,185].
A further point that may add to the variation in reported frequencies of EMAST is the ethnicity of the cohorts in which it is analysed. As discussed in **papers II and III**, is possible that EMAST, as canonical MSI, follows different patterns of prevalence across demographic characteristics [270,271], thus registering varying rates across cohorts. Should EMAST be investigated as a distinct form of microsatellite instability, a standardisation of panels and cut-off is warranted.

In our **paper III**, we identified a series of pre-operatively recorded variables indicative of an elderly, frailer phenotype that associated with EMAST. These included older age, a lower body mass index (BMI) with loss of >5% body weight prior to diagnosis, lower levels of haemoglobin and serum albumin.
Older patients generally constitute a population affected by a range of comorbidities, impairments and syndromes that may have an influence on the outcome of clinical interventions [272]. Frailty is often recognised as an aging-associated syndrome, corresponding to a weakened state and susceptibility to adverse health events [273,274]. Several geriatric assessment and frailty scoring systems have been proposed to date, none however reaching unanimous consensus [272,275,276]. A generally undernourished state, with specific biomarkers including BMI, low levels of albumin and haemoglobin and weight loss have been reported to predict frail status and poor outcomes, even against more extensive geriatric assessments [272,277-282]. Genomic instability – or the accumulation of genetic and epigenetic aberrations - tends to increase and share mechanisms with biological and physiological aging [283,284]. It is thus possible that EMAST, intended as an extension of the genetic perturbance otherwise reflected by canonical MSI, may help drawing frailer subgroups of

patients. EMAST was shown in paper III to correlate with improved RFS, although no differences in disease-specific and overall survival were noted. As >25% of the cohort was older than 79, older age and frailty might constitute modifiers in survival analyses, as indicators of an overall shorter lifespan.

5.2 Role of MSH3

Association of MSH3 loss with EMAST stems from effects seen in experimental studies on cancer cell lines [191,192]. These led to the postulation of nuclear MSH3 protein inactivation by extranuclear translocation as the leading cause of EMAST. In human CRC four studies showed some degree of MSH3 loss found in EMAST-positive cancers [168,171,177,181]. We have however identified issues in the findings in two of the studies. In one report, an initial analysis yielded no difference in median of MSH3-negative cells in EMAST groups, and the cut-off value was then altered thus introducing objective bias in the analysis [177]. In the second report [181], the figure chosen to represent an MSH3-negative case, claimed to correlate with EMAST, showed normal colonic tissue. Normal colon should always express MSH3 unless germline inactivating mutations are present and such findings undermine the assay's validity.

Other groups found no significant association between MSH3 loss and EMAST in CRC [174] and pancreatic ductal adenocarcinoma [166].

In our study on immunohistochemical expression of MSH3 (**paper III**), we aimed at validating whether MSH3 rather than MSH2 and MLH1, commonly involved in MSI, could drive EMAST. No striking downregulation of the protein was however found, with very low degrees of nuclear staining heterogeneity. Contribution to results discrepancy may stem from the fact that all studies employed different antibodies, where reported, and cut-offs to discriminate between MSH3 negative and positive

cases. Again, lack of standardisation in assays and methods results in great variability.

A further issue to consider is that MutSβ (MSH2-MSH3) is erroneously thought to have higher affinity for insertion-deletion loops (IDL) repair than MutSα (MSH2-MSH6) [285]. In reality, while MutSα was shown to have higher affinity for single nucleotides mismatches by the group of Paul Modrich [106], the affinity of either complex was shown to be equivalent for 2-8 nucleotides IDLs [106,186-190].
Both MSH2 and MLH1, the two proteins most reported as dysfunctional in canonical MSI are also integrally part of the MutSβ (MSH2-MSH3) and MutL(α-β-γ) (MLH1-PMS2/PMS1/MLH3) heterodimeric complex (see **Figure 10**). According to the running theory of MSH3-deficiency being the culprit of instability at tetranucleotides, MutSβ and its cooperation with either one of the above-mentioned MutL should therefore be affected. The association, albeit apparently logical, is inconsistent with the role of MSH2 and MLH1 and PMS2, which could equally contribute to the lack of repair of IDLs longer than a single mismatch. Assuming – possibly erroneously – that all human Mut complexes and their functions have been discovered and described, each one of MSH2, MSH3, MLH1, PMS1, PMS2 and MLH3 could equally destabilise repair activity at tetranucleotide repeats.
An additional issue is that the variation in amounts of MutS alpha and beta in human eukaryotic cells are attributed to available pools of MSH3 and MSH6, to which MSH2 binds to form either complex. Overexpression of MSH6 could therefore sequester available amounts of MSH2, stoichiometrically shift binding towards the MutSα complex. In such case, MSH3 would still be detected by IHC (as in our paper III) but would lack functionality unincorporated in MutSβ due to unavailability of MSH2. To date, no report described expression of MSH6 in relation to EMAST.
Moreover, as for MSI, EMAST molecular mechanisms should be held true regardless of tissues. In pancreatic cancer EMAST was found in up to 40% of patients, but no evidence of MSH3 inactivation was found in those tumours [166]. In endometrial cancer, another malignancy known for elevated EMAST frequencies [159], *MSH3* frameshift mutations were found to be not as common as in CRC [111]. Lastly, near-ubiquitous overlap of EMAST with MSI, as pointed out recently in a meta-analysis

[174], suggests overlapping rather than mutually exclusive mechanisms. Taken together, all these observations at least question the proposed exclusive role of MSH3 in EMAST development.
Better insight into the role of each MutS-MutL combination complexes in MMR mechanisms, as well as concomitant analysis of each individual protein in a multi-organ tumour cohort, could shed more light into EMAST relationships with MMR.

5.3 Tumour microenvironment

To be able to reach a level of evidence sufficient for incorporation in clinical guidelines, biomarkers need to be validated thoroughly in randomized controlled trials and subsequently in systematic reviews after "proving their worth" in cohort studies.
Constitutive activation due to mutations in the KRAS gene is to date the most reliable predictive marker in CRC, informing clinicians on the usefulness of anti-EGFR therapies such as cetuximab and panitumumab [286]. Compared to KRAS, evidence on PD-L1 expression as a predictive biomarker is still in its infancy, albeit encouraging. It is therefore valuable to add to the body of evidence surrounding the role of PD-L1 in CRC and interaction with the surrounding microenvironment. The tumour microenvironment was barely considered two decades ago, while search on PubMed for the terms returns >4000 publications in 2018. The reciprocal influence of tumour and its surrounding stroma, particularly components of the immune system, is an important dynamic to elucidate.

In **paper IV,** we described a generally higher immune reaction in the tumour area of EMAST patients, exemplified by higher numbers of CD3 and CD8 cells, and their collective indicator, IS. Higher levels of C-reactive protein (CRP), a systemic

indicator of active immune response, was also described in **paper III.** CRP is an acute phase protein produced by hepatocytes in response to secretion of pro-inflammatory factors such as interleukin 6 [287]. Its levels are increased in plasma during infection, inflammatory disease, trauma and cancer.
These findings were in line with a study reporting higher density of CD8 T cell infiltrations in both tumour centre and stroma of EMAST patients [172].

Expression of PD-L1 on cells from various tumours is believed to be either inducible by IFNγ secreted by tumour-infiltrating lymphocytes, or constitutively activated via oncogenic signalling [205]. These two distinct mechanisms have been termed innate and adaptive immune resistance, respectively [205,209,288]. It was thus proposed that further stratifying tumours for both PD-L1 expression and tumour-infiltrating lymphocytes (termed TILs) may have superior predictive value for response to PD-1 blockade immunotherapies [209,289,290]. In this model, termed Tumor Immunity in the MicroEnvironment (TIME) [290,291], only the group showing both PD-L1-positive tumour cells and presence of TILs (*PD-L1+/TILs+*) would benefit from immunotherapy. The other three groups (*PD-L1+/TILs-*, no immunosurveillance and innate immune resistance; *PD-L1-/TILs+*, TILs not producing IFNγ; and *PD-L1-/TILs-*, lack of immunosurveillance and response to immunotherapy) would not [290].
In **paper IV** we showed a good correlation between EMAST and tumour PD-L1 expression, which coupled with the higher IS (and therefore TILs) could align with the narrative of generalised instability at microsatellites and increased immunogenicity. According to the TIME classification, it may be then speculated that EMAST predicts good response to immune checkpoint blockade. The number of PD-L1-positive tumours were however too low to draw sub-stratification according to both PD-L1 and TILs in our cohort, something that by expanding the analysis to larger samples might be achievable. No clinical trial so far, however, included the TIME classification as a predictive biomarker of immunotherapy response .[125]

We also describe an additional pattern of PD-L1 expression in CRC, with distinct implications for patients' prognosis, namely PD-L1 expressed on cells outside of the tumour. More recent focus emerged on the role of PD-L1 expression in

tumour-infiltrating immune cells [142,292-294]. As part of normal immune dynamics, PD-L1 is expressed on antigen-presenting mononuclear cells such as macrophages. PD-L1 is been described in peri- and intratumorally infiltrating immune cells in CRC, with distinct clinical and pathological connotations than on tumour cells [292]. First, immune cells expression of PD-L1 is consistently higher than in tumour cells [292,293,295-299], even when selecting for MSI patients. There is therefore a substantial proportion of microsatellite-stable CRC that display high levels of PD-L1-expressing immune cells.

Second, PD-L1 expression on immune cells is almost unanimously associated with improved recurrence-free [297,299-301] and overall survival [293,299,300,302]. Association of tumour PD-L1 expression is usually linked to poorer survival, although not consistently [292,300,302,303]. Data on expression of PD-L1 in the immune component of the tumour microenvironment is nonetheless scarce, and prevalently available in Asiatic (Japanese, Korean) cohorts. It is necessary to establish the role of PD-L1-expressing immune cells to better understand the tumour-immune system dynamics. This is especially true for clinical trials of PD-1 blockade, where an anti-PD-L1 antibody would be direct towards both tumour and immune cells expressing its antigen.

It is known that a very small proportion (circa 5%) of metastatic colorectal cancer show MSI [125,211,304]. At present, anti-PD1/PD-L1 therapy is only approved by the FDA as second-line treatment in metastatic MSI CRC (pembrolizumab, nivolumab or nivolumab in combination with anti-CTLA-4 ipilimumab). Only two trials are addressing safety and efficacy of immune checkpoints inhibitors in non-metastatic CRC, NCT03026140 and NCT02912559 (ATOMIC trial). The former, an open label randomised trial [305], already showed promising results with 7/7 major pathological responses and 4/7 complete responses in the MSI group. The ATOMIC study is a phase III randomized trial comparing standard chemotherapy (FOLFOX) alone or in combination with atezolizumab (anti-PD-L1) as adjuvant treatment for stage III MSI colon cancer.

Should a causality be established between immune resistance and progression to

metastases of MSI CRCs and PD-L1 expression, the latter may serve as stratification tool to identify "higher-risk" MSI CRCs. This in turn may strengthen the rationale for the use of anti PD1/PD-L1 immunotherapy in non-metastatic CRC.

Finally, an additional consideration should be made on the underrepresentation of older, frailer patients from clinical trials, as in the case of immune checkpoints inhibitors [306]. As discussed earlier, in **paper III** we found that EMAST patients were associated with indicators of frailty and older age. We are not aware of other studies supporting the association of EMAST with clinical markers of frailty. A recent multicentre retrospective analysis, however, found no difference in survival (overall and progression-free) nor in immune-related toxicity between patients older or younger than 70 years [307]. It thus of interest to gain further insight to gauge whether the benefits of inclusion in such trials might outweigh the risks for older patients.

5.4 Methodological considerations

Several points related to limitations in the presented works need addressing.
First, the size of both cohorts (N<200) are relatively small, especially when considering rare events such as MSH3 negative cells or PD-L1 positive ones. While the cohort used in **paper I** is of archival, retrospective nature, only a sub-cohort of the ACROBATICC study has been employed in the other publications. Since 5-years RFS and OS were the main endpoints considered in our analysis, limiting the sampling to the first 200 patients ensured acceptable median follow up lengths (52.5 months in **paper III** and 68.8 in **paper IV**).
Moreover, it is important to explicit criteria for patient selection to disclose possible

selection bias, a recurrent issue in both clinical trials and cohort studies. By selecting for patients <75 years and excluding stage III (pN+) not fit for adjuvant chemotherapy, as well as those deemed not fit for metastatic surgery, a bias towards the "fittest" patients might have been introduced. Notably, while no change in the treatment recommendations from the Norwegian health authorities took place for rectal cancers between 1996 and 1999, the introduction of neoadjuvant radiotherapy for advanced stage was introduced in the year 2000 [308]. For colonic cancers, on the other side, adjuvant chemotherapy (5FU monotherapy) for stage III patients aged 75 or less was introduced in 1997, with documented benefits in terms of 5-year relative survival [308].
Guidelines for systematic surveillance and regimen specifications for neoadjuvant (rectum) and adjuvant treatment have evolved considerably in the past two decades in Norway [19]. It may be thus further speculated that the discrepancies in survival noted from EMAST in **paper I** (n.s.) and **papers III-IV** (improved RFS) might reflect improvements in standard of care.

Analysis of MSH3 in **paper II** also need commenting. To date, two main functional mechanism of MSH3 inactivation have been proposed, one involving dislocation to the cytoplasm [191,192], and the other via downregulation at the transcription level [183]. Neither could be noted in our results, as MSH3 staining looked strong and specific, and negligible nuclear-cytoplasmic heterogeneity was noted. However, it should be noted how other aberrations might affect MSH3. A 2012 study identified 25 sequence variants (not somatic mutations), of which none resulted in a truncated protein, although inaccurate splicing of MSH3 as a consequence could not be excluded [309]. Moreover, the authors identified two regions where loss of heterozygosity was frequent in the tumours analysed, located at the far end of the coding sequence of the MSH3 gene. The anti-human MSH3 clone EPR4334(2) (AbCam, Cambridge UK) used in the study is a recombinant monoclonal antibody raised in rabbit, targeted to the 50-150 AA. region of the N-terminus of the MSH3 peptide. Of all the antibodies used in previous studies, it was at the time the only commercially available and optimised for IHC by the manufacturer. Being however

raised against the peptides corresponding to the initial sequence, similarly to the antibody used in the afore-mentioned study, downstream protein modifications might have been missed by our antibody.

The analysis of PD-L1 expression in the peritumoral compartment, simplified in the text of **paper IV** as related to immune cells, is incomplete. PD-L1 is known to be expressed in activated cells including T- and B-cells, dendritic cells, macrophages, mast cells, natural killer cells, activated vascular endothelial cells and mesenchymal stem cells [204,208]. Our analysis was however limited to PD-L1, lacking additional cell-specific markers and thus not allowing us to speculate upon which type of non-tumour cell expressed it in our samples. More in-depth marker analysis should be done to resolve the components of the tumour microenvironment in our samples.

Finally, a limitation that deserves note in our paper III is the lack of standardised measure of frailty in our CRC cohort. We have shown the association of EMAST with markers that have been often associated in the literature, although no values pertaining any of the frailty score systems so far proposed were reported. This is partly due, as mentioned earlier, to the lack of consensus on frailty assessment. Only the American Society of Anesthesiologists (ASA) classification system was reported for our cohort, as the one routinely assessed and recorded in preoperative settings. Should the association of EMAST with markers of frailty hold true in expanded and more extensive studies, it should be tested against more standardised scores, such as the geriatric assessment (GA) [310] or the frailty index [276].

6. CONCLUSIONS AND FUTURE PERSPECTIVES

In these works, we investigated EMAST in CRC, for the first time in a Norwegian cohort, and established its relationship with patients' characteristics that are routinely recorded for their value in the clinical setting. The EMAST patients often have instability at mono- other than tetranucleotides, are older and show markers of frailty, although with a better outlook in terms of recurrence-free survival. EMAST was also an indicator of a stronger immunosurveillance activity, exemplified by the increased presence of CD3 and CD8 expressing immune cells. Further, as for MSI tumours, EMAST cancers were linked with immune evasion, by increased PD-L1 expression.

Owing to the great degree of overlap between MSI at mononucleotides and EMAST in our cohorts, the two mixed groups EMAST-/MSI+ and EMAST+/MSI- were underrepresented in our studies. It was thus not possible to achieve a level of statistical significance in order to carry out extensive subgroup analysis. Whether these two types of instability should be studied independently, or as a mutually propagating phenomenon therefore still needs further investigation. To date, EMAST was investigated only once in a large cohort. The ACROBATICC cohort is to date ongoing and counts >1100 patients, with an average of 22 specimens from each patient, including FFPE and fresh tissue, serum and plasma. In operable metastatic patients the number of samples collected doubles, all in highly standardised and optimal conditions. The ACROBATICC biobank represents therefore an invaluable source for future projects.

Statistical power would benefit from the expansion of the analyses performed in this study to increasing number of patients, to finally allow for separation of rare subgroups such as EMAST+/MSI-, EMAST-/MSI+ as well as the tumour PD-L1+. EMAST and MSI showed combined association with improved RFS, CD3 and CD8 immune infiltration and PD-L1 tumour expression. Specific subgroup analyses would

allow to establish whether these associations are dependent or independent from EMAST and MSI reciprocal status.
Moreover, mutational status of markers such as KRAS, BRAF and TP53, commonly reported in clinical CRC studies, was not available outside of the cases where its assessment fell under diagnostic prescription, and is thus not included in the analysis. Genome- and transcriptome-wide association studies are necessary to truly link the clinical, genetic and biological contexture of each CRC patient in the ACROBATICC cohort. With the technical and cost-effectiveness evolution of next-generation sequencing platforms, and expansion of project-allocated resources, these high-throughput analyses are possible.

In the context of the tumour microenvironment, a future goal of this project should be to better characterize the cells infiltrating the EMAST tumour area, such as those expressing PD-L1. IHC has evolved considerably to allow for highly multiplexed staining of tissue samples, by coupling with mass cytometry (CyTOF) [311,312]. The full implementation of digital pathology via digital image analysis algorithms, increasingly developed on the backbone of artificial intelligence [259], is also an exciting near-future perspective. Standardised and reliable biomarker analyses are needed to establish any "common denominator" able to identify subgroups of patients to be investigated prospectively in clinical trials. Examples are the above-mentioned immune checkpoints inhibitor trials in early stage CRC, or trials that are targeted at the improvement of metastatic CRC health care. Of the over 400 people diagnosed with CRC yearly in Norway, circa 20% present with synchronous metastases and an additional 15-20% will develop metachronous ones [19]. Implementation of improved peri- and preoperative techniques increased possibilities for treatment of metastatic CRC considerably in the past decade, although no more than 20-30% of patients are actually eligible for surgery. The present work focused exclusively on stage I-III patients, but planning is underway to extend analyses to ACROBATICC patients with hepatic metastases. These include collaborations with research groups both within and outside of Stavanger University Hospital, improving quality and dissemination of results to a larger audience.

Last, the basic molecular biology research should not be neglected as the understanding of its mechanisms of action is at the base of every reputable biomarker. Patient material is invaluable and modern biomedical sciences offer incredible opportunities. Primary cell cultures can be established *ex vivo* that give the opportunity of studying tumour cells belonging to the very same patient included in the study. This can be done in 3D structures such as spheroids or organoids which, albeit not perfect models of *in vivo* disease, constitute an incredible advancement compared to 2D monolayers cultures. Three-dimensional cultures and co-cultures with cells typically found in the tumour microenvironment offer a better insight into cell-cell interactions, diffusion gradients of nutrients or therapeutics, and real-time cellular changes.

The “clinical endpoint” in translational research is improvement of patient care, and its “surrogate endpoint” is the deepening of our understanding of tumour biology. All the tested hypotheses, analyses, expended funds and hours in a laboratory or a library must converge on this goal.

7. REFERENCES

1. 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: A Cancer Journal for Clinicians.* 2018;68(6):394-424.
2. Norway CRo. *Cancer in Norway 2018: Cancer incidence, mortality, survival and prevalence in Norway.* Oslo: Cancer Registry of Norway; 2019 2019.
3. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nature Reviews Gastroenterology & Hepatology.* 2019.
4. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 8.2 (26.03.2019). Association of the Nordic Cancer Registries. Danish Cancer Society; 2019. Available from http://www.ancr.nu. Accessed 24/06/2019.
5. Safiri S, Sepanlou SG, Ikuta KS, et al. The global, regional, and national burden of colorectal cancer and its attributable risk factors in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Gastroenterology & Hepatology.* 2019;4(12):913-933.
6. O'Keefe SJD. Diet, microorganisms and their metabolites, and colon cancer. *Nature reviews Gastroenterology & hepatology.* 2016;13(12):691-706.
7. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut.* 2017;66(4):683–691.
8. Feng YL, Shu L, Zheng PF, et al. Dietary patterns and colorectal cancer risk: a meta-analysis. *European journal of cancer prevention.* 2017;26(3):201-211.
9. Magalhães B, Peleteiro B, Lunet N. Dietary patterns and colorectal cancer: Systematic review and meta-analysis. *European journal of cancer prevention.* 2011;21:15-23.
10. Potter JD. Colorectal Cancer: Molecules and Populations. *JNCI: Journal of the National Cancer Institute.* 1999;91(11):916-932.
11. Tabung FK, Brown LS, Fung TT. Dietary Patterns and Colorectal Cancer Risk: A Review of 17 Years of Evidence (2000-2016). *Curr Colorectal Cancer Rep.* 2017;13(6):440-454.
12. Drew DA, Cao Y, Chan AT. Aspirin and colorectal cancer: the promise of precision chemoprevention. *Nature Reviews Cancer.* 2016;16:173.
13. Strum WB. Colorectal Adenomas. *New England Journal of Medicine.* 2016;374(11):1065-1075.
14. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet.* 2014;383(9927):1490–1502.
15. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *The Lancet.* 2019;394(10207):1467-1480.
16. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology.* 2010;138(6):2044-2058.

17. de la Chapelle A. Genetic predisposition to colorectal cancer. *Nature Reviews Cancer.* 2004;4(10):769-780.
18. Sinicrope FA. Lynch Syndrome–Associated Colorectal Cancer. *New England Journal of Medicine.* 2018;379(8):764-773.
19. Norwegian health council. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av kreft i tykktarm og endetarm. Available from: http://www.helsedirektoratet.no/retningslinjer. Published 2019. Accessed 09/05/2019.
20. Holme Ø, Løberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial. *JAMA.* 2014;312(6):606-615.
21. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *The New England journal of medicine.* 1993;328(19):1365-1371.
22. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet.* 1996;348(9040):1472-1477.
23. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet.* 1996;348(9040):1467-1471.
24. Tinmouth J, Lansdorp-Vogelaar I, Allison JE. Faecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut.* 2015;64(8):1327-1337.
25. Lauby-Secretan B, Vilahur N, Bianchini F, Guha N, Straif K. The IARC Perspective on Colorectal Cancer Screening. *New England Journal of Medicine.* 2018;378(18):1734-1740.
26. Nagtegaal ID, Quirke P. What Is the Role for the Circumferential Margin in the Modern Treatment of Rectal Cancer? *Journal of Clinical Oncology.* 2008;26(2):303-312.
27. Hwang MR, Park JW, Park S, et al. Prognostic Impact of Circumferential Resection Margin in Rectal Cancer Treated with Preoperative Chemoradiotherapy. *Annals of Surgical Oncology.* 2014;21(4):1345-1351.
28. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet.* 2018;391(10125):1023-1075.
29. Norway CRo. *Cancer in Norway 2017: Cancer incidence, mortality, survival and prevalence in Norway.* Oslo: Cancer Registry of Norway; 2019 2018.
30. Dukes CE. The classification of cancer of the rectum. *The Journal of Pathology and Bacteriology.* 1932;35(3):323-332.
31. Amin MB. *AJCC Cancer Staging Manual, Eight Edition.* 8th ed. New York, NY: Springer; 2017.
32. Quirke P, Williams GT, Ectors N, Ensari A, Piard F, Nagtegaal I. The future of the TNM staging system in colorectal cancer: time for a debate? *The Lancet Oncology.* 2007;8(7):651-657.

33. Nagtegaal ID, Quirke P, Schmoll HJ. Has the new TNM classification for colorectal cancer improved care? *Nat Rev Clin Oncol.* 2012;9(2):119–123.
34. Lea D, Haland S, Hagland HR, Soreide K. Accuracy of TNM staging in colorectal cancer: a review of current culprits, the modern role of morphology and stepping-stones for improvements in the molecular era. *Scandinavian journal of gastroenterology.* 2014;49(10):1153-1163.
35. O'Connell JB, Maggard MA, Ko CY. Colon Cancer Survival Rates With the New American Joint Committee on Cancer Sixth Edition Staging. *JNCI: Journal of the National Cancer Institute.* 2004;96(19):1420-1425.
36. Hari DM, Leung AM, Lee J-H, et al. AJCC-7(TH) Edition Staging Criteria for Colon Cancer: Do the Complex Modifications Improve Prognostic Assessment? *Journal of the American College of Surgeons.* 2013;217(2):181-190.
37. Chen SL, Bilchik AJ. More extensive nodal dissection improves survival for stages I to III of colon cancer: a population-based study. *Ann Surg.* 2006;244(4):602–610.
38. Chang GJ, Rodriguez-Bigas MA, Skibber JM, Moyer VA. Lymph Node Evaluation and Survival After Curative Resection of Colon Cancer: Systematic Review. *JNCI: Journal of the National Cancer Institute.* 2007;99(6):433-441.
39. Soreide K, Nedrebo BS, Soreide JA, Slewa A, Korner H. Lymph node harvest in colon cancer: influence of microsatellite instability and proximal tumor location. *World J Surg.* 2009;33(12):2695–2703.
40. McDonald JR, Renehan AG, O'Dwyer ST, Haboubi NY. Lymph node harvest in colon and rectal cancer: Current considerations. *World J Gastrointest Surg.* 2012;4(1):9–19.
41. Berg M, Guriby M, Nordgard O, et al. Influence of microsatellite instability, KRAS and BRAF mutations on lymph node harvest in stage I-III colon cancers. *Mol Med.* 2013.
42. Shen SS, Haupt BX, Ro JY, Zhu J, Bailey HR, Schwartz MR. Number of Lymph Nodes Examined and Associated Clinicopathologic Factors in Colorectal Carcinoma. *Archives of Pathology & Laboratory Medicine.* 2009;133(5):781-786.
43. Veen T, Nedrebo BS, Stormark K, Soreide JA, Korner H, Soreide K. Qualitative and quantitative issues of lymph nodes as prognostic factor in colon cancer. *Dig Surg.* 2013;30(1):1–11.
44. Carethers JM. One colon lumen but two organs. *Gastroenterology.* 2011;141(2):411-412.
45. de Santa Barbara P, van den Brink GR, Roberts DJ. Development and differentiation of the intestinal epithelium. *Cellular and Molecular Life Sciences CMLS.* 2003;60(7):1322-1332.
46. Barker N, van Es JH, Kuipers J, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature.* 2007;449:1003.
47. Jung P, Sato T, Merlos-Suárez A, et al. Isolation and in vitro expansion of human colonic stem cells. *Nature Medicine.* 2011;17:1225.

48. Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology.* 2013;15:19.
49. Clevers H. The Intestinal Crypt, A Prototype Stem Cell Compartment. *Cell.* 2013;154(2):274-284.
50. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. *Nature reviews Gastroenterology & hepatology.* 2011;8(2):90-100.
51. de Lau W, Barker N, Low TY, et al. Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature.* 2011;476(7360):293-297.
52. Rothenberg ME, Nusse Y, Kalisky T, et al. Identification of a cKit(+) colonic crypt base secretory cell that supports Lgr5(+) stem cells in mice. *Gastroenterology.* 2012;142(5):1195-1205.e1196.
53. Clevers H, Batlle E. SnapShot: The Intestinal Crypt. *Cell.* 2013;152(5):1198-1198.e1192.
54. Sugimoto S, Ohta Y, Fujii M, et al. Reconstruction of the Human Colon Epithelium In Vivo. *Cell Stem Cell.* 2018;22(2):171-176.e175.
55. Sato T, Vries RG, Snippert HJ, et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature.* 2009;459(7244):262-265.
56. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70.
57. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell.* 2011;144(5):646–674.
58. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science.* 2015;347(6217):78.
59. Luzzatto L, Pandolfi PP. Causality and Chance in the Development of Cancer. *New England Journal of Medicine.* 2015;373(1):84-88.
60. Fearon ER, Hamilton SR, Vogelstein B. Clonal analysis of human colorectal tumors. *Science.* 1987;238(4824):193-197.
61. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *The New England journal of medicine.* 1988;319(9):525-532.
62. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990;61(5):759-767.
63. Kuipers EJ, Grady WM, Lieberman D, et al. Colorectal cancer. *Nature Reviews Disease Primers.* 2015;1:15065.
64. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology.* 2008;135(4):1079–1099.
65. Takayama T, Katsuki S, Takahashi Y, et al. Aberrant Crypt Foci of the Colon as Precursors of Adenoma and Cancer. *New England Journal of Medicine.* 1998;339(18):1277-1284.
66. Bettington M, Walker N, Clouston A, Brown I, Leggett B, Whitehall V. The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology.* 2013;62(3):367-386.
67. Patai AV, Molnár B, Tulassay Z, Sipos F. Serrated pathway: alternative route to colorectal cancer. *World journal of gastroenterology.* 2013;19(5):607-615.

68. Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology.* 2010;138(6):2073-2087.e2073.
69. Advani SM, Advani P, DeSantis SM, et al. Clinical, Pathological, and Molecular Characteristics of CpG Island Methylator Phenotype in Colorectal Cancer: A Systematic Review and Meta-analysis. *Transl Oncol.* 2018;11(5):1188-1201.
70. Søreide K, Janssen EA, Søiland H, Körner H, Baak JP. Microsatellite instability in colorectal cancer. *Br J Surg.* 2006;93(4):395–406.
71. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America.* 1999;96(15):8681–8686.
72. Boland CR. Molecular screening for Lynch syndrome. *Nature Clinical Practice Gastroenterology & Hepatology.* 2005;2(9):392-393.
73. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895–2015. *Nature Reviews Cancer.* 2015;15(3):181-194.
74. Budinska E, Popovici V, Tejpar S, et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *The Journal of pathology.* 2013;231(1):63-76.
75. De Sousa E Melo F, Wang X, Jansen M, et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nature Medicine.* 2013;19:614.
76. Marisa L, de Reyniès A, Duval A, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 2013;10(5):e1001453-e1001453.
77. Roepman P, Schlicker A, Tabernero J, et al. Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition. *International journal of cancer.* 2014;134(3):552-562.
78. Sadanandam A, Lyssiotis CA, Homicsko K, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nature medicine.* 2013;19(5):619-625.
79. Schlicker A, Beran G, Chresta CM, et al. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. *BMC Med Genomics.* 2012;5:66-66.
80. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nature Medicine.* 2015;21(11):1350–1356.
81. Dienstmann R, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nature Reviews Cancer.* 2017;17:79.
82. Tóth G, Gáspári Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome research.* 2000;10(7):967-981.
83. Subramanian S, Mishra R, Singh L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology.* 2003;4(2):R13.
84. Dib C, Faure S, Fizames C, et al. A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature.* 1996;380:152–154.

85. Ellegren H. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics.* 2004;5(6):435-445.
86. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature.* 1993;363(6429):558–561.
87. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science (New York, NY).* 1993;260(5109):816–819.
88. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. *Science.* 1993;260(5109):812.
89. Peltomaki P, Aaltonen LA, Sistonen P, et al. Genetic mapping of a locus predisposing to human colorectal cancer. *Science.* 1993;260(5109):810–812.
90. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer research.* 1998;58(22):5248-5257.
91. Rodriguez-Bigas MA, Boland CR, Hamilton SR, et al. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: Meeting Highlights and Bethesda Guidelines. *Journal of the National Cancer Institute.* 1997;89(23):1758–1762.
92. Loeb LA, Monnat Jr RJ. DNA polymerases and human disease. *Nature Reviews Genetics.* 2008;9:594.
93. McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* 2008;18(1):148-161.
94. Sun JX, Helgason A, Masson G, et al. A direct characterization of human mutation based on microsatellites. *Nature Genetics.* 2012;44(10):1161-1165.
95. Gadgil R, Barthelemy J, Lewis T, Leffak M. Replication stalling and DNA microsatellite instability. *Biophysical Chemistry.* 2017;225:38-48.
96. Li G-M. Mechanisms and functions of DNA mismatch repair. *Cell Research.* 2007;18:85.
97. Kunkel TA, Erie DA. DNA MISMATCH REPAIR. *Annual Review of Biochemistry.* 2005;74(1):681-710.
98. Mori Y, Yin J, Rashid A, et al. Instabilotyping. *Cancer research.* 2001;61(16):6046.
99. Kane MF, Loda M, Gaida GM, et al. Methylation of the MLH1 Promoter Correlates with Lack of Expression of hMLH1 in Sporadic Colon Tumors and Mismatch Repair-defective Human Tumor Cell Lines. *Cancer research.* 1997;57(5):808.
100. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the MLH1 Promoter in Colon Cancer with Microsatellite Instability. *Cancer research.* 1998;58(15):3455.
101. Li SKH, Martin A. Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored. *Trends in Molecular Medicine.* 2016;22(4):274-289.
102. Lahue RS, Au KG, Modrich P. DNA mismatch correction in a defined system. *Science.* 1989;245(4914):160.

103. Modrich P. MECHANISMS AND BIOLOGICAL EFFECTS OF MISMATCH REPAIR. *Annual Review of Genetics.* 1991;25(1):229-253.
104. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell.* 1993;75(5):1027–1038.
105. Leach FS, Nicolaides NC, Papadopoulos N, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell.* 1993;75(6):1215–1225.
106. Genschel J, Littman SJ, Drummond JT, Modrich P. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J Biol Chem.* 1998;273(31):19895–19901.
107. Sinicrope FA, Sargent DJ. Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2012;18(6):1506–1512.
108. Pawlik TM, Raut CP, Rodriguez-Bigas MA. Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis Markers.* 2004;20(4-5):199-206.
109. Tomlinson I, Halford S, Aaltonen L, Hawkins N, Ward R. Does MSI-low exist? *The Journal of Pathology.* 2002;197(1):6-13.
110. González-García I, Moreno V, Navarro M, et al. Standardized Approach for Microsatellite Instability Detection in Colorectal Carcinomas. *JNCI: Journal of the National Cancer Institute.* 2000;92(7):544-549.
111. Kim TM, Laird PW, Park PJ. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell.* 2013;155(4):858–868.
112. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *Journal of the National Cancer Institute.* 2004;96(4):261-268.
113. Laiho P, Launonen V, Lahermo P, et al. Low-Level Microsatellite Instability in Most Colorectal Carcinomas. *Cancer research.* 2002;62(4):1166.
114. Suraweera N, Duval A, Reperant M, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology.* 2002;123(6):1804-1811.
115. Soreide K. High-fidelity of five quasimonomorphic mononucleotide repeats to high-frequency microsatellite instability distribution in early-stage adenocarcinoma of the colon. *Anticancer Res.* 2011;31(3):967–971.
116. Luchini C, Bibeau F, Ligtenberg MJL, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. *Annals of Oncology.* 2019;30(8):1232-1243.
117. Popat S, Hubner R, Houlston RS. Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis. *Journal of Clinical Oncology.* 2005;23(3):609-618.
118. Schumacher TN, Scheper W, Kvistborg P. Cancer Neoantigens. *Annual Review of Immunology.* 2018;37(1):173-200.

119. The Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487:330.
120. Fridman WH, Zitvogel L, Sautès–Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. *Nature Reviews Clinical Oncology.* 2017;14:717.
121. Giannakis M, Mu Xinmeng J, Shukla Sachet A, et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Reports.* 2016;15(4):857-865.
122. Saeterdal I, Bjorheim J, Lislerud K, et al. Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America.* 2001;98(23):13255–13260.
123. Linnebacher M, Gebert J, Rudy W, et al. Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens. *International journal of cancer.* 2001;93(1):6–11.
124. Watanabe T, Wu T-T, Catalano PJ, et al. Molecular Predictors of Survival after Adjuvant Chemotherapy for Colon Cancer. *New England Journal of Medicine.* 2001;344(16):1196-1206.
125. Ganesh K, Stadler ZK, Cercek A, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nature Reviews Gastroenterology & Hepatology.* 2019;16(6):361-375.
126. Hou J-T, Zhao L-N, Zhang D-J, et al. Prognostic Value of Mismatch Repair Genes for Patients With Colorectal Cancer: Meta-Analysis. *Technol Cancer Res Treat.* 2018;17:1533033818808507-1533033818808507.
127. Guastadisegni C, Colafranceschi M, Ottini L, Dogliotti E. Microsatellite instability as a marker of prognosis and response to therapy: A meta-analysis of colorectal cancer survival data. *European Journal of Cancer.* 2010;46(15):2788-2798.
128. Tejpar S, Bosman F, Delorenzi M, et al. Microsatellite instability (MSI) in stage II and III colon cancer treated with 5FU-LV or 5FU-LV and irinotecan (PETACC 3-EORTC 40993-SAKK 60/00 trial). *Journal of Clinical Oncology.* 2009;27(15_suppl):4001-4001.
129. Gkekas I, Novotny J, Fabian P, et al. Deficient mismatch repair as a prognostic marker in stage II colon cancer patients. *European Journal of Surgical Oncology.* 2019;45(10):1854-1861.
130. Kim CG, Ahn JB, Jung M, et al. Effects of microsatellite instability on recurrence patterns and outcomes in colorectal cancers. *British journal of cancer.* 2016;115:25.
131. Sinicrope FA, Mahoney MR, Smyrk TC, et al. Prognostic impact of deficient DNA mismatch repair in patients with stage III colon cancer from a randomized trial of FOLFOX-based adjuvant chemotherapy. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2013;31(29):3664-3672.
132. Bertagnolli MM, Niedzwiecki D, Compton CC, et al. Microsatellite instability predicts improved response to adjuvant therapy with irinotecan, fluorouracil, and leucovorin in stage III colon cancer: Cancer and Leukemia Group B

Protocol 89803. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2009;27(11):1814-1821.
133. Soreide K, Slewa A, Stokkeland PJ, et al. Microsatellite instability and DNA ploidy in colorectal cancer: potential implications for patients undergoing systematic surveillance after resection. *Cancer.* 2009;115(2):271–282.
134. Sargent DJ, Marsoni S, Monges G, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol.* 2010;28(20):3219–3226.
135. Sinicrope FA. DNA mismatch repair and adjuvant chemotherapy in sporadic colon cancer. *Nat Rev Clin Oncol.* 2010;7(3):174–177.
136. Ribic CM, Sargent DJ, Moore MJ, et al. Tumor Microsatellite-Instability Status as a Predictor of Benefit from Fluorouracil-Based Adjuvant Chemotherapy for Colon Cancer. *New England Journal of Medicine.* 2003;349(3):247-257.
137. Jensen SA, Vainer B, Kruhøffer M, Sørensen JB. Microsatellite instability in colorectal cancer and association with thymidylate synthase and dihydropyrimidine dehydrogenase expression. *BMC cancer.* 2009;9:25-25.
138. Bendardaf R, Lamlum H, Ristamäki R, Korkeila E, Syrjänen K, Pyrhönen S. Thymidylate synthase and microsatellite instability in colorectal cancer: Implications for disease free survival, treatment response and survival with metastases. *Acta Oncologica.* 2008;47(6):1046-1053.
139. Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. Mismatch repair deficient colorectal cancer in the era of personalized treatment. *Nature Reviews Clinical Oncology.* 2010;7:197.
140. Romiti A, Rulli E, Pilozzi E, et al. Exploring the Prognostic Role of Microsatellite Instability in Patients With Stage II Colorectal Cancer: A Systematic Review and Meta-Analysis. *Clinical Colorectal Cancer.* 2017;16(2):e55-e59.
141. Venderbosch S, Nagtegaal ID, Maughan TS, et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2014;20(20):5322-5330.
142. Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer discovery.* 2015;5(1):43-51.
143. Le DT, Durham JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *New England Journal of Medicine.* 2015;372(26):2509-2520.
144. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science.* 2017;357(6349):409.
145. Terrell RB, Wille AH, Cheville JC, Nystuen AM, Cohen MB, Sheffield VC. Microsatellite instability in adenocarcinoma of the prostate. *Am J Pathol.* 1995;147(3):799–805.

146. Larson AA, Kern S, Sommers RL, Yokota J, Cavenee WK, Hampton GM. Analysis of Replication Error (RER+) Phenotypes in Cervical Carcinoma. *Cancer research.* 1996;56(6):1426.
147. Mao L, Lee DJ, Tockman MS, Erozan YS, Askin F, Sidransky D. Microsatellite alterations as clonal markers for the detection of human cancer. *Proceedings of the National Academy of Sciences of the United States of America.* 1994;91(21):9871–9875.
148. Fong KM, Zimmerman PV, Smith PJ. Microsatellite instability and other molecular abnormalities in non-small cell lung cancer. *Cancer research.* 1995;55(1):28–30.
149. Jönsson M, Johannsson O, Borg Å. Infrequent occurrence of microsatellite instability in sporadic and familial breast cancer. *European Journal of Cancer.* 1995;31(13):2330-2334.
150. Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, Rüschoff J. Diagnostic Microsatellite Instability: Definition and Correlation with Mismatch Repair Protein Expression. *Cancer research.* 1997;57(21):4749.
151. Yeh S-H, Chen P-J, Chen H-L, Lai M-Y, Wang C-C, Chen D-S. Frequent Genetic Alterations at the Distal Region of Chromosome 1p in Human Hepatocellular Carcinomas. *Cancer research.* 1994;54(15):4188.
152. Watson MM, Berg M, Soreide K. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. *British journal of cancer.* 2014;111(5):823-827.
153. Kloor M, Schwitalle Y, Knebel Doeberitz Mv, Wentzensen N. Tetranucleotide repeats in coding regions: no evidence for involvement in EMAST carcinogenesis. *J Mol Med (Berl).* 2006;84(4):329–333.
154. Ahrendt SA, Decker PA, Doffek K, et al. Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. *Cancer research.* 2000;60(9):2488–2491.
155. Xu L, Chow J, Bonacum J, et al. Microsatellite instability at AAAG repeat sequences in respiratory tract cancers. *International journal of cancer.* 2001;91(2):200–204.
156. Woenckhaus M, Stoehr R, Dietmaier W, et al. Microsatellite instability at chromosome 8p in non-small cell lung cancer is associated with lymph node metastasis and squamous differentiation. *Int J Oncol.* 2003;23:1357–1363.
157. Arai H, Okudela K, Oshiro H, et al. Elevated microsatellite alterations at selected tetra-nucleotide (EMAST) in non-small cell lung cancers--a potential determinant of susceptibility to multiple malignancies. *Int J Clin Exp Pathol.* 2013;6(3):395–410.
158. Singer G, Kallinowski T, Hartmann A, et al. Different types of microsatellite instability in ovarian carcinoma. *International journal of cancer.* 2004;112(4):643–646.
159. Choi YD, Choi J, Kim JH, et al. Microsatellite instability at a tetranucleotide repeat in type I endometrial carcinoma. *J Exp Clin Cancer Res.* 2008;27:88.
160. Burger M, Denzinger S, Hammerschmied CG, et al. Elevated microsatellite alterations at selected tetranucleotides (EMAST) and mismatch repair gene expression in prostate cancer. *J Mol Med (Berl).* 2006;84(10):833–841.

161. Azzouzi AR, Catto JW, Rehman I, et al. Clinically localised prostate cancer is microsatellite stable. *BJU Int.* 2007;99(5):1031–1035.
162. Perinchery G, Nojima D, Goharderakhshan R, Tanaka Y, Alonzo J, Dahiya R. Microsatellite instability of dinucleotide tandem repeat sequences is higher than trinucleotide, tetranucleotide and pentanucleotide repeat sequences in prostate cancer. *Int J Oncol.* 2000;16(6):1203–1209.
163. Danaee H, Nelson HH, Karagas, et al. Microsatellite instability at tetranucleotide repeats in skin and bladder cancer. *Oncogene.* 2002;21:4894–4899.
164. Burger M, Burger SJ, Denzinger S, et al. Elevated microsatellite instability at selected tetranucleotide repeats does not correlate with clinicopathologic features of bladder cancer. *Eur Urol.* 2006;50(4):770-775; discussion 776.
165. Temam S, Casiraghi O, Lahaye JB, et al. Tetranucleotide microsatellite instability in surgical margins for prediction of local recurrence of head and neck squamous cell carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2004;10(12 Pt 1):4022–4028.
166. Mori T, Hamaya Y, Uotani T, et al. Prevalence of elevated microsatellite alterations at selected tetranucleotide repeats in pancreatic ductal adenocarcinoma. *PloS one.* 2018;13(12):e0208557-e0208557.
167. Catto JW, Azzouzi AR, Amira N, et al. Distinct patterns of microsatellite instability are seen in tumours of the urinary tract. *Oncogene.* 2003;22:8699–8706.
168. Haugen AC, Goel A, Yamada K, et al. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer research.* 2008;68(20):8465-8472.
169. Devaraj B, Lee A, Cabrera BL, et al. Relationship of EMAST and microsatellite instability among patients with rectal cancer. *J Gastrointest Surg.* 2010;14(10):1521–1528.
170. Yamada K, Kanazawa S, Koike J, et al. Microsatellite instability at tetranucleotide repeats in sporadic colorectal cancer in Japan. *Oncology reports.* 2010;23(2):551-561.
171. Lee SY, Chung H, Devaraj B, et al. Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias. *Gastroenterology.* 2010;139(5):1519–1525.
172. Lee SY, Miyai K, Han HS, et al. Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. *Dig Dis Sci.* 2012;57(1):72–78.
173. Hamaya Y, Guarinos C, Tseng-Rogenski SS, et al. Efficacy of Adjuvant 5-Fluorouracil Therapy for Patients with EMAST-Positive Stage II/III Colorectal Cancer. *PloS one.* 2015;10(5).
174. Venderbosch S, van Lent-van Vliet S, de Haan AF, et al. EMAST is associated with a poor prognosis in microsatellite instable metastatic colorectal cancer. *PloS one.* 2015;10(4):e0124538.
175. Watson MM, Lea D, Rewcastle E, Hagland HR, Søreide K. Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal

cancers with and without high-frequency microsatellite instability: same, same but different? *Cancer Medicine.* 2016;5(7):1580–1587.
176. Koi M, Garcia M, Choi C, et al. Microsatellite Alterations With Allelic Loss at 9p24.2 Signify Less-Aggressive Colorectal Cancer Metastasis. *Gastroenterology.* 2016;150(4):944-955.
177. Lee HS, Park KU, Kim DW, et al. Elevated Microsatellite Alterations at Selected Tetranucleotide Repeats (EMAST) and Microsatellite Instability in Patients with Colorectal Cancer and Its Clinical Features. *Current molecular medicine.* 2016;16(9):829-839.
178. Chen MH, Chang SC, Lin PC, et al. Combined Microsatellite Instability and Elevated Microsatellite Alterations at Selected Tetranucleotide Repeats (EMAST) Might Be a More Promising Immune Biomarker in Colorectal Cancer. *The Oncologist.* 2019.
179. Torshizi Esfahani A, Seyedna SY, Nazemalhosseini Mojarad E, Majd A, Asadzadeh Aghdaei H. MSI-L/EMAST is a predictive biomarker for metastasis in colorectal cancer patients. *Journal of Cellular Physiology.* 2019;234(8):13128-13136.
180. Kuan T-C, Chang S-C, Lin J-K, et al. Prognosticators of Long-Term Outcomes of TNM Stage II Colorectal Cancer: Molecular Patterns or Clinicopathological Features. *World Journal of Surgery.* 2019.
181. Mohammadpour S, Goodarzi HR, Jafarinia M, Porhoseingholi MA, Nazemalhosseini-Mojarad E. EMAST status as a beneficial predictor of fluorouracil-based adjuvant chemotherapy for Stage II/III colorectal cancer. *Journal of Cellular Physiology.* 2019;0(0).
182. Garcia M, Choi C, Kim HR, et al. Association between recurrent metastasis from stage II and III primary colorectal tumors and moderate microsatellite instability. *Gastroenterology.* 2012;143(1):48-50 e41.
183. Li J, Koike J, Kugoh H, et al. Down-regulation of MutS homolog 3 by hypoxia in human colorectal cancer. *Biochim Biophys Acta.* 2012;1823(4):889–899.
184. Wang Y, Vnencak-Jones CL, Cates JM, Shi C. Deciphering Elevated Microsatellite Alterations at Selected Tetra/Pentanucleotide Repeats, Microsatellite Instability, and Loss of Heterozygosity in Colorectal Cancers. *J Mol Diagn.* 2018;20(3):366-372.
185. Koi M, Tseng-Rogenski SS, Carethers JM. Inflammation-associated microsatellite alterations: Mechanisms and significance in the prognosis of patients with colorectal cancer. *World Journal of Gastrointestinal Oncology.* 2018;10(1):1-14.
186. Acharya S, Wilson T, Gradia S, et al. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proceedings of the National Academy of Sciences.* 1996;93(24):13629.
187. Palombo F, Iaccarino I, Nakajima E, Ikejima M, Shimada T, Jiricny J. hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology.* 1996;6(9):1181-1184.
188. Drummond JT, Genschel J, Wolf E, Modrich P. DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSα/hMutSβ ratio and reduces the efficiency of base–base

mismatch repair. *Proceedings of the National Academy of Sciences.* 1997;94(19):10144.
189. Marra G, Iaccarino I, Lettieri T, Roscilli G, Delmastro P, Jiricny J. Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proceedings of the National Academy of Sciences of the United States of America.* 1998;95(15):8568-8573.
190. Umar A, Risinger JI, Glaab WE, Tindall KR, Barrett JC, Kunkel TA. Functional Overlap in Mismatch Repair by Human MSH3 and <em>MSH6</em&gt. *Genetics.* 1998;148(4):1637.
191. Tseng-Rogenski SS, Chung H, Wilk MB, Zhang S, Iwaizumi M, Carethers JM. Oxidative Stress Induces Nuclear-to-Cytosol Shift of hMSH3, a Potential Mechanism for EMAST in Colorectal Cancer Cells. *PloS one.* 2012;7(11).
192. Tseng-Rogenski S, Hamaya Y, Choi DY, Carethers JM. Interleukin 6 Alters Localization of hMSH3, Leading to DNA Mismatch Repair Defects in Colorectal Cancer Cells. *Gastroenterology.* 2015;148(3):579-589.
193. Janeway CA. How the immune system works to protect the host from infection: A personal view. *Proceedings of the National Academy of Sciences.* 2001;98(13):7461.
194. Dranoff G. Cytokines in cancer pathogenesis and cancer therapy. *Nature Reviews Cancer.* 2004;4(1):11-22.
195. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science.* 2015;348(6230):69.
196. Chen DS, Mellman I. Elements of cancer immunity and the cancer–immune set point. *Nature.* 2017;541:321.
197. Golstein P, Griffiths GM. An early history of T cell-mediated cytotoxicity. *Nature Reviews Immunology.* 2018;18(8):527-535.
198. Riella LV, Paterson AM, Sharpe AH, Chandraker A. Role of the PD-1 pathway in the immune response. *Am J Transplant.* 2012;12(10):2575-2587.
199. Park JH, Powell AG, Roxburgh CSD, Horgan PG, McMillan DC, Edwards J. Mismatch repair status in patients with primary operable colorectal cancer: associations with the local and systemic tumour environment. *British journal of cancer.* 2016;114(5):562.
200. Leach DR, Krummel MF, Allison JP. Enhancement of Antitumor Immunity by CTLA-4 Blockade. *Science.* 1996;271(5256):1734.
201. Ishida Y, Agata Y, Shibahara K, Honjo T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *The EMBO Journal.* 1992;11(11):3887-3895.
202. Press release: The Nobel Prize in Physiology or Medicine 2018. NobelPrize.org: Nobel Media AB 2019, Wed. 6 Nov 2019 2018.
203. Boussiotis VA. Molecular and Biochemical Aspects of the PD-1 Checkpoint Pathway. *New England Journal of Medicine.* 2016;375(18):1767-1778.
204. Sun C, Mezzadra R, Schumacher TN. Regulation and Function of the PD-L1 Checkpoint. *Immunity.* 2018;48(3):434-452.
205. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer.* 2012;12(4):252-264.

206. Herbst RS, Soria J-C, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature.* 2014;515(7528):563-567.
207. Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. *Nature.* 2011;480(7378):480-489.
208. Zou W, Chen L. Inhibitory B7-family molecules in the tumour microenvironment. *Nature Reviews Immunology.* 2008;8(6):467-477.
209. Taube JM, Anders RA, Young GD, et al. Colocalization of Inflammatory Response with B7-H1 Expression in Human Melanocytic Lesions Supports an Adaptive Resistance Mechanism of Immune Escape. *Science Translational Medicine.* 2012;4(127):127ra137.
210. Brahmer JR, Tykodi SS, Chow LQM, et al. Safety and Activity of Anti–PD-L1 Antibody in Patients with Advanced Cancer. *New England Journal of Medicine.* 2012;366(26):2455-2465.
211. Oliveira AF, Bretes L, Furtado I. Review of PD-1/PD-L1 Inhibitors in Metastatic dMMR/MSI-H Colorectal Cancer. *Front Oncol.* 2019;9:396-396.
212. Pagès F, Berger A, Camus M, et al. Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer. *New England Journal of Medicine.* 2005;353(25):2654-2666.
213. Ogino S, Nosho K, Irahara N, et al. Lymphocytic Reaction to Colorectal Cancer is Associated with Longer Survival, Independent of Lymph Node Count, MSI and CpG Island Methylator Phenotype. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2009;15(20):6412–6420.
214. Nosho K, Baba Y, Tanaka N, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer and prognosis: cohort study and literature review. *The Journal of pathology.* 2010;222(4):350–366.
215. Pages F, Galon J, Dieu-Nosjean MC, Tartour E, Sautes-Fridman C, Fridman WH. Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene.* 2010;29(8):1093–1102.
216. Galon J, Mlecnik B, Bindea G, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. *The Journal of pathology.* 2014;232(2):199–209.
217. Pages F, Kirilovsky A, Mlecnik B, et al. In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *J Clin Oncol.* 2009;27(35):5944–5951.
218. Galon J, Pagès F, Marincola FM, et al. Cancer classification using the Immunoscore: a worldwide task force. *Journal of translational medicine.* 2012;10:205.
219. Kirilovsky A, Marliot F, El Sissy C, Haicheur N, Galon J, Pagès F. Rational bases for the use of the Immunoscore in routine clinical settings as a prognostic and predictive biomarker in cancer patients. *Int Immunol.* 2016;28(8):373-382.
220. 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. *The Lancet.* 2018;391(10135):2128-2139.

221. Kwak Y, Koh J, Kim D-W, Kang S-B, Kim WH, Lee HS. Immunoscore encompassing CD3+ and CD8+ T cell densities in distant metastasis is a robust prognostic marker for advanced colorectal cancer. *Oncotarget.* 2016;7(49):81778-81790.
222. Wirta E-V, Seppälä T, Friman M, et al. Immunoscore in mismatch repair-proficient and -deficient colon cancer. *J Pathol Clin Res.* 2017;3(3):203-213.
223. Anitei M-G, Zeitoun G, Mlecnik B, et al. Prognostic and Predictive Values of the Immunoscore in Patients with Rectal Cancer. *Clinical Cancer Research.* 2014;20(7):1891.
224. National Cancer Institute. NCI Dictionary of Cancer Terms. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomarker. Accessed 08/11/19, 2019.
225. Group BDW. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics.* 2001;69(3):89-95.
226. Buyse M, Burzykowski T, Carroll K, et al. Progression-free survival is a surrogate for survival in advanced colorectal cancer. *J Clin Oncol.* 2007;25(33):5218-5224.
227. Longo DL. Tumor Heterogeneity and Personalized Medicine. *New England Journal of Medicine.* 2012;366(10):956-957.
228. Heidelberger C, Chaudhuri NK, Danneberg P, et al. Fluorinated Pyrimidines, A New Class of Tumour-Inhibitory Compounds. *Nature.* 1957;179(4561):663-666.
229. Goossens N, Nakagawa S, Sun X, Hoshida Y. Cancer biomarker discovery and validation. *Transl Cancer Res.* 2015;4(3):256-269.
230. Poste G. Bring on the biomarkers. *Nature.* 2011;469(7329):156-157.
231. Selleck MJ, Senthil M, Wall NR. Making Meaningful Clinical Use of Biomarkers. *Biomark Insights.* 2017;12:1177271917715236-1177271917715236.
232. Arrowsmith J. Phase II failures: 2008–2010. *Nature Reviews Drug Discovery.* 2011;10(5):328-329.
233. Harrison RK. Phase II and phase III failures: 2013–2015. *Nature Reviews Drug Discovery.* 2016;15:817.
234. Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? *Nature Reviews Drug Discovery.* 2011;10(9):712-712.
235. Begley CG, Ellis LM. Raise standards for preclinical cancer research. *Nature.* 2012;483(7391):531-533.
236. Turner JR. Rigor, Reproducibility, and Responsibility: A Quantum of Solace. *Cell Mol Gastroenterol Hepatol.* 2019;7(4):869-871.
237. Doucet M, Becker KF, Björkman J, et al. Quality Matters: 2016 Annual Conference of the National Infrastructures for Biobanking. *Biopreserv Biobank.* 2017;15(3):270-276.
238. Robbe P, Popitsch N, Knight SJL, et al. Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the

100,000 Genomes Project. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2018;20(10):1196–1205.
239. Jones W, Greytak S, Odeh H, et al. Deleterious effects of formalin-fixation and delays to fixation on RNA and miRNA-Seq profiles. *Sci Rep.* 2019;9(1):6980.
240. Ilie M, Hofman V, Dietel M, Soria J-C, Hofman P. Assessment of the PD-L1 status by immunohistochemistry: challenges and perspectives for therapeutic strategies in lung cancer patients. *Virchows Archiv.* 2016;468(5):511-525.
241. Padmanabhan S. Chapter 15 - Clinical Trials in Pharmacogenomics and Stratified Medicine. In: Padmanabhan S, ed. *Handbook of Pharmacogenomics and Stratified Medicine.* San Diego: Academic Press; 2014:309-320.
242. Chavalarias D, Wallach JD, Li AHT, Ioannidis JPA. Evolution of Reporting P Values in the Biomedical Literature, 1990-2015. *JAMA.* 2016;315(11):1141-1148.
243. Amrhein V, Trafimow D, Greenland S. Inferential Statistics as Descriptive Statistics: There Is No Replication Crisis if We Don't Expect Replication. *The American Statistician.* 2019;73(sup1):262-270.
244. Bussolati G, Leonardo E. Technical pitfalls potentially affecting diagnoses in immunohistochemistry. *Journal of Clinical Pathology.* 2008;61(11):1184.
245. Cartun RWT, Clive R.; Dabbs, David J. Techniques of immunohistochemistry: Principles, Pitfalls, and Standardization. In: Dabbs DJ, ed. *Diagnostic Immunohistochemistry* 5th ed. Philadelphia, PA, USA: Elsevier; 2019.
246. Pantanowitz LR, David L. Imaging and Quantitative Immunohistochemistry. In: Dabbs DJ, ed. *Diagnostic Immunohistochemistry* 5th ed. Philadelphia, PA, USA: Elsevier; 2019.
247. Stålhammar G, Robertson S, Wedlund L, et al. Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology.* 2018;72(6):974–989.
248. EQUATOR network. www.equator-network.org/. Accessed 01-12-2019.
249. McShane LM, Altman DG, Sauerbrei W, et al. REporting recommendations for tumour MARKer prognostic studies (REMARK). *British journal of cancer.* 2005;93(4):387-391.
250. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *BMC Med.* 2012;10:51-51.
251. Sauerbrei W, Taube SE, McShane LM, Cavenagh MM, Altman DG. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration. *Journal of the National Cancer Institute.* 2018;110(8):803-811.
252. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLoS Med.* 2007;4(10):e297.
253. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453-1457.

254. Moore HM, Kelly AB, Jewell SD, et al. Biospecimen Reporting for Improved Study Quality (BRISQ). *Journal of Proteome Research.* 2011;10(8):3429-3438.
255. Soreide K, Watson MM, Lea D, Nordgard O, Soreide JA, Hagland HR. Assessment of clinically related outcomes and biomarker analysis for translational integration in colorectal cancer (ACROBATICC): study protocol for a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastasis. *Journal of translational medicine.* 2016;14(1):192.
256. Hewitt SM, Baskin DG, Frevert CW, Stahl WL, Rosa-Molinar E. Controls for immunohistochemistry: the Histochemical Society's standards of practice for validation of immunohistochemical assays. *J Histochem Cytochem.* 2014;62(10):693-697.
257. Barnes M, Srinivas C, Bai I, et al. Whole tumor section quantitative image analysis maximizes between-pathologists' reproducibility for clinical immunohistochemistry-based biomarkers. *Laboratory Investigation.* 2017;97(12):1508-1515.
258. Aeffner F, Wilson K, Martin NT, et al. The Gold Standard Paradox in Digital Image Analysis: Manual Versus Automated Scoring as Ground Truth. *Archives of Pathology & Laboratory Medicine.* 2017;141(9):1267-1275.
259. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nature Reviews Clinical Oncology.* 2019;16(11):703-715.
260. John T, Liu G, Tsao MS. Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors. *Oncogene.* 2009;28(1):S14-S23.
261. Hirsch FR, Varella-Garcia M, Bunn PA, et al. Epidermal Growth Factor Receptor in Non–Small-Cell Lung Carcinomas: Correlation Between Gene Copy Number and Protein Expression and Impact on Prognosis. *Journal of Clinical Oncology.* 2003;21(20):3798-3807.
262. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc.* 1998;11(2):155-168.
263. Rizzardi AE, Johnson AT, Vogel RI, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. *Diagnostic Pathology.* 2012;7(1):42.
264. Bayes T. An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* 1974;53:370-418.
265. Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft. http://www.helsedirektoratet.no/retningslinjer. Published 2018. Accessed May 9, 2019.
266. Retningslinjer for brystkræft. http://www.dbcg.dk/. Published 2017. Accessed May 9, 2019.

267. Kvalitetsdokument för patologi. http://www.svfp.se/foreningar/uploads/L15178/kvast/brostpatologi/KVASTbrostcancer2018.pdf. Published 2018. Accessed May 9, 2019.
268. Stalhammar G, Robertson S, Wedlund L, et al. Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology.* 2018;72(6):974-989.
269. Griffin J, Treanor D. Digital pathology in clinical use: where are we now and what is holding us back? *Histopathology.* 2017;70(1):134-145.
270. Carethers JM, Jung BH. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. *Gastroenterology.* 2015;149(5):1177-1190.e1173.
271. Soliman AS, Bondy ML, El-Badawy SA, et al. Contrasting molecular pathology of colorectal carcinoma in Egyptian and Western patients. *British journal of cancer.* 2001;85(7):1037-1046.
272. Shahrokni A, Tin A, Alexander K, et al. Development and Evaluation of a New Frailty Index for Older Surgical Patients With Cancer. *JAMA network open.* 2019;2(5):e193545.
273. Bergman H, Ferrucci L, Guralnik J, et al. Frailty: an emerging research and clinical paradigm--issues and controversies. *J Gerontol A Biol Sci Med Sci.* 2007;62(7):731-737.
274. Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. *J Am Med Dir Assoc.* 2013;14(6):392-397.
275. Fried LP, Tangen CM, Walston J, et al. Frailty in Older Adults: Evidence for a Phenotype. *The Journals of Gerontology: Series A.* 2001;56(3):M146-M157.
276. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ.* 2005;173(5):489-495.
277. Soysal P, Isik AT, Arik F, Kalan U, Eyvaz A, Veronese N. Validity of the Mini-Nutritional Assessment Scale for Evaluating Frailty Status in Older Adults. *J Am Med Dir Assoc.* 2019;20(2):183-187.
278. Ng T-P, Feng L, Niti M, Yap KB. Albumin, haemoglobin, BMI and cognitive performance in older adults. *Age and Ageing.* 2008;37(4):423-429.
279. Huisingh-Scheetz M, Walston J. How should older adults with cancer be evaluated for frailty? *Journal of Geriatric Oncology.* 2017;8(1):8-15.
280. Lai C-C, You J-F, Yeh C-Y, et al. Low preoperative serum albumin in colon cancer: a risk factor for poor outcome. *International journal of colorectal disease.* 2011;26(4):473–481.
281. Palmer K, Vetrano DL, Marengoni A, et al. The Relationship Between Anaemia and Frailty: A Systematic Review and Meta-Analysis of Observational Studies. *The journal of nutrition, health & aging.* 2018;22(8):965-974.
282. Saedi AA, Feehan J, Phu S, Duque G. Current and emerging biomarkers of frailty in the elderly. *Clin Interv Aging.* 2019;14:389-398.
283. Aunan JR, Watson MM, Hagland HR, Søreide K. Molecular and biological hallmarks of ageing. *BJS (British Journal of Surgery).* 2016;103(2):e29-e46.
284. Aunan JR, Cho WC, Søreide K. The Biology of Aging and Cancer: A Brief Overview of Shared and Divergent Molecular Hallmarks. *Aging and disease.* 2017;8(5):628–642.

285. Carethers JM, Koi M, Tseng-Rogenski SS. EMAST is a Form of Microsatellite Instability That is Initiated by Inflammation and Modulates Colorectal Cancer Progression. *Genes.* 2015;6(2):185-205.
286. Sepulveda AR, Hamilton SR, Allegra CJ, et al. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *Journal of Clinical Oncology.* 2017;35(13):1453-1486.
287. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. *The Journal of clinical investigation.* 2003;111(12):1805-1812.
288. Topalian SL, Hodi FS, Brahmer JR, et al. Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer. *New England Journal of Medicine.* 2012;366(26):2443-2454.
289. Sanmamed MF, Chen L. Inducible expression of B7-H1 (PD-L1) and its selective role in tumor site immune modulation. *Cancer J.* 2014;20(4):256-261.
290. Zhang Y, Chen L. Classification of Advanced Human Cancers Based on Tumor Immunity in the MicroEnvironment (TIME) for Cancer Immunotherapy. *JAMA Oncology.* 2016;2(11):1403-1404.
291. Hamada T, Soong TR, Masugi Y, et al. TIME (Tumor Immunity in the MicroEnvironment) classification based on tumor CD274 (PD-L1) expression status and tumor-infiltrating lymphocytes in colorectal carcinomas. *Oncoimmunology.* 2018;7(7):e1442999-e1442999.
292. Ho H-L, Chou T-Y, Yang S-H, et al. PD-L1 is a double-edged sword in colorectal cancer: the prognostic value of PD-L1 depends on the cell type expressing PD-L1. *Journal of Cancer Research and Clinical Oncology.* 2019;145(7):1785-1794.
293. Lee KS, Kwak Y, Ahn S, et al. Prognostic implication of CD274 (PD-L1) protein expression in tumor-infiltrating immune cells for microsatellite unstable and stable colorectal cancer. *Cancer Immunology, Immunotherapy.* 2017;66(7):927-939.
294. Teng MWL, Ngiow SF, Ribas A, Smyth MJ. Classifying Cancers Based on T-cell Infiltration and PD-L1. *Cancer research.* 2015;75(11):2139-2145.
295. Kim JH, Park HE, Cho N-Y, Lee HS, Kang GH. Characterisation of PD-L1-positive subsets of microsatellite-unstable colorectal cancers. *British journal of cancer.* 2016;115(4):490-496.
296. Wang L, Ren F, Wang Q, et al. Significance of Programmed Death Ligand 1 (PD-L1) Immunohistochemical Expression in Colorectal Cancer. *Molecular Diagnosis & Therapy.* 2016;20(2):175-181.
297. Koganemaru S, Inoshita N, Miura Y, et al. Prognostic value of programmed death-ligand 1 expression in patients with stage III colorectal cancer. *Cancer Science.* 2017;108(5):853-858.
298. Korehisa S, Oki E, Iimori M, et al. Clinical significance of programmed cell death-ligand 1 expression and the immune microenvironment at the invasive front of colorectal cancers with high microsatellite instability. *International journal of cancer.* 2018;142(4):822-832.

299. Yomoda T, Sudo T, Kawahara A, et al. The Immunoscore is a Superior Prognostic Tool in Stages II and III Colorectal Cancer and is Significantly Correlated with Programmed Death-Ligand 1 (PD-L1) Expression on Tumor-Infiltrating Mononuclear Cells. *Annals of Surgical Oncology.* 2019;26(2):415-424.
300. Lee KS, Kim BH, Oh H-K, et al. Programmed cell death ligand-1 protein expression and CD274/PD-L1 gene amplification in colorectal cancer: Implications for prognosis. *Cancer Science.* 2018;109(9):2957-2969.
301. Lee SJ, Jun S-Y, Lee IH, et al. CD274, LAG3, and IDO1 expressions in tumor-infiltrating immune cells as prognostic biomarker for patients with MSI-high colon cancer. *Journal of Cancer Research and Clinical Oncology.* 2018;144(6):1005-1014.
302. Droeser RA, Hirt C, Viehl CT, et al. Clinical impact of programmed cell death ligand 1 expression in colorectal cancer. *European Journal of Cancer.* 2013;49(9):2233-2242.
303. Li Y, Liang L, Dai W, et al. Prognostic impact of programed cell death-1 (PD-1) and PD-ligand 1 (PD-L1) expression in cancer cells and tumor infiltrating lymphocytes in colorectal cancer. *Molecular Cancer.* 2016;15(1):55.
304. Tintelnot J, Stein A. Immunotherapy in colorectal cancer: Available clinical evidence, challenges and novel approaches. *World journal of gastroenterology.* 2019;25(29):3920-3928.
305. Chalabi M, Fanchi LF, Van den Berg JG, et al. LBA37_PRNeoadjuvant ipilimumab plus nivolumab in early stage colon cancer. *Annals of Oncology.* 2018;29(suppl_8).
306. Kanesvaran R, Cordoba R, Maggiore R. Immunotherapy in Older Adults With Advanced Cancers: Implications for Clinical Decision-Making and Future Research. *American Society of Clinical Oncology Educational Book.* 2018(38):400-414.
307. Corbaux P, Maillet D, Boespflug A, et al. Older and younger patients treated with immune checkpoint inhibitors have similar outcomes in real-life setting. *European Journal of Cancer.* 2019;121:192-201.
308. Nedrebo BS, Soreide K, Eriksen MT, et al. Survival effect of implementing national treatment strategies for curatively resected colonic and rectal cancer. *Br J Surg.* 2011;98(5):716–723.
309. Plaschke J, Preußler M, Ziegler A, Schackert HK. Aberrant protein expression and frequent allelic loss of MSH3 in colorectal cancer with low-level microsatellite instability. *International Journal of Colorectal Disease.* 2012;27(7):911-919.
310. Wildiers H, Heeren P, Puts M, et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2014;32(24):2595–2603.
311. Zhang T, Lv J, Tan Z, et al. Immunocyte Profiling Using Single-Cell Mass Cytometry Reveals EpCAM+ CD4+ T Cells Abnormal in Colon Cancer. *Frontiers in Immunology.* 2019;10(1571).

312. Balaji S, Li H, Steen E, Keswani SG. Considerations for Immunohistochemistry. In: Kennedy G, Gosain A, Kibbe M, LeMaire SA, eds. *Success in Academic Surgery: Basic Science.* Cham: Springer International Publishing; 2019:105-144.

8. FIGURES CREDITS

Figure 1. Reprinted by permission from Springer Nature, Nature Reviews Gastroenterology & Hepatology, Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies NaNa Keum et al. © 2019

Figure 2. Reprinted from The Lancet, Vol 394/10207, Evelien Dekker et al. Pages No.1467-1480, Copyright (2019), with permission from Elsevier.

Figure 3. Used with the permission of the American College of Surgeons. Edge, S., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F., Trotti, A. (Eds.) AJCC Cancer Staging Manual, 7th Ed. Springer New York, 2010.

Figure 4. By courtesy of Encyclopaedia Britannica, Inc., copyright 2013; used with permission (top panel). Reprinted by permission from Springer Nature, Nature Reviews Molecular Cell Biology, "Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration", Nick Barker, ©2013 (bottom panel)

Figure 5. Reprinted from Cell, Vol 144/5, Douglas Hanahan,Robert A. Weinberg Hallmarks of Cancer: The Next Generation, 646-674, Copyright (2011), with permission from Elsevier

Figure 6. Reprinted by permission from Springer Nature, Nature Reviews Disease Primers, "Colorectal cancer", Ernst J. Kuipers et al ©2015

Figure 7. Reprinted from *The Lancet*, Vol 383/9927, Hermann Brenner et al. Pages No.1490-1502, Copyright (2014), with permission from Elsevier.

Figure 8. Reprinted by permission from Springer Nature, *Nature Medicine*, "The consensus molecular subtypes of colorectal cancer", Justin Guinney et al ©2015 (top panel). Reprinted by permission from Springer Nature, *Nature Reviews Cancer*, "Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer", Rodrigo Dienstmann et al ©2017 (bottom panel).

Figure 9. Reprinted by permission from Springer Nature, *Nature Reviews Cancer*, "Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer", Rodrigo Dienstmann et al ©2017 (top panel). Reprinted from *The Lancet*, Vol 394/10207, Evelien Dekker et al. Pages No.1467-1480, Copyright (2019), with permission from Elsevier. (bottom panel)

Figure 10. Reprinted by permission from Springer Nature, *Nature Reviews Cancer*, "Milestones of Lynch syndrome: 1895–2015", Henry T. Lynch et al ©2015

Figure 11. Reprinted from *Clinical Cancer Research* ©2012, Volume 18/6, A. Sinicrope & Daniel J. Sargent, "Molecular Pathways: Microsatellite Instability in Colorectal Cancer: Prognostic, Predictive, and Therapeutic Implications", with permission from AACR

Figure 12. Reprinted by permission from Springer Nature, *Nature Reviews Gastroenterology & Hepatology*, "Immunotherapy in colorectal cancer: rationale, challenges and potential", Karuna Ganesh et al ©2019

Figure 13. Reprinted from Subramanian et al, *Genome Biology*, 2003 under creative commons licence attribution 4.0 International (CC BY 4.0) (https://creativecommons.org/licenses/by/4.0/)

Figure 15. Reprinted by permission from Springer Nature, *Nature Reviews Cancer*, "Cytokines in cancer pathogenesis and cancer therapy", Glenn Dranoff ©2004

Figure 16. Reproduced with permission from Sinicrope FA., "Lynch Syndrome–Associated Colorectal Cancer". *New England Journal of Medicine.* 2018;379(8):764-773, Copyright Massachusetts Medical Society

Figure 17. Reprinted by permission from Oxford University Press, *International Immunology*, "Rational bases for the use of the Immunoscore in routine clinical settings as a prognostic and predictive biomarker in cancer patients", Amos Kirilovsky and Florence Marliot, ©2016

9. ERRATA

9.1 In paper I:

- *Patients and methods*, *DNA extraction and fragment analysis* sub-chapter. Second paragraph (5 lines from the top) should read:

"DNA extracted from tumor tissue and their corresponding normal tissue (from surgical resection margins) was then PCR- amplified with five tetranucleotide microsatellites primer pairs (EMAST: D20S85, D20S82, D9S242, D8S321, MYCL1, 5` fluorescently labeled) and five mono**nucleotide ~~and dinucleotid~~e** microsatellite primer pairs (MSI: NR- 27, NR- 21, NR- 24, BAT- 25, BAT- 26, 5` fluorescently labeled)."

- *Discussion,* first paragraph should read*:*

"In this cohort study, we found EMAST+ CRC to largely overlap with features associated with MSI+ cancers, including a predominant location in the colon, association with **high**-grade histology, larger tumor size, and advanced depth of growth (T3- 4)."

- *Figure I,* the significance (p) value for the MSI/EMAST association table was given as $p = 0.001$, while it should read $p < 0.001$.

10. PAPERS I-IV

I

Cancer Medicine

Open Access

ORIGINAL RESEARCH

Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal cancers with and without high-frequency microsatellite instability: same, same but different?

Martin M. Watson[1,2], Dordi Lea[2,3], Emma Rewcastle[3], Hanne R. Hagland[2,4] & Kjetil Søreide[1,2,5]

[1]Department of Gastrointestinal Surgery, Stavanger University Hospital, Stavanger, Norway
[2]Gastrointestinal Translational Research Unit, Molecular Laboratory, Hillevåg, Stavanger University Hospital, Stavanger, Norway
[3]Department of Pathology, Stavanger University Hospital, Stavanger, Norway
[4]Centre for Organelle Research (CORE), University of Stavanger, Stavanger, Norway
[5]Department of Clinical Medicine, University of Bergen, Bergen, Norway

Keywords
Colorectal cancer, elevated microsatellite alterations, microsatellite instability, node status, survival

Correspondence
Kjetil Søreide, Department of Gastrointestinal Surgery, Stavanger University Hospital, POB 8100, N-4068 Stavanger, Norway.
Tel: +47 5151 8330; Fax: +47 5151 9919;
E-mail: ksoreide@mac.com

Funding Information
Funded in part by grants from the Folke Hermansen Cancer Fund (424510, 424512) and Mjaaland Cancer Fund (424508). The funding source had no influence on any parts of the planning, conduction, analyses, interpretation, or presentation of this study.

Received: 8 January 2016; Revised: 26 February 2016; Accepted: 29 February 2016

***Cancer Medicine* 2016; 5(7):1580–1587**

doi: 10.1002/cam4.709

Abstract

Microsatellite instability (MSI) is associated with better prognosis in colorectal cancer (CRC). Elevated microsatellite alterations at selected tetranucleotides (EMAST) is a less-understood form of MSI. Here, we aim to investigate the role of EMAST in CRC±MSI related to clinical and tumor-specific characteristics. A consecutive, population-based series of stage I–III colorectal cancers were investigated for MSI and EMAST using PCR primers for 10 microsatellite markers. Of 151 patients included, 33 (21.8%) had MSI and 35 (23.2%) were EMAST+, with an overlap of 77% for positivity, (odds ratio [OR] 61; $P < 0.001$), and 95% for both markers being negative. EMAST was more prevalent in colon versus rectum (86% vs. 14%, $P = 0.004$). EMAST+ cancers were significantly more frequent in proximal colon (77 vs. 23%, $P = 0.004$), had advanced t-stage (T3–4 vs. T1–2 in 94% vs. 6%, respectively; $P = 0.008$), were larger (≥5 cm vs. <5 cm in 63% and 37%, respectively; $P = 0.022$) and had poorly differentiated tumor grade (71 vs. 29%, $P < 0.01$). Furthermore, EMAST+ tumors had a higher median number of harvested lymph nodes than EMAST− (11 vs. 9 nodes; $P = 0.03$). No significant association was found between EMAST status and age, gender, presence of distant metastases or metastatic lymph nodes, and overall survival. A nonsignificant difference toward worse survival in node-negative colon cancers needs confirmation in larger cohorts. EMAST+ cancers overlap and share features with MSI+ in CRC. Overall, survival was not influenced by the presence of EMAST, but may be of importance in subgroups such as node-negative disease of the colon.

Introduction

Colorectal cancer (CRC) remains a formidable global health burden and represents one of the most frequent tumors in both genders [1]. Prognosis and treatment decisions are still largely based on the TNM system, but despite revisions to improve its predictive and prognostic value, this system is still under debate [2, 3]. Among the strongest prognostic factors is lymph node status, with node-positivity usually indicating a less favorable prognosis and a need for adjuvant chemotherapy after surgery. However, even the role of lymph nodes has been debated [4], as this is a fairly rough quality indicator and fails to avoid under- and overtreatment. The growing evidence for the role of genetic variability in cancer behavior and disease outcome has therefore called for a stratified approach to cancer care based on specific molecular traits.

The last decades have shed light on several important molecular mechanisms of CRC, allowing for useful clinical subtyping and making CRC a useful model for the

understanding of cancer initiation and progression [5]. Microsatellite instability (MSI) is one such important feature and has been associated with better prognosis and tumor-specific characteristics [6]. First described in the hereditary proportion of CRCs and associated with the Lynch syndrome, MSI also occurs in about 15% of sporadic CRCs.

MSI represents a pathway of carcinogenesis that runs parallel to that of chromosomal instability and is of acknowledged prognostic, predictive, and potentially therapeutic relevance [6]. Instability at mono- and dinucleotide microsatellites is today included in the clinical and biological definition of MSI, for example, by the Bethesda criteria for MSI testing and definitions [7]. However, in a rapidly increasing number of studies, instability at tetranucleotides has been described and considered as a particular subtype of MSI over a wide range of tumor types, from those originating in the aerodigestive organs to the gastrointestinal tract [8]. This newly described form of microsatellite instability was named "elevated microsatellite alterations at selected tetranucleotides" (EMAST).

In CRC, several recent findings have suggested potential molecular mechanisms underlying EMAST [9–13]. However, the clinicopathological relevance and difference with "canonical" MSI in CRC is still poorly investigated. Thus, we aim to investigate the role of EMAST in relation to clinical- and tumor-specific data, including MSI status, and analyze the effect on survival.

Patients and Methods

Study cohort

The study cohort represents consecutive patients with non-metastatic colorectal cancer (stage I–III) who were <75 years of age at diagnosis and who entered into an in-hospital, surgeon-led systematic surveillance program per national standards at the time[14], and as previously described at the time [15–17]. All included patients presented between 1996 and 1999 and underwent curative surgery for colorectal cancer at the Department of Gastrointestinal Surgery, Stavanger University Hospital, Norway. Clinicopathological information was recorded, and follow-up was updated as of July 23rd, 2011, thus providing up to 15 years follow-up after surgery [16].

Notably, as this cohort represents patients who were eligible for a systematic surveillance program at the time of surgery, patient >75 years and stage III (pN+) not fit for adjuvant chemotherapy were excluded [15, 16]. Patients with stage III disease and who were otherwise fit were offered adjuvant chemotherapy according to national guidelines at the time, typically consisting of 5-fluorouracil and leucovorin (5-FU/LV) [18]. Thus, elderly patients and those deemed unfit for adjuvant chemotherapy or, in the case of distant recurrence, deemed not fit for a second surgery were not included in this cohort.

From the above-described initial cohort (n = 196), there were 151 specimens (98 from colon and 53 from rectum) with available tissue for DNA extraction from formalin-fixed, paraffin-embedded (FFPE) tumor and tumor-free resection margin tissues for this study.

Ethics

The study was approved as a quality assurance project (REK#2010/3414) by the Regional Ethics Committee of the Health Trust of Western Norway.

Gross and histopathological assessment

All tumors were assessed for gross and histomorphological characteristics, and staged according to TNM-classifications per routine at the time. For the current analyses, a pathologist reviewed the slides to ensure appropriate selection of tumor tissue and blocks with appropriate high tumor content (>50% viable tumor tissue) per block used for DNA extraction.

DNA extraction and fragment analysis

Following inspection by an experienced pathologist, four consecutive tumor and tumor-free 10 μm sections were cut from FFPE blocks for DNA extraction, using the Tissue DNA E.Z.N.A. kit (Omega BioTech®, Norcross, GA, USA) according to manufacturer's instructions. DNA extracted from tumor tissue and their corresponding normal tissue (from surgical resection margins) was then PCR-amplified with five tetranucleotide microsatellites primer pairs (EMAST: D20S85, D20S82, D9S242, D8S321, MYCL1, 5` fluorescently labeled) and five mono- and dinucleotide microsatellite primer pairs (MSI: NR-27, NR-21, NR-24, BAT-25, BAT-26, 5` fluorescently labeled). PCR conditions were as follows: initial denaturation step of 5′ at 95°C, followed by 37 cycles of denaturation (30″ at 95°C), annealing (90″ at 55°C), and extension (30″ at 72°C), and concluded by a final elongation step (30′ at 60°C). The primers sequences, expected amplicon sizes and fluorescent dyes are provided in Table 1.

The PCR products were analyzed for fragment lengths on a 3130xl GeneticAnalyzer (Applied Biosystems, Foster City, CA, USA), with GeneMapper v3.7 software (Applied Biosystems, Foster City, CA, USA). Tumor samples were compared with their corresponding normal samples. Those showing any number of extra peaks at $\pm 4n$ ($n \neq 0$) (tetranucleotides markers, EMAST), and/or $\pm 1n$ or $2n$ ($n \neq 0$) (mono- and dinucleotide markers, respectively, MSI) were scored as unstable for that marker.

Table 1. Name, size, fluorescent label, and primer sequences of the microsatellite markers investigated.

Marker	Amplicon size (bp)	Label	Forward primer	Reverse primer
EMAST primers				
MYCL1	181	6-FAM	TGGCGAGACTCCATCAAAG	CCTTTTAAGCTGCAACAATTTC
D20S85	146	NED	GAGTATCCAGAGAGCTATTA	ATTACAGTGTGAGACCCTG
D8S321	237	VIC	GATGAAAGAATGATAGATTACAG	ATCTTCTCATGCCATATCTGC
D20S82	249	6-FAM	GCCTTGATCACACCACTACA	GTGGTCACTAAAGTTTCTGCT
D9S242	178	PET	GTGAGAGTTCCTTCTGGC	ACTCCAGTACAAGACTCTG
MSI primers				
NR-27	89	VIC	AACCATGCTTGCAAACCACT	CGATAATACTAGCAATGACC
NR-21	110	6-FAM	GAGTCGCTGGCACAGTTCTA	CTGGTCACTCGCGTTTACAA
NR-24	128	PET	GCTGAATTTTACCTCCTGAC	ATTGTGCCATTGCATTCCAA
BAT-25	152	VIC	TACCAGGTGGCAAAGGGCA	TCTGCATTTTAACTATGGCTC
BAT-26	182	NED	CTGCGGTAATCAAGTTTTTAG	AACCATTCAACATTTTTAACCC

EMAST, Elevated microsatellite alterations at selected tetranucleotides; MSI, Microsatellite instability.

Definition of EMAST and MSI

To detect EMAST either direct sequencing or fragment analysis are generally used, with most laboratories adopting a panel of five tetranucleotide polymorphic markers (at least two unstable markers to score EMAST positivity). In CRC, up to seven microsatellite markers have been reportedly used, with EMAST considered present (EMAST+) when at least one marker was found unstable. In this study, we adopt the most used definitions of at least two out of five tetranucleotide markers unstable to confirm EMAST.

Samples showing instability in at least two out of five markers (40%) were recorded as EMAST-positive and/or microsatellite instability-high (MSI-H), while instability of one out of five markers was scored as EMAST-negative and/or microsatellite instability-low (MSI-L). If no unstable markers were found, the specimens were considered as microsatellite-stable (MSS). MSI analysis was done as previously described [17, 19]. Two investigators completed the scoring process independently, blinded to each other's results. Discordance among investigators' scoring was addressed by rerunning the samples by PCR followed by rescoring.

Statistical analysis

All statistical analyses were performed on IBM® SPSS® Statistics for Mac and Windows, version 23 (Armonk, NY, USA). Continuous variables were tested for normality by the Shapiro–Wilks test and for comparison by Mann–Whitney U test. Relationships between categorical variables were investigated via Fischer's exact and Chi-square tests, as appropriate. Overall and recurrence-free survival was assessed by Kaplan–Meier analysis using the log rank test. All tests are two-tailed and statistical significance was set at $P < 0.050$.

Results

Of the 151 patients included, the age and gender distribution together with other clinicopathological characteristics are presented in Table 2. The frequencies of MSI-H and EMAST were of 33 and 35 (22 and 23%) out of 151 patients included, respectively (Table 2). Seventy-seven percent of EMAST cases (27/35) were also MSI-H for an odds ratio (OR) of 61.9, (95% CI: 19.8–193.3; $P < 0.001$). The distribution of MSI- and EMAST-positive tumors across the different sections of the large intestine are presented in Figure 1. Dual positive cases (both EMAST+ and MSI+; $n = 27$), were predominantly located in the colon ($n = 25$; 92.6%) compared to rectum ($n = 2$; 7.8%). The ascending and transverse colon had the highest number of dual positive cases, for seven and eight each (25.9% and 29.6%, respectively).

Of the 53 rectal and 98 colon tumors, five (9.4%) and 30 (30.6%) were positive for EMAST, respectively (cumulative: 86% colon, 14% rectum, $P = 0.004$). EMAST+ tumors had a higher prevalence in proximal versus distal colon (77% vs. 23%, $P = 0.004$) and were also associated with advanced t-stage in both EMAST (OR 6.0, 95% CI: 1.4–26.6; $P = 0.008$) and MSI cancers (OR 5.5, 95% CI: 1.3–24.5; $P = 0.013$), respectively.

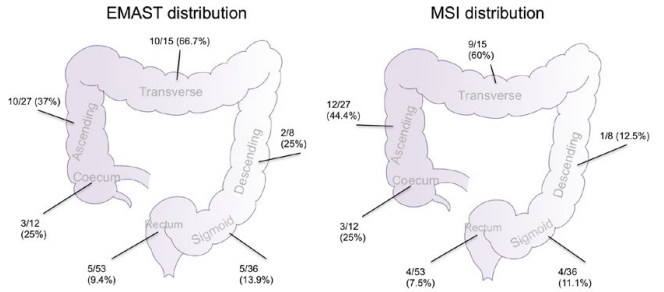
EMAST+ tumors had a higher median number of harvested lymph nodes than EMAST– (11 vs. 9 nodes; $P = 0.029$; Fig. 2), but no difference in the number of lymph nodes positive for tumor cells infiltration was found.

A total of 38 (25.2%) patients developed metastases and died from CRC in this cohort. Neither EMAST nor MSI predicted risk of development of distant metastases, nor was EMAST predictive for disease-specific and for overall long-term survival in this cohort (Fig. 3A). However, on subanalyses of colon cancers only, as these harbor a higher frequency of EMAST-positive cases in comparison with rectal cancers, a nonsignificant difference in

Table 2. Characteristics of patient and tumors according to EMAST status

		n (%)		*P*	*n* (%)		*P*
n = 151		EMAST–	EMAST+		MSS	MSI	
Age (years)	<65	58 (73.4)	21 (26.6)	*0.299*	57 (72.2)	22 (27.8)	*0.062*
	≥65	58 (80.6)	14 (19.4)		61 (84.7)	11 (15.3)	
Gender	M	73 (79.3)	19 (20.7)	*0.358*	75 (81.5)	17(18.5)	*0.210*
	F	43 (72.9)	16 (27.1)		43 (72.9)	16 (27.1)	
Tumor location	Colon	68 (69.4)	30 (30.6)	*0.003*	69 (70.4)	29 (29.6)	*0.002*
	Rectum	48 (90.6)	5 (9.4)		49 (92.5)	4 (7.5)	
Tumor stage	$T_{1–2}$	31 (93.9)	2 (6.1)	*0.008*	31 (93.9)	2 (6.1)	*0.013*
	$T_{3–4}$	85 (72.0)	33 (28.0)		87 (73.7)	31 (26.3)	
Tumor grade	Poor/mucinous	10 (50.0)	10 (50.0)	*0.002*	8 (40.0)	12 (60.0)	*<0.001*
	Moderate/well	106 (80.9)	25 (19.1)		110 (84.0)	21 (16.0)	
Tumor size[1]	≥5 cm	44 (68.8)	20 (31.2)	*0.022*	43 (67.2)	21 (32.8)	*<0.001*
	<5 cm	67 (84.8)	12 (15.2)		72 (91.1)	7 (8.9)	

EMAST, Elevated microsatellite alterations at selected tetranucleotides; MSI, microsatellite instability; MSS, microsatellite stable. Values highlighted in bold indicate significance level of $p < 0.05$.
[1]Size missing in eight samples (5.3%)



	EMAST –	EMAST +	Total
MSS	110 (95%)	8 (23%)	118
MSI-H	6 (5%)	27 (77%)	33
Total	116	35	151

P = 0.001

Figure 1. Elevated microsatellite alterations at selected tetranucleotides (EMAST) and microsatellite instability (MSI) cancer distribution in colon and rectum. EMAST denotes elevated microsatellite alterations at selected tetranucleotides; MSI denotes microsatellite instability.

long-term cancer-specific survival was noted, particularly for the node-negative (stage I–II) patients (Fig. 3B). Furthermore, these depicted an apparent worse outcome for EMAST+ (Fig. 3C) compared to microsatellite-stable cancers, and those with either one form of microsatellite instability only.

Discussion

In this cohort study, we found EMAST+ CRC to largely overlap with features associated with MSI+ cancers, including a predominant location in the colon, association with low-grade histology, larger tumor size, and advanced depth of growth (T3-4). Despite a higher number of lymph nodes sampled for EMAST+ cancers, there was no difference in the number of malignant nodes (neither in actual numbers nor in the rate of pN+ cases) and no statistically significant effect on survival could be found. The nonsignificant yet apparent difference in survival curves between groups depicted in Figure 3 with a trend toward reduced survival in EMAST+ cancers, specifically for node-negative colon cancers needs verification in larger cohorts.

Several findings need comment in this study. We found a significantly lower prevalence of EMAST in

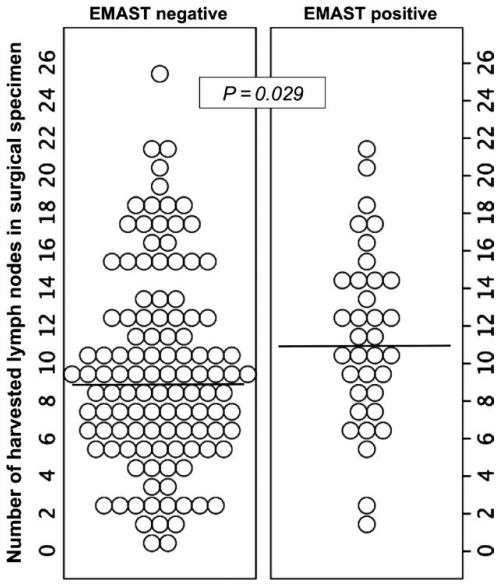


Figure 2. Number of lymph nodes found in the resected specimen, according to elevated microsatellite alterations at selected tetranucleotides (EMAST) status. *P*-value for difference in median number between groups.

our sample (23%) when compared with the majority of the studies available in the literature, reporting EMAST frequencies in CRC ranging between 33% and 64.8% [9, 12, 13, 20–22]. As previously reviewed [8], such variation could be due to the type and number of markers and the thresholds used in EMAST analyses, with the highest (60% and over) frequencies reported by groups using a less stringent approach to define EMAST [22]. The degree of overlap between MSI cases and EMAST (77% in this study), is in line with studies that report between 67% and 100% overlap, and thus, the selected cohort should be within the range of variation as reported elsewhere in the literature. The fact that most, and in some cases all, the MSI-H tumors are also EMAST+, and that the latter tumors are generally more prevalent, could suggest a cause–consequence relationship. It is also interesting to see the prognostic features shared by MSI and EMAST-positive cancers (Table 2), including larger size, poor differentiation, and depth of growth. While these features would normally be associated with a worse prognosis, some emerging data suggest that MSI and possibly EMAST cancers may be associated with specific immune reactions and T-lymphocyte infiltrations associated with a more favorable outcome [10].

Neither EMAST nor MSI was found to significantly correlate with survival (neither overall nor recurrence-free). Of the four studies investigating survival, specifically in EMAST+ and EMAST− tumors currently available in the scientific literature, two found no significant difference in overall survival [9, 12]. However, one study of metastatic CRC observed that MSI-H tumors that also displayed EMAST+ had significantly worse overall survival, compared to non-EMAST cancers with MSI-H [13]. This is in line with the nonsignificant findings in this study, albeit in early-stage colon cancers. A further study found a significantly reduced recurrence-free survival (RFS) in EMAST+ cancers, when compared to MSI-H, but the degree of overlapping EMAST/MSI was not disclosed [23]. In both studies the MSI-H/EMAST- group was composed of a limited number of individuals.

As demonstrated in the current cohort, both MSI and EMAST produce larger tumors. If this is because of EMAST/MSI tumors developing more quickly due to a much higher rate of mutation that a defective MMR system confers to the nest of cancer-initiating cells remains speculative, but warrants investigation. Conversely, EMAST/MSI may be genetic events that occur as a side effect of other drivers of carcinogenesis, or merely reflects a high turnover and induction of genetic errors during rapid growth of the tumor cells. Some studies look into mechanisms of EMAST [11, 22, 24, 25], and point to a role of hypoxia, oxidative stress, and DNA repair mechanisms [26]. However, overall data are scarce and further understanding is thus needed.

Some limitations should be considered when comparing our study cohort with other patient series. First, we included only patients who were <75 years and who entered a systematic surveillance program after surgery [15], and excluded elderly patients or those with comorbidities who were unfit for adjuvant chemotherapy or unlikely to tolerate metastatic surgery. Thus, we have introduced a clinical bias towards younger, fitter patients with stage I–III colorectal cancers in this series. This should be taken into account when interpreting our findings, as the results could thus not apply to a more general patient cohort with higher age and that frequently included stage IV cancers. The latter may also be the reason for a nonsignificant trend in the analyses, as others have found EMAST status of importance in stage IV and metastatic disease [13, 23, 27]. For example, Venderbosch et al.[13] found a statistically significant difference in survival between patients with and without EMAST. Notably, the included patients all came from clinical phase III trials (the Dutch CAIRO and CAIRO2 studies) of metastatic CRC, and thus all patients in the cohort had an unfavorable outcome [13]. Indeed, EMAST may be an accumulated effect of worse biology, higher mutational load, and thus play a

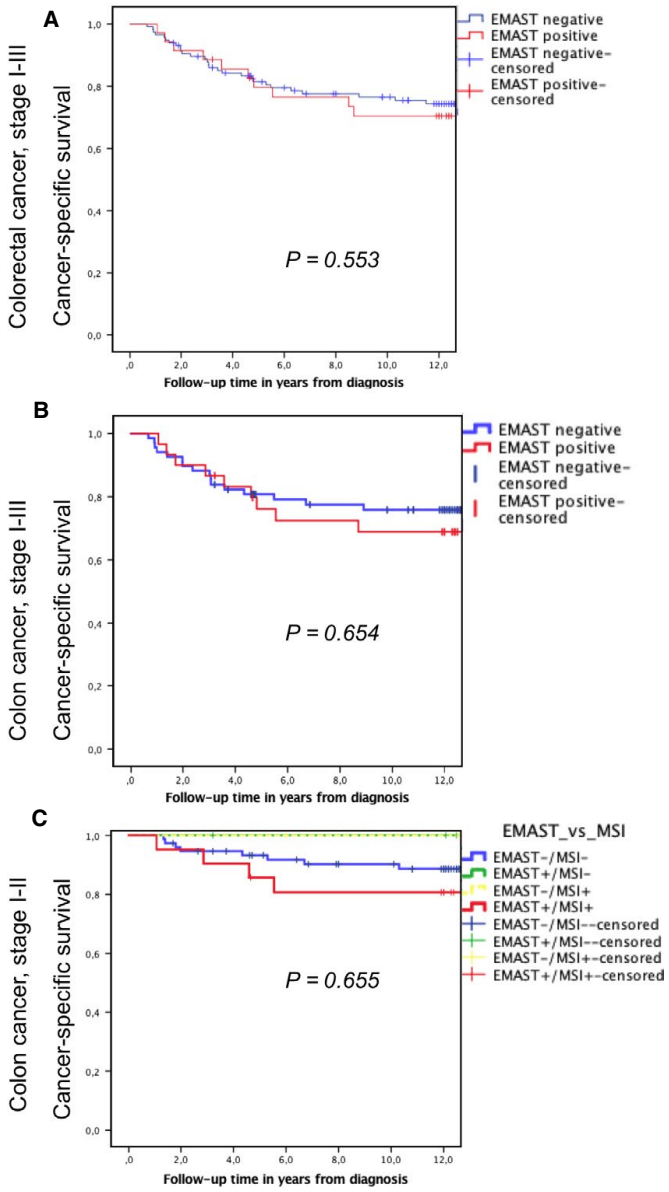


Figure 3. Cancer-specific survival according to elevated microsatellite alterations at selected tetranucleotides (EMAST) status, stages, and location. (A) overall cancer-specific survival for all colorectal stage I–III cancers, with no significant difference, yet a somewhat poorer outcome in EMAST + cancers. (B) cancer-specific survival for colon cancers only, again with nonsignificant poorer survival in EMAST+ patients. (C) outcome for node-negative (stage I–II) colon cancers, split into patients with microsatellite stable (MSS; blue line), microsatellite instability (MSI)-/EMAST+ (green), MSI+/EMAST– (yellow), and the dual positive cancers of EMAST+ and MSI+ (red), with poorest outcome.

more prominent role in biology in late stage (e.g., metastatic disease) compared with early (stage I–III) disease. Evidence that EMAST may act as a potential biological modulator among the different types of molecular classes (e.g., microsatellite instability, epigenetics, and chromosomal instability) involved in CRC have been proposed [28], and is further suggested in a combined series of metastatic disease in CRC [27]. Thus, EMAST may be more specific for tumor biology and disease outcome in late-stage groups [29], such as colorectal liver metastasis, but further studies need to corroborate these findings. Notably, several aspects in clinical practice, including a higher frequency of metastatic surgery and extending adjuvant chemotherapy to elderly patients have occurred since the late 1990s, so clinical differences in practice may have introduced selection and outcome bias in this cohort compared with more recent cohorts. However, long-term follow-up would not be possible with more recent cohorts, so the true eventual outcome of the patients (e.g., death from disease or other cause; still alive with no evidence of disease etc.) is likely to have been captured accurately in this series. The small cohort prevents from robust subgroup analysis and these should therefore be interpreted with caution. An apparent prognostic role in stage I–II CRC warrants further investigation. Finally, how EMAST should be determined lacks firm definition in the current literature, possibly explaining why our results deviate from others based on the choice of defined markers and numbers used for positivity. This methodological issue must be solved through further clarification of biological mechanisms and ability for robust and valid tests of selected markers or panels of markers.

While the clinical role of EMAST in CRC is still being investigated, the biological implications of recent investigations may yield findings of new mechanisms that have a clinical relevance in selected patients at both extremes of presentation, either as early cancers or as metastatic disease. Thus, further investigation into the biological mechanisms and their potential clinical implications should be pursued. Whether EMAST is an epiphenomenon or a specific genetic trait warrants further investigation.

Conflict of Interest

None declared.

References

1. Brenner, H., M. Kloor, and C. P. Pox. 2014. Colorectal cancer. Lancet 383:1490–1502.
2. Quirke, P., G. T. Williams, N. Ectors, A. Ensari, F. Piard, and I. Nagtegaal. 2007. The future of the TNM staging system in colorectal cancer: time for a debate? Lancet Oncol. 8:651–657.
3. Lea, D., S. Haland, H. R. Hagland, and K. Søreide. 2014. Accuracy of TNM staging in colorectal cancer: a review of current culprits, the modern role of morphology and stepping-stones for improvements in the molecular era. Scand. J. Gastroenterol. 49:1153–1163.
4. Veen, T., B. S. Nedrebø, K. Stormark, J. S. Søreide, H. Kørner, and K. Søreide. 2013. Qualitative and quantitative issues of lymph nodes as prognostic factor in colon cancer. Dig. Surg. 30:1–11.
5. Grady, W. M., and J. M. Carethers. 2008. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 135:1079–1099.
6. Sinicrope, F. A., and D. J. Sargent. 2012. Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. Clin. Cancer Res. 18:1506–1512.
7. Umar, A., C. R. Boland, J. P. Terdiman, S. Syngal, A. de la Chapelle, J. Rüschoff, et al. 2004. Revised Bethesda Guidelines for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. J. Natl Cancer Inst. 96:261–268.
8. Watson, M. M., M. Berg, and K. Soreide. 2014. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. Br. J. Cancer 111:823–827.
9. Yamada, K., S. Kanazawa, J. Koike, H. Sugiyama, C. Xu, K. Funahashi, et al. 2010. Microsatellite instability at tetranucleotide repeats in sporadic colorectal cancer in Japan. Oncol. Rep. 23:551–561.
10. Lee, S. Y., K. Miyai, H. S. Han, D. Y. Hwang, M. K. Seong, H. Chung, et al. 2012. Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. Dig. Dis. Sci. 57:72–78.
11. Tseng-Rogenski, S. S., H. Chung, M. B. Wilk, S. Zhang, M. Iwaizumi, and J. M. Carethers. 2012. Oxidative stress induces nuclear-to-cytosol shift of hMSH3, a potential mechanism for EMAST in colorectal cancer cells. PLoS ONE 7:e50616.
12. Hamaya, Y., C. Guarinos, S. S. Tseng-Rogenski, M. Iwaizumi, R. Das, R. Jover, et al. 2015. Efficacy of Adjuvant 5-Fluorouracil Therapy for Patients with EMAST-Positive Stage II/III Colorectal Cancer. PLoS ONE 10(5):e0127591.
13. Venderbosch, S., Vliet S. van Lent-van, A. F. de Haan, M. J. Ligtenberg, M. Goossens, C. J. Punt, et al. 2015. EMAST is associated with a poor prognosis in microsatellite instable metastatic colorectal cancer. PLoS ONE 10:e0124538.
14. Søreide, K., J. H. Traeland, P. J. Stokkeland, T. Glomsaker, J. A. Søreide, and H. Kørner. 2012.

Adherence to national guidelines for surveillance after curative resection of nonmetastatic colon and rectum cancer: a survey among Norwegian gastrointestinal surgeons. Colorectal Dis. 14:320–324.
15. Kørner, H., K. Søreide, P. J. Stokkeland, and J. A. Søreide. 2005. Systematic follow-up after curative surgery for colorectal cancer in Norway: a population-based audit of effectiveness, costs, and compliance. J. Gastrointest. Surg. 9:320–328.
16. Veen, T., K. Stormark, B. S. Nedrebø, M. Berg, J. A. Søreide, H. Kørner, et al. 2015. Long-Term Follow-Up and Survivorship After Completing Systematic Surveillance in Stage I-III Colorectal Cancer: who Is Still at Risk? J. Gastrointest. Cancer 46:259–266.
17. Søreide, K., A. Slewa, P. J. Stokkeland, B. van Diermen, E. A. Janssen, J. A. Søreide, et al. 2009. Microsatellite instability and DNA ploidy in colorectal cancer: potential implications for patients undergoing systematic surveillance after resection. Cancer 115:271–282.
18. Dahl, O., O. Fluge, E. Carlsen, J. N. Wiig, H. E. Myrvold, B. Vonen, et al. 2009. Final results of a randomised phase III study on adjuvant chemotherapy with 5 FU and levamisol in colon and rectum cancer stage II and III by the Norwegian Gastrointestinal Cancer Group. Acta Oncol. 48:368–376.
19. Søreide, K. 2011. High-fidelity of five quasimonomorphic mononucleotide repeats to high-frequency microsatellite instability distribution in early-stage adenocarcinoma of the colon. Anticancer Res. 31:967–971.
20. Devaraj, B., A. Lee, B. L. Cabrera, K. Miyai, L. Luo, S. Ramamoorthy, et al. 2010. Relationship of EMAST and microsatellite instability among patients with rectal cancer. J. Gastrointest. Surg. 14:1521–1528.
21. Lee, S. Y., H. Chung, B. Devaraj, M. Iwaizumi, H. S. Han, D. Y. Hwang, et al. 2010. Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias. Gastroenterology 139:1519–1525.
22. Haugen, A. C., A. Goel, K. Yamada, G. Marra, T. P. Nguyen, T. Nagasaka, et al. 2008. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. Cancer Res. 68:8465–8472.
23. Garcia, M., C. Choi, H. R. Kim, Y. Daoud, Y. Toiyama, M. Takahashi, et al. 2012. Association between recurrent metastasis from stage II and III primary colorectal tumors and moderate microsatellite instability. Gastroenterology 143:48–50. e1.
24. Tseng-Rogenski, S. S., Y. Hamaya, D. Y. Choi, and J. M. Carethers. 2015. Interleukin 6 Alters Localization of hMSH3, Leading to DNA Mismatch Repair Defects in Colorectal Cancer Cells. Gastroenterology 148:579–589.
25. Campregher, C., G. Schmid, F. Ferk, S. Knasmüller, V. Khare, B. Kortüm, et al. 2012. MSH3-deficiency initiates EMAST without oncogenic transformation of human colon epithelial cells. PLoS ONE 7:e50541.
26. Carethers, J. M., M. Koi, and S. S. Tseng-Rogenski. 2015. EMAST is a Form of Microsatellite Instability That is Initiated by Inflammation and Modulates Colorectal Cancer Progression. Genes (Basel) 6:185–205.
27. Koi, M., M. Garcia, C. Choi, H. R. Kim, J. S. Koike, H. Hemmi, et al. 2016. Microsatellite Alterations With Allelic Loss at 9p24.2 Signify Less-Aggressive Colorectal Cancer Metastasis. Gastroenterology. doi: http://dx.doi.org/10.1053/j.gastro.2015.12.032. [Epub ahead of print].
28. Carethers, J. M., and B. H. Jung. 2015. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. Gastroenterology 149:1177–1190. e3.
29. Søreide, K., M. M. Watson, and H. R. Hagland. 2016. Deciphering the molecular code to colorectal liver metastasis biology through microsatellites and allelic loss: the good, the bad and the ugly. Gastroenterology. doi: http://dx.doi.org/10.1053/j.gastro.2016.02.060. [Epub ahead of print].

II

Translational Oncology
www.transonc.com

Volume 12 Number 12 December 2019 pp. 1583–1588 1583

Elevated Microsatellite Alterations at Selected Tetranucleotides (EMAST) Is Not Attributed to MSH3 Loss in Stage I-III Colon cancer: An Automated, Digitalized Assessment by Immunohistochemistry of Whole Slides and Hot Spots[1,2]

Martin M Watson*,†,‡, Dordi Lea*,‡,§, Hanne R. Hagland*,¶ and Kjetil Søreide*,†,‡

*Gastrointestinal Translational Research Unit, Laboratory for Molecular Medicine, Stavanger University Hospital, Gerd Ragna Bloch Thorsens Gate 8, 4011, Stavanger, Norway; †Department of Gastrointestinal Surgery, Stavanger University Hospital, Gerd Ragna Bloch Thorsens Gate 8, 4011, Stavanger, Norway; ‡Department of Clinical Medicine, University of Bergen, Bergen, Norway; §Department of Pathology, Stavanger University Hospital, Gerd Ragna Bloch Thorsens Gate 8, 4011, Stavanger, Norway; ¶Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, Norway

Abstract

INTRODUCTION: EMAST is a poorly understood form of microsatellite instability (MSI) in colorectal cancer (CRC) for which loss of MSH3 has been proposed as the underlying mechanism, based on experimental studies. We aimed to evaluate whether MSH3 loss is associated with EMAST in CRC. **METHODS:** A consecutive cohort of patients with stage I-III CRC. Digital image analysis using heatmap-derived hot spots investigated MSH3 expression by immunohistochemistry. Fragment analysis of multiplex PCR was used to assess MSI and EMAST, and results cross-examined with MSH3 protein expression. **RESULTS:** Of 152 patients, EMAST was found in 50 (33%) and exclusively in the colon. Most EMAST-positive cancers had instability at all 5 markers, and EMAST overlapped with MSI-H in 42/50 cases (84%). The most frequently altered tetranucleotide markers were D8S321 (38.2% of tumors) and D20S82 (34.4%). Subjective evaluation of MSH3 expression by IHC in tumor found ≤10% negative tumor cells in all samples, most being ≤5% negative. Digital analysis improved the detection but showed a similar spread of MSH3 loss (range 0.1—15.7%, mean 2.2%). Hotspot MSH3 negativity ranged between 0.1 to 95.0%, (mean 8.6%) with significant correlation with the whole slide analysis (Spearman's rho = 0.677 $P < .001$). Loss of MSH3 expression did not correlate with EMAST. **CONCLUSIONS:** In a well-defined cohort of patients with CRC, loss of MSH3 was not associated with EMAST. Further investigation into the mechanisms leading to EMAST in CRC is needed.

Translational Oncology (2019) 12, 1583–1588

Introduction

Microsatellite instability (MSI) is caused by defects in the mismatch-repair (MMR) family of proteins [1]. This results in mosaic populations of cells bearing microsatellite loci with diverse numbers of repeats due to uncorrected slippages during DNA replication. Silent or deleterious consequences arise according to the microsatellites affected and their location within the genome. In colorectal cancer (CRC), MSI is recognized as an alternative carcinogenic pathway to the chromosomal instability model, with a series of clinical and pathological implications [2]. MSI continues to be debated as a prognostic factor in CRC [3—5], and is implicated

Address all correspondence to: Martin M Watson; Department of Gastrointestinal Surgery, Stavanger University Hospital, Gerd Ragna Bloch Thorsens gate 8, 4011, Stavanger, Norway. E-mail: martin.watson@uib.no

[1] Conflict of interest: The authors declare no conflict of interest.

[2] The study was funded by the Folke Hermansen Fond, by an unrestricted grant from Mjaaland foundation and from the University fund at University of Stavanger (UiS).
Received 22 June 2019; Revised 8 August 2019; Accepted 14 August 2019

1936-5233/19
https://doi.org/10.1016/j.tranon.2019.08.009

in the "hypermutated" or "immunogenic" consensus molecular subtype [6].

Elevated microsatellite alterations at selected tetranucleotides (EMAST) is a variant of MSI described in lung, skin, prostate, and other cancers, including CRC [7]. While MSI was initially defined as instability at mono- and dinucleotide repeats (e.g. CA_n) [8], today commonly measured in a panel consisting exclusively of mononucleotides [9], the definition of EMAST is based on instability found in tetranucleotides (e.g. $AAAG_n$).

MSI in CRC commonly displays loss of expression of MMR proteins such as MLH1, MSH2, MSH6 and PMS2. The MSH2 member of the MMR family can dimerize with either MSH6 or MSH3 to form the MutSα or MutSβ complexes [1]. The latter is believed to have a higher affinity for repair of longer IDLs and mismatched sequences occurring during replication, such as tetranucleotides. MSH3 has therefore been implicated as a potential candidate to explain instability at longer microsatellites, as found in EMAST.

In vitro, MSH3 dysfunction was associated to instability at several tetranucleotide loci in MLH1- and MSH3-deficient CRC cell lines via whole chromosome transfer, as well as silencing/knockdown studies [10–12]. Additionally, it has been suggested that activity of MSH3 could be impaired by its dislocation from the nucleus to the cytosol, a process possibly mediated by interleukin-6 in a context of oxidative stress in CRC cell lines [12,13]. Furthermore, the cancer genome atlas (TCGA) consortium described *MSH3* frameshift mutations—and not point mutations—as common (40%) in a subclass of CRCs defined as hypermutated and microsatellite-unstable [14]. Later, it was shown how *MSH3*, specifically in CRC, represents a frequent target of frameshift mutation, as opposed to the promoter hypermethylation that occurs at *MLH1* in MSI CRCs [15]. The fact that the *MSH3* gene contains a mononucleotide-repeat locus could suggest that frameshift mutations in *MSH3* are a consequence of instability at mononucleotides initiated by loss of MLH1. In the mentioned studies it was not reported whether the frameshift mutations found in *MSH3* were silent or non-silent, and their effect on functionality of the protein can therefore not be inferred. Should MSH3 be proven as the biological driver of EMAST, a causal relationship between MSI and EMAST could therefore be speculated. Thus, the relationship between MSH3 and EMAST need to be investigated in clinical cohorts. However, to date only 3 studies in human tissue have investigated immunohistochemical (IHC) staining of MSH3 in patients, and are discordant in the association between MSH3 expression with EMAST [10,16,17].

The aim of this study was to assess if MSH3 loss could explain EMAST in colorectal cancer and, if so, to develop a standardized method to more accurately assess protein loss in the samples.

Materials and Methods

The patient cohort was derived from the ACROBATICC project [18] (clinicaltrials.gov identifier: NCT01762813) and is conducted in accordance to national regulations and approved by regional ethics committee (REK Helse Vest, #2012/742). Written informed consent was obtained from each participant prior to inclusion in the study.

Patient Material

Formalin-fixed, paraffin-embedded (FFPE) tumor and normal tissue derived from stage I-III surgically removed CRC was used in this study. Appropriate slides were assessed by a certified pathologist and representative tissue blocks selected for DNA extraction, fragment analysis and immunohistochemistry.

EMAST and MSI Analyses

FFPE blocks were selected by an experienced pathologist and 4 × 10 μm sections were cut at a microtome. Automated DNA extraction was carried out using AllPrep DNA/RNA FFPE kit (Qiagen, Hilden, Germany) on a QiaCUBE instrument (Qiagen) according to manufacturer's instructions. Nucleic acid concentration and purity were measured on a NanoDrop 2000 (ThermoFischer scientific, Waltham, USA).

Multiplex PCR reactions (one for each MSI and EMAST) were set up for tumor and normal DNA from each patient. TypeIT microsatellite (Qiagen) master mix, together with a blending of 5 × 5′-fluorescently labeled primer pairs was used for each reaction. PCR conditions were as follows: 5′ at 95 °C (initial denaturation and enzyme activation), followed by 37 cycles of 30″ at 95 °C (denaturation), 90″ at 55 (MSI) or 57 °C (EMAST, annealing) and 30″ at 72 °C (extension). A final extension step for 30′ at 60 °C. The primers for EMAST were specific to the tetranucleotide loci MYCL1, D20S85, D20S82, D9S242 and D8S321 [19]. The primers for MSI were specific for BAT-26, NR-21, NR-24 and NR-27 [9,20], which are all quasimonomorphic mononucleotide repeats with a high fidelity to high-frequency MSI (MSI-H) as shown previously [21]. To define a tumor as EMAST and/or MSI-H, at least 2/5 markers needed to be unstable in their respective panels.

MSH3 Immunohistochemistry

Antigen retrieval and antibody dilution were optimized prior to the study onset. From FFPE blocks, 2 μm sections were cut and mounted onto Superfrost Plus slides (Menzel, Braunschweig, Germany). The sections were incubated at 60 °C for 1 h and then placed in the Dako Omnis autostainer (DAKO Agilent, Santa Clara, CA, USA). Automated protocol from the manufacturer was followed. Following deparaffinization and rehydration, antigen retrieval was performed at 97 °C for 30 minutes, and the slides were then incubated with the primary anti-MSH3 antibody (rabbit monoclonal anti-human MSH3; AbCam, Cambridge UK), clone EPR4334 (2), diluted 1:100 for 1 h. A peroxidase-DAB detection kit (Envision+, DAKO) was used to visualize the immune-complex. Sections were then counterstained with hematoxylin, dehydrated in increasing concentrations of ethanol and mounted manually.

Subjective IHC Score

Slides were evaluated and scored by an experienced pathologist for nuclear positivity of MSH3 (given as per cent, %) blinded to MSI and EMAST status of each case. A composite high-resolution image at 20× magnification of each slide was obtained with a Leica SCN400 scanner and uploaded onto an internal digital image hub for image analysis.

Digital Image Analysis

To increase scoring sensitivity of MSH3 expression, digitalized whole-slide and hotspots scoring of positive–negative nuclei in the tumor portion of the sections was performed with the aid of Visiopharm Integrator System software (VIS; Visiopharm A/S, Hoersholm, Denmark). An image analysis algorithm using Bayesian classification methods was built in an app-based tool which allowed for identification of tumor tissue within the scanned slides, and for the highlighting of its contours (Figure 1, *A–C*). Manual revision of each slide ensured then rigorous exclusion of tissue folds, stroma, necrotic areas, immune and

blood cells, and normal tissue from the analysis. A second, app-based algorithm was then developed to allow for the marking of positive and negative cells with colored labels (Figure 1, *D*–*K*). A heatmap based on the label associated with MSH3-negativity was created for each of the whole slides (Figure 1, E-F), for unbiased placement of one 0.8 mm^2 hotspots on the areas where the concentration of MSH3-negative cells was highest within the slides (Figure 1, G-H). Relative MSH3 negativity (both whole-slide and hotspots) was then derived from the ratio of negative label area and total negative and positive label areas, as calculated via the developed classifier.

Statistical Analysis

All statistical analyses were carried out using IBM SPSS statistics v. 25. Chi-square or Fishers Exact test were used for categorical variables. Relationship between different operators (pathologist/digitalized whole-slide/digitalized hotspot) in the scoring of MSH3 expression were tested using the Spearman's rank order correlation. All tests were two-tailed with statistical significance set at $P < .050$.

Results

Patients Characteristics

Median age was 71.5 years (range 37.0–92.0), female patients were 85 (56%). There were 31 (20%) rectum and 121 (80%) colon cancers, 71 (59%) of which were in the proximal tract.

Tumor stages were equally represented with 51, 51 (34%) and 50 (33%) cases for Stage I, II and III, respectively.

EMAST and MSI Analysis

EMAST was present in 50 (33%, Figure 2, *A* and *B*) and MSI-H in 44 (28%) of 152 tumors, all of which in the colon (none in the rectum) and 90% in the proximal part of the colon. EMAST was positively associated with MSI (42/50 EMAST we also MSI-H, 84% overlap; $P < .001$), but not with tumor stage.

Almost half of the cohort (45%) showed no instability at any EMAST marker, while most EMAST-positive patients had instability at all 5 markers (Figure 2, *C* and *D*). The most frequently mutated marker was D8S321 (38% of tumors), followed by D20S82 (34%). The marker with the least events in microsatellite-stable tumors (most specific to EMAST status) was D20S85 (Table 1). Thirty-four (34) tumors had one unstable EMAST marker and were thus classified as EMAST-negative (Figure 2*D*).

MSH3 Analysis

Whole slide microscopical evaluation of tumor area for MSH3 loss found only <10% of the tumor nuclei with negative stain in all samples, most being ≤5% negative (Figure 3, *A* and *B*).

Digital analysis (Figure 3) showed a similar spread (range 0.1–15.7%, mean 2.2%). Hotspot MSH3 negativity ranged between 0.1 to 95.0%, (mean 8.6%) with a significant correlation between the two sets of measurements (Spearman's rho = 0.677 $P < .001$), indicating that the measurements in the hotspots are indicative of the rest of the tumor.

MSH3 loss in hotspots was categorized in subclasses according to different cut-off points (1%, 5%, 10%, 25% and 50%) to establish whether a certain degree of protein loss could account for EMAST presence. None of the subclasses correlated with EMAST status, nor MSI-H or disease recurrence.

Heterogeneous expression of MSH3 was noted in some cases, with nuclei staining only partially positive for the protein (Figure 3*C*). However, these findings were not related to EMAST nor MSI.

Discussion

Based on a robust, automated, and digitalized protocol of IHC assessment with a verified MSH3 antibody, this study could not demonstrate an association between loss of MSH3 and EMAST in CRC. While other mechanistic contributions of MSH3 to EMAST

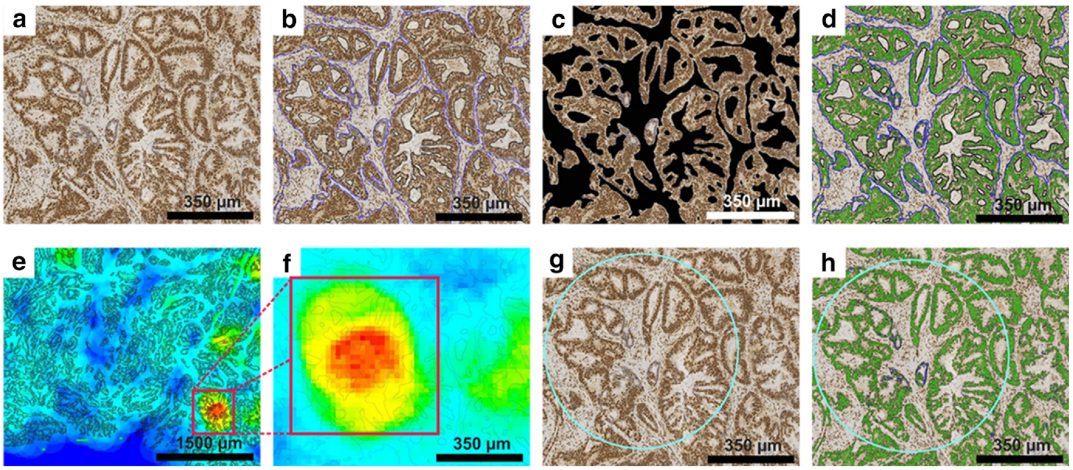


Figure 1. MSH3 immunohistochemistry virtual image analysis process.(**A-C**) An initial gross exclusion of stroma, tissue folds and normal tissue is carried out, selecting a "work area" where an app-based algorithm is then run to specifically mark tumor cells and exclude stroma. This results in a highlighted region of interest (ROI, marked in blue); (**D**) A second app-based algorithm classifies cells on the basis of their positivity (green) or negativity (blue) for the MSH3 protein; (**E-F**) A heatmap is created to highlight areas on the whole slide where the highest concentration of negative cells are located; (**G**) Based on the heatmap created in (**E**), a 0.8 mm^2 round ROI (hotspot, light blue circle) is placed and becomes the focus of the analysis; (**H**) Fully classified, hotspot-derived ROI.

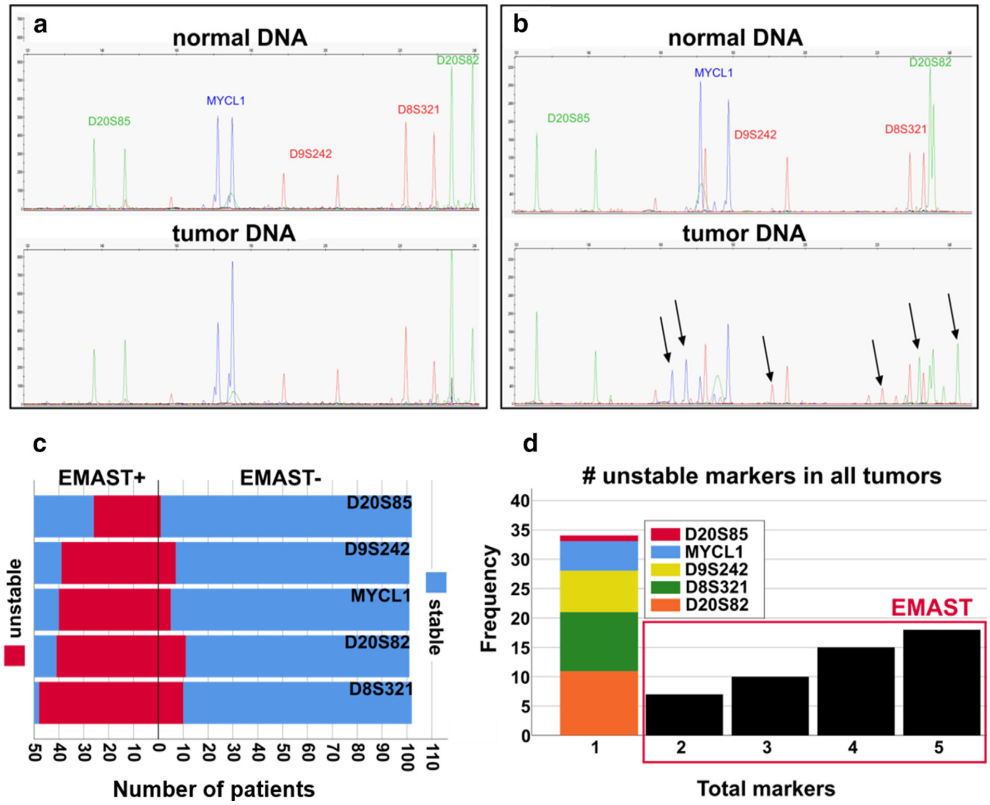


Figure 2. EMAST analysis. Electropherograms of multiplex PCR fragment analysis are shown for (**A**) an EMAST-negative and (**B**) an EMAST-positive patient. Arrows indicate extra PCR products at +/− 4_n bp in unstable markers. (**C**) Stacked bar population graph showing frequency of instability (red) at each EMAST marker, in EMAST-positive and EMAST-negative populations. (**D**) bar chart showing proportions of patients grouped by total number of unstable EMAST markers. For patients bearing only 1 marker mutated (not EMAST, according to our thresholds) the bar is stacked to specify each marker's abundance.

cannot be ruled out based on the current experiment, this study suggests that neither protein loss, nor protein translocation (e.g. from cytosol to nuclei) is likely to be the cause of EMAST in patients with CRC. Thus, other mechanisms to EMAST must be considered beyond the notion that loss of MSH3 is essential to EMAST development [22]. Several points warrant further discussion.

The role of MSH3 expression in EMAST is debated and not yet fully understood. Some previous mechanistic investigations have based their experimental studies on effects seen in cancer cell lines [12,13]. In human CRC, however, one study showed high degree of MSH3 loss found in EMAST-positive cancers [16], but others have found no significant association between MSH3 loss and EMAST [17].

One possible explanation for the discordance between the current findings and previous results could be the source and type of antibody used for MSH3. One previous study used a clone to MSH3 that is no longer on the market [17], while neither the source nor clonality of the antibody used in a second study [16] could be reproduced. A third study used an affinity-purified rabbit polyclonal antibody [10]. The antibody used in the current study was developed and tested for human immunohistochemistry protocols, thus thoroughly validated using appropriate positive and negative controls.

Table 1. EMAST markers, genomic loci (GRCh38/hg38 assembly), repeat type (− strand), and frequency of mutation.

MARKER	LOCUS	REPEAT	EMAST+	EMAST-	MSI	MSS	all tumors
MYCL1 *, n (%)	1p34.3	AAAG	40 (88.9)	5 (11.1)	38 (84.4)	7 (15.6)	45 (29.8)
D9S242 *, n (%)	9q33.3	AAAG	39 (84.8)	7 (15.2)	38 (82.6)	8 (17.4)	46 (30.5)
D20S82 *, n (%)	20p12.3	AAAG	41 (78.8)	11 (21.2)	36 (69.2)	16 (30.8)	52 (34.4)
D8S321, n (%)	8q24.21	AAAG	48 (82.8)	10 (17.2)	40 (69.0)	18 (31.0)	58 (38.2)
D20S85, n (%)	20q12	AAAG	26 (96.3)	1 (3.7)	25 (92.6)	2 (7.4)	27 (17.8)

* N = 151.

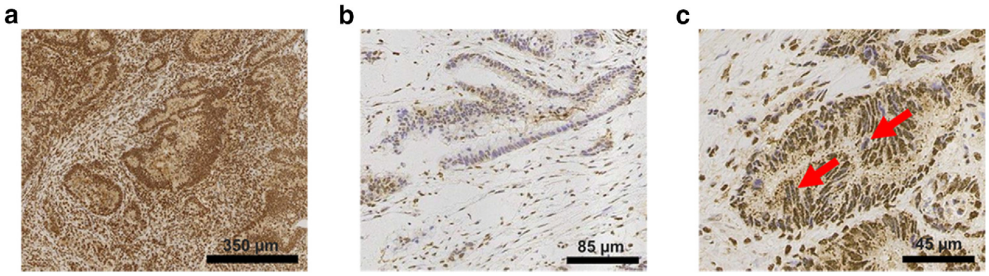


Figure 3. D**igitalized evaluation of MSH3 expression.**(**A**) The entire cohort was scored by an experienced pathologist as >90% positive for MSH3 (majority >95%). (**B**) With the help of digitalized hotspot analysis, small areas where a higher proportion of negative cells were found in all the slides. (**C**) Within most MSH3-positive cells, some nuclei exhibited heterogeneous staining.

Subjective selection and cherry picking of regions of interest to measure IHC stains may be another issue across studies. In the current study, in order to account for variation and involuntary selection bias in the analysis of MSH3 expression, we constructed a heatmap-based hotspot analysis in addition to evaluation of whole slides. The hotspot approach is presently proving useful in the evaluation of prognostic biomarkers with thresholds of low expression, such as Ki67 in breast cancer [23], and has been implemented in the Norwegian, Danish and Swedish pathology guidelines [24–26]. However, results will still largely depend on the robustness of the antibodies used and their reproducibility. For MSH3, the previous studies used antibodies that are no longer on the market, thus difficult to reproduce.

Immunohistochemistry is a complex multi-step procedure, performed by laboratories on a range of instruments and interpreted subjectively by pathologists according to guidelines that are often designed for a specific diagnostic test. Autostainer instruments and digital image analysis have reduced the discrepancies originating from manual experiments [27], although this technique is still often performed manually in research labs and as such open to a multitude of sources of variation [28,29]. To reduce bias and misinterpretation of results to the best of our abilities, this study employed automated staining together with objective, automated assessment.

We found no association between MSH3 expression and EMAST in CRC in the present study. The findings may have implications beyond CRC as EMAST has been described in tumors other than CRC [7]. A recent study in patients with pancreatic cancer that described EMAST in up to 40% of patients found no inactivation of MSH3 in the tumors [30], which is in line with the current results, although in a different cancer type. Furthermore, *MSH3* frameshift mutations were found to be specific to CRC, but not for endometrial cancer [15], which is also known for harboring high levels of EMAST [31]. These observations, taken together with the results of this study and the general disagreement in the literature, discourage the proposed role of MSH3 loss as a universal biological mechanism underlying EMAST.

The large overlap of MSI and EMAST in CRC found in the current study correlates well with results from the literature [17]. EMAST could represent an exacerbation of MSI rather than a separate occurrence. In the current study we found EMAST to occur in colon cancers and to be associated with MSI features, as described previously [19]. Shared features between MSI and EMAST include prevalence in the proximal part of the colon and, in the case of the present study, an absence or low prevalence in rectal cancers. Two reports [16,32] describe a high prevalence of EMAST among rectal cancer patients, however in one (a rectum-only study) EMAST prevalence was higher in African American than Caucasian patients [32]. In the second report, describing 61% of rectal cancers as EMAST, the cohort was based on patients of Asian ethnicity [16]. In contrast, the present cohort consists predominantly of Caucasian patients. Considering the documented variation of MSI-H across demographic factors such as gender and ethnicity [33], EMAST might well follow a similar pattern, thus explaining discrepancy in rectal distributions. The current results are nonetheless consistent with other studies [16,17,34] finding EMAST prevalently in colon, and more specifically in the proximal part of the colon.

In conclusion, the mechanism leading to genomic instability in tetranucleotides and expressed as EMAST is still largely unknown at present. Indeed, EMAST may, as suggested for conventional MSI, be related to epigenetic mechanisms that occur with aging (e.g. epigenetic loss of DNA repair mechanisms). EMAST may thus represent an epiphenomenon of age and tumorigenesis rather than a specific tumor-driving trait per se. On the other hand, MSI-induced frameshifts could affect MSH3 functionality—and lead to EMAST—in ways that the antibody used in this study could not highlight. Further investigations into the mechanisms of EMAST is warranted.

References

[1] Boland CR and Goel A (2010). Microsatellite instability in colorectal cancer. *Gastroenterology* **138**(6), 2073–2087. e2073.

[2] Soreide K, Janssen EA, Soiland H, Korner H and Baak JP (2006). Microsatellite instability in colorectal cancer. *Br J Surg* **93**(4), 395–406.

[3] Dienstmann R, Mason MJ, Sinicrope FA, Phipps AI, Tejpar S, Nesbakken A, Danielsen SA, Sveen A, Buchanan DD and Clendenning M, et al (2017). Prediction of overall survival in stage II and III colon cancer beyond TNM system: a retrospective, pooled biomarker study. *Ann Oncol* **28**(5), 1023–1031.

[4] Domingo E, Camps C, Kaisaki PJ, Parsons MJ, Mouradov D, Pentony MM, Makino S, Palmieri M, Ward RL and Hawkins NJ, et al (2018). Mutation burden and other molecular markers of prognosis in colorectal cancer treated with curative intent: results from the QUASAR 2 clinical trial and an Australian community-based series. *Lancet Gastroenterol Hepatol* **3**(9), 635–643.

[5] Kang S, Na Y, Joung SY, Lee SI, Oh SC and Min BW (2018). The significance of microsatellite instability in colorectal cancer after controlling for clinicopathological factors. *Medicine (Baltimore)* **97**(9). e0019.
[6] Guinney J, Dienstmann R, Wang X, Reyniès Ad, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G and Angelino P, et al (2015). The consensus molecular subtypes of colorectal cancer. *Nat Med* **21**(11), 1350−1356.
[7] Watson MM, Berg M and Soreide K (2014). Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. *Br J Cancer* **111**(5), 823−827.
[8] Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R and Ranzani GN, et al (1998). A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* **58**(22), 5248−5257.
[9] Suraweera N, Duval A, Reperant M, Vaury C, Furlan D, Leroy K, Seruca R, Iacopetta B and Hamelin R (2002). Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* **123**(6), 1804−1811.
[10] Haugen AC, Goel A, Yamada K, Marra G, Nguyen T-P, Nagasaka T, Kanazawa S, Koike J, Kikuchi Y and Zhong X, et al (2008). Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer Res* **68**(20), 8465−8472.
[11] Campregher C, Schmid G, Ferk F, Knasmüller S, Khare V, Kortüm B, Dammann K, Lang M, Scharl T and Spittler A, et al (2012). MSH3-deficiency initiates EMAST without oncogenic transformation of human colon epithelial cells. *PLoS ONE* **7**(11). e50541.
[12] Tseng-Rogenski SS, Chung H, Wilk MB, Zhang S, Iwaizumi M and Carethers JM (2012). Oxidative stress induces nuclear-to-cytosol shift of hMSH3, a potential mechanism for EMAST in colorectal cancer cells. *PLoS ONE* **7**(11). e50616.
[13] Tseng-Rogenski S, Hamaya Y, Choi DY and Carethers JM (2015). Interleukin 6 alters localization of hMSH3, leading to DNA mismatch repair defects in colorectal cancer cells. *Gastroenterology* **148**(3), 579−589.
[14] The Cancer Genome Atlas Network, Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, Kovar CL, Lewis LR and Morgan MB, et al (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330.
[15] Kim T-M, Laird Peter W and Park Peter J (2013). The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell* **155**(4), 858−868.
[16] Lee SY, Chung H, Devaraj B, Iwaizumi M, Han HS, Hwang DY, Seong MK, Jung BH and Carethers JM (2010). Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias. *Gastroenterology* **139**(5), 1519−1525.
[17] Venderbosch S, van Lent-van Vliet S, de Haan AF, Ligtenberg MJ, Goossens M, Punt CJ, Koopman M and Nagtegaal ID (2015). EMAST is associated with a poor prognosis in microsatellite instable metastatic colorectal cancer. *PLoS One* **10**(4). e0124538.
[18] Soreide K, Watson MM, Lea D, Nordgard O, Soreide JA and Hagland HR (2016). Assessment of clinically related outcomes and biomarker analysis for translational integration in colorectal cancer (ACROBATICC): study protocol for a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastasis. *J Transl Med* **14**(1), 192.
[19] Watson MM, Lea D, Rewcastle E, Hagland HR and Søreide K (2016). Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal cancers with and without high-frequency microsatellite instability: same, same but different? *Cancer Med* **5**(7), 1580−1587.
[20] Buhard O, Suraweera N, Lectard A, Duval A and Hamelin R (2004). Quasimonomorphic mononucleotide repeats for high-level microsatellite instability analysis. *Disease Markers* **20**(4-5), 251−257.
[21] Soreide K (2011). High-fidelity of five quasimonomorphic mononucleotide repeats to high-frequency microsatellite instability distribution in early-stage adenocarcinoma of the colon. *Anticancer Res* **31**(3), 967−971.
[22] Koi M, Tseng-Rogenski SS and Carethers JM (2018). Inflammation-associated microsatellite alterations: mechanisms and significance in the prognosis of patients with colorectal cancer. *World Journal of Gastrointestinal Oncology* **10**(1), 1−14.
[23] Stalhammar G, Robertson S, Wedlund L, Lippert M, Rantalainen M, Bergh J and Hartman J (2018). Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology* **72**(6), 974−989.
[24] Norwegian Health Council (2018). Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft. 2018. http://www.helsedirektoratet.no/retningslinjer. [Accessed 9 May 2019].
[25] Danish Breast Cancer Cooperative Group (2017). Retningslinjer for brystkræft. http://www.dbcg.dk/. [Accessed 9 May 2019].
[26] Swedish Society of Pathology (2018). Kvalitetsdokument för patologi. http://www.svfp.se/foreningar/uploads/L15178/kvast/brostpatologi/KVASTbrostcancer2018.pdf. [Accessed 9 May 2019].
[27] Pantanowitz LR and David L (2019). Imaging and Quantitative Immunohistochemistry. In: Dabbss David J, editor. Diagnostic Immunohistochemistry. 5th ed. Philadelphia, PA, USA: Elsevier; 2019.
[28] Bussolati G and Leonardo E (2008). Technical pitfalls potentially affecting diagnoses in immunohistochemistry. *J Clin Pathol* **61**(11), 1184.
[29] Cartun RWT, Dabbs Clive R and David J (2019). Techniques of immunohistochemistry: Principles, Pitfalls, and Standardization. In: Dabbss David J, editor. Diagnostic Immunohistochemistry. 5th ed. Philadelphia, PA, USA: Elsevier; 2019.
[30] Mori T, Hamaya Y, Uotani T, Yamade M, Iwaizumi M, Furuta T, Miyajima H, Osawa S and Sugimoto K (2018). Prevalence of elevated microsatellite alterations at selected tetranucleotide repeats in pancreatic ductal adenocarcinoma. *PLoS One* **13**(12). e0208557.
[31] Choi YD, Choi J, Kim JH, Lee JS, Lee JH, Choi C, Choi HS, Lee MC, Park CS and Juhng SW, et al (2008). Microsatellite instability at a tetranucleotide repeat in type I endometrial carcinoma. *J Exp Clin Cancer Res* **27**, 88.
[32] Devaraj B, Lee A, Cabrera BL, Miyai K, Luo L, Ramamoorthy S, Keku T, Sandler RS, McGuire KL and Carethers JM (2010). Relationship of EMAST and microsatellite instability among patients with rectal cancer. *J Gastrointest Surg* **14**(10), 1521−1528.
[33] Carethers JM and Jung BH (2015). Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* **149**(5), 1177−1190. e1173.
[34] Lee HS, Park KU, Kim DW, Lhn MH, Kim WH, Seo AN, Chang HE, Nam SK, Lee SY and Oh HK, et al (2016). Elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) and microsatellite instability in patients with colorectal cancer and its clinical features. *Curr Mol Med* **16**(9), 829−839.

IV

Cancer Immunology, Immunotherapy
https://doi.org/10.1007/s00262-020-02573-0

ORIGINAL ARTICLE

Prevalence of PD-L1 expression is associated with EMAST, density of peritumoral T-cells and recurrence-free survival in operable non-metastatic colorectal cancer

Martin M. Watson[1,2] · Dordi Lea[1,2,3] · Einar Gudlaugsson[3] · Ivar Skaland[3] · Hanne R. Hagland[1,4] · Kjetil Søreide[1,2,5]

Received: 19 December 2019 / Accepted: 7 April 2020

Abstract

Introduction Microsatellite instability (MSI) predict response to anti-PD1 immunotherapy in colorectal cancer (CRC). CRCs with MSI have higher infiltration of immune cells related to a better survival. Elevated Microsatellite Alterations at Tetranucleotides (EMAST) is a form of MSI but its association with PD-L1 expression and immune-cell infiltration is not known.

Methods A consecutive, observational cohort of patients undergoing surgery for CRC. EMAST and clinicopathological characteristics were investigated against PD-L1, as well as CD3 and CD8 expression in the invasive margin or tumour centre (Immunoscore). Difference in survival between groups was assessed by log rank test.

Results A total of 149 stage I–III CRCs patients, with a median follow up of 60.1 months. Patients with PD-L1+ tumours (7%) were older (median 79 vs 71 years, $p=0.045$) and had EMAST+ cancers (OR 10.7, 95% CI 2.2–51.4, $p=0.001$). Recurrence-free survival was longer in cancers with PD-L1+ immune cells (HR 0.35, 95% CI 0.16–0.76, $p=0.008$, independent of EMAST) and high Immunoscore (HR 0.10, 95% CI 0.01–0.72, $p=0.022$). Patients expressing PD-L1 in immune cells had longer disease-specific survival (HR 0.28, 95% CI 0.10–0.77, $p=0.014$).

Conclusions Higher Immunoscore (CD3/CD8 cells) and expression of tumour PD-L1 is found in CRCs with EMAST. Lymphocytic infiltrate and peritumoral PD-L1 expression have prognostic value in CRC.

Keywords Colorectal cancer · EMAST · PD-L1 · Immunoscore · Survival · Recurrence

Abbreviations

CRC	Colorectal cancers
EMAST	Elevated Microsatellite Alterations at Selected Tetranucleotides
MSI	Microsatellite instability

Introduction

Colorectal cancers (CRCs) with deficient mismatch repair (MMR) are often hypermutated, have microsatellite instability (MSI), are associated with improved prognosis and is defined to the 'immunogenic' class of consensus molecular subtypes [1]. Notably, MSI is determined by a panel of microsatellite markers, commonly mononucleotides, according to established guidelines [2]. However, an alternative form of MSI is found in tetranucleotide-based microsatellites and labelled Elevated Microsatellite Alterations at Selected Tetranucleotides (EMAST) [3, 4]. Currently, the prognostic value, molecular mechanisms and clinical implications of EMAST are unclear. EMAST was linked in vitro with

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00262-020-02573-0) contains supplementary material, which is available to authorized users.

✉ Kjetil Søreide
ksoreide@mac.com

1 Gastrointestinal Translational Research Unit, Stavanger University Hospital, Stavanger, Norway

2 Department of Clinical Medicine, University of Bergen, Bergen, Norway

3 Department of Pathology, Stavanger University Hospital, Stavanger, Norway

4 Department of Chemistry, Bioscience and Environmental Engineering, University of Stavanger, Stavanger, Norway

5 Department of Gastrointestinal Surgery, Stavanger University Hospital, Stavanger, Norway

Published online: 20 April 2020

downregulation of MSH3, a member of MMR specifically implicated with repair of long indels [4]. This proposed mechanism has not been confirmed across patient series, with a previous study from our group refuting an association between MSH3 and EMAST [5]. Prognostic data on EMAST is also scarce. In a previous study, we found that patients with EMAST+ were older, frailer and less likely to have recurrence from CRC [6].

CRCs with MSI are associated with a higher production of neoantigens and consequent immune system activation [7]. The understanding of host immune system and its relevance for cancer control has evolved across several tumour types yet with varying potential for therapeutic intervention and effect on disease trajectory [7]. In colorectal cancer (CRC), data suggest that type and density of immune cells are related to survival and may be used to improve TNM-staging by incorporating an Immunoscore [8, 9]. Hence, the immune cells infiltrating in the tumour microenvironment have a functional role in CRC, although understanding of associated factors related to this peritumoral activation is poor at present.

Cancer immunosurveillance of the adaptive immune system may be disturbed through various mechanisms [10]. One example is the activation of immune checkpoints such as the receptor-ligand complex PD-1/PD-L1 that dampens the immune response and cause T-cell exhaustion [7, 11, 12]. Data suggesting that PD-1 blockade therapy potentially benefits the MMR/MSI subsets of CRCs and other cancers [13–15], introduced immunotherapy for clinical use [16]. However, selection of patients who may benefit and respond is currently uncertain. Further, scarce evidence exists to date on the association of PD-L1 expression and prognosis and survival, both within and outside the predictive subsets of CRC. Data regarding the relationship between EMAST and PD-L1 expression and the associated T-cell infiltration are lacking.

The aim of the present study was thus to describe the prevalence of PD-L1 expression, Immunoscore, their relationship with MSI/EMAST and their relevance towards clinical outcomes in a well-defined, consecutive series of operable CRCs.

Materials and methods

Study population and design

Patients were consecutively recruited during the 01/2013–05/2015 period at Stavanger University Hospital (SUH), Norway. Norway has a universal health care coverage for all citizens and the university hospital serves a primary catchment region of about 370,000 inhabitants. With no selection or referral bias in the health care system, the study cohort can be considered as population representative and generalizable to similar regions in Northern Europe.

The present study cohort is part of an ongoing prospective project (ACROBATICC) approved by the regional ethics committee (REK Helse Vest: 2012/742) and registered on clinicaltrials.gov (NCT01762813) [17]. All consecutive patients amenable to curative intent surgery, aged ≥ 18 years of age and who could provide written informed consent were eligible for inclusion into ACROBATICC. This observational cohort study of patients presenting with operable stage I–III disease and is reported according to the STROBE [18] and the REMARK [19] guidelines for biomarker studies.

Histopathology

All cancers were staged by an experienced pathologist following guidelines published in the 7th edition of the AJCC staging manual [20]. Proximal tumour location is intended as the region between caecum and transverse colon, while distal is intended as the region between the splenic flexure and sigmoid colon.

EMAST and MSI analysis

Analyses of EMAST and MSI, including primer sequences and PCR conditions, are described previously [21, 22]. Briefly, formalin-fixed paraffin blocks selected by an experienced pathologist were sectioned for DNA extraction. Macrodissection of areas indicated by the pathologist was employed where necessary to enrich for tumour cells. Automated DNA extraction was carried out using AllPrep DNA/RNA FFPE kit (Qiagen, Hilden, Germany) on a QiaCUBE instrument (Qiagen), according to manufacturer's instructions. Nucleic acid concentration and purity were measured on a NanoDrop 2000 (ThermoFischer scientific, Waltham, USA). Two separate multiplex PCR reactions (one for each MSI and EMAST) were set up for tumour and normal DNA in each patient. TypeIT microsatellite (Qiagen) master mix, together with a blending of $5 \times 5'$-fluorescently labelled primer pairs was used for each reaction. The primers for MSI were specific for the quasimonomorphic mononucleotides BAT-26, NR-21, NR-24 and NR-27, while the EMAST marker panel consisted of MYCL1, D8S321, D9S242, D20S82, and D20S85. To define a tumour as MSI-H, at least 2/5 markers needed to be unstable in their respective panels.

Immunohistochemistry

Paraffin sections consecutive to the haematoxylin–eosin (H&E) sections were cut to 2 µm and mounted onto Superfrost Plus slides (Menzel, Braunschweig, Germany). Antigen retrieval and antibody dilution were optimised for each induvial staining. All antibody protocols were optimized

before study onset. Paraffin sections consecutive to the haematoxylin–eosin (H&E) sections were cut to 2 µm and mounted onto Superfrost Plus slides (Menzel, Braunschweig, Germany). Slides were incubated at 60 °C for 1 h and then transferred to a Dako Omnis (Dako, Glostrup, Denmark) instrument. CD3 (Dako, Clone F7.2.38) was used at a dilution of 1:75 and visualised by EnVision FLEX, High pH (Dako Omnis) (GV80011-2). CD8 (Dako, Clone C8/144B) was used at a dilution of 1:50 and visualized by EnVision FLEX, High pH (Dako Omnis) (GV80011-2).with EnVision FLEX+ Mouse LINKER (Dako Omnis) (GV82111-2) signal amplification.

EnVision FLEX Antibody Diluent (Dako, K800621-2) was used as diluent. Pre-treatment time was 20 min at 97 °C using EnVision FLEX Target Retrieval Solution, High pH (50×) (Dako Omnis) (GV800). Both antibodies were incubated for 20 min. Hematoxylin (Dako Omnis) (GC80811-2) was used as counterstain.

PD-L1 IHC 22C3 pharmDx (Dako SK00621-2) was used strictly according to manufacturer's recommendation on a Dako Autostainer Link 48 instrument.

Scoring of PD-L1 expression

PD-L1 expression (Fig. 1) was assessed independently by two experienced pathologists (DL and EG), blinded to patients' other characteristics and each other results, on whole sections. Membranous staining was regarded as positive, and staining intensity was not evaluated. PD-L1 in tumour cells was scored as positive or negative using ≥ 5% positive as cut-off, based on previous studies [14, 23].

For PD-L1 expression on peritumoral immune cells, the percentage of positive cells were evaluated in the visually most positive area of 1 mm^2 in the invasive margins of the tumour on the scanned slides (same area for both pathologists). In cases with > 10% discordance between the pathologists, the slides were reviewed together, until consensus was reached. For expression in less than 10% of the immune cells, discordance of < 5% was accepted.

The cut-off for positive or negative classifications of patients based on PD-L1 expression in peritumoral immune cells, was determined experimentally.

Receiver operator characteristics (ROC) curve analysis was used to determine cut-offs for PD-L1 expression

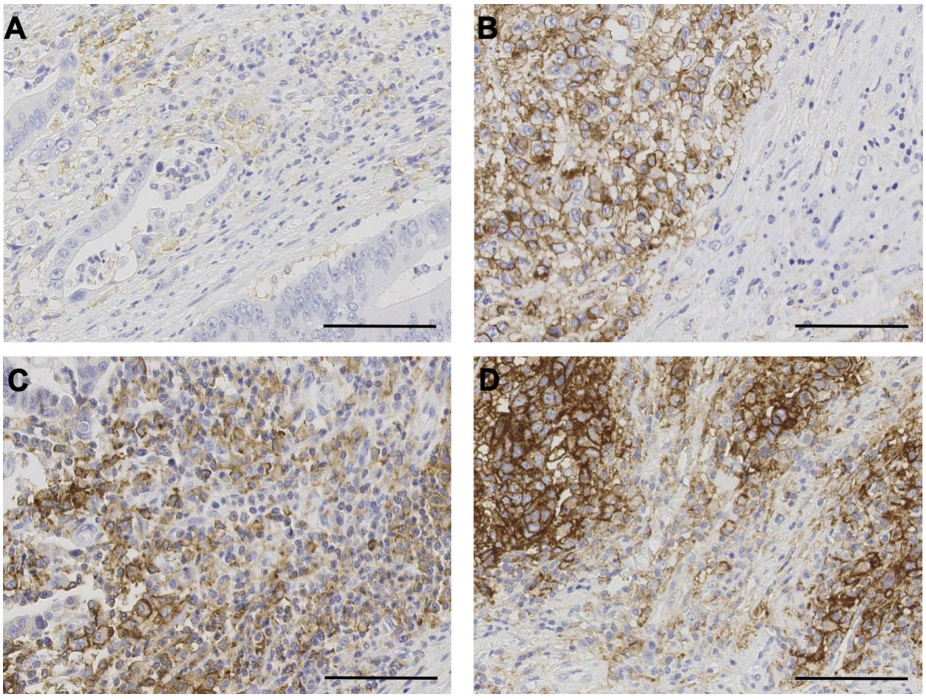


Fig. 1 Immunohistochemistry of PD-L1. 20X magnification view of **a** immune PD-L1−/tumour PD-L1−. **b** Immune PD-L1−/tumour PD-L1+. **c** Immune PD-L1+/tumour PD-L1−. **d** Immune PD-L1+ / tumour PD-L1+. Scale bar represents 100 µm

in immune cells, with disease-specific death and disease recurrence as the endpoints. The optimal cut-off point in both ROC curve analysis corresponded to the 25th percentile and was therefore chosen as the discriminant cut-off to dichotomize expression of immune cells PD-L1 into positive/negative.

Immune scoring of CD3 and CD8 markers

All the sections were scanned at 40 × magnification using Leica SCN400 slide scanner (Leica Microsystems, Wetzlar, Germany) and uploaded onto the image analysis software Visiopharm® (Hoersholm, Denmark). Tumour centre and invasive margin areas were marked manually on whole slide images and the same areas were used for the CD3 and CD8 stained sections. Using Bayesian optimisations [24], an algorithm was developed to identify and label CD3+ and CD8+ T-cells in both regions.

Relative quantification of positive cells was obtained by dividing the Visiopharm-measured area of positive label by the estimation of mean area of a lymphocyte (60 μm^2), thereby approximating the number of CD3+ and CD8+ T-cells per square millimetres (cells/mm^2). All the cases were inspected, and unspecific staining and artefacts were manually removed from the analyses where appropriate. For individual CD3/CD8 analysis, patients were assigned either a "low" or "high" score for each individual staining (CD3 and CD8), in each tumour location (tumour centre and invasive margin and IM), using the 75th percentile as a threshold. This created four categories (CD3 and CD8, in tumour centre and invasive margin as either high or low).

The Immunoscore was calculated as described elsewhere [8]. Briefly, the densities (in cells/mm^2) of CD3+ and CD8+ cells in both tumour centre and invasive margin were first converted into percentiles, and then the mean value of the four percentiles calculated. An Immunoscore of "Low", "Intermediate" or "High" was then assigned to each patient according to their mean percentile scores, with cut-offs as 0–25%, 25–70%, and 70–100% respectively, as described in [8].

Collection of clinical data and follow up

Clinical measurements as well as follow up data (cause and date of death, date of recurrence) were retrieved from the electronic patient records. Patients' surveillance after surgery was according to the national guidelines as an interval-based serum CEA (quarterly) and imaging (e.g. a biannual CT-chest and US liver for the first 3 years, then annually) for up to 5 years after surgery. Colon cancers were usually followed up by general practitioners while rectal cancers were seen by gastrointestinal surgeons in the hospital outpatient clinics. Any suspected recurrence or deviation on imaging were worked up in-hospital and consulted in multidisciplinary team meetings, where applicable. The patients' electronic health records were queried for any documented events, and follow-up for this study was completed as of 24th September 2019.

Definition of survival endpoints

Recurrence-free survival (RFS) was defined as time from primary surgery until first clinical evidence (histologically confirmed or image-based) of recurrent disease. Disease-specific survival (DSS) was defined as time from primary surgery and death imputable to CRC. Survival was assessed for overall survival (OS) defined as time from primary surgery to death of any cause.

Statistical analyses

All statistical tests were done using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corporation, Armonk, NY, USA). Associations between categorical variables were tested with Chi-square (or Fischer's exact test, where appropriate) method and reported with odds ratios and 95% confidence interval (95% CI). Spearman's rho or Pearson tests were used for correlations between continuous/ordinal variables, where appropriate. Mann–Whitney *U* test was used to compare differences in continuous or ordinal variables between groups. Inter-coder reliability score for PD-L1 evaluation was estimated using the KALPHA extension for SPSS and expressed as Krippendorff's alpha (α).

The Kaplan–Meier method with log rank comparison of factors was used to investigate survival curves differences between groups and are given as (months difference [95% CI]). Univariable proportional hazards are given in hazard ratios (HR) with 95% CI. All tests were two-tailed and a p value < 0.050 considered as statistically significant.

Results

The study cohort included 149 stage I–III CRC patients who underwent surgery with curative intent (Fig. 2). Patients' descriptive parameters are included in Table 1.

PD-L1 expression and EMAST

Of the 11 patients classified as PD-L1+ in tumour cells, nine were diagnosed with right-side CRC (82%, no rectum, $p = 0.111$) and were EMAST+ (82%; Table 2). Inter-coder reliability score for PD-L1 expression in tumour cells was high (Krippendorff's $\alpha = 0.93$; 95% CI 0.83–0.99). A weak correlation was also seen between expression of tumoral

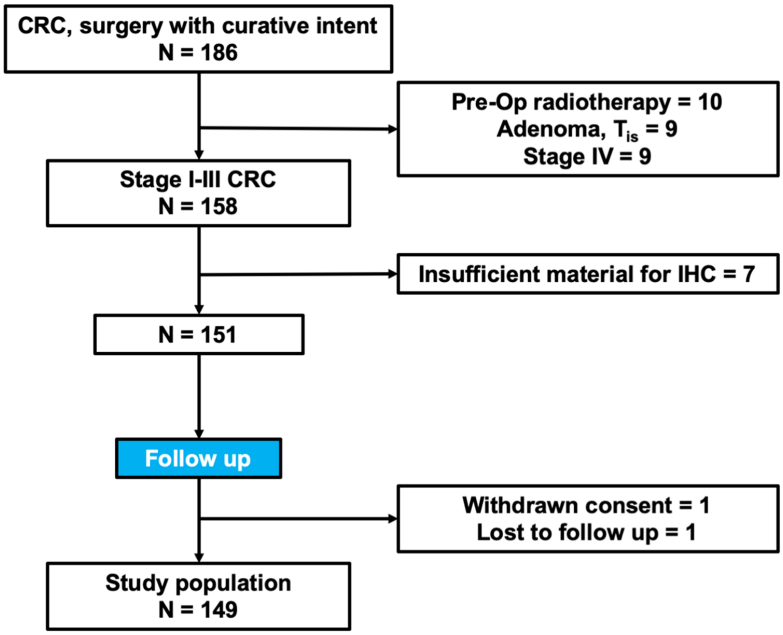


Fig. 2 Flowchart of inclusion/exclusion criteria. CRC denotes colorectal cancer; IHC denotes immunohistochemistry

Table 1 Variables associated with EMAST status

	Total *n*, (%)	EMAST− *n* = 99 (66)	EMAST+ *n* = 50 (34)	*p*
Age				**< 0.001**
Median (range)	72 (37–92)	70 (37–91)	77.5 (50–92)	
≤ 72	75 (50)	60 (80)	15 (20)	
> 72	74 (50)	39 (53)	35 (47)	
Sex				**< 0.001**
Male	65 (44)	54 (83)	11 (17)	
Female	84 (56)	45 (54)	39 (46)	
Localisation				**< 0.001**
Colon	124 (83)	74 (60)	50 (40)	
Rectum	25 (17)	25 (100)	0 (0)	
Within colon				**< 0.001**
Right	70 (57.5)	25 (36)	45 (64)	
Left	54 (43.5)	49 (90)	5 (10)	
Grade*				**< 0.001**
High	39 (26)	15 (38)	24 (62)	
Low	109 (74)	83 (76)	26 (24)	
Stage				0.234
I	51 (34)	30 (59)	21 (41)	
II	50 (34)	33 (66)	17 (34)	
III	48 (32)	36 (75)	12 (25)	
MSI				**< 0.001**
MSS	105 (70.5)	97 (92)	8 (8)	
MSI-H	44 (29.5)	2 (5)	42 (95)	

N = 149

Bold values indicate statistical significance ($P < 0.050$)

*One missing

Table 2 Associations with immune markers and EMAST status

N = 149	Total *n*, (%)	EMAST− *n* = 99 (66)	EMAST+ *n* = 50 (34)	OR (95% CI)	*p*
PD-L1 in tumour cells				10.7 (2.2–51.4)	**0.001**
Low	138 (93)	97 (98)	41 (82)		
High	11 (7)	2 (2)	9 (18)		
PD-L1 in immune cells				1.0 (0.5–2.2)	0.973
Low	39 (26)	26 (26)	13 (26)		
High	110 (74)	73 (74)	37 (74)		
Immune cells in tumour centre					
CD3+				2.37 (1.1–5.1)	**0.025**
Low	112 (75)	80 (81)	32 (64)		
High	37 (25)	19 (19)	18 (36)		
CD8+				2.4 (1.1–5.1)	**0.025**
Low	112 (75)	80 (81)	32 (64)		
High	37 (25)	19 (19)	18 (36)		
Immune cells in invasive margin					
CD3+				3.22 (1.5–7.0)	**0.002**
Low	112 (75)	82 (83)	30 (60)		
High	37 (25)	17 (17)	20 (40)		
CD8+				2.4 (1.1–5.1)	**0.025**
Low	112 (75)	80 (81)	32 (64)		
High	37 (25)	19 (19)	18 (36)		
Immunoscore				n.c.	**0.020**
Low	31 (21)	24 (24)	7 (14)		
Interm	79 (53)	56 (57)	23 (46)		
High	39 (26)	19 (19)	20 (40)		

Bold values indicate statistical significance ($P < 0.050$)

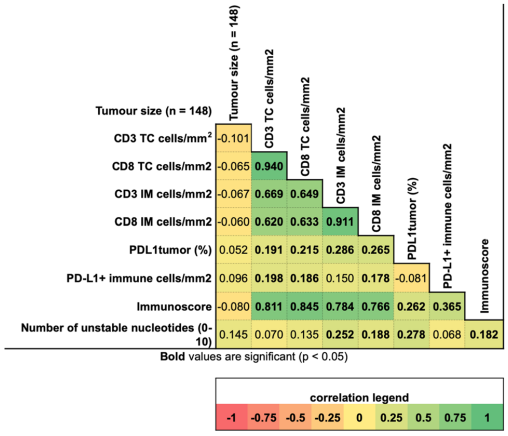


Fig. 3 Correlation matrix of immune-related variables. For pure ordinal variables (marked by the notation "cells/mm2"), Pearson correlation coefficient is shown. For all other variable Spearman R-O tests were used. Bold correlation coefficients are significant ($p < 0.05$)

PD-L1 and total number of unstable markers analysed for both EMAST and MSI (Fig. 3; $p = 0.001$). A higher number of markers from the two panels combined were indeed unstable in PD-L1+ tumours (median 9/10 vs 1/10 markers, $p = 0.001$), when dichotomised accordingly (Suppl. Table 1). Tumour PD-L1+ patients were significantly older (79 vs 71 years, $p = 0.045$) and had lower preoperative levels of serum albumin (33.6 vs 38.1 g/L, $p = 0.011$) (Suppl. Table 1). All PD-L1+ tumours (11/11, 100%) were in the colon, while none of the 25 rectum tumours scored positive ($p = 0.212$).

In peritumoral infiltrating immune cells, the rate of PD-L1 expression was higher than in tumour cells (Fig. 1). No statistically significant association was found between expression of PD-L1 in immune cells and patients' age, EMAST status or number of unstable markers (Suppl. Table 1). Again, a significant but small correlation was found between % PD-L1 and CD3/CD8 in immune cells, albeit not in the case of CD3 in the invasive margin. The two ROC analyses for determination of ideal cut-off value of % PD-L1 positive immune cells had AUC = 0.698, $p = 0.012$ with disease-specific death and AUC = 0.648, $p = 0.018$ with disease recurrence as endpoint (data not shown). The 25th percentile cut-off showed no difference in the distribution among colon and rectum cancers, of which 73% and 76%

showed PD-L1-positive immune cells, respectively. Intercoder reliability score for PD-L1 expression in immune cells was relatively high (Krippendorff's $\alpha = 0.81$; 95% CI 0.72–0.88).

Immune cell types, Immunoscore and EMAST status

Higher density of CD3+ and CD8+ cells in tumour centre and invasive margins were found in EMAST-positive patients (Table 2).

Immunoscore was distributed into low ($n = 31$, 21%), intermediate (79, 53%) and high ($n = 39$, 26%) categories, respectively. EMAST-positive patients were proportionally more represented in the higher Immunoscore subclasses (Table 2). As expected, Immunoscore correlated strongly with each individual CD3+ and CD8+ tally. A stronger relationship between Immunoscore and % of PD-L1+ in immune (Spearman 0.365, $p < 0.001$) rather than in tumour cells (0.262, $p = 0.001$) was found (Fig. 3).

Tumours with PD-L1+ tumours had significantly higher counts of CD3 and CD8 in the invasive margin, as well as CD8, but not CD3 in the tumour centre (Suppl. Table 1).

Patients with PD-L1+ immune cells had significantly higher counts of CD3 and CD8 in both the invasive margin and tumour centre (Suppl. Table 1). Both immune PD-L1+ ($p < 0.001$) and tumour PD-L1+ ($p = 0.037$) patients were significantly associated with a high Immunoscore.

Risk of recurrence and recurrence-free survival

During the follow up period, a total of 26 (17.4%) patients experienced recurrent disease. Eight recurrences (31%) were in the liver, eight in the lungs, 7 (27%) were local recurrences, and one (4%) each for bone, peritoneum and brain.

Generally, higher numbers of infiltrating lymphocytes correlated with lower risk of disease recurrences (Suppl. Figure 1).

A significant difference was found between the categories of the Immunoscore for RFS (Fig. 4; Table 3). No

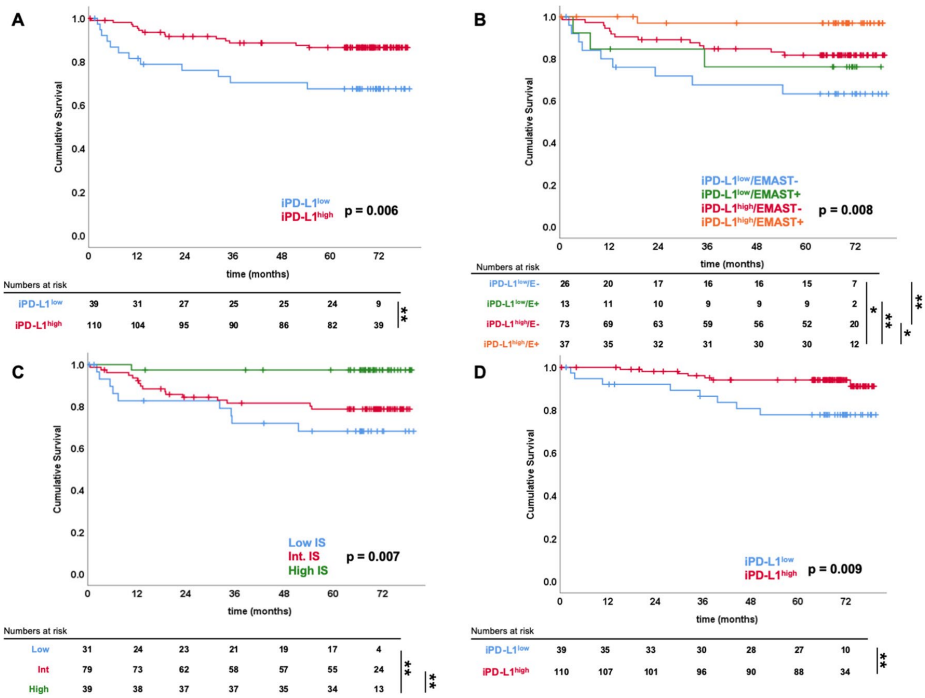


Fig. 4 Survival analyses comparing prognostic groups. Kaplan-Meier analysis of **a** recurrence-free survival (RFS) of immune PD-L1+/− groups. **b** RFS of immune PD-L1+/− groups, stratified for EMAST status. **c** RFS of Immunoscore groups (low-intermediate-high). **d** Disease-specific survival (DSS) of immune PD-L1+/− groups

Table 3 Univariate analyses for survival

Term	HR	95% CI	*p*
Recurrence-free survival (26/149)*			
PD-L1 in immune cells	0.35	0.16–0.76	**0.008**
PD-L1 in tumour cells	0.53	0.07–3.87	0.527
Immunoscore (int. + low vs high)	0.10	0.01–0.72	**0.022**
pN (N0 vs N+)	6.94	2.91–16.52	**< 0.001**
EMAST	0.35	0.12–1.02	0.054
Disease-specific survival (15/149)*			
PD-L1 in immune cells	0.28	0.10–0.77	**0.014**
PD-L1 in tumour cells	0.98	0.13–7.44	0.982
Immunoscore (int. + low vs high)	0.03	0.00–3.37	0.145
pN (N0 vs N+)	229.76	1.44–36788.19	**0.036**
EMAST	0.51	0.14–1.79	0.289
Overall survival (35/149)*			
PD-L1 in immune cells	0.61	0.30–1.23	0.165
PD-L1 in tumour cells	1.78	0.63–5.05	0.277
Immunoscore (int. + low vs high)	0.42	0.16–1.08	0.073
pN (N0 vs N+)	3.17	1.62–6.19	**0.001**
EMAST	1.04	0.52–2.09	0.915

Bold values indicate statistical significance ($P < 0.050$)

*Numbers in parentheses are events/total number of cases

significant association between tumour PD-L1 and rate ($p = 0.690$) or time to recurrence ($p = 0.520$) were recorded. Of the patients with negative immune-PD-L1 expression, 12 (31%) presented with recurrent disease patients against 14 (13%) of those with immune-PD-L1+. Patients with PD-L1+ immune cells had longer estimated RFS (72 [68–75] vs 59 [49–69] months, $p = 0.006$) than immune-PD-L1− patients, independently of EMAST status ($p = 0.041$ and 0.021 in EMAST− and EMAST+ cases, respectively) (Fig. 4).

Overall and disease-specific survival

At the time of final follow-up, a total of 35 (23%) patients had died. Of those, 15 (43%) were CRC-related deaths. Median follow up length was 68.8 months (range 0.4–79.6) from primary surgery to death or right-censoring.

Only nodal status (pN0 vs pN+ or stage I–II vs stage III) was associated with OS in univariate analyses (Table 3). When stratified according to the three Immunoscore levels, a high Immunoscore had significantly longer overall survival than low (74 vs 60 [50–70] months, $p = 0.008$), but not intermediate (74 [69–79] vs 68 [63–72] months, $p = 0.192$). No difference was noted in survival time when patients were divided according to tumour PD-L1 expression, whilst patients with a higher PD-L1 proportion in immune cells had longer DSS (log rank $p = 0.009$; Fig. 4; Table 3). When stratified for EMAST status, patients with PD-L1+ in immune cells had better DSS in the EMAST-negative group (log rank $p = 0.033$) but not in the EMAST+ group (log rank $p = 0.107$).

Discussion

In the current study, CRC having EMAST correlated with a higher count of intra- and peritumoral CD3+ and CD8+ T-cells and a higher Immunoscore compared to CRC cancers with no EMAST. Also, PD-L1 expression occurred both in immune cells and in tumour cells in CRCs, specifically those with EMAST and MSI.

While the patterns of expression in tissue does not directly translate into functional ability, there are several observations we would point out as being of interest.

First, expression of PD-L1 showed a dual role according to its localisation in this study. Tumour cell-confined PD-L1 correlated with EMAST and generally increasing degree of MSI, while immune cells PD-L1 did not. EMAST independently correlated with a generally higher immunogenicity, with higher levels of CD3+, CD8+ and PD-L1+ in tumour cells. This is generally in accordance with the relationship between MSI, high mutational burden, generation of tumour neoantigens, and activation of the immune system [7, 25]. Notably, one study previously reported an association with EMAST and CD8+ but not CD4+ T-cells infiltration in tumour [26]. In a previous report, a link between EMAST and older age and a frailer phenotype in patients with EMAST positive cancers was found [6]. These observations, pertinent to a cumulative increase in genetic abnormalities (e.g. EMAST, MSI, mutation burden) during physiological and cellular senescence, may weigh in on the picture of a neoantigen-rich tumour microenvironment. Of note, while the distribution of immune cells expressing PD-L1 seemed comparable between colon and rectum cancers, PD-L1-expressing tumours were exclusively found in colon cancers. This may further confirm a relationship between PD-L1 and instability at microsatellites, as both MSI and EMAST are more prevalent in the colon.

A direct relationship between high Immunoscore and high PD-L1 expression in both tumour and immune cells was also shown in the present study. Tumours with low counts for CD3+ and CD8+ cells are associated with less overall (tumour/immune) PD-L1 expression. PD-L1 was here found to be rarely expressed in tumour cells, and strictly connected to EMAST status, while more diffuse in infiltrating immune cells. This is concordant with recent reports placing tumour PD-L1 rates generally under 15–20% of CRCs, and immune PD-L1 consistently higher [27–32]. These observations may suggest that induction of PD-L1 is regulated

by different pathways in immune and tumour cells. On one side, EMAST (as MSI) tumours, due to their higher load of tumour neoantigens are possibly subject to a more vigorous cytotoxic immune response, and endogenously expressing PD-L1 to counteract it. In non-EMAST tumours with high Immunoscore, otherwise, modulation of immune response is achieved by expression of PD-L1 on immune cells, in a mechanism also referred to as adaptive immune resistance [33, 34]. Finally, tumours having both low Immunoscore and PD-L1 expression on immune cells lack an immune reaction in the tumour microenvironment and present with higher rate of recurrences, sooner.

Prognostically, only expression of PD-L1 in the peritumoral cells proved discriminant in both rate of- and time to recurrences, as well as for disease-specific survival. In terms of RFS, the association was independent of EMAST status and comparable to that of Immunoscore, suggesting that immune expression of PD-L1 contributes to the protective effect of tumour immunosurveillance. Tumour PD-L1 was not associated with any of the survival endpoints examined. In contrast to Immunoscore, there is scarce data on the prognostic significance of PD-L1 expression in CRC. The relationship between tumour-expressed PD-L1 and tumour-infiltrating lymphocytes is being investigated in multiple cancers [35, 36]. The focus is however usually on tumour-expressed PD-L1, because of its predictive value for immunotherapy, while there is discordance on its prognostic role [32, 37, 38]. PD-L1 positivity on peritumoral immune cells, on the other hand, is generally a sign of an active immune response and thus associated with improved survival [31, 32, 39, 40].

A limitation of the present study is the limited size of the cohort, with only 11 patients scoring positive for tumour PD-L1, therefore limiting the statistical power. However, the idea of modern personalized medicine is to identify particular subgroups with potential for refined therapy. Prevalence of tumour PD-L1 in MSI CRCs and the low (15–20%) incidence of the subgroup, warrant expansion of the cohort in order to investigate the findings in larger cohorts and refined sub populations. A further limit is the cut-off determination for immune PD-L1 expression. Derived from ROC analysis for recurrence and disease-specific death, which are time-dependent variable, this method may only apply to the present cohort, in the elapsed follow up time. Indeed, a range of methods of PD-L1 scoring and subgrouping are described in the literature, including variation in antibodies used, without a generalised consensus. In the present study, the cut-off value used for PD-L1 expression in tumour cells (5%) was based on previous studies [14, 23, 29], including original anti-PD-1 immunotherapy clinical trials.

The current study correlates PD-L1 expression in tumour cells with EMAST. Moreover, the findings add to the mounting data on PD-L1 expression in peritumoral immune infiltrate and Immunoscore as prognostic factors in CRC. Finally, this study supports the differentiation between tumour- and immune cell expression of PD-L1 as representative of two distinct mechanisms of immune resistance.

Acknowledgements Open Access funding provided by University of Bergen.

Author contributions MMW, DL, HRH and KS contributed to the study conception and design. Material preparation, data collection and analysis were performed by MMW, DL, EG and KS. Technical supervision was provided by IS and KS. The first draft of the manuscript was written by MMW and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Funded in part by the Folke Hermansen Fond, Mjaaland foundation, Universitetsfondet from the University of Stavanger, and intramural grants from Stavanger University Hospital.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Research involving human participants and/or animals This study received pre-emptive approval by the regional ethics committee (REK Helse Vest: 2012/742). All included patients were > 18 years old and signed an informed consent form.

References

1. Guinney J, Dienstmann R, Wang X et al (2015) The consensus molecular subtypes of colorectal cancer. Nat Med 21:1350–1356
2. Umar A, Boland CR, Terdiman JP et al (2004) Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96:261–268
3. Watson MM, Berg M, Soreide K (2014) Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. Br J Cancer 111:823–827. https://doi.org/10.1038/bjc.2014.167
4. Koi M, Tseng-Rogenski SS, Carethers JM (2018) Inflammation-associated microsatellite alterations: mechanisms and significance in the prognosis of patients with colorectal cancer. World J Gastrointest Oncol 10:1–14. https://doi.org/10.4251/wjgo.v10.i1.1

5. Watson MM, Lea D, Hagland HR, Søreide K (2019) Elevated Microsatellite Alterations at Selected Tetranucleotides (EMAST) is not attributed to MSH3 loss in stage I–III colon cancer: an automated, digitalized assessment by immunohistochemistry of whole slides and hot spots. Transl Oncol 12:1583–1588. https://doi.org/10.1016/j.tranon.2019.08.009
6. Watson MM, Kanani A, Lea D, Khajavi RB, Søreide JA, Kørner H, Hagland HR, Søreide K (2019) Elevated Microsatellite Alterations at Selected Tetranucleotides (EMAST) in colorectal cancer is associated with an elderly, frail phenotype and improved recurrence-free survival. Ann Surg Oncol. https://doi.org/10.1245/s10434-019-08048-6
7. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G (2017) The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol 14:717. https://doi.org/10.1038/nrclinonc.2017.101
8. Pagès F, Mlecnik B, Marliot F et al (2018) International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 391:2128–2139. https://doi.org/10.1016/S0140-6736(18)30789-X
9. Galon J, Mlecnik B, Bindea G et al (2014) Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 232:199–209. https://doi.org/10.1002/path.4287
10. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646–674. https://doi.org/10.1016/j.cell.2011.02.013
11. Chen DS, Mellman I (2017) Elements of cancer immunity and the cancer–immune set point. Nature 541:321. https://doi.org/10.1038/nature21349
12. Oliveira AF, Bretes L, Furtado I (2019) Review of PD-1/PD-L1 inhibitors in metastatic dMMR/MSI-H colorectal cancer. Front Oncol 9:396. https://doi.org/10.3389/fonc.2019.00396
13. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA (2019) Immunotherapy in colorectal cancer: rationale, challenges and potential. Nat Rev Gastroenterol Hepatol 16:361–375. https://doi.org/10.1038/s41575-019-0126-x
14. Le DT, Durham JN, Wang H et al (2015) PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 372:2509–2520. https://doi.org/10.1056/NEJMoa1500596
15. Le DT, Durham JN, Smith KN et al (2017) Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 357:409. https://doi.org/10.1126/science.aan6733
16. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB (2019) Colorectal cancer. Lancet 394:1467–1480. https://doi.org/10.1016/S0140-6736(19)32319-0
17. Soreide K, Watson MM, Lea D, Nordgard O, Soreide JA, Hagland HR (2016) Assessment of clinically related outcomes and biomarker analysis for translational integration in colorectal cancer (ACROBATICC): study protocol for a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastasis. J Transl Med 14:192. https://doi.org/10.1186/s12967-016-0951-4
18. Ev E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ (Clin Res Ed) 335:806–808. https://doi.org/10.1136/bmj.39335.541782.AD
19. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, Statistics Subcommittee of the NCIEWGoCD (2005) Reporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer 93:387–391. https://doi.org/10.1038/sj.bjc.6602678
20. Byrd DR, Edge SB, Compton CC, Fritz AG, Greene FL, Trotti A (2010) AJCC cancer staging manual, 7th edn. Springer-Verlag, New York
21. Soreide K (2011) High-fidelity of five quasimonomorphic mononucleotide repeats to high-frequency microsatellite instability distribution in early-stage adenocarcinoma of the colon. Anticancer Res 31:967–971
22. Watson MM, Lea D, Rewcastle E, Hagland HR, Søreide K (2016) Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal cancers with and without high-frequency microsatellite instability: same, same but different? Cancer Med 5:1580–1587. https://doi.org/10.1002/cam4.709
23. Topalian SL, Hodi FS, Brahmer JR et al (2012) Safety, activity, and immune correlates of anti–PD-1 antibody in cancer. N Engl J Med 366:2443–2454. https://doi.org/10.1056/NEJMoa1200690
24. Vantaram SR, Saber E (2012) Survey of contemporary trends in color image segmentation. J Electron Imaging 21:1–28
25. Giannakis M, Mu Xinmeng J, Shukla Sachet A et al (2016) Genomic correlates of immune-cell infiltrates in colorectal carcinoma. Cell Rep 15:857–865. https://doi.org/10.1016/j.celrep.2016.03.075
26. Lee SY, Miyai K, Han HS et al (2012) Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. Dig Dis Sci 57:72–78. https://doi.org/10.1007/s10620-011-1825-5
27. Kim JH, Park HE, Cho N-Y, Lee HS, Kang GH (2016) Characterisation of PD-L1-positive subsets of microsatellite-unstable colorectal cancers. Br J Cancer 115:490–496. https://doi.org/10.1038/bjc.2016.211
28. Lee LH, Cavalcanti MS, Segal NH et al (2016) Patterns and prognostic relevance of PD-1 and PD-L1 expression in colorectal carcinoma. Mod Pathol 29:1433–1442. https://doi.org/10.1038/modpathol.2016.139
29. Rosenbaum MW, Bledsoe JR, Morales-Oyarvide V, Huynh TG, Mino-Kenudson M (2016) PD-L1 expression in colorectal cancer is associated with microsatellite instability, BRAF mutation, medullary morphology and cytotoxic tumor-infiltrating lymphocytes. Mod Pathol 29:1104–1112. https://doi.org/10.1038/modpathol.2016.95
30. Inaguma S, Lasota J, Wang Z, Felisiak-Golabek A, Ikeda H, Miettinen M (2017) Clinicopathologic profile, immunophenotype, and genotype of CD274 (PD-L1)-positive colorectal carcinomas. Mod Pathol 30:278–285. https://doi.org/10.1038/modpathol.2016.185
31. Lee KS, Kwak Y, Ahn S et al (2017) Prognostic implication of CD274 (PD-L1) protein expression in tumor-infiltrating immune cells for microsatellite unstable and stable colorectal cancer. Cancer Immunol Immunother 66:927–939. https://doi.org/10.1007/s00262-017-1999-6
32. Berntsson J, Eberhard J, Nodin B, Leandersson K, Larsson AH, Jirström K (2018) Expression of programmed cell death protein 1 (PD-1) and its ligand PD-L1 in colorectal cancer: relationship with sidedness and prognosis. Oncoimmunology 7:e1465165. https://doi.org/10.1080/2162402X.2018.1465165
33. Pardoll DM (2012) The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 12:252–264. https://doi.org/10.1038/nrc3239
34. Taube JM, Galon J, Sholl LM et al (2018) Implications of the tumor immune microenvironment for staging and therapeutics. Mod Pathol 31:214–234. https://doi.org/10.1038/modpathol.2017.156
35. Hamada T, Soong TR, Masugi Y et al (2018) TIME (Tumor Immunity in the MicroEnvironment) classification based on tumor CD274 (PD-L1) expression status and tumor-infiltrating lymphocytes in colorectal carcinomas. Oncoimmunology 7:e1442999. https://doi.org/10.1080/2162402X.2018.1442999
36. Teng MWL, Ngiow SF, Ribas A, Smyth MJ (2015) Classifying cancers based on T-cell infiltration and PD-L1. Cancer Res 75:2139–2145. https://doi.org/10.1158/0008-5472.CAN-15-0255

37. Eriksen AC, Sørensen FB, Lindebjerg J, Hager H, dePont Christensen R, Kjær-Frifeldt S, Hansen TF (2019) Programmed Death Ligand-1 expression in stage II colon cancer-experiences from a nationwide populationbased cohort. BMC Cancer 19:142. https://doi.org/10.1186/s12885-019-5345-6
38. Yang L, Xue R, Pan C (2019) Prognostic and clinicopathological value of PD-L1 in colorectal cancer: a systematic review and meta-analysis. Onco Targets Ther 12:3671–3682. https://doi.org/10.2147/OTT.S190168
39. Kong P, Wang J, Song Z et al (2019) Circulating lymphocytes, PD-L1 expression on tumor-infiltrating lymphocytes, and survival of colorectal cancer patients with different mismatch repair gene status. J Cancer 10:1745–1754. https://doi.org/10.7150/jca.25187
40. Yomoda T, Sudo T, Kawahara A et al (2019) The immunoscore is a superior prognostic tool in atages II and III colorectal cancer and is significantly correlated with programmed death-ligand 1 (PD-L1) expression on tumor-infiltrating mononuclear cells. Ann Surg Oncol 26:415–424. https://doi.org/10.1245/s10434-018-07110-z

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	5, 8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	-
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5, 8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-9
Bias	9	Describe any efforts to address potential sources of bias	6-7
Study size	10	Explain how the study size was arrived at	Fig.2
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6-9
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	6-9
		(*c*) Explain how missing data were addressed	-
		(*d*) If applicable, explain how loss to follow-up was addressed	Fig.1
		(*e*) Describe any sensitivity analyses	7
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	10, Fig.2
		(b) Give reasons for non-participation at each stage	-
		(c) Consider use of a flow diagram	Fig.2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5, Tab.1
		(b) Indicate number of participants with missing data for each variable of interest	Tab.1
		(c) Summarise follow-up time (eg, average and total amount)	11
Outcome data	15*	Report numbers of outcome events or summary measures over time	11, Tab.3

Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	11-12 Tab.3 6-8 -
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11, 7
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	13-14
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	1

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

The REMARK checklist

	Item to be reported	Page no.
INTRODUCTION		
1	State the marker examined, the study objectives, and any pre-specified hypotheses.	3-4
MATERIALS AND METHODS		
Patients		
2	Describe the characteristics (e.g., disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.	5
3	Describe treatments received and how chosen (e.g., randomized or rule-based).	-
Specimen characteristics		
4	Describe type of biological material used (including control samples) and methods of preservation and storage.	5
Assay methods		
5	Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.	5-9
Study design		
6	State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g., by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.	4, 8
7	Precisely define all clinical endpoints examined.	8
8	List all candidate variables initially examined or considered for inclusion in models.	8
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.	-
Statistical analysis methods		
10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.	8-9
11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.	6-9
RESULTS		
Data		
12	Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.	Fig. 2
13	Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.	Tab.1
Analysis and presentation		
14	Show the relation of the marker to standard prognostic variables.	11-12
15	Present univariable analyses showing the relation between the marker and outcome, with the estimated effect (e.g., hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan-Meier plot is recommended.	Tab.3 Fig.3
16	For key multivariable analyses, report estimated effects (e.g., hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.	-
17	Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.	11
18	If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.	7
DISCUSSION		
19	Interpret the results in the context of the pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.	13-14
20	Discuss implications for future research and clinical value.	15

Source: McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM: Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* 2005; 97: 1180-1184.

11. APPENDIX

1. **Watson, M**., Berg, M. & Søreide, K. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. Br J Cancer 111, 823–827 (2014) doi:10.1038/bjc.2014.167

2. Søreide, K., **Watson, M.M**., Lea, D. et al. Assessment of clinically related outcomes and biomarker analysis for translational integration in colorectal cancer (ACROBATICC): study protocol for a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastasis. J Transl Med 14, 192 (2016) doi:10.1186/s12967-016-0951-4

MINI REVIEW

BJC

British Journal of Cancer (2014) 111, 823–827 | doi: 10.1038/bjc.2014.167

Keywords: Tetranucleotide repeats; microsatellite repeats; prevalence; microsatellite instability; DNA mismatch repair; solid tumours; TP53; MSH3; T cells; cancer early detection; prognosis

Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer

M M C Watson[1,2], M Berg[1,2,3] and K Søreide*[,1,2,4]

[1]Department of Gastrointestinal Surgery, Stavanger University Hospital, Stavanger 4068, Norway; [2]Gastrointestinal Surgical Research Unit, Molecular Lab, Hillevåg, Stavanger University Hospital, Stavanger 4068, Norway; [3]Centre of Organelle Research, University of Stavanger, Stavanger 4068, Norway and [4]Department of Clinical Medicine, University of Bergen, Bergen 5020, Norway

Elevated microsatellite alterations at selected tetranucleotides (EMAST), a variation of microsatellite instability (MSI), has been reported in a variety of malignancies (e.g., neoplasias of the lung, head and neck, colorectal region, skin, urinary tract and reproductive organs). EMAST is more prominent at organ sites with potential external exposure to carcinogens (e.g., head, neck, lung, urinary bladder and colon), although the specific molecular mechanisms leading to EMAST remain elusive. Because it is often associated with advanced stages of malignancy, EMAST may be a consequence of rapid cell proliferation and increased mutagenesis. Moreover, defects in DNA mismatch repair enzyme complexes, TP53 mutation status and peritumoural inflammation involving T cells have been described in EMAST tumours. At various tumour sites, EMAST and high-frequency MSI share no clinicopathological features or molecular mechanisms, suggesting their existence as separate entities. Thus EMAST should be explored, because its presence in human cells may reflect both increased risk and the potential for early detection. In particular, the potential use of EMAST in prognosis and prediction may yield novel types of therapeutic intervention, particularly those involving the immune system. This review will summarise the current information concerning EMAST in cancer to highlight the knowledge gaps that require further research.

Microsatellites are repeating units of one to six base pairs, which are ubiquitous, abundant and repeated several times in eukaryotic genomes. These repeats are prone to errors during DNA replication, and the failure to correct such errors results in microsatellite instability (MSI). Since its initial description as a form of genetic instability in colorectal cancer (CRC), MSI has been linked to a number of phenotypic characteristics and clinicopathological features of tumours (Sinicrope and Sargent, 2012). The prognostic and predictive relevance of MSI in CRC has been well documented (Sinicrope and Sargent, 2012). In CRC, MSI is the hallmark of forms of hereditary non-polyposis CRC (HNPCC; or Lynch syndrome) caused by germline mutations in mismatch repair (MMR) genes. MSI is also detected in 15% of sporadic CRC cases, usually as a consequence of epigenetically silenced MMRs (Sinicrope and Sargent, 2012). Furthermore, the presence of MSI has been demonstrated in other cancers, including endometrial, ovarian and urinary tract cancers. Currently, MSI appears to be a genetic aberration found in a wide range of human solid tumours that is best defined in the context of colorectal neoplasia; at tumour sites other than CRC, the clinical relevance and frequency of MSI are variable (Catto *et al*, 2003; Pal *et al*, 2008; Diaz-Padilla *et al*, 2013). In the clinical setting, such as when screening for CRC, HNPCC or sporadic MSI, MSI is usually classified as mononucleotide and dinucleotide repeats (Umar *et al*, 2004). Recently, a number of reports have also documented instability at specific tetranucleotide repeats, a phenomenon termed elevated microsatellite alterations at selected tetranucleotide repeats (EMAST). As an emerging specific form of genetic alteration, EMAST remains poorly described. However, EMAST may hold clues for deeper insights into the roles of MSI in cancer

*Correspondence: Professor K Søreide; E-mail: ksoreide@mac.com

Revised 1 March 2014; accepted 5 March 2014; published online 1 April 2014

biology, thus facilitating improved preventive, therapeutic and prognostic/predictive strategies in human cancer. This review aims to concisely present the current knowledge concerning EMAST in human solid cancers; we briefly present the proposed molecular mechanisms and describe the potential clinical implications for human cancer.

MICROSATELLITES, REPEAT SIZE AND INSTABILITY

During DNA replication, microsatellites are prone to erroneous duplication, which involves the addition or deletion of one or several repeated units, and the failure of DNA proofreading mechanisms to correct such errors results in instability at microsatellites. Microsatellite alterations, such as the insertion/deletion of a repeat, can produce deleterious effects via the induction of strand slippage and frameshift mutations when a protein-coding region is affected (Saeterdal *et al*, 2001; Kim *et al*, 2013). The results of such effects range from changes in the physical and chemical properties of the translated proteins to interference with transcription at promoter sites or the translation of truncated and dysfunctional proteins.

Generally, repetitive elements can occur in a variety of patterns and distributions (i.e., tandem, dispersed and paralogues). These elements constitute as much as 50% of the genome, and approximately one-fourth of empirically defined human promoters are surrounded by or contain clustered repetitive elements. However, the genomic distribution of microsatellites is rather random; mononucleotides, dinucleotides and tetranucleotides are primarily found in intronic and intergenic noncoding regions, whereas other microsatellites (trinucleotides and hexanucleotides) are located more abundantly within sequences involved in gene transcription (Subramanian *et al*, 2003).

Previous attempts at defining MSI have not produced straightforward results. Following the first MSI international workshop at the National Cancer Institute (NCI) at Bethesda in 1998, the 2004 revised NCI/Bethesda microsatellite screening panel (Umar *et al*, 2004), which is based on five mononucleotide and dinucleotide microsatellites, defined 'classical' MSI as high-frequency (MSI-H; $\geqslant$2 out of 5 markers are unstable or $\geqslant$40%), low-frequency (MSI-L; 1 out of 5 markers is unstable or $\geqslant$20% and $<$40%) or microsatellite-stable (MSS; no unstable markers detected). In CRC, there is a consensus on the definitions of MSI for clinical use (Umar *et al*, 2004), although defining the different degrees of instability remains a matter of controversy. Based on the Bethesda panel, MSS, MSI-L and MSI-H tumours contain different numbers of altered microsatellites, suggesting that these forms of MSI represent separate phenomena, at least at the molecular level. However, MSS tumours and MSI-L unstable tumours are typically grouped together, because both present similar clinical forms and gross abnormalities. Notably, MSI-L is generally detected only when dinucleotide markers are used, and some authors have disputed the existence of MSI-L completely (Kim *et al*, 2013). MSI is found with the highest frequency in mononucleotide repeats, which are also the most frequently occurring nucleotide repeats throughout the genome. Notably, tetranucleotide repeats occur at a much lower frequency than mononucleotide repeats, and it is not known whether forms of MSI involving nucleotide repeats of different sizes share common underlying mechanisms. However, the most clinically relevant microsatellites studied in cancer are composed of mononucleotide (e.g., G_n), dinucleotide (e.g., CA_n) or tetranucleotide (e.g., $AAAG_n$) repeats. Variations in the number of repeats in microsatellite sequences among the population are defined as microsatellite polymorphisms, as opposed to MSI, which refers to such alterations within multiple cells of the same individual (e.g., when comparing tumour and normal tissues). Instability at tetranucleotides represents a variation of MSI that was not addressed by the Bethesda definitions (Umar *et al*, 2004), which are based on mononucleotide and dinucleotide markers. However, in most EMAST-associated cancers, instability occurs at loci containing $AAAG_n$ or $ATAG_n$ repeats.

PREVALENCE AND RELEVANCE OF EMAST IN HUMAN CANCERS

EMAST has been reported in several cancers of solid organs (Figure 1), including the colorectum (Haugen *et al*, 2008; Devaraj *et al*, 2010; Lee *et al*, 2010; Lee *et al*, 2012), lungs (Ahrendt *et al*, 2000; Xu *et al*, 2001; Arai *et al*, 2013), ovaries (Singer *et al*, 2004), bladder (Xu *et al*, 2001; Danaee *et al*, 2002; Catto *et al*, 2003; Burger *et al*, 2006a), prostate (Perinchery *et al*, 2000; Burger *et al*, 2006b; Azzouzi *et al*, 2007), kidney (Xu *et al*, 2001; Catto *et al*, 2003), head and neck (Xu *et al*, 2001; Temam *et al*, 2004), non-melanoma skin (Danaee *et al*, 2002) and uterus (Choi *et al*, 2008). The estimated prevalence of EMAST according to previously reported studies is presented in Table 1. In fact, the reported prevalence and relationship between EMAST and clinicopathological features and molecular mechanisms vary considerably across tumour sites (Figure 1).

CRCs. Presumably because of the great impact that MSI has had on the understanding of human cancer biology, the vast majority of EMAST-related reports in the scientific literature concern CRC. EMAST is found most frequently ($>$60%) in advanced-stage CRC with poor tumour differentiation (Lee *et al*, 2010). The documented presence of EMAST in adenomas has been inconsistent, with reports ranging from 0% (Lee *et al*, 2012) to 33% (Lee *et al*, 2010). The overall frequency of EMAST in CRC is much higher (up to 60%) than the frequency of classical MSI-H (usually 15–20%) (Yamada *et al*, 2010), and EMAST occurs at high rates in rectal cancers (Devaraj *et al*, 2010). EMAST is associated with the MSI-L classification and demonstrates a worse prognosis with a shorter time to recurrence and development of distant metastasis in stage II and III CRC (Garcia *et al*, 2012). This worse prognosis is observed regardless of the association of EMAST with greater infiltration of CD8+ T cells in both the tumour and surrounding stroma (Devaraj *et al*, 2010; Lee *et al*, 2012), a feature usually associated with better prognosis in CRC. Thus the dual role of inflammation in this setting warrants further investigation.

Respiratory tract cancers. The reported frequency of EMAST in lung cancers ranges from 32% to 64.5% (Ahrendt *et al*, 2000; Xu *et al*, 2001; Woenckhaus *et al*, 2003; Arai *et al*, 2013). However, there appears to be a distinction between EMAST and MSI-H, and mutations in *TP53* are common in EMAST cancers (Ahrendt *et al*, 2000; Xu *et al*, 2001; Woenckhaus *et al*, 2003). EMAST was found to be most prevalent in cases of the squamous subtype and early stage disease (Xu *et al*, 2001; Woenckhaus *et al*, 2003; Arai *et al*, 2013), although it was associated with higher rates of lymph node metastasis in another study (Woenckhaus *et al*, 2003). Additionally, one study found a higher incidence of previous cancer history in patients with EMAST (Arai *et al*, 2013), potentially suggesting a 'field defect' or cancer susceptibility related to EMAST.

Kidney and urinary tract cancers. The frequency of EMAST in cancers of urogenital organs is markedly lower compared with that in other solid tumours and is higher in the lower urinary tract (urinary bladder) than the upper urinary tract (kidneys, ureters and renal pelvis) (see Table 1, Figure 1).

In bladder cancer, EMAST-positive tumours (approximately one-third of bladder tumours) have been associated with *TP53* mutations and non-invasive lesions (Danaee *et al*, 2002). However, a much lower EMAST frequency (5.3%) was demonstrated in a large, non-selected cohort of urothelial bladder cancers, and

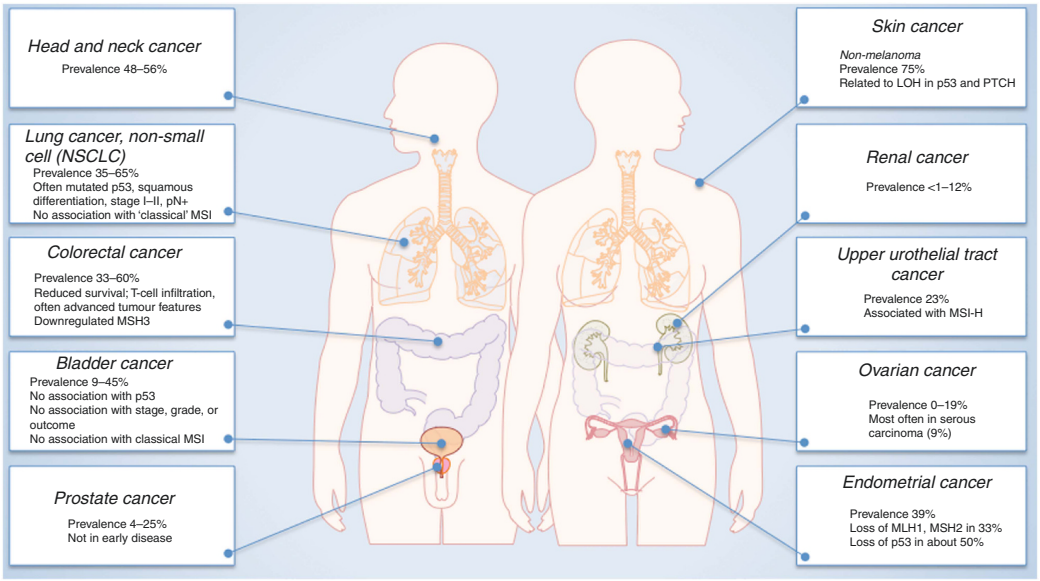


Figure 1. Prevalence and main observed features of EMAST in human solid cancers.

Table 1. Prevalence of EMAST among human solid malignancies

Site	Author, year	Cohort size (*n*)	Definitions	Prevalence (%)
Bladder	Xu *et al*, 2001	38	⩾1/12 markers unstable	21%
	Burger *et al*, 2006a, b	117	⩾1/10 markers unstable	9%
	Catto *et al*, 2003	89	⩾1/8 markers unstable	45%
	Danaee *et al*, 2002	57	⩾1/7 markers unstable	44%
Colorectal	Lee *et al*, 2010	108	⩾2/5 markers	50%
	Lee *et al*, 2012	108	⩾2/5 markers	50%
	Devaraj *et al*, 2010	147 (only rectal)	⩾2/5 markers	33%
	Haugen *et al*, 2008	117	⩾1/7 markers unstable	60%
NSCLC	Xu *et al*, 2001	47	⩾1/12 markers unstable	51%
	Arai *et al*, 2013	65	⩾1/10 markers unstable	65%
	Ahrendt *et al*, 2000	88	⩾1/13 markers unstable	35%
Prostate	Burger *et al*, 2006a, b	81	⩾1/10 markers unstable	5%
	Azzouzi *et al*, 2007	50	⩾2/4 markers	4%
	Perinchery *et al*, 2000	40	⩾1/4 markers	25%
Renal	Xu *et al*, 2001	25	⩾1/12 markers unstable	12%
	Catto *et al*, 2003	71 (upper urinary tract)	⩾1/8 markers unstable	23%
Head and neck	Xu *et al*, 2001	18	⩾1/12 markers unstable	56%
	Temam *et al*, 2004	54	⩾1/5 markers	48%
Non-melanoma skin	Danaee *et al*, 2002	61	⩾1/7 markers unstable	75%
Ovarian	Singer *et al*, 2004	53	⩾1/6 markers unstable	19%
Endometrial	Choi *et al*, 2008	39	⩾1/6 markers unstable and no instability in mononucleotides and dinucleotides	39%

Abbreviation: NSCLC = non-small cell lung cancer.

EMAST was not found to be associated with TP53 status (Burger *et al*, 2006a). In the urinary tract, EMAST appears to be unrelated to classical MSI and occurs less frequently in the renal pelvis (Catto *et al*, 2003). No firm associations between EMAST and clinicopathological features or clinical outcomes have been reported (Danaee *et al*, 2002; Catto *et al*, 2003; Burger *et al*, 2006a). The reported EMAST frequency in renal cell carcinomas is very low, ranging from < 1% (Stoehr *et al*, 2012) to 12% (Xu *et al*, 2001). Based on the limited available evidence, it appears that the prevalence of EMAST decreases from the upper to the lower urinary tract, with no solid clinical relationships or molecular mechanisms having been established in this area.

Reproductive organs. In cancers of the reproductive organs, EMAST occurs slightly more often in malignancies of the female reproductive organs, such as the ovaries (13%, Singer *et al*, 2004)

and uterus (38.5%, Choi *et al*, 2008), compared with the male reproductive organs, although the overall available literature remains limited.

Classical MSI is noted in approximately 12% of ovarian cancers (Pal *et al*, 2008) and is most frequently associated with the non-serous histological type. In contrast, EMAST appears predominantly in the serous subtype of ovarian cancers; no EMAST has been found in non-serous carcinomas (Singer *et al*, 2004).

Approximately 15% of endometrial cancers exhibit MSI-H, but there are conflicting reports of the relation to clinical outcome (Diaz-Padilla *et al*, 2013). EMAST has been reported at a higher frequency (39%) than classical MSI, but these two markers showed no correlation in type I endometrial carcinomas. Furthermore, most (70%) EMAST-positive tumours exhibited normal expression of both MSH2 and MLH1 in immunohistochemical analyses (Choi *et al*, 2008).

In prostate cancer, MSI is generally prevalent at dinucleotide tandem repeats and is less common in trinucleotide and tetranucleotide repeats (Perinchery *et al*, 2000). According to most reports, the frequency of EMAST in prostate cancer is $<5\%$ (Terrell *et al*, 1995; Burger *et al*, 2006b; Azzouzi *et al*, 2007). In prostate cancer, EMAST is related neither to MSI nor to defects in MMR proteins, *TP53* mutational status or any histopathological features (Burger *et al*, 2006b).

Other malignancies. Instabilities in tetranucleotide repeats have been reported in both the MSI-H and MSI-L types of gastric tumours, although EMAST has not been specifically investigated in this setting (Kim *et al*, 2001). In a small study of 22 pheochromocytomas (Kupka *et al*, 2008), 2 patients demonstrated MSI in 1 and 2 of three tetranucleotide repeats (D2S443, D16S752 and D21S1436) investigated; both cancers were defined as MSI-L by the Bethesda panel. None of the endocrine tumours in that study exhibited instability in the tetranucleotide markers used, although 13% were classified as MSI-H (Kupka *et al*, 2008).

EMAST was also analysed in one study of non-melanoma skin cancer, where it was found at a high incidence (75%) and was associated with *TP53* or *PTCH* (patched gene) loss of heterozygosity (Danaee *et al*, 2002). PTCH acts as receptor in the sonic hedgehog signalling pathway, and loss-of-function mutations in PTCH contribute to skin cancer development. The PTCH gene harbours mutational hot-spot residues and regions, including a slippage-sensitive sequence at the N-terminus (Lindstrom *et al*, 2006).

EMAST frequencies in head and neck cancers have been reported to range from 48% to 56% (Xu *et al*, 2001; Temam *et al*, 2004). Of clinical relevance was one report showing that patients with histologically proven ($=$R0) radical and curative surgery, but who displayed EMAST at the resection margins, had a higher risk of tumour recurrence (Temam *et al*, 2004). 'Molecular-positive' resection margins may thus be proposed as grounds for more aggressive treatment or surveillance if validated as a broadened concept.

More recently, the length of an AAAG tetranucleotide repeat tract and polymorphisms in a tetranucleotide repeat, both located within the KCNQ1OT1 gene, have been shown to correlate with an increased risk of breast cancer (Karimi *et al*, 2013) and hepatocellular carcinoma (Wan *et al*, 2013), respectively.

CONCLUSIONS

The molecular mechanisms of EMAST have yet to be clearly unravelled. EMAST induction and its contribution to carcinogenesis may stem from both exposure to external mutagens and malfunctioning intrinsic cellular mechanisms. Experiments have suggested that long microsatellites with higher numbers of repeats may be prone to replication errors at higher frequencies compared with the shorter ones. However, the consequences of EMAST development in carcinogenesis require further investigation. The results from human solid cancers have led to proposed mechanisms involving DNA repair by TP53 and MSH3 and have suggested the immunological involvement of T cells, at least in CRC, although precise mechanistic understanding clearly requires further investigation.

The potential role of EMAST in early detection, prognostication and prediction has been poorly investigated. However, some findings have revealed the potential future use of this marker in cancer. Variations of specific lengths in certain polymorphic tetranucleotide repeats are associated with an increased cancer risk, as recently reported for breast cancer and hepatocellular carcinomas (Karimi *et al*, 2013; Wan *et al*, 2013). The high prevalence of EMAST in the early stages of some cancers, particularly those of the upper respiratory organs and lower urinary tract, may suggest EMAST as a potentially exploitable marker for preventive/early detection screening. In fact, in biopsies from endoscopically normal colons, MSI in mononucleotide and dinucleotide markers has been observed earlier than neoplastic changes (Tug *et al*, 2012). Although markers suggestive of EMAST have yet to be explored in the same manner, they EMAST markers may be found more abundantly, because EMAST occurs at higher frequencies. EMAST may also be generally related to environmental carcinogen exposure and may therefore serve as a marker of exposure or risk in certain cancers.

EMAST in the tissues of apparently tumour-free marginal sites of surgical resections has been correlated with disease recurrence (Temam *et al*, 2004). Therefore, EMAST in the tissues of histomorphologically normal (R0 resected) surgical cancer specimens could be used as a prognostic predictor of disease recurrence in head and neck cancers. Optimal marker selection may also lead to the feasibility of urine analysis for the early detection or postsurgical surveillance of patients with bladder cancer (van Tilborg *et al*, 2012).

Finally, the prognostic value of different forms of MSI in CRC warrants further research (Garcia *et al*, 2012). The 5-year recurrence-free survival trends for EMAST-positive and MSI-L tumours are worse compared with those of MSI-H tumours, suggesting differences in tumour biology that are currently not fully appreciated and that may have consequences for the use of the other available predictive and prognostic markers.

Taken together, these data suggest that EMAST should be further investigated for its potential use in screening and/or as a prognostic and predictive biomarker. In the fields of clinical and translational oncology, important steps need to be taken to improve the understanding of the role of EMAST in cancer. In our view, one unresolved issue remains regarding whether EMAST represents an 'innocent bystander' in cancer or an 'active partner' in carcinogenesis. Based on the emerging findings and cumulative reports discussed above, EMAST may have diverse roles in carcinogenesis and tumour biology at different tumour sites. Nevertheless, common underlying principles may be crucial for understanding cancer behaviour and may yield new tools for prevention, early detection or personalised treatment.

ACKNOWLEDGEMENTS

Work in our laboratory was supported by grants from the Folke Hermansens Cancer Fund, the Mjaaland Research Fund and the Institute for Surgical Sciences, University of Bergen, Norway. We apologise to those colleagues whose work could not be cited due to space limitations.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Ahrendt SA, Decker PA, Doffek K, Wang B, Xu L, Demeure MJ, Jen J, Sidransky D (2000) Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. *Cancer Res* **60**(9): 2488–2491.

Arai H, Okudela K, Oshiro H, Komitsu N, Mitsui H, Nishii T, Tsuboi M, Nozawa A, Noishiki Y, Ohashi K, Inui K, Masuda M (2013) Elevated microsatellite alterations at selected tetra-nucleotide (EMAST) in non-small cell lung cancers—a potential determinant of susceptibility to multiple malignancies. *Int J Clin Exp Pathol* **6**(3): 395–410.

Azzouzi AR, Catto JW, Rehman I, Larre S, Roupret M, Feeley KM, Cussenot O, Meuth M, Hamdy FC (2007) Clinically localised prostate cancer is microsatellite stable. *BJU Int* **99**(5): 1031–1035.

Burger M, Burger SJ, Denzinger S, Wild PJ, Wieland WF, Blaszyk H, Obermann EC, Stoehr R, Hartmann A (2006a) Elevated microsatellite instability at selected tetranucleotide repeats does not correlate with clinicopathologic features of bladder cancer. *Eur Urol* **50**(4): 770–775.

Burger M, Denzinger S, Hammerschmied CG, Tannapfel A, Obermann EC, Wieland WF, Hartmann A, Stoehr R (2006b) Elevated microsatellite alterations at selected tetranucleotides (EMAST) and mismatch repair gene expression in prostate cancer. *J Mol Med (Berl)* **84**(10): 833–841.

Catto JW, Azzouzi AR, Amira N, Rehman I, Feeley KM, Cross SS, Fromont G, Sibony M, Hamdy FC, Cussenot O, Meuth M (2003) Distinct patterns of microsatellite instability are seen in tumours of the urinary tract. *Oncogene* **22**(54): 8699–8706.

Choi YD, Choi J, Kim JH, Lee JS, Lee JH, Choi C, Choi HS, Lee MC, Park CS, Juhng SW, Nam JH (2008) Microsatellite instability at a tetranucleotide repeat in type I endometrial carcinoma. *J Exp Clin Cancer Res* **27**: 88.

Danaee H, Nelson HH, Karagas MR, Schned AR, Ashok TD, Hirao T, Perry AE, Kelsey KT (2002) Microsatellite instability at tetranucleotide repeats in skin and bladder cancer. *Oncogene* **21**(32): 4894–4899.

Devaraj B, Lee A, Cabrera BL, Miyai K, Luo L, Ramamoorthy S, Keku T, Sandler RS, McGuire KL, Carethers JM (2010) Relationship of EMAST and microsatellite instability among patients with rectal cancer. *J Gastrointest Surg* **14**(10): 1521–1528.

Diaz-Padilla I, Romero N, Amir E, Matias-Guiu X, Vilar E, Muggia F, Garcia-Donas J (2013) Mismatch repair status and clinical outcome in endometrial cancer: a systematic review and meta-analysis. *Crit Rev Oncol Hematol* **88**(1): 154–167.

Garcia M, Choi C, Kim HR, Daoud Y, Toiyama Y, Takahashi M, Goel A, Boland CR, Koi M (2012) Association between recurrent metastasis from stage II and III primary colorectal tumors and moderate microsatellite instability. *Gastroenterology* **143**(1): 48–50 e41.

Haugen AC, Goel A, Yamada K, Marra G, Nguyen TP, Nagasaka T, Kanazawa S, Koike J, Kikuchi Y, Zhong X, Arita M, Shibuya K, Oshimura M, Hemmi H, Boland CR, Koi M (2008) Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer Res* **68**(20): 8465–8472.

Karimi P, Hematti S, Safari F, Tavassoli M (2013) Polymorphic AAAG repeat length in estrogen-related receptor gamma (ERRgamma) and risk of breast cancer in Iranian women. *Cancer Invest* **31**(9): 600–603.

Kim HS, Lee BL, Woo DK, Bae SI, Kim WH (2001) Assessment of markers for the identification of microsatellite instability phenotype in gastric neoplasms. *Cancer Lett* **164**(1): 61–68.

Kim TM, Laird PW, Park PJ (2013) The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell* **155**(4): 858–868.

Kupka S, Haack B, Zdichavsky M, Mlinar T, Kienzle C, Bock T, Kandolf R, Kroeber SM, Konigsrainer A (2008) Large proportion of low frequency microsatellite-instability and loss of heterozygosity in pheochromocytoma and endocrine tumors detected with an extended marker panel. *J Cancer Res Clin Oncol* **134**(4): 463–471.

Lee SY, Chung H, Devaraj B, Iwaizumi M, Han HS, Hwang DY, Seong MK, Jung BH, Carethers JM (2010) Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias. *Gastroenterology* **139**(5): 1519–1525.

Lee SY, Miyai K, Han HS, Hwang DY, Seong MK, Chung H, Jung BH, Devaraj B, McGuire KL, Carethers JM (2012) Microsatellite instability, EMAST, and morphology associations with T cell infiltration in colorectal neoplasia. *Dig Dis Sci* **57**(1): 72–78.

Lindstrom E, Shimokawa T, Toftgard R, Zaphiropoulos PG (2006) PTCH mutations: distribution and analyses. *Hum Mutat* **27**(3): 215–219.

Pal T, Permuth-Wey J, Kumar A, Sellers TA (2008) Systematic review and meta-analysis of ovarian cancers: estimation of microsatellite-high frequency and characterization of mismatch repair deficient tumor histology. *Clin Cancer Res* **14**(21): 6847–6854.

Perinchery G, Nojima D, Goharderakhshan R, Tanaka Y, Alonzo J, Dahiya R (2000) Microsatellite instability of dinucleotide tandem repeat sequences is higher than trinucleotide, tetranucleotide and pentanucleotide repeat sequences in prostate cancer. *Int J Oncol* **16**(6): 1203–1209.

Saeterdal I, Bjorheim J, Lislerud K, Gjertsen MK, Bukholm IK, Olsen OC, Nesland JM, Eriksen JA, Moller M, Lindblom A, Gaudernack G (2001) Frameshift-mutation-derived peptides as tumor-specific antigens in inherited and spontaneous colorectal cancer. *Proc Natl Acad Sci USA* **98**(23): 13255–13260.

Singer G, Kallinowski T, Hartmann A, Dietmaier W, Wild PJ, Schraml P, Sauter G, Mihatsch MJ, Moch H (2004) Different types of microsatellite instability in ovarian carcinoma. *Int J Cancer* **112**(4): 643–646.

Sinicrope FA, Sargent DJ (2012) Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. *Clin Cancer Res* **18**(6): 1506–1512.

Stoehr C, Burger M, Stoehr R, Bertz S, Ruemmele P, Hofstaedter F, Denzinger S, Wieland WF, Hartmann A, Walter B (2012) Mismatch repair proteins hMLH1 and hMSH2 are differently expressed in the three main subtypes of sporadic renal cell carcinoma. *Pathobiology* **79**(3): 162–168.

Subramanian S, Mishra RK, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol* **4**(2): R13.

Temam S, Casiraghi O, Lahaye JB, Bosq J, Zhou X, Julieron M, Mamelle G, Lee JJ, Mao L, Luboinski B, Benard J, Janot F (2004) Tetranucleotide microsatellite instability in surgical margins for prediction of local recurrence of head and neck squamous cell carcinoma. *Clin Cancer Res* **10**(12 Pt 1): 4022–4028.

Terrell RB, Wille AH, Cheville JC, Nystuen AM, Cohen MB, Sheffield VC (1995) Microsatellite instability in adenocarcinoma of the prostate. *Am J Pathol* **147**(3): 799–805.

Tug E, Balaban YH, Sahin EK (2012) Mapping of microsatellite instability in endoscopic normal colon. *Genet Test Mol Biomarkers* **16**(5): 388–395.

Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S (2004) Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* **96**(4): 261–268.

van Tilborg AA, Kompier LC, Lurkin I, Poort R, El Bouazzaoui S, van der Keur K, Zuiverloon T, Dyrskjot L, Orntoft TF, Roobol MJ, Zwarthoff EC (2012) Selection of microsatellite markers for bladder cancer diagnosis without the need for corresponding blood. *PLoS One* **7**(8): e43345.

Wan J, Huang M, Zhao H, Wang C, Zhao X, Jiang X, Bian S, He Y, Gao Y (2013) A novel tetranucleotide repeat polymorphism within KCNQ1OT1 confers risk for hepatocellular carcinoma. *DNA Cell Biol* **32**(11): 628–634.

Woenckhaus M, Stoehr R, Dietmaier W, Wild PJ, Zieglmeier U, Foerster J, Merk J, Blaszyk H, Pfeifer M, Hofstaedter F, Hartmann A (2003) Microsatellite instability at chromosome 8p in non-small cell lung cancer is associated with lymph node metastasis and squamous differentiation. *Int J Oncol* **23**(5): 1357–1363.

Xu L, Chow J, Bonacum J, Eisenberger C, Ahrendt SA, Spafford M, Wu L, Lee SM, Piantadosi S, Tockman MS, Sidransky D, Jen J (2001) Microsatellite instability at AAAG repeat sequences in respiratory tract cancers. *Int J Cancer* **91**(2): 200–204.

Yamada K, Kanazawa S, Koike J, Sugiyama H, Xu C, Funahashi K, Boland CR, Koi M, Hemmi H (2010) Microsatellite instability at tetranucleotide repeats in sporadic colorectal cancer in Japan. *Oncol Rep* **23**(2): 551–561.

Søreide *et al. J Transl Med (2016) 14:192*
DOI 10.1186/s12967-016-0951-4

Journal of Translational Medicine

PROTOCOL

Open Access

Assessment of clinically related outcomes and biomarker analysis for translational integration in colorectal cancer (ACROBATICC): study protocol for a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastasis

Kjetil Søreide[1,2,3*], Martin M. Watson[1,2], Dordi Lea[2,4], Oddmund Nordgård[5], Jon Arne Søreide[1,3], Hanne R. Hagland[2,6] and ACROBATICC collaborators

Abstract

Background: More accurate predictive and prognostic biomarkers for patients with colorectal cancer (CRC) primaries or colorectal liver metastasis (CLM) are needed. Outside clinical trials, the translational integration of emerging pathways and novel techniques should facilitate exploration of biomarkers for improved staging and prognosis.

Methods: An observational study exploring predictive and prognostic biomarkers in a population-based, consecutive cohort of surgically treated colorectal cancers and resected colorectal liver metastases. Long-term outcomes will be cancer-specific survival, recurrence-free survival and overall survival at 5 years from diagnosis. Beyond routine clinicopathological and anthropometric characteristics and laboratory and biochemistry results, the project allows for additional blood samples and fresh-frozen tumour and normal tissue for investigation of circulating tumour cells (CTCs) and novel biomarkers (e.g. immune cells, microRNAs etc.). Tumour specimens will be investigated by immunohistochemistry in full slides. Extracted DNA/RNA will be analysed for genomic markers using specific PCR techniques and next-generation sequencing (NGS) panels. Flow cytometry will be used to characterise biomarkers in blood. Collaboration is open and welcomed, with particular interest in mutual opportunities for validation studies.

Status and perspectives: The project is ongoing and recruiting at an expected rate of 120–150 patients per year, since January 2013. A project on circulating tumour cells (CTCs) has commenced, with analysis being prepared. Investigating molecular classes beyond the TNM staging is under way, including characteristics of microsatellite instability (MSI) and elevated microsatellite alterations in selected tetranucleotides (EMAST). Hot spot panels for known mutations in CRC are being investigated using NGS. Immune-cell characteristics are being performed by IHC and flow cytometry in tumour and peripheral blood samples. The project has ethical approval (REK Helse Vest, #2012/742), is financially supported with a Ph.D.-Grant (EMAST project; Folke Hermansen Cancer Fund) and a CTC-project (Norwegian Research Council; O. Nordgård). The ACROBATICC clinical and molecular biobank repository will serve as a

*Correspondence: ksoreide@mac.com
[1] Department of Gastrointestinal Surgery, Stavanger University Hospital, POB 8100, 4068 Stavanger, Norway
Full list of author information is available at the end of the article

long-term source for novel exploratory analysis and invite collaborators for mutual validation of promising biomarker results. The project aims to generate results that can help better discern prognostic groups in stage II/III cancers; explore prognostic and predictive biomarkers, and help detail the biology of colorectal liver metastasis for better patient selection and tailored treatment. The project is registered at http://www.ClinicalTrials.gov NCT01762813.

Keywords: Biomarker, Cancer, Population-based, Translational research, Colorectal cancer, Liver metastasis, Circulating tumour cells, Genetics

Background

Colorectal cancer represents a formidable health burden worldwide with an expected 60 % increase towards 2030 [1]. Currently, CRC ranges as the second most frequent cancer in both genders in the Western world. Despite an increasingly favourable prognosis due to stepwise progression in surgical and oncological management [2], still about 40–50 % will develop metastasis and die from the disease. The liver is the most frequent site for metastasis, followed by the lungs, and is also the rate-limiting organ step for long-term survival. For non-metastatic disease, prognosis is guided through the tumour-node-metastasis (TNM)-system, which heavily relies on the status of lymph nodes for current staging [3, 4]. Further, node status may vary with the underlying molecular composition of primary tumour [5, 6]. Also, more refined node-examination including ultrastaging by immunohistochemistry, sentinel node techniques or use of molecular markers to identify malignant cells have not yielded a higher precision overall [7, 8]. Furthermore, other methods and techniques of staging patients, such as the use of "liquid biopsies" i.e. by investigating circulating tumour cells (CTCs) or other tumour constituents in peripheral blood (e.g. microRNAs), may prove to have higher prognostic and predictive value in both primary and metastatic CRC [9–12]. Notably, well-described molecular routes of progression in CRC have been linked to specific prognosis and outcomes, including microsatellite instability (MSI), CpG-island methylator phenotype (CIMP) and chromosomal instability (CIN) [13–17].

While the TNM is the best staging system at hand for clinical decision making, the TNM system is known to be imperfect [4], and substantial over- and undertreatment results from failure to accurately predict disease outcomes. Indeed, increased knowledge of cancer heterogeneity has led researchers and clinicians alike to pursue better ways of stratifying therapy to individual risk and effects response and efficacy of therapy [18]. One suggested consensus taxonomy has emerged for novel risk-groups [19], however these have yet to be implemented in clinical practice. Variation in definition of common denominators for disease stratification may be due to a number of reasons, including heterogeneous patient groups investigated; investigations done on patients recruited to randomized trial with strict inclusion criteria; variation in tumour sampling, and; variation in molecular analyses and techniques, to mention but a few [14, 20–22]. Thus, exploring biomarkers and the described genetic and epigenetic pathways in CRC [23, 24] by well-defined population cohorts with access to biobanking beyond routine samples is crucial (Fig. 1).

The aim of this prospective project is to facilitate prospective accrual of patients with snap-frozen blood samples and fresh-frozen tissue outside routine clinical care, for more refined molecular evaluation (Fig. 1). For one, we will explore the role of microsatellite instability (MSI) and specifically a form found in tetranucleotide repeats (elevated microsatellite alterations in selected tetranucleotides; EMAST), which is found at varying frequencies in several cancers [25]. EMAST is a less-well described molecular trait in CRC, however, recent data point to a prognostic role and as a potential modulator of cancer biology [26–28]. The role of MSI and EMAST in relation to presently well-described prognostic mutations, such as KRAS and BRAF [29, 30], is not well described. Also, investigation into putative mechanisms leading to EMAST and the consequence for a predictive and prognostic role is warranted. Second, a cohort of consecutive patients will be evaluated for CTCs in resectable stage I–III CRC for its prognostic value. Third, as the role of the immune system is increasingly recognized as an integral component to carcinogenesis, cancer biology and patient prognosis [31–33], the project will investigate elements of the immune system in tumour samples as well as in peripheral blood. Finally, additional expansion of related projects are emerging with novel techniques, particularly with the availability of next-generation sequencing. Here we present the study design and protocol for a population-based translational cancer research project.

Methods

Study ethics approval

The project and research biobank has been approved by the Regional Ethics Committee of the Western Health Authority (REK Helse Vest, #2012/742) and by the Institutional review board (Helse Stavanger HF Protocol Record #29034/2012). The project has been registered at www.ClinicalTrials.gov NCT#01762813.

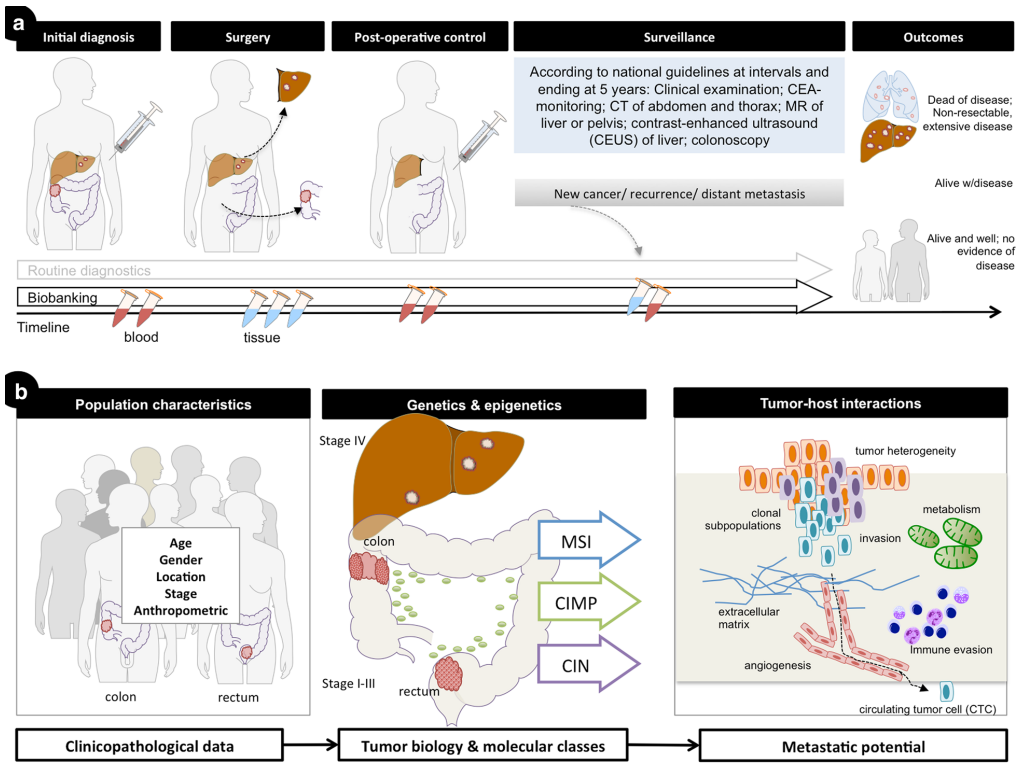


Fig. 1 The ACROBATICC project flow sequence and rationale for cancer biology investigation. **a** Illustrated is a simple workflow of patients' recruitment and samples of blood (*red vials*) and tissues (*blue vials*) from initial diagnosis, before and after surgery and during follow up. Overall, disease-free and cancer-specific survival will be analysed at 5 years. **b** Illustrated are the specific levels of patient information gathered for prognostic and predictive use, ranging from clinicopathological characteristics (such as sex, age, body weight and height) to genetic and epigenetic mechanisms (including microsatellite instability; CpG-island methylator phenotypes and chromosomal instability) and specific tumour-host interactions (such as immune-response in tumor; cancer metabolism and role of circulating tumor cells). With access to newer techniques and development of novel hypothesis, the project will allow for exploration of other predictors, as well as serve as external validation cohort in collaborative research

Consenting

Patients are informed and consented in the surgery outpatient clinic or, if directly admitted, in the surgical ward, before admission for surgery. A trained research nurse certified in Good Clinical Practice does consenting of patients and registration.

Study population

Stavanger University Hospital serves as the only hospital for a population of about 350,000 inhabitants in the South-Western part of Norway. The population is predominant of Caucasian origin, the average population age slightly younger than the national average. Socioeconomic differences are not extreme in the country; the life expectancy is just below 80 years for men and about 81 years for women. With no other competing hospitals in the region and a social security system ensuring equal care for all patients, the study population allows for reliable, unselected, population-based and representative data sampling with little risk of bias.

We have previously reported epidemiological characteristics to other disease categories based on the same non-selected, unbiased conditions, which should validate the methodology to the current population-based perspective [34–39].

With low migration in the region, long-term follow up is feasible and allows for high precision in catching new events (disease related or other) with impact on

outcomes, as previously reported [40, 41]. The risk of loosing patients to long-term follow-up is thus minimal.

All Norwegian citizens have a unique 11-digit social security number that allows identification through the hospital electronic hospital records with national registries, including the Norwegian Patient Register and the Cancer Registry of Norway. All CRC patients are registered to the National Colorectal Cancer Registry via an electronic template record form. The study will comply to the strengthening the reporting of observational studies in epidemiology (STROBE) recommendations on what should be included in an accurate and complete report of any observational study [42].

Inclusion and exclusion criteria

All consecutive patients diagnosed with operable primary or metastatic colorectal cancer and able to provide informed, written consent are included. Participants can at any time withdraw from the study without need for providing any explanation for the withdrawal, upon which the records will be destroyed and deleted. Excluded are patients presenting as emergencies and unable to consent, or patients unable to understand oral and written Norwegian or, patients whose cognitive status does not allow for informed consent.

Study period

The study commenced in January 2013 and recruitment is ongoing, for an expected closure of recruitment in 2018, with final 5-year follow-up to be completed in 2022.

Study number and sample size

For prognostic information and estimated sufficient numbers of patients and events, we expect to recruit about 150 patients per year. With an expected 35–45 % recurrence rate within 5 years of primary diagnosis—which is expected from previous regional and national data of curatively resected CRC in Norway [2, 40, 43]—and, an accrual commencing over at least 5 years (2013–2018) we expect to have about 750 patients with operable primary and/or metastatic CRC for evaluation at the end of the period. In stage I-III CRC, the recurrence rate at 35–45 % by 5 years [2, 40, 43] should yield appropriate number of events (cancer-specific survival) for creation of test-sets, validation-sets and prospective evaluation. Recurrence rates are expected to be even higher in colorectal liver metastasis (>80 % recurrence within 5 years), suggesting fewer patients are needed to evaluate the endpoint.

Clinical work-up and care

The Department of Gastrointestinal Surgery provide all clinical work-up and surgical care for patients with colorectal cancer and subsequent evaluation for metastases, resectable or non-resectable. Oncologic care is provided at the Department of Oncology. All radiologic work-up (except PET/CT scans) are performed at the Department of Radiology. Specimen evaluation and tissue blocks preparation for routine diagnostics are performed at the Department of Pathology. Routine blood tests are analysed at the Department of Clinical Chemistry. Storage of research samples (fresh-frozen) are archived in an intramural research biorepository at the Stavanger University Hospital. Subsequent DNA and RNA retrieval and elaborate laboratory work outside routine diagnostics are performed at the Laboratory for Molecular Biology, except if otherwise stated.

All patient care are performed under the recommended national guidelines issued by the multidisciplinary Norwegian Gastrointestinal Cancer Group (NGICG), for both colorectal cancer (NGICG-CRC) and liver metastasis (NGICG-HPB), respectively.

Collaboration

Interested collaborators are welcome to make contact. Discussion is extant with other groups and thus has the potential to generate an international cohort for comparison and validation of results. No external collaboration is yet confirmed but discussion in progress.

The cohort material will similarly also be available for cross-evaluation with other cohorts generated elsewhere, and will be beneficial for external validation purposes and hypothesis-generating experimental studies.

Samples

Issues to attend to for biomarker research and reporting have been addressed in several leading journals [21, 22]. Consequently, the current study will seek to comply and report according to the biospecimen reporting for improved study quality (BRISQ) recommendations for all tissue sampling and storage in the study [44]. For parts of the study relating to clinical prognosis, we will aim to address the REMARK guidelines for biomarker research [45].

Tissue samples

Formalin-fixed paraffin-embedded tumour/normal

Resected specimens are handled at the Department of Pathology according to protocol. An electronic template is followed and applied for gross examination and microscopic description of pathologic features and data for staging. Staging is done per the TNM-system (AJCC 7th edition). Representative tissue slides (resection ends or normal tissue distant from primary tumour; several tumour slides including most invasive front; all sampled lymph nodes) are formalin-fixed paraffin-embedded

(FFPE) for routine H&E diagnostics and microscopy. Lymph nodes are sampled per protocol and aimed to achieve at least 12 nodes and, if less, a 'lymph node revealing solution' (a mixture composed of 95 % ethanol, diethyl ether, glacial acetic acid, and buffered formalin; also called GWEF) is applied to mesenteric fat in order to enhance node recovery [46].

Frozen fresh tumour tissue and normal sample

Before formalin fixation and immediate upon retrieval of the specimen, representative fresh tumour samples are obtained (at least three per tumour), stored in meticulously marked vials, and frozen in liquid nitrogen. Time is kept to a minimum between resection and delivery at Pathology in order to minimize loss of RNA quality [47, 48], usually delivered by an orderly within 15 min of retrieval from the operating room to the laboratory.

For the rare occasional procedures commencing or proceeding outside opening hours of the Department of Pathology, the surgeon in charge samples the fresh tumour biopsies per protocol and provides this in a portable insulated box container with dry ice for storage, normally <12 h. Samples are then collected by a technician and processed per protocol as early as possible the next morning.

Blood samples

Peripheral blood samples are drawn (usually) from the antecubital vein on admission before surgery and on the outpatient follow-up appointment, approximately 4 weeks after surgery. Subsequent blood is drawn if patient is readmitted for new surgery for recurrence or metastatic disease, and then again if a second curative-intent surgery (e.g. resection of new large bowel tumour; local recurrence; or, metastasectomy) is planned. Blood samples are processed to serum and plasma by centrifugation and two vials of full blood (EDTA-containers) are frozen and stored in −80 °C freezers together with processed samples.

Circulating tumour cell (CTC) detection

Peripheral blood samples (9 ml) are collected in EDTA tubes and subjected to density centrifugation within 20 h (preferably 2 h) from the collection time. RNA is isolated from the peripheral blood mononuclear cell fraction and reverse transcribed. Circulating tumour cells (CTCs) are then detected indirectly by measuring epithelial-specific mRNAs, which are not present in normal blood cells, as surrogate markers [12, 49, 50]. mRNA concentrations are measured by quantitative reverse-transcription PCR. The background levels in blood samples from healthy control persons are utilized as a reference material to determine which patient samples are positive for CTCs.

DNA/RNA extraction

DNA and RNA are extracted from freshly frozen tumour and normal (surgical resection margins or normal tissue sampled distant from primary tumour) using the QIACUBE (Qiagen) instrument and dedicated reagents and kits, according to manufacturers instructions. Weighted 15–20 mg of tissue are resuspended in lysis buffer and homogenized in the presence of 5 mm Ø steel beads, in a TissueLyser LT (Qiagen), at 50 Hz, for 4 min. Two consecutive protocols are then used on the QIACUBE instrument to extract DNA first, and RNA later (from flow through of first protocol) via the use of AllPrep DNA/RNA/miRNA Universal Kit. Concentration, purity (A260/280) and presence of phenol and protein contaminants in the eluted sample (A260/230) are measured and noted with a NanoDrop (ThermoFischer) intrument. Extracted DNA and RNA are labelled and stored at −80 °C in the aforementioned intramural biobanking facility.

Employed molecular techniques

Mentioned examples here are not exclusive, but include:

Immunohistochemistry (IHC)

Antigen retrieval and antibody dilution are optimized prior to the study onset for the different antibodies. To ensure uniform handling of samples, all sections are processed simultaneously. Paraffin sections adjacent to the haematoxylin-eosin (H&E) sections used for histology are mounted onto Superfrost Plus slides and dried overnight at 37 °C followed by 1 h at 60 °C. Sections are deparaffinised in xylene and rehydrated in decreasing concentrations of alcohol. Antigen is retrieved using Tris–EDTA (pH 9.0) as the retrieval buffer. Endogenous peroxidase activity is blocked with a peroxidase-blocking reagent. The immune complex is visualized with the Dako REAL EnVision Detection System, Peroxidase/DAB, Rabbit/Mouse (K5007; Dako). Sections are incubated with EnVision/HRP, Rabbit/Mouse for 30 min and diaminobenzidine (DAB+) chromogen. The sections are counterstained with haematoxylin, dehydrated, and mounted. All steps are performed using DakoAutostainer and TBS (S1968; Dako) with 0.05 % Tween 20 as wash buffer. Quality assessment and scoring of the samples are executed by an experienced pathologist and with use of digital pathology software (Visiopharm) for some antibodies.

Candidate markers for investigation are in development, and includes (but not limited to) suggested markers for immune cells (e.g. CD4+, CD8+, CD45RO+) and as suggested in the Immunoscore [51], potential markers of differentiation (e.g. CDX2) in CRC stage subtypes [52], and markers related to mechanistic insight, such as

MSH3 and its relation to EMAST [53, 54]. Other markers will be employed for specific subprojects as needed, e.g. for validating protein expression for gene variations.

Flowcytometry

Freshly drawn blood is collected in EDTA coated blood vacuum containers. Percentages of human CD4+ and CD8+ T-lymphocytes in erythrocyte-lysed whole blood are determined by flow cytometry. The antibody kit is acquired from BD Biosciences (Cat no: 342417) with CD3 FITC/CD8 PE/CD45 PerCP/CD4 APC conjugated staining. Cells are prepared according to protocol and 100 µl whole blood is used. BD Pharm Lyse (Cat no: 555899) is used at appropriate dilution to lyse erythrocytes. The samples are run on an Accuri C6 (BD biosciences) or a Cytoflex (Beckman-Coulter) flow cytometer both systems equipped with a blue and red laser, two light scatter detectors, and four fluorescence detectors with optical filters optimized for the detection of FITC, PE, PerCP, and APC. Further analysis of results is being done using the corresponding software (BD Biosciences Accuri C6 analysis software and CytExpert Beckman-Coulter).

Next generation sequencing (NGS)

Template preparation and chip loading is carried out using the Ion Chef™ System. With the use of CRC-specific, validated custom and commercially available panels, targeted DNA ion semiconductor sequencing of tumour material is performed on the Ion Torrent™ Personal Genome Machine® (PGM, ThermoFischer) platform. Data analysis against the reference (human) genome is executed in-house with the aid of the Torrent Suite™ and Ion Reporter™ softwares.

Statistical analysis and endpoints

The project is exploratory and thus no formal statistical power has been done. The population-based, observational, real-life, non-selected cohort will allow for adequate power based on the expected recruitment of a cohort size of 750 patients, of which an expected one-third (about 250) will have recurrence or death from disease. With the high number of events, this will allow for reasonable multivariable adjustments for outcomes. For smaller samples in subgroups (e.g. only stage II; or, only patients with resected liver metastasis), hypothesis-generating results will be pursued with appropriate sized prospective cohort samples with internal and external validation cohorts, where available.

The main endpoints will be cancer-specific, recurrence-free and overall survival, which will be analysed with Kaplan–Meier figures and log rank test. Exploratory analyses will be done using descriptive techniques for hypothesis-generating results. For laboratory values without established cut-offs, we will apply receiving-operator characteristics (ROC) analyses for optimal cut-off determination [55]. For prognostic factors, we will apply appropriate multivariable regression analyses for appropriately adjusted analyses.

Discussion

The ACROBATICC projects aim to integrate the routine clinical work-up and treatment of patients with primary CRC and resectable liver metastasis with state-of-the-art molecular technology investigations of blood samples and tumor tissues. The aim is to explore and identify better predictive and prognostic biomarkers that may eventually help in clinical decision-making for more precise, personalized and tailored treatment. The ACROBATICC clinical and molecular biobank repository will serve as a long-term source for novel exploratory analysis and invite collaborators for mutual validation of promising biomarker results. The project aims to generate results that can help better discern prognostic groups in stage II/III cancers; explore prognostic and predictive biomarkers, and help detail the biology of colorectal liver metastasis for better patient selection and tailored treatment [23, 28, 56].

The role of population-based cancer biobanking is increasingly recognized as important for exploratory and confirmatory studies at an unselected, population-level. While regular diagnostic biobanking [i.e. formalin-fixed and paraffin embedded (FFPE) tissue blocks] allow for valuable analyses outside routine descriptive data, such repositories may have a number of medicolegal and laboratory limitations which may be overcome by specific research-driven projects. The current translational cancer research project will allow for further in-depth analyses into cancer biology otherwise not available from material obtained by routine care.

The use of "liquid biopsy" has gained considerable attention as a novel source of biomarkers. Blood-based biomarkers could prove to be practical tools for CRC detection, as the monitoring of biomarkers in biological fluids offers many advantages, including minimal invasiveness and easy accessibility [57]. In the current study, we will have the opportunity to explore for tumour-specific markers in blood and tissue that may be related to prediction and prognosis of outcome.

Lack of uniform research designs, poor quality control and large variation in reporting have hampered biomarker research and comparison of data in the past. This has invariably led to a number of promising but non-validated biomarkers in past studies. Currently, a large number of guidelines and recommendations are available to instruct, inform and impede better and more uniform reporting of results. However, the number of such

guidelines is increasing rapidly [58], with some suggesting there be too many guidelines to possibly comply to. However, we believe that a core set of important guides help set useful framework for reporting and help avoid huge deviation from recommended practice. Evaluation of compliance to such guidelines suggest that considerable deviation and lack of reporting core data still exist in biomarker research studies [59]. Thus, we would seek to adhere and comply with the recommendations addressed in the protocol and any other relevant recommendations, as issued by the EQUATOR network (http://www.equator-network.org).

Project status

The project is currently recruiting patients and laboratory work on the CTCs is ongoing, as well as laboratory work on MSI and EMAST in primary tumours. Hot spot panels of known CRC mutations with NGS technology is being prepared. A pilot, feasibility study to test for same-time comparison of patients' circulating immune-cells in peripheral blood and comparison to tumour-infiltrating cells in the cancer specimen is currently being conducted.

Future aspects

We envision several add-ons to be possible with increases resources and manpower in the project. For one, patient reported outcomes (PROs) is an increasing are of interest and would yield yet another dimension to the clinical-translational aspect of the project [60]. Also, the sampled biopsies will allow for a number of other experiments and analyses, such as exosomic DNA, microRNA and other emerging biomarkers. Further, other sampling techniques and specimens would be feasible in the future, such as sampling and investigating faeces for both genetic and epigenetic biomarkers [24], but also investigating the microbiome for its putative role in carcinogenesis but also possible influence on cancer biology [61, 62]. Last, but not least, we would pursue international collaboration for mutual validation of similar ongoing biomarker projects [63]. The prospective cohort results will seek collaboration for external validation studies but may also serve as an external validation cohort for other research groups interested in collaboration.

Abbreviations

CRC: colorectal cancer; CLM: colorectal liver metastasis; CTC: circulating tumour cell; NGS: next-generation sequencing; EMAST: elevated microsatellite alterations in selected tetranucleotides; MSI: microsatellite instability; IHC: immunohistochemistry; TNM: tumour-node-metastasis; CIMP: CpG-island methylator phenotype; CIN: chromosomal instability; SUH: Stavanger University Hospital; NGICG: Norwegian Gastrointestinal Cancer Group; FFPE: formalin-fixed paraffin-embedded; GWEF: glacial acetic acid, water, ethanol, formalin; ROC: receiving-operator characteristics; H&E: haematoxylin and eosin; PROs: patient reported outcomes.

Authors' contributions

KS is the primary investigator of ACROBATICC, conceived the idea and project design, applied for regional ethics committee and institutional review board approval and obtained funding. KS prepared the first manuscript draft, including tables/figures. All authors contributed to specific parts of the design of the study and co-wrote sections of the paper. MMW wrote section on DNA extraction and NGS, DL wrote section on IHC, ON wrote section on CTC and HRH wrote section on flow cytometry. MMW is PhD candidate on the EMAST project with financial support from the Folke Hermansen Cancer Fund and main responsible for DNA/RNA related techniques. DL is a trained pathologist and co-responsible for the tissue processing and quality assurance in the biorepository, as well as the histopathology and immunohistochemistry processing. ON obtained research funding for the circulating tumour cell (CTC) project and is the lead investigator for the CTC study and related analyses. JAS is group leader of the Surgical Research Group and is co-responsible for project flow, project supervision, project funding and patient care. HRH is responsible for flow cytometry and related experiments. All authors contributed to rounds of revisions and critical assessment of the paper content. All authors read and approved the final manuscript.

Author details

[1] Department of Gastrointestinal Surgery, Stavanger University Hospital, POB 8100, 4068 Stavanger, Norway. [2] Gastrointestinal Translational Research Unit, Laboratory for Molecular Biology, Stavanger University Hospital, Stavanger, Norway. [3] Department of Clinical Medicine, University of Bergen, Bergen, Norway. [4] Department of Pathology, Stavanger University Hospital, Stavanger, Norway. [5] Department of Haematology and Oncology, Stavanger University Hospital, Stavanger, Norway. [6] Centre of Organelle Research (CORE), University of Stavanger, Stavanger, Norway.

Acknowledgements

We would like to thank the patients and the staff at the Department of Gastrointestinal Surgery at SUH, particularly for securing tissue biobanking outside office hours. We would like to acknowledge the work of the research nurse in the project.

Competing interests

The authors declared that they have no competing interests.

Funding

The project is supported financially in parts by the Folke Hermansen Cancer Fund; the Mjaaland Cancer Fund; the University Fund at University of Stavanger; intramural funding from the Stavanger University Hospital; intramural funds from K1, University of Bergen; and, in part, by a grant from the Norwegian Research Council.

Received: 7 March 2016 Accepted: 21 June 2016
Published online: 29 June 2016

References

1. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2016. doi:10.1136/gutjnl-2015-310912.
2. Nedrebo BS, Soreide K, Eriksen MT, Dorum LM, Kvaloy JT, Soreide JA, et al. Survival effect of implementing national treatment strategies for curatively resected colonic and rectal cancer. Br J Surg. 2011;98(5):716–23. doi:10.1002/bjs.7426.
3. Veen T, Nedrebo BS, Stormark K, Soreide JA, Korner H, Soreide K. Qualitative and quantitative issues of lymph nodes as prognostic factor in colon cancer. Dig Surg. 2013;30(1):1–11. doi:10.1159/000349923.
4. Lea D, Haland S, Hagland HR, Soreide K. Accuracy of TNM staging in colorectal cancer: a review of current culprits, the modern role of morphology and stepping-stones for improvements in the molecular era. Scand J Gastroenterol. 2014;49(10):1153–63. doi:10.3109/00365521.2014.950692.
5. Soreide K, Nedrebo BS, Soreide JA, Slewa A, Korner H. Lymph node harvest in colon cancer: influence of microsatellite instability and proximal tumor location. World J Surg. 2009;33(12):2695–703. doi:10.1007/s00268-009-0255-4.

6. Berg M, Guriby M, Nordgard O, Nedrebo BS, Ahlquist TC, Smaaland R, et al. Influence of microsatellite instability and KRAS and BRAF mutations on lymph node harvest in stage I-III colon cancers. Mol Med. 2013;19:286–93. doi:10.2119/molmed.2013.00049.
7. Oltedal S, Gilje B, Korner H, Aasprong OG, Tjensvoll K, Heikkila R, et al. Detection of occult metastases in sentinel lymph nodes from colon cancer patients by K-ras mutation peptide nucleic acid clamp PCR. Ann Surg. 2010;251(6):1087–91. doi:10.1097/SLA.0b013e3181dae1bc.
8. Nordgard O, Oltedal S, Aasprong OG, Soreide JA, Soreide K, Tjensvoll K, et al. Prognostic relevance of occult metastases detected by cytokeratin 20 and mucin 2 mRNA levels in sentinel lymph nodes from colon cancer patients. Ann Surg Oncol. 2012;19(12):3719–26. doi:10.1245/s10434-012-2454-8.
9. Seeberg LT, Waage A, Brunborg C, Hugenschmidt H, Renolen A, Stav I, et al. Circulating tumor cells in patients with colorectal liver metastasis predict impaired survival. Ann Surg. 2015;261(1):164–71. doi:10.1097/sla.0000000000000580.
10. Huang X, Gao P, Song Y, Sun J, Chen X, Zhao J, et al. Relationship between circulating tumor cells and tumor response in colorectal cancer patients treated with chemotherapy: a meta-analysis. BMC Cancer. 2014;14:976. doi:10.1186/1471-2407-14-976.
11. Huang X, Gao P, Song Y, Sun J, Chen X, Zhao J, et al. Meta-analysis of the prognostic value of circulating tumor cells detected with the cell search system in colorectal cancer. BMC Cancer. 2015;15:202. doi:10.1186/s12885-015-1218-9.
12. Iinuma H, Watanabe T, Mimori K, Adachi M, Hayashi N, Tamura J, et al. Clinical significance of circulating tumor cells, including cancer stem-like cells, in peripheral blood for recurrence and prognosis in patients with Dukes' stage B and C colorectal cancer. J Clin Oncol. 2011;29(12):1547–55. doi:10.1200/jco.2010.30.5151.
13. Soreide K, Janssen EA, Soiland H, Korner H, Baak JP. Microsatellite instability in colorectal cancer. Br J Surg. 2006;93(4):395–406. doi:10.1002/bjs.5328.
14. Berg M, Hagland HR, Soreide K. Comparison of CpG island methylator phenotype (CIMP) frequency in colon cancer using different probe- and gene-specific scoring alternatives on recommended multi-gene panels. PLoS One. 2014;9(1):e86657. doi:10.1371/journal.pone.0086657.
15. Berg M, Nordgaard O, Korner H, Oltedal S, Smaaland R, Soreide JA, et al. Molecular subtypes in stage II-III colon cancer defined by genomic instability: early recurrence-risk associated with a high copy-number variation and loss of RUNX3 and CDKN2A. PLoS One. 2015;10(4):e0122391. doi:10.1371/journal.pone.0122391.
16. Soreide K, Soreide JA, Korner H. Prognostic role of carcinoembryonic antigen is influenced by microsatellite instability genotype and stage in locally advanced colorectal cancers. World J Surg. 2011;35(4):888–94. doi:10.1007/s00268-011-0979-9.
17. Soreide K, Slewa A, Stokkeland PJ, van Diermen B, Janssen EA, Soreide JA, et al. Microsatellite instability and DNA ploidy in colorectal cancer: potential implications for patients undergoing systematic surveillance after resection. Cancer. 2009;115(2):271–82. doi:10.1002/cncr.24024.
18. Linnekamp JF, Wang X, Medema JP, Vermeulen L. Colorectal cancer heterogeneity and targeted therapy: a case for molecular disease subtypes. Cancer Res. 2015;75(2):245–9. doi:10.1158/0008-5472.can-14-2240.
19. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350–6. doi:10.1038/nm.3967.
20. Betge J, Kerr G, Miersch T, Leible S, Erdmann G, Galata CL, et al. Amplicon sequencing of colorectal cancer: variant calling in frozen and formalin-fixed samples. PLoS One. 2015;10(5):e0127146. doi:10.1371/journal.pone.0127146.
21. McShane LM, Hayes DF. Publication of tumor marker research results: the necessity for complete and transparent reporting. J Clin Oncol. 2012;30(34):4223–32. doi:10.1200/jco.2012.42.6858.
22. Simeon-Dubach D, Burt AD, Hall PA. Quality really matters: the need to improve specimen quality in biomedical research. J Pathol. 2012;228(4):431–3. doi:10.1002/path.4117.
23. Carethers JM, Jung BH. Genetics and genetic biomarkers in sporadic colorectal cancer. Gastroenterology. 2015;149(5):1177–1190.e3. doi:10.1053/j.gastro.2015.06.047.
24. Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. Gastroenterology. 2015;149(5):1204–1225.e12. doi:10.1053/j.gastro.2015.07.011.
25. Watson MM, Berg M, Soreide K. Prevalence and implications of elevated microsatellite alterations at selected tetranucleotides in cancer. Br J Cancer. 2014;111(5):823–7. doi:10.1038/bjc.2014.167.
26. Koi M, Garcia M, Choi C, Kim HR, Koike J, Hemmi H, et al. Microsatellite alterations with allelic loss at 9p24.2 signify less-aggressive colorectal cancer metastasis. Gastroenterology. 2016. doi:10.1053/j.gastro.2015.12.032.
27. Watson MM, Lea D, Rewcastle E, Hagland HR, Søreide K. Elevated microsatellite alterations at selected tetranucleotides in early-stage colorectal cancers with and without high-frequency microsatellite instability: same, same but different? Cancer medicine. 2016. doi:10.1002/cam4.709.
28. Søreide K, Watson MM, Hagland HR. Deciphering the molecular code to colorectal liver metastasis biology through microsatellite alterations and allelic loss: the good, the bad, and the ugly. Gastroenterology. 2016;150(4):811–4. doi:10.1053/j.gastro.2016.02.060.
29. Soreide K, Sandvik OM, Soreide JA. KRAS mutation in patients undergoing hepatic resection for colorectal liver metastasis: a biomarker of cancer biology or a byproduct of patient selection? Cancer. 2014;120(24):3862–5. doi:10.1002/cncr.28979.
30. Berg M, Soreide K. EGFR and downstream genetic alterations in KRAS/BRAF and PI3K/AKT pathways in colorectal cancer: implications for targeted therapy. Discov Med. 2012;14(76):207–14.
31. Maby P, Tougeron D, Hamieh M, Mlecnik B, Kora H, Bindea G, et al. Correlation between density of CD8+ T-cell infiltrate in microsatellite unstable colorectal cancers and frameshift mutations: a rationale for personalized immunotherapy. Cancer Res. 2015;75(17):3446–55. doi:10.1158/0008-5472.can-14-3051.
32. Steinert G, Scholch S, Niemietz T, Iwata N, Garcia SA, Behrens B, et al. Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer. Cancer Res. 2014;74(6):1694–704. doi:10.1158/0008-5472.can-13-1885.
33. Pentheroudakis G, Raptou G, Kotoula V, Wirtz RM, Vrettou E, Karavasilis V, et al. Immune response gene expression in colorectal cancer carries distinct prognostic implications according to tissue, stage and site: a prospective retrospective translational study in the context of a hellenic cooperative oncology group randomised trial. PLoS One. 2015;10(5):e0124612. doi:10.1371/journal.pone.0124612.
34. Sandvik OM, Soreide K, Gudlaugsson E, Kvaloy JT, Soreide JA. Epidemiology and classification of gastroenteropancreatic neuroendocrine neoplasms using current coding criteria. Br J Surg. 2016;103(3):226–32. doi:10.1002/bjs.10034.
35. Sandvik OM, Soreide K, Kvaloy JT, Gudlaugsson E, Soreide JA. Epidemiology of gastrointestinal stromal tumours: single-institution experience and clinical presentation over three decades. Cancer Epidemiol. 2011;35(6):515–20. doi:10.1016/j.canep.2011.03.002.
36. Meling T, Harboe K, Soreide K. Incidence of traumatic long-bone fractures requiring in-hospital management: a prospective age- and gender-specific analysis of 4890 fractures. Injury. 2009;40(11):1212–9. doi:10.1016/j.injury.2009.06.003.
37. Reite A, Soreide K, Ellingsen CL, Kvaloy JT, Vetrhus M. Epidemiology of ruptured abdominal aortic aneurysms in a well-defined Norwegian population with trends in incidence, intervention rate, and mortality. J Vasc Surg. 2015;61(5):1168–74. doi:10.1016/j.jvs.2014.12.054.
38. Thorsen K, Soreide JA, Kvaloy JT, Glomsaker T, Soreide K. Epidemiology of perforated peptic ulcer: age- and gender-adjusted analysis of incidence and mortality. World J Gastroenterol. 2013;19(3):347–54. doi:10.3748/wjg.v19.i3.347.
39. Soreide K, Kruger AJ, Vardal AL, Ellingsen CL, Soreide E, Lossius HM. Epidemiology and contemporary patterns of trauma deaths: changing place, similar pace, older face. World J Surg. 2007;31(11):2092–103. doi:10.1007/s00268-007-9226-9.
40. Veen T, Stormark K, Nedrebo BS, Berg M, Soreide JA, Korner H, et al. Long-term follow-up and survivorship after completing systematic surveillance in stage I–III colorectal cancer: who is still at risk? J Gastrointest Cancer. 2015;46(3):259–66. doi:10.1007/s12029-015-9723-2.
41. Korner H, Soreide K, Stokkeland PJ, Soreide JA. Systematic follow-up after curative surgery for colorectal cancer in Norway: a population-based audit of effectiveness, costs, and compliance. J Gastrointest Surg. 2005;9(3):320–8. doi:10.1016/j.gassur.2004.09.023.
42. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in

epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008;61(4):344–9. doi:10.1016/j.jclinepi.2007.11.008.
43. Stormark K, Søreide K, Søreide JA, Kvaløy JT, Pfeffer F, Eriksen MT, et al. Nationwide implementation of laparoscopic surgery for colon cancer: short-term outcomes and long-term survival in a population-based cohort. Surg Endosc. 2016. doi:10.1007/s00464-016-4819-8.
44. Moore HM, Kelly AB, Jewell SD, McShane LM, Clark DP, Greenspan R, et al. Biospecimen reporting for improved study quality (BRISQ). J Proteome Res. 2011;10(8):3429–38. doi:10.1021/pr200021n.
45. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer. 2005;93(4):387–91. doi:10.1038/sj.bjc.6602678.
46. Koren R, Kyzer S, Paz A, Veltman V, Klein B, Gal R. Lymph node revealing solution: a new method for detection of minute axillary lymph nodes in breast cancer specimens. Am J Surg Pathol. 1997;21(11):1387–90.
47. Bao WG, Zhang X, Zhang JG, Zhou WJ, Bi TN, Wang JC, et al. Biobanking of fresh-frozen human colon tissues: impact of tissue ex vivo ischemia times and storage periods on RNA quality. Ann Surg Oncol. 2013;20(5):1737–44. doi:10.1245/s10434-012-2440-1.
48. Bray SE, Paulin FE, Fong SC, Baker L, Carey FA, Levison DA, et al. Gene expression in colorectal neoplasia: modifications induced by tissue ischaemic time and tissue handling protocol. Histopathology. 2010;56(2):240–50. doi:10.1111/j.1365-2559.2009.03470.x.
49. Alix-Panabieres C, Pantel K. Challenges in circulating tumour cell research. Nat Rev Cancer. 2014;14(9):623–31. doi:10.1038/nrc3820.
50. Tjensvoll K, Nordgard O, Smaaland R. Circulating tumor cells in pancreatic cancer patients: methods of detection and clinical implications. Int J Cancer. 2014;134(1):1–8. doi:10.1002/ijc.28134.
51. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the Immunoscore in the classification of malignant tumours. J Pathol. 2014;232(2):199–209. doi:10.1002/path.4287.
52. Dalerba P, Sahoo D, Paik S, Guo X, Yothers G, Song N, et al. CDX2 as a prognostic biomarker in stage II and stage III colon cancer. N Engl J Med. 2016;374(3):211–22. doi:10.1056/NEJMoa1506597.
53. Tseng-Rogenski SS, Hamaya Y, Choi DY, Carethers JM. Interleukin 6 alters localization of hMSH3, leading to DNA mismatch repair defects in colorectal cancer cells. Gastroenterology. 2015;148(3):579–89. doi:10.1053/j.gastro.2014.11.027.
54. Campregher C, Schmid G, Ferk F, Knasmuller S, Khare V, Kortum B, et al. MSH3-deficiency initiates EMAST without oncogenic transformation of human colon epithelial cells. PLoS One. 2012;7(11):e50541. doi:10.1371/journal.pone.0050541.
55. Soreide K. Receiver-operating characteristic curve analysis in diagnostic, prognostic and predictive biomarker research. J Clin Pathol. 2009;62(1):1–5. doi:10.1136/jcp.2008.061010.
56. Berg M, Soreide K. Genetic and epigenetic traits as biomarkers in colorectal cancer. Int J Mol Sci. 2011;12(12):9426–39. doi:10.3390/ijms12129426.
57. Yoruker EE, Holdenrieder S, Gezer U. Blood-based biomarkers for diagnosis, prognosis and treatment of colorectal cancer. Clin Chim Acta. 2016;455:26–32. doi:10.1016/j.cca.2016.01.016.
58. Vandenbroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK… and CONSORT: for whom does the guideline toll? J Clin Epidemiol. 2009;62(6):594–6. doi:10.1016/j.jclinepi.2008.12.003.
59. Mallett S, Timmer A, Sauerbrei W, Altman DG. Reporting of prognostic studies of tumour markers: a review of published articles in relation to REMARK guidelines. Br J Cancer. 2010;102(1):173–80. doi:10.1038/sj.bjc.6605462.
60. Soreide K, Soreide AH. Using patient-reported outcome measures for improved decision-making in patients with gastrointestinal cancer—the last clinical frontier in surgical oncology? Front Oncol. 2013;3:157. doi:10.3389/fonc.2013.00157.
61. Hagland HR, Soreide K. Cellular metabolism in colorectal carcinogenesis: influence of lifestyle, gut microbiome and metabolic pathways. Cancer Lett. 2015;356(2 Pt A):273–80. doi:10.1016/j.canlet.2014.02.026.
62. Mima K, Sukawa Y, Nishihara R, Qian ZR, Yamauchi M, Inamura K, et al. Fusobacterium nucleatum and T Cells in Colorectal Carcinoma. JAMA Oncol. 2015;1(5):653–61. doi:10.1001/jamaoncol.2015.1377.
63. Soreide K, Alderson D, Bergenfelz A, Beynon J, Connor S, Deckelbaum DL, et al. Strategies to improve clinical research in surgery through international collaboration. Lancet. 2013;382(9898):1140–51. doi:10.1016/s0140-6736(13)61455-5.

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230849965 (print)
9788230859704 (PDF)