

Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer

Florian Posch^{a,†}, Karina Silina^{b,†}, Sebastian Leibl^c, Axel Mündlein^d, Holger Moch^c, Alexander Siebenhüner^e, Panagiotis Samaras^e, Jakob Riedl^a, Michael Stotz^a, Joanna Szkandera^a, Herbert Stöger^a, Martin Pichler^{a,g}, Roger Stupp^e, Maries van den Broek^b, Peter Schraml^c, Armin Gerger^a, Ulf Petrausch^{e,f,†}, and Thomas Winder^{d,e,†}

^aDivision of Oncology, Department of Internal Medicine, Medical University of Graz, Auenbruggerplatz 15, Graz, Austria; ^bTumor Immunology Research Unit, Institute of Experimental Immunology, University of Zurich, Winterthurerstrasse 190, Zurich, Switzerland; ^cDepartment of Pathology and Molecular Pathology, University Hospital Zurich, Schmelzbergstrasse 12, Zürich, Switzerland; ^dVorarlberg Institute for Vascular Investigation and Treatment (VIVIT), Carinagasse 47, Feldkirch, Austria; ^eDepartment of Oncology, University Hospital Zurich, Rämistrasse 100, Zürich, Switzerland; ^fSwiss Tumor Immunology Institute, OnkoZentrum Zürich, Seestrasse 259, Zürich, Switzerland; ^gDepartment of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

ABSTRACT

Tertiary lymphoid structures (TLS) are associated with favorable outcome in non-metastatic colorectal carcinoma (nmCRC), but the dynamics of TLS maturation and its association with effective anti-tumor immune surveillance in nmCRC are unclear. Here, we hypothesized that not only the number of TLS but also their composition harbors information on recurrence risk in nmCRC. In a comprehensive molecular, tissue, laboratory, and clinical analysis of 109 patients with stage II/III nmCRC, we assessed TLS numbers and degree of maturation in surgical specimens by multi-parameter immunofluorescence of follicular dendritic cell (FDC) and germinal center (GC) markers. TLS formed in most tumors and were significantly more prevalent in highly-microsatellite-instable (MSI-H) and/or BRAF-mutant nmCRC. We could distinguish three sequential TLS maturation stages which were characterized by increasing prevalence of FDCs and mature B-cells: [1] Early TLS, composed of dense lymphocytic aggregates without FDCs, [2] Primary follicle-like TLS, having FDCs but no GC reaction, and [3] Secondary follicle-like TLS, having an active GC reaction. A simple integrated TLS immunoscore reflecting these parameters identified a large subgroup of nmCRC patients with a very low risk of recurrence independently of clinical co-variables such as ECOG performance status, age, stage, and adjuvant chemotherapy. We conclude that (1) mismatch repair and BRAF mutation status are associated with the formation of TLS in nmCRC, (2) TLS formation in nmCRC follows sequential maturation steps, culminating in germinal center formation, and (3) this maturation process harbors important prognostic information on the risk of disease recurrence.

ARTICLE HISTORY

Received 18 July 2017
Revised 7 September 2017
Accepted 8 September 2017

KEYWORDS

colorectal cancer; Crohn's-like reaction; germinal center; immunoscore; recurrence; risk factor; structural equation model; tertiary lymphoid structures

Introduction


Colorectal Cancer (CRC) is the third most frequent malignancy world-wide in both men and women.¹ Currently, an estimated 1.4 million new colorectal cancer cases and 693,900 deaths are annually recorded.² Around 30–40% of patients with non-metastatic stage II and III CRC (nmCRC) develop recurrence after surgical removal of the primary tumor with curative intent and these recurrences are the major determinant of CRC mortality.³ Therefore, improved prognostication of disease recurrence in nmCRC is important in order to select patients with the highest risk of recurrence for adjuvant chemotherapy while sparing low-recurrence-risk patients from unnecessary and oftentimes toxic overtreatment.⁴ The immune system is a critical determinant of tumor control and disease recurrence. Several authors have proposed that immunologic biomarkers harbor prognostic

information on recurrence risk beyond traditional parameters such as staging.^{5,6} Tertiary lymphoid structures (TLS) are ectopic lymphoid organs that develop at sites of infection or chronic inflammation including cancer.⁷ In analogy to lymph nodes, TLS are thought to provide important lymphocytic functional environments for both cellular and humoral immunity.⁸ In autoimmunity, TLS harboring germinal centers (GC) contribute to disease severity by producing auto-reactive B cells and facilitating lymphocyte infiltration via high endothelial venules.⁷ In various cancer types, TLS correlate with improved survival including localized and metastatic CRC,^{9,10} and might represent important sites of lymphocyte recruitment into tumors as well as adaptive immune response activation against malignant cells.^{6,8,11}

The formation of TLS is a complex process and is orchestrated by several lympho-organogenic chemokines in response to

CONTACT Armin Gerger, MD, MBA, Associate Professor of Medicine  armin.gerger@medunigraz.at  Division of Oncology, Department of Internal Medicine, Medical University of Graz, Auenbruggerplatz 15, 8036 Graz.

[†]These authors contributed equally.

 Supplemental data for this article can be accessed on the [publisher's website](#).

© 2018 Florian Posch, Karina Silina, Sebastian Leibl, Axel Mündlein, Holger Moch, Alexander Siebenhüner, Panagiotis Samaras, Jakob Riedl, Michael Stotz, Joanna Szkandera, Herbert Stöger, Martin Pichler, Roger Stupp, Maries van den Broek, Peter Schraml, Armin Gerger, Ulf Petrausch and Thomas Winder. Published with license by Taylor & Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

inflammatory stimuli.¹² The availability of cognate antigens is crucial for the generation of a GC reaction, however, it is not known whether antigen load has an impact also on the induction and/or maintenance of TLS. Mismatch-repair-deficiency in CRC causes microsatellite instability (MSI) and increases the number of neo-antigens comparing to microsatellite-stable (MSS) tumors. As a consequence patients with MSI mCRC respond to treatment with checkpoint inhibitors.¹³ The development of CRC-associated TLS tends to be increased in MSI tumors.¹⁴ Furthermore, *BRAF* and *KRAS* are frequently mutated in CRC and such alterations can be recognized by patient's T cells.^{15,16} Nevertheless, TLS predict improved survival in untreated CRC patients independently of patient characteristics, MSI and other molecular variables.⁹ Silina et al. demonstrated that the development of TLS in lung squamous cell carcinoma (LSCC) follows sequential maturation stages leading to the formation of GCs. The development of GCs was necessary for the prognostic potential of tumor-associated TLS and was the most significant independent prognostic marker in untreated LSCC patients.¹⁷ It is not known, however, whether TLS maturation has the same relevance and follows the same process in other tumor types. Furthermore, it is unknown whether the increased antigen load in MSI and oncogene-mutated tumors influences development and maturation of TLS. Finally, the relevance of TLS maturation as prognosticator for relapse has not been studied.

To address these questions, we explored TLS maturation stages and their association with patient characteristics, mutation status, and blood based biomarkers of systemic inflammatory response in a cohort of patients with stage II and III colorectal cancer. We examined the prognostic relevance of TLS maturation and developed a TLS immunoscore that integrates TLS numbers and maturation parameters. We found that the TLS immunoscore predicts the risk of recurrence in nmCRC, thus enabling the identification of patients with a high or low anticipated benefit from adjuvant chemotherapy.

Results

Analysis at baseline and crude event rates

One hundred nine patients with non-metastatic CRC were included in this retrospective cohort study at the time of surgery (Table 1). Ten (24%) of the 42 patients with UICC stage II disease and 62 (90%) of the 69 patients with UICC stage III disease received adjuvant chemotherapy ($p < 0.0001$). A highly- unstable microsatellite status (MSI-H) was present in 10 patients (9%), and the *BRAF* V600E mutation was detected in 10 patients (9%). A strong association between *BRAF* mutation and microsatellite status was observed, with 5 (50%) out of the 10 patients with the *BRAF* mutation also having an MSI-H status and vice versa ($\chi^2 p < 0.0001$).

During the 3-year follow-up period, we observed 17 recurrences (i.e. the primary endpoint, comprising 3 local recurrences and 14 occurrences of distant metastasis) and 10 patients died. Six of these deaths were adjudicated to tumor and 4 deaths to other causes. Sixty-eight (77%) of the patients that did not develop recurrence or die were followed-up for the full 3-year period, the remaining 20 patients had a median follow-up of 1.4 years. In competing risk analysis, the cumulative 1-year, 2-year, and 3-year incidences of recurrence were 8.4%

Table 1. Baseline characteristics of the study population (n = 109). Summary estimates represent medians [25th-75th percentile] for continuous variables and absolute counts (%) for categorical variables. Abbreviations: BMI – Body mass index, ECOG – Eastern Cooperative Oncology Group, L1 – Lymphatic invasion, V1 – Vascular invasion, MSI – Microsatellite instability status, UICC – Union Internationale pour le lutte contre le cancer TLS/mm – TLS count, ECOG – Eastern Cooperative Oncology Group, % E-TLS – Proportion of Early tertiary lymphoid structure (TLS), Q1 – 25th percentile of the variable's distribution (i.e. quartile 1 (Q1)), % PFL-TLS – Proportion of primary follicle-like TLS, % SFL-TLS – Proportion of secondary follicle-like TLS, GC – Germinal center, EB – Empirical Bayes, ITIS – Integrated TLS ImmunoScore, CART – Classification and Regression Tree, G/L – 109/liter.

Variable	N (%missing)	Summary estimate
Demographic parameters		
Female sex	109 (0.0%)	52 (47.7%)
BMI	103 (5.5%)	25.9 [23.0–28.2]
Age at study entry	109 (0.0%)	64.8 [57.3–74.9]
Family history of cancer	77 (29.4%)	5 (6.5%)
Diabetes	108 (0.9%)	11 (10.2%)
Smoking	80 (26.6%)	13 (16.3%)
Karnofsky index	106 (2.7%)	N/A
—100%	N/A	72 (67.9%)
—90%	N/A	23 (21.7%)
—80%	N/A	9 (8.5%)
—70%	N/A	1 (0.9%)
—60%	N/A	1 (0.9%)
ECOG performance status > 0	106 (2.7%)	34 (32.1%)
Tumor characteristics		
Tumor localization	109 (0.0%)	N/A
—Cecum, Appendix, Ascending Colon, Right Flexure, Right Transverse Colon	N/A	36 (33.0%)
—Left Transverse Colon, Left Flexure, Descending Colon, Sigma	N/A	28 (25.7%)
—Rectum	N/A	38 (34.9%)
—Two or more synchronous CRCs	N/A	7 (6.4%)
Right sided tumor	109 (0.0%)	36 (33.0%)
L1	109 (0.0%)	19 (17.4%)
V1	109 (0.0%)	12 (11.0%)
MSI high	108 (0.9%)	10 (9.3%)
B-RAF mutant	107 (1.8%)	10 (9.4%)
Tumor grade	109 (0.0%)	N/A
—G1	N/A	6 (5.5%)
—G2	N/A	52 (47.7%)
—G3	N/A	51 (46.8%)
Tumor stage	109 (0.0%)	N/A
—UICC Stage II	N/A	41 (37.6%)
—UICC Stage III	N/A	68 (62.4%)
Adjuvant chemotherapy	109 (0.0%)	71 (65.1%)
TLS parameters		
TLS/mm	109 (0.0%)	0.5 [0.2–0.9]
% E-TLS	106 (2.7%)	56 [40–78]
% PFL-TLS	106 (2.7%)	20 [6–36]
% SFL-TLS	106 (2.7%)	15 [0–32]
≥1 GC-harboring TLS	109 (0.0%)	68 (62%)
Immune_context EB prediction	109 (0.0%)	−0.06 [−0.93–0.61]
ITIS	106 (2.7%)	−0.1 [−15.4–11.1]
ITIS ≤ CART cut-off	106 (2.7%)	33 (31.1%)
Laboratory parameters		
Absolute leukocyte count (G/L)	109 (0.0%)	7.1 [5.9–8.5]
Absolute platelet count (G/L)	109 (0.0%)	290 [242–360]
Absolute neutrophil count (G/L)	105 (3.7%)	4.9 [3.9–6.0]
Absolute lymphocyte count (G/L)	105 (3.7%)	1.3 [1.1–1.7]
Neutrophil-lymphocyte-ratio (units)	105 (3.7%)	3.4 [2.6–4.8]

(95%CI: 4.1–14.6), 15.6% (9.4–23.3), and 16.7% (10.2–24.6), respectively (Supplementary Figure 1). The rate of recurrence peaked around 1 year after surgery and then rapidly declined (Supplementary Figure 2).

TLS form in most CRCs

TLS formed in tumor periphery (Fig. 1A) in 97% of patients (Fig. 1B). The number of TLS per millimeter of tumor invasive

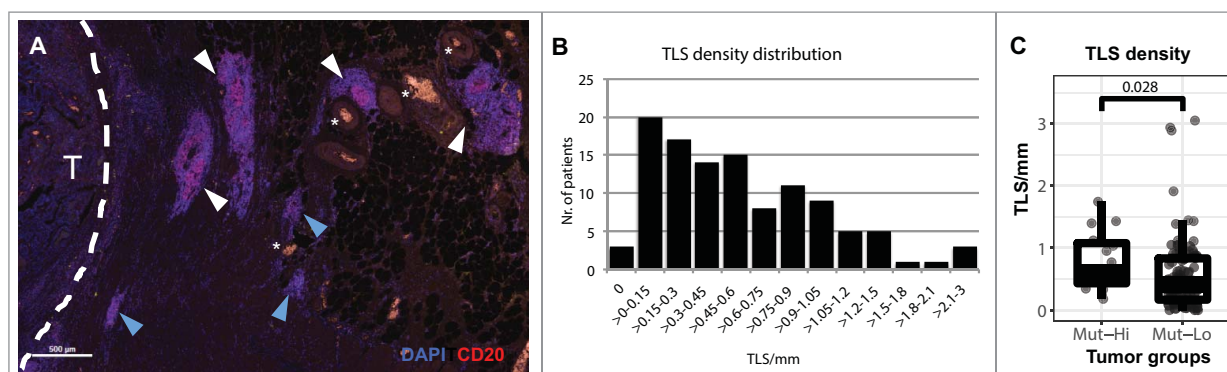


Figure 1. Tertiary lymphoid structures in stage II and III colorectal cancer. (A) The presence of TLS was analyzed in 109 CRC tissues by immunofluorescence and identified as dense B cell (CD20) aggregates. TLS near the invasive margin (white dashed line) of a CRC (T) and blood vessels (white stars). TLS with a central GC morphology (white arrowheads) or without (blue arrowheads) were counted. (B) The density of TLS in each patient was determined as the number of TLS per millimeter of tumor invasive front. (C) Comparison of TLS density in CRC patients with known MSI and *BRAF* mutation status. MSI-H and/or *BRAF* mutated CRCs (Mut-Hi, n = 15) were compared versus patients with wild type *BRAF* and MSS status (Mut-L0, n = 94) by Wilcoxon rank-sum test.

front, here referred to as TLS density, ranged from 0 to 3 (median: 0.5 TLS/mm [25th-75th percentile: 0.2–0.9]). To determine if genetic alterations bearing immunogenic potential affect TLS development in CRC, we compared TLS density in tumors with high microsatellite instability (MSI-H) and/or *BRAF* mutations versus tumors not harboring these features. We observed a significantly higher TLS density in tumors with these molecular features (median: 0.61 vs. 0.45, rank-sum $p = 0.03$, Fig. 1C). Further, TLS density positively correlated with the absolute peripheral lymphocyte count (Spearman's $\rho = 0.27$, $p = 0.006$), and negatively correlated with the blood-based neutrophil-lymphocyte-ratio (NLR), a marker of systemic chronic inflammation ($\rho = -0.26$, $p = 0.007$). We did not observe differences in TLS density with respect to tumor grade and stage, performance status, or age (Supplementary Table 1).

Characteristics of TLS maturation in CRC

We detected three distinct phenotypes of TLS representing different stages of maturation in CRC: (1) early TLS (E-TLS, dense lymphocytic aggregates without differentiated FDCs), (2) primary follicle-like TLS (PFL-TLS, B cell clusters with FDC network but without germinal centers (GCs)), and (3) secondary follicle-like TLS (SFL-TLS, with GCs) (Fig. 2A). We calculated the proportion of each TLS maturation stage in each patient and correlated with different clinical parameters. The proportion of mature SFL-TLS correlated with the total TLS density (Spearman $\rho = 0.41$, $p < 0.0001$). NLR negatively correlated with the SFL-TLS proportion ($\rho = -0.31$, $p = 0.002$) while positively correlating with the proportion of the more immature E-TLS ($\rho = 0.20$, $p = 0.04$). We studied in detail the differences of TLS maturation in patients with different TLS densities and found that TLS maturation was arrested at the E-TLS stage in tumors with low TLS density (Fig. 2B, left). This in turn resulted in significantly decreased proportions of the PFL and SFL-TLS stages (Fig. 2B, middle and right).

Because the total TLS density was associated with tumor mutation status (Fig. 1C), we compared the density of each maturation stage in these patient groups. Here, patients with

MSI-H and/or *BRAF* mutated tumors showed higher numbers of all TLS stages, especially SFL-TLS (Fig. 2C). Associations were similar when considering each molecular characteristic separately (Supplemental Figure 3).

TLS characteristics predict CRC recurrence

We used all TLS variables, namely TLS density and the proportions of E-TLS, PFL-TLS, and SFL-TLS, to generate distinct patient subgroups in order to compare the cumulative incidence of recurrence. In the absence of validated cut-offs, these variables were dichotomized into binary variables using an empirical cut-off at the 25th percentile of their distribution (Q1), with values \leq this cut-off defining the “low” group. For E-TLS, we used a cut-off at the 75th percentile (Q3) to define a “high” group, because high rather than low proportion of E-TLS indicates to impaired TLS maturation. For all time-to-event analyses, we used proportions rather than crude numbers of E-TLS, PFL-TLS, and SFL-TLS, because we considered proportions more reflective of the relative TLS composition within patients.

In univariable competing risk analysis, TLS/mm, E-TLS, and PFL-TLS showed a strong trend of association with recurrence risk (Fig. 3A–C, Table 2), while the cumulative risk of recurrence was significantly higher in patients with low SFL-TLS. In detail, CRC recurrence risk was 31.0% and 9.5% in patients with low and high SFL-TLS proportion (Gray's test $p = 0.006$, Fig. 3D), respectively. Similar data were obtained in univariable competing risk regression (Table 3). Here, the relative risk of recurrence was 3.7-fold higher in patients with low SFL-TLS (SHR = 3.73, 95%CI: 1.39–10.00, $p = 0.009$), respectively.

After multivariable adjustment for age, ECOG performance status, stage and adjuvant chemotherapy, low SFL-TLS was associated with a 4-fold higher risk of recurrence (adjusted SHR for low SFL-TLS = 3.99, 95%CI: 1.30–12.20, $p = 0.015$, multivariable model #4 in Table 4). In multivariable analysis, all other TLS parameters were also strongly associated with recurrence risk (multivariable models #1–#3 in Table 4). Among the included covariates, both low ECOG performance status and high age were independently associated with recurrence. Here, an ECOG performance status > 0 and age \geq

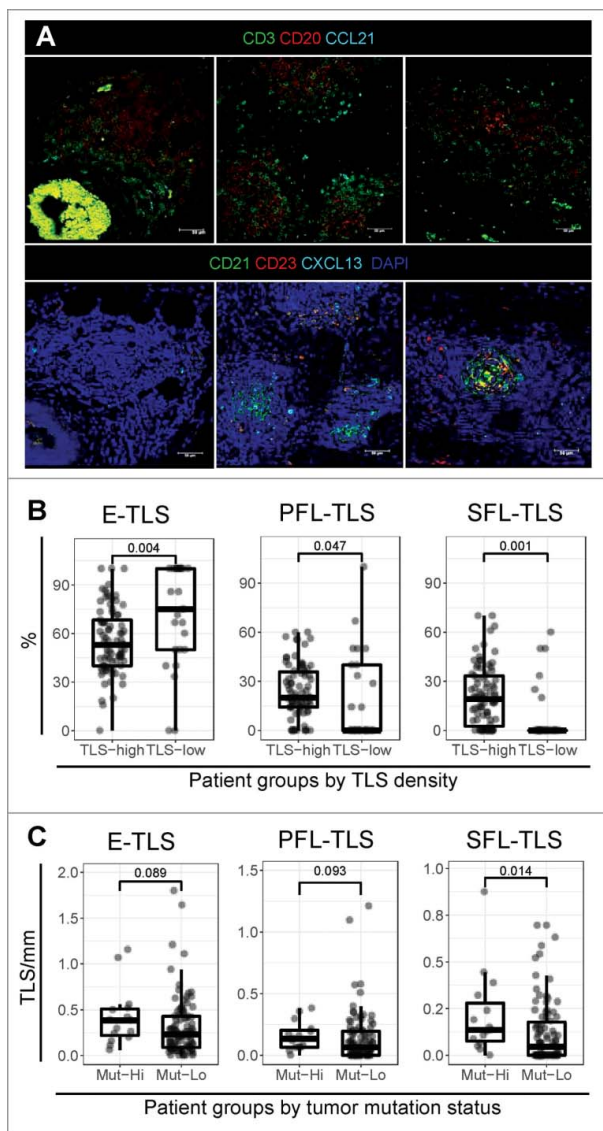


Figure 2. Tertiary lymphoid structure maturation in stage II and III colorectal cancer. (A) The composition of TLS was analyzed by multi-parameter immunofluorescence using serial tissue sections of CRC. Two sets of antibodies were combined to visualize (1) the spatial organization of B cells (CD20), T cells (CD3), and CCL21 (top row), and (2) TLS maturation by the presence of FDCs (CD21), germinal center (GC) B cells (CD23), and CXCL13 (bottom row). (B) Different maturation stages were assessed by multiparameter immunofluorescence in CRC tissues. TLS in each maturation stage (E-TLS – dense lymphocytic clusters: CD21⁺CD23⁻, PFL-TLS – clusters with an FDC network: CD21⁺CD23⁻, SFL-TLS – clusters with active GC reaction: CD21⁺CD23⁺) were counted for each patient and expressed as the proportion of the total TLS count. The proportions of the three maturation stages were compared in patients with high ($n = 84$) and low ($n = 25$) TLS density (cut-off at the 25th percentile of TLS density distribution (i.e. Q1)) by a Wilcoxon rank-sum test. (C) The density of each TLS maturation stage was compared for patients with mutated *BRAF* and/or MSI-H status (Mut-Hi, $n = 15$) versus patients with wild type *BRAF* and MSS status (Mut-Lo, $n = 94$) by a two-tailed Mann-Whitney U test.

75 years were associated with a 5.8-fold higher and 4.9-fold higher risk of recurrence, respectively (multivariable model #4 in Table 4). TLS parameters were also independently prognostic after adjusting for the MSI and *BRAF* mutation status (Supplementary Table 2).

Our analysis treats E-TLS, SFL-TLS, PFL-TLS as proportions reflecting TLS maturity. To investigate whether not only these relative measures of TLS maturation but also absolute TLS maturity parameters harbor prognostic information on

CRC recurrence, we investigated the association between the presence of at least one TLS harboring an active GC reaction and CRC recurrence risk. In detail, the absolute 3-year risk of CRC recurrence was 9.5% in the 68 patients (62%) who had at least one GC-harboring TLS, as compared to 28.5% in the 41 patients (38%) who did not have any GC-harboring TLS (Fig. 4; corresponding to a 70% lower relative risk of recurrence: SHR = 0.30 (95%CI: 0.11–0.79, $p = 0.015$; Table 3)). This association also fully prevailed upon multivariable adjustment for age, performance status, adjuvant chemotherapy, and tumor stage (Adjusted SHR = 0.28, 95%CI: 0.09–0.84, $p = 0.024$; multivariable model #5 in Table 4).

An Integrated TLS ImmunoScore (ITIS) predicts the recurrence risk in CRC

Based on these results, we hypothesized that the integration of TLS density and TLS maturation would provide a higher predictive power for CRC recurrence risk than each TLS parameter separately. We used structural equation modeling and developed an integrated TLS immunoScore (ITIS) to analyze TLS density and relative abundance of maturation stages as a joint parameter. In detail, we assumed a latent variable (immune_context), which predicts for 3-year time-to-recurrence (time2-mets), and is informed by four continuous measures, namely total TLS density (TLSmm), and the proportions of E-TLS (E_TLS), PFL-TLS (PFL_TLS), and SFL-TLS (SFL_TLS, Fig. 5). In this model, all four measures significantly loaded onto the latent variable (TLSmm $p = 0.01$, all others $p < 0.0001$), with E-TLS being associated with a decrease in the latent variable and all other variables being associated with an increase in the latent variable (Supplementary Table 3). Further, the latent variable (distribution reported in Table 1) significantly predicted for a lower risk of recurrence (Hazard ratio per 1 standard deviation increase = 0.58, 95%CI: 0.38–0.89, $p = 0.013$).

We then applied the path coefficients of the model's measurement component (Supplementary Table 3) to construct the ITIS (distribution reported in Table 1 and Supplementary Fig. 4), defined as the sum of the products of the path coefficients and their observations in our cohort, standardized by their intercept (Equation 1):

$$ITIS = \frac{0.11 * TLSmm}{0.63} - \frac{18.68 * E_TLS}{56.12} + \frac{9.76 * PFL_TLS}{24.33} + \frac{8.92 * SFL_TLS}{19.56}$$

Canceling out the denominators, the ITIS reduces to Equation 2:

$$ITIS = 0.18 * TLSmm - 0.33 * E_TLS + 0.40 * PFL_TLS + 0.46 * SFL_TLS$$

This score aligned well with the latent “immune_context” variable (Spearman's rho = 0.996, $p < 0.0001$, Supplementary Figure 5).

The cumulative 3-year risks of recurrence were 31.7% (95%CI: 17.2–47.3), 9.8% (2.5–23.3), and 9.4% (2.4–22.4) in

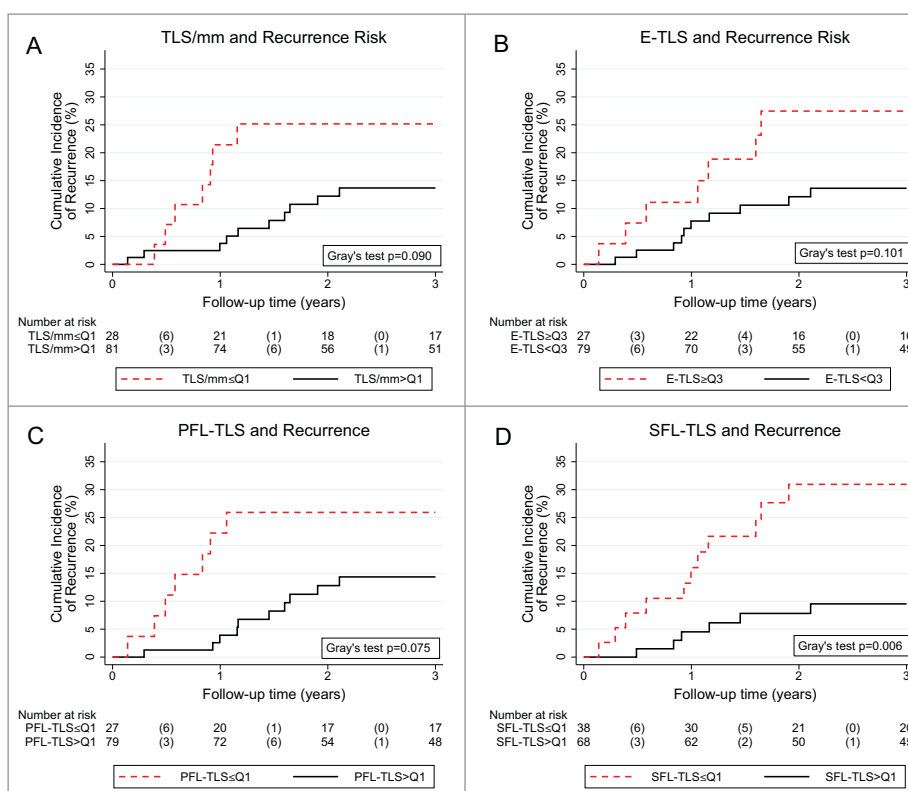


Figure 3. Tertiary lymphoid structure parameters and 3-year risk of colorectal cancer recurrence. Patients with low TLS counts (A), high proportions of E-TLS (B), and low proportions of PFL-TLS (C) showed a tendency towards numerically higher risks of recurrence. Patients with low SFL-TLS proportion (D) experienced a significantly higher risk of recurrence. Absolute risk estimates corresponding to these curves are reported in Table 2. Recurrence risks were estimated with competing risk analysis, treating death-from-any-cause as the competing event of interest. In the risk table, numbers in brackets represent the number of recurrences in the respective time interval. Abbreviations: Q1 – Cut-off at the 25th percentile of the variables' distribution (i.e. quartile 1), Q3 – Cut-off at the 75th percentile of the variable's distribution (i.e. quartile 3).

patients in the 1st, 2nd, and 3rd tertile of the ITIS distribution, respectively (Gray's test $p = 0.01$, Fig. 6A). These tertile cut-offs were chosen empirically in the absence of established cut-offs. In competing risk regression, a higher ITIS (as a continuous variable) was significantly associated with a lower risk of the primary endpoint (SHR for 10 units increase = 0.73, 95%CI: 0.55–0.95, $p = 0.02$), and this association prevailed in multivariable analysis adjusting for age, ECOG performance status, stage, and adjuvant chemotherapy (multivariable model #6 in Table 4).

Further, we did not observe an interaction between ITIS and stage (UICC II vs. UICC III, p for interaction = 0.74), suggesting that the ITIS may be comparably relevant for both stages. To further investigate this, we performed a subgroup analysis by stage where we Z-standardized the ITIS within stage. ITIS showed a similar relative association with recurrence risk in patients with stage II disease (SHR per 1 SD increase = 0.48, 95%CI: 0.19–1.22, $p = 0.12$) and stage III disease (SHR per 1

SD increase = 0.60, 95%CI: 0.33–1.07, $p = 0.08$). Although the subgroup analyses were not significant due to lower numbers of cases within stages, the comparable magnitude of the regression coefficients suggest that ITIS may be similarly relevant in stage II and stage III patients. Moreover, we also did not observe an interaction with adjuvant chemotherapy ($p = 0.97$).

Potential time-dependent associations of the ITIS were investigated by fitting interactions between the ITIS and linear-follow-up time, which suggested that the protective association of a high TLS immunoscore slightly weakened with time (p for interaction = 0.009). In detail, the SHRs for 10 units increase in ITIS were estimated at 0.50 (95%CI: 0.28–0.88, $p = 0.02$), 0.62 (95%CI: 0.44–0.89, $p = 0.01$), 0.69 (95%CI: 0.50–0.96, $p = 0.05$), and 0.70 (95%CI: 0.53–0.93, $p = 0.01$) after 6, 12, 18, and 24 months of follow-up, respectively.

Finally, a classification and regression tree (CART) analysis was performed to investigate the strongest recurrence risk

Table 2. TLS variables and 3-year recurrence risk. Risks were estimated with competing risk cumulative incidence estimators, treating death-from-any-cause as the competing event of interest. * \geq and < Q3 for % E-TLS. Abbreviations: Q1 – 25th percentile of the variable's distribution (i.e. quartile 1 (Q1)), TLS/mm – TLS count, % E-TLS – Proportion of Early tertiary lymphoid structure (TLS), % PFL-TLS – Proportion of primary follicle-like TLS, % SFL-TLS – Proportion of secondary follicle-like TLS.

Variable	3-year recurrence risk in patients \leq Q1 (95%CI)*	3-year recurrence risk in patients $>$ Q1 (95%CI)*	Gray's test p-value
TLS/mm	25.2% (11.1–42.0)	13.7% (7.0–22.6)	0.090
% E-TLS*	27.5% (12.1–45.4)	13.6% (7.0–22.5)	0.101
% PFL-TLS	25.9% (11.5–43.1)	14.4% (7.3–23.7)	0.075
% SFL-TLS	31.0% (16.7–46.4)	9.5% (3.9–18.3)	0.006

Table 3. Univariable competing risk regression models of 3-year recurrence risk. Results were estimated with univariable Fine & Gray competing risk regression models, treating death-from-any-cause as the competing event of interest. Abbreviations: SHR – Subdistribution hazard ratio, 95%CI – 95% confidence interval, BMI – Body mass index, ECOG – Eastern Cooperative Oncology Group, L1 – Lymphatic invasion, V1 – Vascular invasion, MSI – Microsatellite instability status, UICC – Union Internationale pour le lutte contre le cancer TLS/mm – TLS count, ECOG – Eastern Cooperative Oncology Group, % E-TLS – Proportion of Early tertiary lymphoid structure (TLS), Q1 – 25th percentile of the variable's distribution (i.e. quartile 1 (Q1)), % PFL-TLS – Proportion of primary follicle-like TLS, % SFL-TLS – Proportion of secondary follicle-like TLS, GC – Germinal center, EB – Empirical Bayes, ITIS – Integrated TLS ImmunoScore, CART – Classification and Regression Tree, G/L – 10⁹/liter.

Variable	SHR	95%CI	p
Demographic variables			
Female sex	0.58	0.22–1.57	0.283
BMI (per 5 kg/m ² increase)	0.57	0.29–1.12	0.105
Age at study entry ≥ 75 years	3.97	1.54–10.21	0.004
Smoking	0.92	0.20–4.24	0.916
ECOG performance status > 0	7.85	2.54–24.28	<0.0001
Tumor variables			
Right sided tumor	1.14	0.42–3.08	0.794
L1	1.38	0.47–4.01	0.554
V1	1.08	0.23–5.13	0.925
MSI high	0.54	0.07–3.90	0.538
B-RAF V600E mutant	0.54	0.08–3.67	0.532
Tumor grade G3	2.14	0.80–5.76	0.130
UICC tumor stage III	2.02	0.65–6.25	0.222
Adjuvant chemotherapy	1.23	0.42–3.59	0.710
Immune contexture variables			
TLS/mm ≤ Q1	2.24	0.84–5.92	0.106
% E-TLS ≥ Q3	2.21	0.85–5.77	0.104
% PFL-TLS ≤ Q1	2.24	0.83–6.06	0.111
% SFL-TLS ≤ Q1	3.73	1.39–10.00	0.009
≥1GC-harboring TLS	0.30	0.11–0.79	0.015
Immune_context EB prediction(per 1 unit increase)	0.54	0.34–0.86	0.010
ITIS (per 10 units increase)	0.73	0.55–0.95	0.020
ITIS ≤ CART cut-off	4.83	1.82–12.85	0.002
Laboratory parameters			
Absolute leukocyte count (per 1 G/L increase)	0.91	0.78–1.07	0.264
Absolute platelet count (per 50 G/L increase)	0.79	0.62–1.00	0.052
Absolute neutrophil count (per 1 G/L increase)	0.94	0.81–1.08	0.368
Absolute lymphocyte count (per 1 G/L increase)	0.41	0.19–0.86	0.019
Neutrophil-lymphocyte-ratio (per 5 units increase)	1.00	0.73–1.35	0.975

predictors in our dataset and explore optimal cut-offs for defining patients with a low ITIS. The CART algorithm selected ECOG status and ITIS as the two strongest predictors of recurrence risk, and suggested a cut-off for low ITIS at -12 ITIS points (Supplementary Figure 6). In multivariable competing risk regression, patients with a low ITIS had an eight-fold higher risk of recurrence than patients above this cut-off, respectively (multivariable model #7 in Table 4). The corresponding 3-year recurrence risk was 34.7% (95%CI: 18.9–51.1) for patients with low ITIS, and 9.1% (95%CI: 3.7–17.6) for patients with high ITIS, respectively (Gray's test $p = 0.02$, Fig. 6B). Thus, the ITIS is a stronger indicator of recurrence risk than the analyzed TLS parameters individually.

Discussion

It is of paramount clinical interest to predict the risk of recurrence in nmCRC in order to personalize the selection of patients for adjuvant chemotherapy. The determinants of an effective anti-tumor immune-surveillance leading to long-term freedom from recurrent disease after complete cancer resection in nmCRC are not yet fully understood.¹⁸ The composition of the immune cell infiltrate has repeatedly been shown to be prognostic in a variety of cancer types.¹⁹ Specifically for CRC, the Immunoscore derived from the number of tumor-infiltrating T cells⁵ has been validated in a large multicenter study as a significant prognosticator of stage I/II/III disease.²⁰ Further, the presence of lymphatic vessels and cytotoxic T cells rather

than the genetic alterations of primary CRCs predict the development of metastatic disease.²¹ The number of TILs significantly correlates with TLS in CRC,^{8,22,23} however it is still not clear whether the two parameters are interdependent.

Several TLS-associated cell types like mature DCs, B cells, T follicular helper cells and high endothelial venules have been used as surrogates to quantify TLS in cancer tissues. No matter the method, the large majority of studies report a positive association between TLS and survival in CRC and other tumor types.²⁴ Here we applied a quantitative pathology approach to detect TLS as a whole structure rather than using surrogate markers. We observed that peritumoral TLS exhibited strong phenotypic differences both within and between patients based on their capacity to recruit FDCs and generate GCs, which supports the concept of sequential TLS maturation steps with increasing B-cell activity. This is in agreement with the observations in LSCC (Silina et al. in revision) and suggests that TLS maturation follows a similar process in different cancers.

In exploring the prognostic potential of the different TLS maturation populations, we studied their association with disease recurrence after surgery in curative intent in nmCRC. Here, we observed that patients with a higher density of mature TLS had a four-fold lower risk of recurrence. Further, the proportion of TLS with FDCs and GCs showed a stronger association with recurrence risk than mere TLS density. In addition to these proportions, also the presence of at least on GC-harboring TLS was associated with a much lower risk of CRC recurrence. These data are in line with the survival study

Table 4. Multivariable competing risk regression models of 3-year recurrence risk. Results were estimated with multivariable Fine & Gray competing risk regression models, treating death-from-any-cause as the competing event of interest. Abbreviations: SHR – Subdistribution hazard ratio, 95% CI – 95% confidence interval, TLS/mm – TLS count, ECOG – Eastern Cooperative Oncology Group, UICC – Union Internationale pour le lutte contre le cancer, % E-TLS – Proportion of Early tertiary lymphoid structure (TLS), Q1 – 25th percentile of the variable's distribution (i.e. quartile 1 (Q1)), % PFL-TLS – Proportion of primary follicle-like TLS, % SFL-TLS – Proportion of secondary follicle-like TLS, GC – Germinal center, ITIS – Integrated TLS ImmunoScore, CART – Classification and Regression Tree.

Model	Variable	SHR	95%CI	p
Multivariable Model #1	TLS/mm \leq Q1	4.09	1.32–12.71	0.015
	Age \geq 75 years	5.40	1.71–17.01	0.004
	ECOG performance status $>$ 0	8.06	2.69–24.10	$<$ 0.0001
	UICC tumor stage III	2.32	0.76–7.05	0.138
	Adjuvant chemotherapy	2.05	0.57–7.36	0.269
Multivariable Model #2	% E-TLS \geq Q3	4.01	1.39–11.55	0.010
	Age \geq 75 years	4.85	1.90–12.40	0.001
	ECOG performance status $>$ 0	7.39	2.16–25.34	0.001
	UICC tumor stage III	1.90	0.48–7.60	0.362
	Adjuvant chemotherapy	1.21	0.30–4.81	0.786
Multivariable Model #3	% PFL-TLS \leq Q1	4.28	1.59–11.54	0.004
	Age \geq 75 years	3.94	1.26–12.27	0.018
	ECOG performance status $>$ 0	6.94	1.66–28.92	0.008
	UICC tumor stage III	2.16	0.58–8.01	0.250
	Adjuvant chemotherapy	1.80	0.48–6.77	0.385
Multivariable Model #4	% SFL-TLS \leq Q1	3.99	1.30–12.20	0.015
	Age \geq 75 years	4.88	1.75–13.58	0.002
	ECOG performance status $>$ 0	5.77	1.90–17.54	0.002
	UICC tumor stage III	3.50	0.96–12.71	0.057
	Adjuvant chemotherapy	0.92	0.29–2.94	0.889
Multivariable Model #5	\geq 1 GC-harboring TLS	0.28	0.09–0.84	0.024
	Age \geq 75 years	4.61	1.69–12.54	0.003
	ECOG performance status $>$ 0	6.98	2.25–21.68	0.001
	UICC tumor stage III	2.96	0.86–10.12	0.084
	Adjuvant chemotherapy	1.09	0.37–3.23	0.871
Multivariable Model #6	ITIS (per 10 units increase)	0.67	0.49–0.92	0.012
	Age \geq 75 years	4.86	1.76–13.42	0.002
	ECOG performance status $>$ 0	6.03	1.85–19.67	0.003
	UICC tumor stage III	2.10	0.42–10.51	0.366
	Adjuvant chemotherapy	1.29	0.26–6.40	0.753
Multivariable Model #7	Low ITIS (i.e. \leq CART cut-off)	8.41	2.73–25.87	$<$ 0.0001
	Age \geq 75 years	6.06	2.18–16.85	0.001
	ECOG performance status $>$ 0	8.80	2.69–28.78	$<$ 0.0001
	UICC tumor stage III	1.47	0.36–6.03	0.594
	Adjuvant chemotherapy	1.31	0.31–5.50	0.709

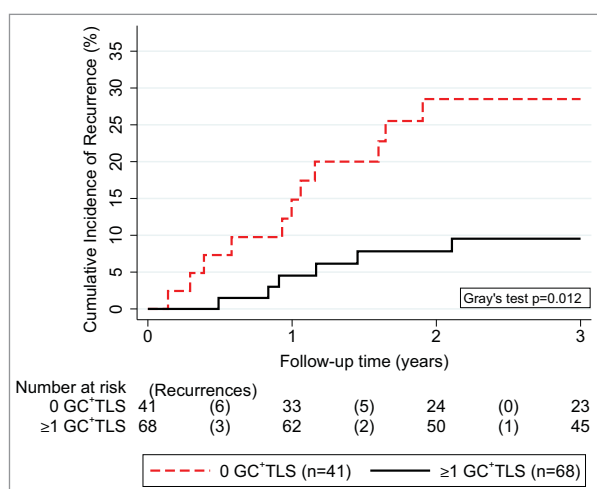


Figure 4. Competing risk analysis of germinal-center-harboring TLS and 3-year risk of colorectal cancer recurrence. Cumulative incidences of recurrence were derived with competing risk analysis, accounting for death-from-any-cause as the competing risk of interest. Patients were dichotomized into two groups according to whether they had at least one TLS with an active GC reaction ($n = 68$) or not ($n = 41$). Risktables (“Number at risk”) report the number of patients in each group at risk of recurrence at the start of each time interval. The number of recurrences in each time interval (“Recurrences”) are reported in brackets.

in neo-adjuvant chemotherapy treated LSCC patients. There, TLS had almost no GCs and TLS counts had no prognostic significance in comparison to untreated patients where TLS frequently had GCs and were the most important independent prognostic marker.¹⁷ These data suggest further hypotheses: (1) TLS with a GC reaction represents the most “functional” subtype, (2) GC development indicates towards tumor-specific B cell and CD4 T cell priming, and (3) TLS maturation harbors additional information to mere TLS enumerations for the prognosis of CRC recurrence.

To address the latter, we summarized both the number and maturity of peritumoral TLS into an integrated TLS immuno-score (ITIS). Elevated ITIS identified a large patient subgroup (two thirds of our study cohort) with a very low risk of recurrence independently of the current most relevant predictors such as ECOG performance status, pathologic tumor stage, adjuvant chemotherapy, MSI status and age. These data suggest a potential relevance of both TLS numbers and maturation for the immune control of nmCRC. If validated in external prospective cohorts, the ITIS could have important implications for identifying an nmCRC subpopulation in which adjuvant chemotherapy may be safely avoided. We did not examine the association between TLS parameters and survival, because the event rate for mortality was low and diluted by deaths from

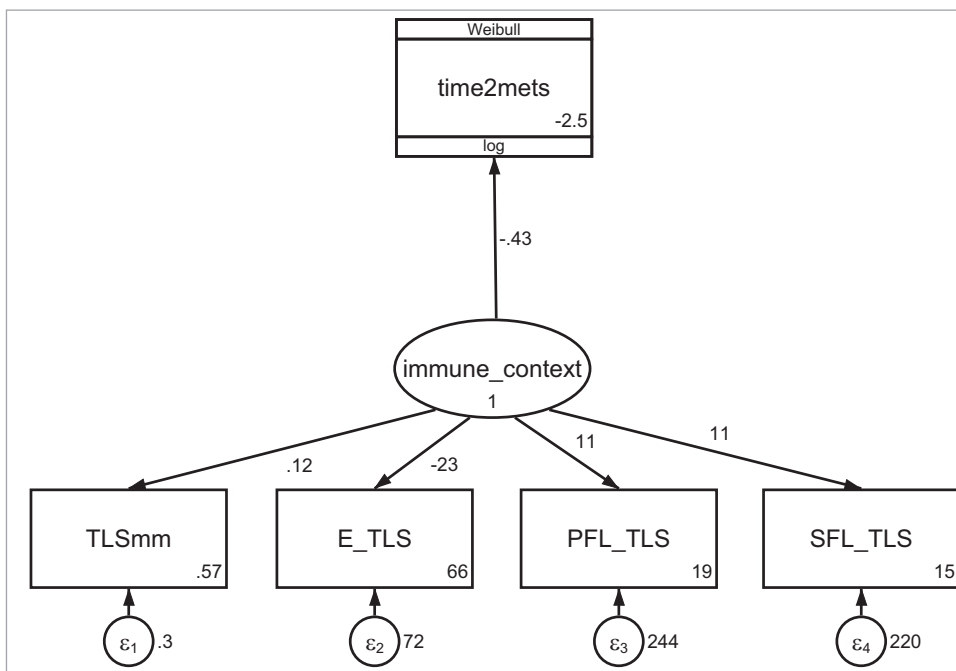


Figure 5. Path diagram of the structural equation model for estimating the relationship between TLS parameters, a latent TLS immune contexture variable, and colorectal cancer recurrence. Numbers to the right of the path represent path coefficients (or the log(hazard ratio) for the path from “immune_context” to “time2mets”). Numbers in the right bottom of the square boxes represent the intercepts of the path coefficients. Numbers adjacent to the round error terms represent the errors of the measurements. Abbreviations: TLSmm – TLS count, E-TLS – Early TLS proportion, PFL-TLS – primary follicle like TLS proportion, SFL-TLS – secondary follicle like TLS proportion.

causes other than CRC disease progression in our cohort. Further, studying the relationship between ITIS and the Immunoscore will allow assessing their interdependence and potentially yield improved immune-based prognostic tools.

We applied a statistical technique called generalized structural equation modeling (SEM) to generate ITIS. SEM is widely used in the field of psychology to test whether highly correlated data are consistent with an underlying, abstract, unobservable or difficult-to-observe “latent” process.²⁵ For example, the intelligence quotient (IQ) is a typical latent variable generated by SEM based on test participants’ answers to the different questions of an IQ test. Likewise, we estimated the ITIS as a

latent variable reflecting TLS formation and maturation by our patients’ TLS counts and relative proportions of E-TLS, PFL-TLS, and SFL-TLS. Our data show that SEM is a powerful approach for analyzing joint signatures of biomarkers in immunology. Studies in the field of immunology often assess multiple immune markers at the same time, which may not only yield more insights but can also inflate type I error rates and make a joint interpretation of results challenging. SEM can estimate latent variables that summarize these multiple immune markers, and study their joint association with other (non-) immune variables and clinical outcomes such as cancer recurrence or disease progression. The concept of using SEM for

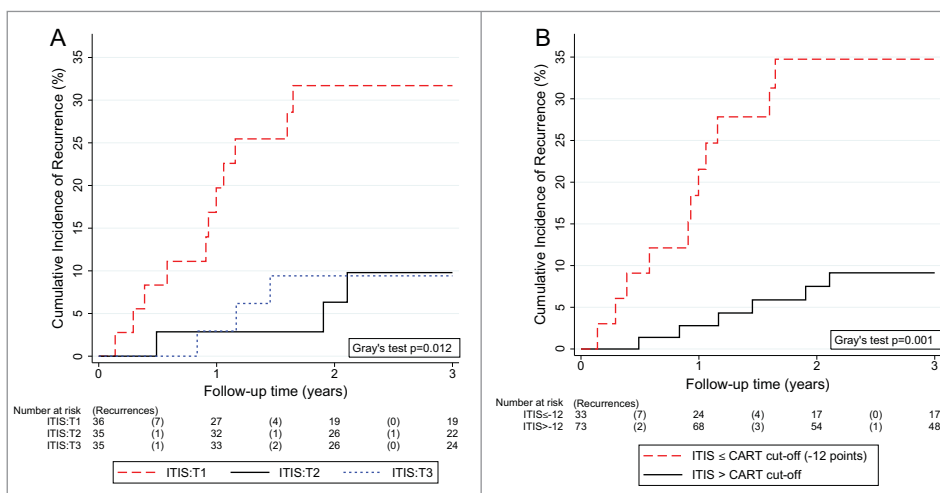


Figure 6. Competing risk analysis of the Integrated TLS ImmunoScore (ITIS) and 3-year risk of colorectal cancer recurrence. Cumulative incidences of recurrence were derived with competing risk analysis, accounting for death-from-any-cause as the competing risk of interest. (A) The ITIS was categorized empirically into tertiles (T1, T2, and T3). (B) The ITIS was dichotomized into a binary variable at the cut-off suggested by the Classification And Regression Tree (CART) analysis at -12 points. Risktables (“Number at risk”) report the number of patients in each group at risk of recurrence at the start of each time interval. The number of recurrences in each time interval (“Recurrences”) are reported in brackets.

immunology research was first proposed in 2011 by Brown and colleagues.²⁶ However, only a few studies have made use of this technique so far. For example, Baltar et al. used SEM for examining the relationships between B vitamins and immune markers in lung cancer pathogenesis.²⁷ Recent software developments in the field will considerably improve the accessibility of SEM methods to a wide group of scientists beyond biostatisticians.²⁸ Therefore, our analysis could serve as a template for future studies assessing prognostic and predictive biomarker signatures in the field of immuno-oncology.

In CRC patients, the mutation load influences tumor antigenicity and is associated with tumor immune recognition and efficacy of immunotherapy.^{29,30} MSI is associated with increased tumor antigen load and increased tumor immunogenicity that is crucial for the success of anti-PD-1 therapy in mCRC.¹³ However, also a part of MSS CRC patients respond to such therapy by combining different treatment strategies.³¹ Apart from MSI, also mutations in oncogenes such as *BRAF* and *KRAS* yield immunogenic epitopes.^{15,16} However, the relevance of antigen load for the induction or maintenance of tumor-associated TLS is not clear. To test this we compared TLS parameters in patients with MSI-H status and/or *BRAF* mutations assuming such tumors are enriched with immunogenic antigens. We observed that TLS were not only more frequent but also more mature in patients with these genetic alterations. This suggests that TLS development is facilitated by the availability of tumor antigens in nmCRC. Whether TLS can serve as a marker for increased tumor antigenicity and thus better identify patients who respond to immunotherapy should be addressed in future clinical trials. One drawback of our study is the lack of the analysis of *KRAS* mutations, which might have provided a more precise definition of the low antigen load tumor group.

Apart from the antigenic load, the local and/or systemic immunosuppressive milieu affects tumor immunogenicity.³² An elevated neutrophil-lymphocyte ratio (NLR) is an indication of systemic inflammatory response, and is an independent poor prognostic marker in multiple tumor types including CRC.³³⁻³⁶ We observed that high NLR significantly correlated with low TLS density and decreased maturation. It is not known whether the increased neutrophil numbers have a direct impact or are an indication of other processes leading to impaired TLS development in patients with increased NLR. Nevertheless, we propose that the presence of GC-positive TLS reflects an immunogenic tumor microenvironment. Although beyond the data of our study and the CRC setting, we carefully speculate that future studies could focus on investigating TLS maturation parameters and the ITIS as putative predictive biomarkers for clinical response to checkpoint inhibitors.

When considering potential future clinical applications of the ITIS and the TLS maturation concept, it is important to note that a low ITIS (as defined by the 1st tertile of the score) could identify patients with a very high risk of CRC recurrence, whereas the CRC recurrence risk was low in patients above the 1st ITIS tertile and did not differ between patients in the 2nd and 3rd ITIS tertile. This “non-linearity” suggests that a low ITIS may be most useful as a “high-risk” biomarker for identifying a third of stage II/III CRC patients with a strong

propensity for disease recurrence, whereas the ITIS does not appear to contribute prognostic resolution in the other two thirds of patients who already have a very low risk of recurrence.

Some limitations should be considered. First, our study would have benefited from a larger sample size. In detail, some associations, which did not meet the pre-specified criteria for statistical significance of association ($p \leq 0.05$), may have crossed this significance threshold with a slightly higher sample size and event rate. Second, we did not externally validate our findings in a separate cohort of CRC patients. Importantly, external validation of both the ITIS score and the proposed cut-offs will be necessary before advancing the concept of TLS maturation into the clinic. Third, the ITIS was developed for the stage II and III CRC setting, and does not necessarily generalize to other tumor types, which may feature other absolute and relative TLS parameter counts. Fourth, the association between *BRAF* V600E mutation status and increased TLS formation may have been confounded by the strong associations between *BRAF*, MSI, and TLS formation. Although some reports support an involvement of *BRAF* mutation status in adverse prognosis in the adjuvant CRC setting,^{37,38} our data on *BRAF* mutation as a determinant of TLS formation hence have to be considered as preliminary and require external validation. Finally, all TLS immunopathology analyses in this study were performed by a single scientist (KS) at a single academic center (University Hospital Zurich, Switzerland) who was fully blinded to the clinical characteristics and outcomes of patients. Although this can be considered as strength, it still needs to be demonstrated that the TLS parameter enumeration approach proposed in this study can successfully be performed in other laboratories and settings.

Conclusion

In conclusion, this comprehensive analysis of clinical, molecular, tissue, and blood-based parameters in patients with stage II and III CRC revealed three novel findings. First, our data demonstrate the presence of distinct maturation subtypes of TLS in nmCRC. Second, TLS maturation is an important parameter of tumor immune contexture and bears significant prognostic and potential predictive value. Notably, the prognostic potential of TLS maturation for CRC recurrence may be higher than the potential of plain TLS enumeration. Third, using SEM for the integration of TLS parameters into a joint immunoscore (ITIS) allowed identification of patient subgroups at high and low risk of tumor recurrence independent of established prognostic markers. If validated in external clinical studies, our data may have important implications for targeted indication of adjuvant chemotherapy in this highly prevalent tumor entity.

Patients & methods

Study design & patient population

This study represents a single-center, observational, retrospective cohort study. The study population includes all patients with histologically-confirmed, UICC stage II or III adenocarcinoma of the colon and rectum who were seen at a large

tertiary academic oncology center in Middle Europe (Division of Oncology, Medical University of Graz, Austria) between Jan, 1996 and Jun, 2011 after having had undergone surgery in curative intent. Patients with evidence of distant metastasis at the time of diagnosis were excluded, even if a (neoadjuvant) treatment plan in curative intent was followed.

All patients were routinely examined every 3 months during the first 3 years after surgery within the aftercare program of the Medical University of Graz' Division of Oncology. During each of these visits a detailed medical history, physical examination and full laboratory investigation was performed. At every second visit a chest-X-ray and abdominal ultrasound was performed. For the current study we retrospectively included only those patients who had a valid local biobank consent permitting the unconditional storage and subsequent analysis of their primary tumor tissue samples, which resulted in 111 patients. TLS analysis was unsuccessful in 2 of these patients, resulting in a final study population of 109 patients.

Baseline and outcome data were collected retrospectively from our in-house electronic healthcare database and from written documents of the oncologic outpatient department as previously reported.^{39,40} The primary endpoint of this study was 3-year disease recurrence, defined as the cumulative incidence of local recurrence and/or distant metastasis within the first 3 years after surgery accounting for death-from-any-cause as a competing event. This research was approved by the Ethics Committee and the biobank at the Medical University of Graz (vote number: 26-196 ex 13/14). The Ethics Committee granted us a "waiver of consent" for this retrospective study consistent with national legal regulations, i.e. no individual consent was required from included patients.

Analysis of TLS maturation stages

The sample analyst (KS) was fully blinded to the clinical characteristics and outcomes of patients. FFPE tumor blocks were used to prepare 2- μ m thick serial sections and used for multiparameter immunofluorescence (IF) staining. TLS maturation stages were analyzed in all patients by the detection of FDCs (CD21), GCs (CD23) and CXCL13 using the thymidine signal amplification (TSA) approach according to manufacturer's protocol (PerkinElmer). IF was performed using the Discovery Ultra automated systems (Ventana Medical Systems, USA) according to the standard operating procedure at the University Hospital Zurich. Slides were mounted using Vectashield Antifade Mounting Medium with DAPI (Vector Laboratories).

The slides were scanned using the automated multispectral microscopy system Vectra 3.0 (PerkinElmer). Whole slide scan was performed at 100x magnification and multispectral high power fields were imaged at 200x. An unstained nmCRC slide was used to generate the spectral profile of autofluorescence in nmCRC tissues, and single stained slides were used to generate the spectral profiles of the used fluorophores. The Inform software (PerkinElmer) was used to unmix the spectra for each specific fluorophore from the autofluorescence. More than 90% of all lymphocytic aggregates were imaged in each patient in high power, excluding Payer's patches. TLS were enumerated in the whole slide including tumor tissue and adjacent normal tissue. TLS density was calculated as the number of TLS per

mm of tumor-invasive front in peri- and intratumoral regions. The peritumoral region was defined as the immediate adjacent non-tumoral tissue within a 7 mm radius of the tumor front. For each patient, the TLS density estimate was derived from the tissue block with the highest TLS density. TLS maturation stages were assessed as described previously by co-staining for CD21, CD23, and CXCL13:¹⁷ (1) early stage (E-TLS), characterized by dense lymphocytic aggregates without CD21 and CD23 expression, (2) primary follicle-like stage (PFL-TLS), characterized by lymphocytic clusters with central network of mature FDCs (CD21⁺), but no GC reaction (CD23⁻), and (3) secondary follicle-like stage (SFL-TLS), characterized by lymphocytic clusters with GC reaction (CD21⁺CD23⁺). The numbers of TLS in each maturation stage were counted and expressed as a proportion from all TLS within each patient or as normalized counts per mm of tumor perimeter. We termed the total number of TLS per mm of tumor perimeter as TLS density. The tumor perimeter was measured in the whole slide image as the gross length of tumor invasive front facing adjacent normal tissue by the Fiji software.

Additionally, serial sections from 5 patients with high numbers of peritumoral TLS were used to stain for B cells (CD20), T cells (CD3) and CCL21 to assess the lymphocytic organization of TLS in CRC. IF was performed by a manual protocol. In detail, after heating of slides for 2 hours at 55°C, slides were incubated with the Trilogy pretreatment solution (CellMarque) for 15 minutes in a pressure cooker. After cooling for 15 minutes and washing in milli-Q water, slides were treated with 3% H₂O₂, washed with 0.1% Triton X-100 in PBS, and blocked BSA/0.1% Triton X-100/PBS. Antibodies were diluted in 1% BSA / 0.1% Triton X-100 / PBS. After overnight incubation of primary antibodies with slides at 4°C, slides were washed and incubated for 1 hour at room temperature with secondary antibodies. Further, slides were incubated with DAPI (Invitrogen 0.5 μ g/mL) and mounted with the ProlongDiamond medium (both Life Technologies). Slides were imaged using the laser-scanning confocal microscope Leica SP8 (Leica). See Supplementary Table 4 for all used antibodies and protocols.

Assessment of BRAF mutation and mismatch-repair status

BRAF V600E mutation status was determined using an allele-specific real-time PCR protocol providing a mutation detection limit of 10 copies of mutant DNA in proportions as low as 1% of the total DNA as described previously.⁴¹ In brief, two separate PCRs were performed to assess mutation status: An allele-specific PCR utilizing an allele-specific PCR primer for mutant allele detection and a mutation unspecific PCR to generate a reference amplicon. Target amplification during real-time PCR was detected by a TaqMan[®] probe. Additionally, an exogenous, internal control PCR product, a fragment in the CYP17 promoter region, was co-amplified in each reference and allele-specific PCR. Used primer sequences are given in Supplementary Table 5. Reference PCR was performed in a 25 μ l reaction volume with 1x TaqMan[®] Genotyping Master Mix (ThermoFisher Scientific), 900nM of *BRAF* mutation-unspecific primers (forward and reverse), 100nM *BRAF* probe, 112.5nM of each internal control primer, 25nM of internal control probe, and 5 μ l DNA of varying concentration. Allele-specific PCRs were

performed according to the same protocol, but using a concentration of 450nM of allele-specific primer. All real-time PCRs were carried out on a LightCycler[®] 480 Real-Time PCR System (Roche Diagnostics, Vienna, Austria) under the following thermocycling conditions: 95°C for 10 min followed by 50 cycles of 90°C for 15 sec and 60°C for 1min. Cycle thresholds (Ct) were recorded for reference PCR as well as for each allele-specific PCR and corresponding Δ Ct values (i.e. allele-specific Ct minus reference Ct) were calculated. A Δ Ct value of 9 was used as cut-off point for mutation detection.

MSI analysis was performed immunohistochemically using antibodies for the mismatch repair proteins MLH1 (clone G168–15, BD Pharmingen, San Diego, California, USA), PMS2 (clone A16–4, BD Pharmingen, San Diego, California, USA), MSH2 (clone G219–1129, Cell Marque, Rocklin, California, USA) and MSH6 (clone SP93, Cell Marque, Rocklin, California, USA). The histological features of the tumours and the immunohistochemical staining were assessed by a Pathologist (SL) unaware of the patients' clinical data. Loss of the respective mismatch repair protein was recorded when nuclear staining was absent from all tumour cells but preserved in normal epithelial and stromal cells.

Statistical methods

All statistical analyses were performed using Stata (Windows version 14.0, Stata Corp., Houston, TX, USA). The statistical analyst (FP) did not have access to tissue samples or primary data from tissue analyses. Continuous variables were summarized as medians [25th–75th percentile], whereas count data were reported as absolute frequencies (%). The association between two categorical variables was assessed with χ^2 -tests or Fisher's exact tests. Means between two or more groups were compared with Wilcoxon rank-sum tests and Kruskal-Wallis tests. Correlations between two continuous variables were computed with Spearman's rank-based correlation coefficient.

Follow-up was defined as the time from the day of surgery until the occurrence of recurrence, death, or censoring alive 3 years after surgery. The median follow-up time was estimated with the method of Schemper & Smith.⁴² Flexible parametric modeling of the hazard function was applied for estimating recurrence rates at different time points of follow-up (Stata routine *stpm2*).⁴³ A competing risk cumulative incidence estimator, treating death-from-any-cause as a competing risk, was used for the calculation of 3-year recurrence risk (Stata routine *stcompet*).^{44,45} Gray's test was used to compare the 3-year cumulative incidences of recurrence between two or more groups (self-written Stata routine *stgrays* using the R library *cmprsk* via *rsource*).⁴⁶ In the absence of validated cut-offs for TLS density and TLS maturation parameters, we pre-specified to select an empirical cut-off at the 25th percentile to dichotomize patients into subgroups with low relative levels of immune contexture parameters. Fine & Gray proportional subdistribution hazards models were fitted for uni- and multivariable time-to-recurrence regression (Stata routine *stcrreg*).⁴⁷ The proportional subdistribution hazards assumption was assessed by including products of linear follow-up time and the variable of interest in the respective models (*tvc* option in *stcrreg*).

A generalized structural equation model with time-to-recurrence as a Weibull time-to-event response, linked to a latent variable ($\mu = 0, \sigma^2 = 1$) via a log-link, was constructed graphically (Stata's GSEM builder), and then estimated with a non-adaptive Gauss Hermite quadrature algorithm (Stata routine *gsem*).^{25,28} The latent variable reflects both TLS density and maturation stages, and in order to simplify its calculation in practice, we expressed it as an "Integrated TLS immunoscore (ITIS)." This was achieved by computing the ITIS as the sum of the products of the path coefficients for the latent variable and their observation in a patient, standardized by the path coefficients' intercepts (Equation 1 in the results section).

In sensitivity analyses, we (1) performed a classification and regression tree analysis (CART, Stata routine *cart*) to explore the strength of recurrence risk predictors and potential ITIS cut-off points within a non-parametric fashion,⁴⁸ and (2) fitted interactions between the ITIS, tumor stage and adjuvant chemotherapy to gauge potential effect modification. The code for all analyses is available on request from FP.

Author contributions

Conceived and designed the study: AG UP TW. Performed immunology analyses: KS. Performed statistical analyses: FP. Contributed patients and clinical data: JR MS JS HS MP. Drafted the manuscript: FP KS TW. Critically reviewed and revised the draft version of the manuscript: All authors. Agree with the manuscript's results and conclusions: All authors. ICJME criteria for authorship met: All authors.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Susanne Dettwiler (lead bioanalytical technician of the University of Zurich biobank) for valuable technical assistance. Further, we are grateful to Dr. Sylvia Gusel (Division of Oncology, Medical University of Graz) for help with electronic clinical data management, and Dr. Sebastian Fuchs (Biobank, Medical University of Graz) for assistance with biospecimen retrieval.

Funding details

This work was supported by the Forschungskredit of the University of Zurich (project number K-84901-01-01), the Swiss National Science Foundation (SNSF), the Sciex Foundation, the Science Foundation for Oncology (SFO), the Cancer League Zurich, the Novartis Research Foundation, the University Research Priority Program (URPP) "Translational Cancer Research", Swiss Cancer League (reference number F-87701-31-01), and the Cancer League Zurich.

References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin.* 2017;67:7-30. doi:10.3322/caac.21387. PMID:28055103
2. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends—An Update. *Cancer Epidemiol Biomarkers Prev.* 2016;25:16-27. doi:10.1158/1055-9965.EPI-15-0578. PMID:26667886
3. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet (London, England).* 2014;383:1490-502. doi:10.1016/S0140-6736(13)61649-9. PMID:24225001

4. Lech G, Slotwinski R, Slodkowski M, Krasnodebski IW. Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. *World J Gastroenterol.* 2016;22:1745-55. doi:10.3748/wjg.v22.i5.1745. PMID:26855534
5. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, Lugli A, Zlobec I, Hartmann A, Bifulco C, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. *J Pathol.* 2014;232:199-209. doi:10.1002/path.4287. PMID:24122236
6. Fridman WH, Zitvogel L, Sautes-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol.* 2017;14:717-734. doi:10.1038/nrclinonc.2017.101. PMID:28741618
7. Neyt K, Perros F, GeurtsvanKessel CH, Hammad H, Lambrecht BN. Tertiary lymphoid organs in infection and autoimmunity. *Trends Immunol.* 2012;33:297-305. doi:10.1016/j.it.2012.04.006. PMID:22622061
8. Hiraoka N, Ino Y, Yamazaki-Itoh R. Tertiary Lymphoid Organs in Cancer Tissues. *Frontiers Immunol.* 2016;7:244. doi:10.3389/fimmu.2016.00244.
9. Ogino S, Noshko K, Irahara N, Meyerhardt JA, Baba Y, Shima K, et al. Lymphocytic reaction to colorectal cancer is associated with longer survival, independent of lymph node count, microsatellite instability, and CpG island methylator phenotype. *Clin Cancer Res.* 2009;15:6412-20. doi:10.1158/1078-0432.CCR-09-1438. PMID:19825961
10. Meshcheryakova A, Tamandl Z, Bajna E, Stift J, Mittlboeck M, Svoboda M, Heiden D, Stremtizer S, Jensen-Jarolim E, Grünberger T, et al. B cells and ectopic follicular structures: novel players in anti-tumor programming with prognostic power for patients with metastatic colorectal cancer. *PLoS One.* 2014;9:e99008. doi:10.1371/journal.pone.0099008. PMID:24905750
11. Germain C, Gnjjatic S, Tamzalit F, Knockaert S, Remark R, Goc J, Lepelletier A, Becht E, Katsahian S, Bizouard G, et al. Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. *Am J Respiratory Critical Care Med.* 2014; 189:832-44. doi:10.1164/rccm.201309-1611OC.
12. van de Pavert SA, Mebius RE. New insights into the development of lymphoid tissues. *Nat Rev Immunol.* 2010; 10:664-74. doi:10.1038/nri2832. PMID:20706277
13. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med.* 2015;372:2509-20. doi:10.1056/NEJMoa1500596. PMID:26028255
14. Alexander J, Watanabe T, Wu TT, Rashid A, Li S, Hamilton SR. Histopathological identification of colon cancer with microsatellite instability. *Am J Pathol.* 2001;158:527-35. doi:10.1016/S0002-9440(10)63994-6. PMID:11159189
15. Andersen MH, Fensterle J, Ugurel S, Reker S, Houben R, Guldborg P, Berger TG, Schadendorf D, Trefzer U, Bröcker EB, et al. Immunogenicity of constitutively active V599EBRAF. *Cancer Res.* 2004;64:5456-60. doi:10.1158/0008-5472.CAN-04-0937. PMID:15289355
16. Tran E, Ahmadzadeh M, Lu YC, Gros A, Turcotte S, Robbins PF, Gartner JJ, Zheng Z, Li YF, Ray S, et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science (New York, NY).* 2015;350:1387-90. doi:10.1126/science.aad1253.
17. Silina K, Soltermann A, Attar FM, Casanova R, Uckelely ZM, Thut H, et al. Corticosteroids impair tertiary lymphoid structures in lung squamous cell carcinoma. *Cancer Res.* 2017; [accepted]
18. Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoeediting. *Annual Rev Immunol.* 2004; 22:329-60. doi:10.1146/annurev.immunol.22.012703.104803.
19. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer.* 2012;12:298-306. doi:10.1038/nrc3245. PMID:22419253
20. Galon J. Validation of the Immunoscore (IM) as a prognostic marker in stage I/II/III colon cancer: Results of a worldwide consortium-based analysis of 1,336 patients. *J Clin Oncol.* 2016;34:15_suppl. 3500-3500.
21. Mlecnik B, Bindea G, Kirilovsky A, Angell HK, Obenauf AC, Tosolini M, Church SE, Maby P, Vasaturo A, Angelova M, et al. The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. *Sci Transl Med.* 2016;8:327ra26. doi:10.1126/scitranslmed.aad6352. PMID:26912905
22. Di Caro G, Bergomas F, Grizzi F, Doni A, Bianchi P, Malesci A, et al. Occurrence of tertiary lymphoid tissue is associated with T-cell infiltration and predicts better prognosis in early-stage colorectal cancers. *Clin Cancer Res.* 2014; 20:2147-58. doi:10.1158/1078-0432.CCR-13-2590. PMID:24523438
23. Vayrynen JP, Kantola T, Vayrynen SA, Klintrup K, Bloigu R, Karhu T, Mäkelä J, Herzig KH, Karttunen TJ, Tuomisto A, et al. The relationships between serum cytokine levels and tumor infiltrating immune cells and their clinical significance in colorectal cancer. *Int J Cancer.* 2016;139:112-21. doi:10.1002/ijc.30040. PMID:26874795
24. Sautes-Fridman C, Lawand M, Giraldo NA, Kaplon H, Germain C, Fridman WH, Dieu-Nosjean MC. Tertiary Lymphoid Structures in Cancers: Prognostic Value, Regulation, and Manipulation for Therapeutic Intervention. *Frontiers Immunol.* 2016;7:407. doi:10.3389/fimmu.2016.00407.
25. Bielby WT, Hauser RM. Structural Equation Models. *Annual Rev Sociol.* 1977;3:131-61. doi:10.1146/annurev.so.03.080177.001033.
26. Brown JR, Stafford P, Johnston SA, Dinu V. Statistical methods for analyzing immunosignatures. *BMC Bioinformatics.* 2011;12:349. doi:10.1186/1471-2105-12-349. PMID:21854615
27. Baltar VT, Xun WW, Johansson M, Ferrari P, Chuang SC, Relton C, Ueland PM, Midttun Ø, Slimani N, Jenab M, et al. A structural equation modelling approach to explore the role of B vitamins and immune markers in lung cancer risk. *Eur J Epidemiol.* 2013; 28:677-88. doi:10.1007/s10654-013-9793-z. PMID:23532743
28. Huber C. Introduction to Structural Equation Modeling Using Stata. 2014. <http://www.cair.org/wp-content/uploads/sites/474/2015/07/HuberC-SEM-Workshop.pdf>
29. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science (New York, NY).* 2016;351:1463-9. doi:10.1126/science.aaf1490.
30. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science (New York, NY).* 2015;348:124-8. doi:10.1126/science.aaa1348.
31. Bendell JK TW, Goh BC, Wallin J, Oh DY, Han SW, Lee CB, Hellmann MD, Desai J, Lewin JH, Solomon BJ, Quan Man Chow L, Miller WH, Gainer JF, Flaherty K, Infante JR, Das-Thakur M, Foster P, Cha E, Bang YJ. Clinical activity and safety of cobimetinib (cobi) and atezolizumab in colorectal cancer (CRC). *J Clin Oncol.* 2016;34:15_suppl. 3502-3502
32. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoeediting: integrating immunity's roles in cancer suppression and promotion. *Science (New York, NY).* 2011;331:1565-70. doi:10.1126/science.1203486.
33. Absenger G, Szkandera J, Pichler M, Stotz M, Arminger F, Weissmueller M, Schaberl-Moser R, Samonigg H, Stojakovic T, Gerger A. A derived neutrophil to lymphocyte ratio predicts clinical outcome in stage II and III colon cancer patients. *Br J Cancer.* 2013;109:395-400. doi:10.1038/bjc.2013.346. PMID:23820252
34. Guthrie GJ, Charles KA, Roxburgh CS, Horgan PG, McMillan DC, Clarke SJ. The systemic inflammation-based neutrophil-lymphocyte ratio: experience in patients with cancer. *Critical Rev Oncol/Hematol.* 2013;88:218-30. doi:10.1016/j.critrevonc.2013.03.010.
35. Paramanathan A, Saxena A, Morris DL. A systematic review and meta-analysis on the impact of pre-operative neutrophil lymphocyte ratio on long term outcomes after curative intent resection of solid tumours. *Surg Oncol.* 2014; 23:31-9. doi:10.1016/j.suronc.2013.12.001. PMID:24378193
36. Chen ZY, Raghav K, Lieu CH, Jiang ZQ, Eng C, Vauthey JN, Chang GJ, Qiao W, Morris J, Hong D, et al. Cytokine profile and prognostic significance of high neutrophil-lymphocyte ratio in colorectal cancer. *British J Cancer.* 2015;112:1088-97. doi:10.1038/bjc.2015.61.
37. Roth AD, Tejpar S, Delorenzi M, Yan P, Fiocca R, Klingbiel D, Dietrich D, Biesmans B, Bodoky G, Barone C, et al. Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-

- 00 trial. *J Clin Oncol*. 2010;28:466-74. doi:10.1200/JCO.2009.23.3452. PMID:20008640
38. Lochhead P, Kuchiba A, Imamura Y, Liao X, Yamauchi M, Nishihara R, Qian ZR, Morikawa T, Shen J, Meyerhardt JA, et al. Microsatellite instability and BRAF mutation testing in colorectal cancer prognostication. *J Natl Cancer Inst*. 2013;105:1151-6. doi:10.1093/jnci/djt173. PMID:23878352
 39. Posch F, Leitner L, Bergovec M, Bezan A, Stotz M, Gerger A, Pichler M, Stöger H, Liegl-Atzwanger B, Leithner A, et al. Can multistate modeling of local recurrence, distant metastasis, and death improve the prediction of outcome in patients with soft tissue Sarcomas? *Clin Orthopaedics Related Res*. 2017;475:1427-35. doi:10.1007/s11999-017-5232-x.
 40. Stotz M, Liegl-Atzwanger B, Posch F, Mrcic E, Thalhammer M, Stojakovic T, Bezan A, Pichler M, Gerger A, Szkandera J. Blood-Based biomarkers are associated with disease recurrence and survival in gastrointestinal stroma tumor patients after surgical resection. *PloS One*. 2016;11:e0159448. doi:10.1371/journal.pone.0159448. PMID:27454486
 41. Lang AH, Drexel H, Geller-Rhomberg S, Stark N, Winder T, Geiger K, Muendlein A. Optimized allele-specific real-time PCR assays for the detection of common mutations in KRAS and BRAF. *J Mol Diagn*. 2011;13:23-8. doi:10.1016/j.jmoldx.2010.11.007. PMID:21227391
 42. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Controlled Clin Trials*. 1996; 17:343-6. doi:10.1016/0197-2456(96)00075-X. PMID:8889347
 43. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics Med*. 2002;21:2175-97. doi:10.1002/sim.1203.
 44. Panotopoulos J, Posch F, Alici B, Funovics P, Stihsen C, Amann G, Brodowicz T, Windhager R, Ay C. Hemoglobin, alkaline phosphatase, and C-reactive protein predict the outcome in patients with liposarcoma. *J Orthop Res*. 2015;33:765-70. doi:10.1002/jor.22827. PMID:25641201
 45. Coviello V, Boggess M. Cumulative incidence estimation in the presence of competing risks. *Stata J*. 2004;4:103-12
 46. Gray RJ. A Class of K-Sample tests for comparing the cumulative incidence of a competing risk. *Annals Statistics*. 1988; 16:1141-54. doi:10.1214/aos/1176350951.
 47. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Statistical Association*. 1999;94:496-509. doi:10.1080/01621459.1999.10474144.
 48. van Putten W. CART: Stata module to perform Classification And Regression Tree analysis. 2006. <https://EconPapers.repec.org/RePEc:boc:bocode:s456776>.