Human Pathology (2019) **89**, 81–89

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Budding, tumor-infiltrating lymphocytes, gland formation: scoring leads to new prognostic groups in World Health Organization low-grade colorectal cancer with impact on survival ☆

Corinna Lang-Schwarz MD [a,*,1], Balint Melcher MD [a,1], Franziska Haumaier MSc [a,1], Anna Schneider-Fuchs PhD, MSc [a,1], Klaus Lang-Schwarz MD [b,2], Jens Krugmann MD [a,1], Michael Vieth MD [a,1], William Sterlacci MD [a,1]

[a] *Institute of Pathology, Klinikum Bayreuth, 95445 Bayreuth, Germany*
[b] *Department of Anesthesiology, Klinikum Bayreuth, 95445 Bayreuth, Germany*

Received 13 February 2019; revised 11 April 2019; accepted 25 April 2019

Keywords:
Gland formation;
Budding;
Tumor-infiltrating lymphocytes (TILs);
Grading;
Colorectal cancer

Summary Grading for colorectal carcinoma (CRC) is traditionally based on the percentage of gland formation. In recent years, high-grade CRC has become subject to more precise molecular grading strategies. Most, however, are low-grade cases according to the World Health Organization (WHO) with inhomogenous outcomes due to still insufficient characterization. On the other hand, budding and tumor-infiltrating lymphocytes have developed as interesting additive prognostic factors in CRC. Especially budding has been very well defined by the International Tumor Budding Consensus Conference recently. We analyzed a large collective of 576 WHO low-grade CRC cases, stages I to IV, diagnosed between 2005 and 2016 in terms of gland formation, budding, and tumor-infiltrating lymphocytes and developed a new, morphology-based risk score, taking into account each of the 3 parameters. For each parameter, 1 to 2 points were given, resulting in a sum score, dividing the CRC cases into a low-, an intermediate-, and a high-risk group. By our score, 179 (34.9%) of the cases were grouped as low risk, 241 (53.5) as intermediate risk, and 92 (35.5%) as high risk. The 3 groups differed significantly in pT, pN, and M as well as tumor stages, lymphatic vessel invasion, venous invasion, and overall survival (0.;$P < .001$ for low risk versus high risk, $P = .038$ for low versus intermediate risk, and $P = .036$ for intermediate versus high risk; log rank: median, 94.0 months [95% confidence interval {CI}, 74.9-113.1] for low risk; median, 63.0 months [95% CI, 44.0-82.0] for intermediate risk; and median, 40.0 months [95% CI, 23.4-56.7] for high risk) in Kaplan-Meier-analysis. Our proposed Bayreuth score enables separating the large group of WHO low-grade CRC cases into subgroups, which differ significantly in outcome.

☆ Disclosures: The authors declare that they have no conflict of interest. No funding or support is associated with this study.

* Corresponding author. Institute of Pathology, Klinikum Bayreuth, Preuschwitzer Str. 101, 95445 Bayreuth, Germany
E-mail address: Corinna.Lang-Schwarz@klinikum-bayreuth.de (C. Lang-Schwarz).
[1] Tel.: +49 921 400 5602; fax: +49 921 400 5609.
[2] Tel.: +49 921 400 5102; fax: ++49 921 400 5109.

https://doi.org/10.1016/j.humpath.2019.04.006
0046-8177/

1. Introduction

Colorectal cancer (CRC) is one of the most common cancer types. In 2018, approximately 1.8 million new cases of CRC occurred worldwide, representing the third most commonly diagnosed malignancy in men and the second most in women. With almost 900,000 people dying from CRC in 2018, it is the third most common cause of cancer deaths worldwide [1].

Histopathologic grading is a significant prognostic factor for CRC, independent of the TNM stage [2,3].

Grading for CRC is based on the percentage of gland formation. Traditionally, it was 3-tiered: well-differentiated (grade 1) lesions showed glandular structures in more than 95% of the tumor and moderately differentiated (grade 2) CRC in 50% to 95%, and high-grade (grade 3) CRC showed glandular structures in <50% of the tumor [4]. As gland formation–based grading suffers from significant interobserver variability [5,6], in the current third volume of the fourth World Health Organization (WHO) edition, well- and moderately differentiated CRCs are summarized as one low-grade group because of similar behavior and better interobserver agreement [7].

Grading is also meant to lend prognostic information regarding tumor biology. A subgroup of morphologically high-grade CRCs, like medullary carcinoma, have significant better outcome because of mismatch repair (MMR) deficiency [8-10]. Also, a subgroup of the high-grade CRCs of no special type and some special subtypes like mucinous CRC, graded as grade 3 in conventional grading, have significant better prognosis if they show high microsatellite instability (MSI) [11]. Therefore, high-grade CRC and special subtypes are already subject to a more precise molecular grading based on their MMR status [12].

In studies with large patient collectives, most CRC cases, though belong to the large and biologically heterogenous group of WHO low-grade CRCs, and more patients die of low-grade CRC than of high-grade CRC, simply because it is much more common and still not sufficiently characterized [13,14].

On the other hand, more and more is known about promising new morphologic additive prognostic markers like budding and tumor-infiltrating lymphocytes (TILs). In our previous work, we could already show that the combination of both parameters allows to establish subgroups of CRC cases with impact on prognosis [14].

For a long time, it has been daily routine to apply grading systems to some cancer types that are based on more than one morphologic parameter, resulting in a summary score, for example, the Elston and Ellis grading in breast cancer [15] or the Fédération Nationale des Centres de Lutte Contre le Cancer grading [16] in sarcoma. Especially in breast cancer grading, tubulus formation is only 1 of 3 parameters, although it is also an adenocarcinoma.

Because high-grade CRC is subject to molecular analyses, we aimed to put the focus on the large, inhomogeneous group of WHO low-grade cases, to determine if the combination of the amount of gland formation, budding, and TILs, which were each analyzed on hematoxylin and eosin (H&E)–stained slides, allows to further characterize this large, inhomogeneous group. Also, we analyzed if summarizing 1 traditional grading parameter and 2 new promising additive prognostic markers into one grading system provides benefits for CRC-patients concerning outcome and survival.

2. Materials and methods

2.1. Case selection

A research in our institutional database provided 783 cases of CRC, stages I IV, diagnosed between 2005 and 2016. Five hundred seventy-six (73.6%) were WHO low grade (grades 1 and 2) and included in the study. One hundred eighty (23%) cases were WHO high grade (grade 3) and were not included. Sixty-four cases (8.2%) showed special subtypes, such as mucinous, serrated, solid medullary carcinoma, or Crohn-like inflammation, and were also not included because they were graded molecularly. Neoadjuvant treatment modalities were also an exclusion criterion, which explains the low rate of rectal carcinomas in our collective. Patients with inflammatory bowel disease–related carcinoma were also not included.

Our WHO low-grade cohort included 576 cases, with 317 men and 259 women. The mean age at diagnosis was 72.8 years (range, 44-97 years). Median follow-up was 42 months (range, 0-138 months). Follow-up data were provided from the local tumor registry in Bayreuth. A complete follow-up was available for 513 cases.

Two hundred seventy-two patients were alive at study end, and 241 died. Further tumor characteristics and TNM stages are listed in Table 1. The ethics commission approved the study (study number 239_18 Bc).

2.2. Histological evaluation

H&E-stained tumor slides of all patients were retrieved from our archives. The slides were reevaluated independently in terms of budding according to the defined criteria of the International Tumor Budding Consensus Conference by 2 different pathologists (C. L. S., B. M.) using Olympus BX 53 (C. L. S.) and BX 46 (B. M.) microscopes (Olympus, Hamburg, Germany), respectively, as described [17]. Budding was reported as proposed: low budding, 0-4 buds; intermediate budding, 5 to 9 buds; and high budding, >10 buds. Furthermore, the percentage of tumor-associated stromal lymphocytic infiltration was also semiquantitatively estimated on the same slides by the 2 pathologists, analog to the defined criteria by Salgado et al [18] in breast cancer. To receive higher interobserver accordance, a subset of the cases (n = 50) served as training slides first and were viewed together between both observers on a multihead microscope.

Table 1 Summary of patient and tumor characteristics

Features	Frequency, n (%)
Age (y; n = 576), mean	72.8 (44-97)
Sex (n = 576)	
Male	317 (55.0)
Female	259 (45.0)
pT (n = 576)	
pT1	38 (6.6)
pT2	97 (16.8)
pT3	340 (59.0)
pT4	101 (17.5)
pN (n = 576)	
pN0	350 (60.8)
pN1	145 (25.2)
pN2	81 (14.1)
M (n = 576)	
M0	489 (84.9)
M1	87 (15.1)
TNM stage (n = 576)	
I	116 (20.1)
II	221 (38.4)
III	147 (25.5)
IV	92 (16.0)
Tumor location (right/left; n = 576)	
Right	365 (63.4)
Left	211 (36.6)
Grading (WHO 2010; n = 576)	
1	29 (5.0)
2	547 (95.0)
Venous invasion (n = 576)	
V0	469 (81.4)
V1	103 (17.9)
V2	4 (0.7)
Lymphatic invasion (n = 576)	
L0	377 (65.5)
L1	199 (34.5)
MMR status (n = 36)	
MMR proficient	27 (75.0)
MMR deficient	9 (25.0)
RAS (n = 112)	
Wild type	57 (50.9)
Mutated	55 (49.1)

The observers were trained to find a consensus on the percentage of TILs. When a consensus was found in most cases, each observer viewed the slides on his own in the second step. All tumor slides were scanned in a 200-fold magnification (ocular ×10, objective ×20), and the average percentage amount of stromal TILs was reported as a continuous variable, including TILs within the border of invasion. A full assessment of average TILs in the tumor area was used without focusing on hotspots. Neutrophilic infiltrates were not taken into account. For statistical analyses, a cutoff at 5% TILs was used as a discrimination threshold between 2 patient groups, for which we could show prognostic impact by receiver operating characteristic curve analysis in our previous work [14]. The slides were also reevaluated concerning traditional, gland formation–based grading, to exclude all cases that did not fulfill the inclusion criteria (C. L. S., B. M.).

2.3. Molecular analyses

MMR protein expression was assessed for a subset of the cases (n = 36; 6.3%) by immunohistochemistry for MLH1, PMS2, MSH2, and MSH6. Tumors were considered MMR proficient if tumor nuclear staining was positive, and MMR deficient if tumor nuclear staining was negative (<10%) with positive onslide controls.

RAS, EGFR, and BRAF mutation analyses were performed for a subset of the cases (n = 112 [19.4%], n = 29 [5.0%], n = 78 [13.5%]) on the Illumina MiSeq next generation sequencer using the TruSight Tumor 15 Assay from Illumina (Illumina, San Diego, USA).

2.4. Evaluation of proposed Bayreuth score

For the 3 parameters gland formation, budding, and TILs, score points were given as follows: >95%, 1 point, and 95% to 50%, 2 points, for gland formation (traditional WHO grading); and low, 1 point, and intermediate or high, 2 points, for budding; and >5%, 1 point, and ;≤5%, 2 points, for TILs.

In the second step, a summary score was established by adding the 3 single-score points resulting in 3 groups: summary score 3 to 4, which we termed "low risk"; score 5, which we termed "intermediate risk"; and score 6, which we termed "high risk."

Cases with 3 and 4 score points were summarized as one low-risk group because of the low number of cases with 3 points. The algorithm and stepwise explanation of our proposed score is shown in Figs. 1 and 2 and also illustrated by images in Fig. 3.

2.5. Statistics

Statistical analyses were performed with the statistics program SPSS 21 (IBM, Armonk, NY). Pearson χ^2 test was used to test the relationship between different parameters. Univariate survival analysis was carried out using the Kaplan-Meier method and log-rank test. Multivariate survival analysis was performed using the Cox regression analysis. Hazard ratios and 95% confidence interval (CI) were used to determine effect size. *P* values 0;of less than .05 were considered statistically significant.

3. Results

3.1. Budding

We found low budding in 391 (67.9%), intermediate budding in 151 (26.2), and high budding in 34 (5.9%) of the

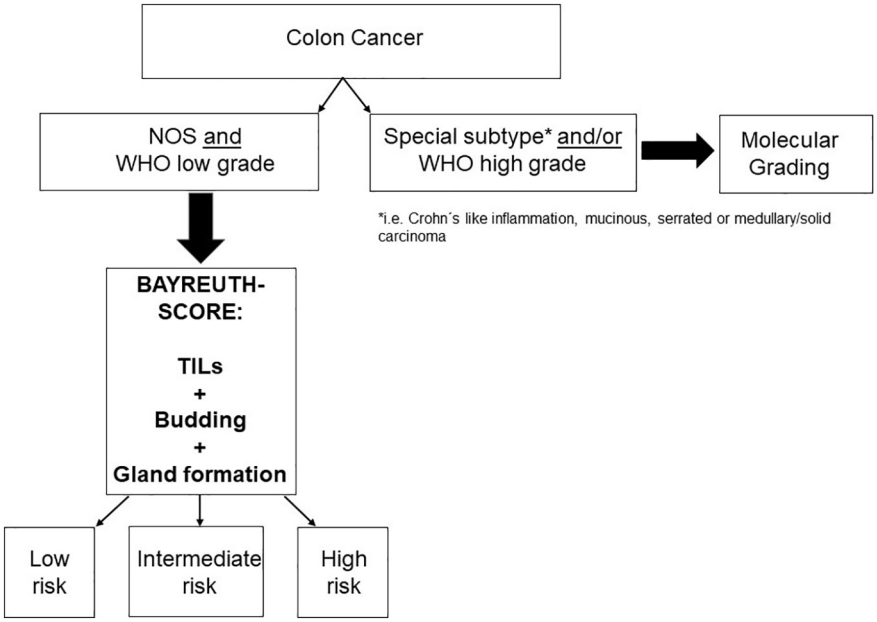

Fig. 1 Algorithm of proposed Bayreuth score. In the first step, cases are graded according to WHO. WHO high-grade cases are further characterized by molecular grading. WHO low-grade cases are subject to Bayreuth score.

cases. Higher budding, as expected, correlated with higher pT stages ($P = .008$), pN stages ($P < .001$), M stages ($P = .010$), TNM stages (0.;$P < .001$), overall survival (OS; $P = .004$), and lymphatic vessel invasion (0.;$P < .001$). No correlation was found with venous invasion ($P = .212$). As in our previous study, cases with intermediate and high budding were summarized as one high-grade group because of the low number of cases with high budding. Cases with low budding showed a trend to longer OS in Kaplan-Meier-analysis ($P = .063$, log rank) compared with the group of cases with intermediate or high budding. Cox regression analysis revealed a risk ratio of 0.245 for cases with intermediate or high budding versus cases with low budding (95% CI, 0.984-1.657; $P = .069$).

3.2. Gland formation

Because the aim of this study was to focus on WHO low-grade CRC cases, first of all, all 783 cases were reevaluated in terms of gland formation according to the old, 3-tiered, as well as the current, 2-tiered WHO grading scheme by the 2 observers (C. L. S./B. M.), to sort out high-grade (G3) cases

BAYREUTH-SCORE

Gland formation		Budding		TILs	
Gland formation in %	Score-points	Amount of budding according to ITBCC criteria	Score-points	Percentage of TILs	Score-points
>95 (WHO grade 1)	1	low (Bd 1)	1	>5	1
95-50% (WHO grade 2)	2	Intermediate (Bd2) or high (Bd3)	2	≤5%	2

Total score: 3 or 4 → low risk

Total score: 5 → intermediate risk

Total score: 6 → high risk

Fig. 2 Grading, budding, and TILs: 1 to 2 score points are given for each parameter resulting in a sum score that defines the risk group.

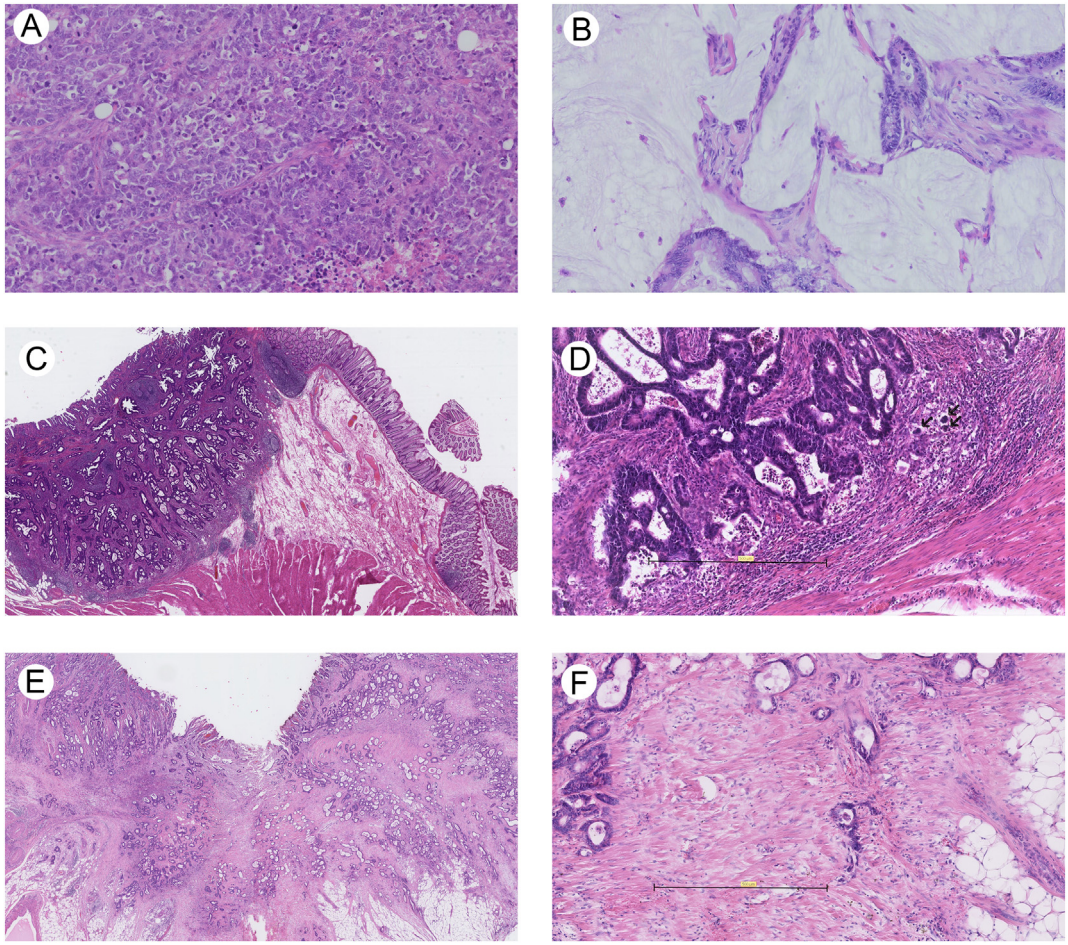

Fig. 3 Representative H&E images of our proposed algorithm. WHO high-grade cases like grade 3, not otherwise specified (NOS) carcinoma (A) or special subtypes like mucinous carcinoma (B) are excluded from further analysis and are subject to molecular grading by microsatellite analysis (magnification ×203, ocular ×10, objective ×20.3). C and E, Overviews of WHO low-grade (grade 2) not otherwise specified (NOS)-CRC (magnification ×14.7, ×ocular 10, objective ×1.47). D and F, Higher magnification of the 2 carcinomas with obvious differences in TIL amount as well as budding between the 2 cases (D: TILS >5%, buds are highlighted by black arrows; F: TILs; ≤5%, no budding; magnification ×120, ocular ×10, objective ×12).

from further analyses. Scoring all 783 CRC cases between the observers according to the 3-tiered WHO grading scheme (G1-G3) resulted in a κ value of 0.560 (Pearson, $P < .001$) corresponding to moderate concordance. Comparing only G1 versus G2 cases resulted in a κ value of 0.268 (Pearson, $P < .001$), which is fair concordance. Comparing the cases according to the current WHO grading scheme (low grade [G1/G2] versus high grade [G3]) resulted in a κ value of 0.733 (Pearson, 0.;$P < .001$), which is substantial concordance.

Focusing on the 576 WHO low-grade cases, 547 (95%) of the cases were moderately differentiated (WHO grade 2), and 29 cases (5%) were well differentiated (WHO grade 1). On univariate analysis, WHO grade 1 was significantly associated with better pT stages (0.;$P < .001$), pN stages ($P = .044$), TNM stage (0.;$P < .001$), lymphatic vessel invasion ($P = .010$), and especially M stages (0% distant metastases in grade 1 cases compared with n = 87/460 in grade 2 cases). As with budding, no correlation was found with venous invasion ($P = .249$). Kaplan-Meier analysis did not reveal better survival for the WHO grade 1 cases compared with WHO grade 2 ($P = .330$, log rank). Cox regression analysis revealed a risk ratio of 0.276 for WHO grade 2 versus WHO grade 1 (95% CI, 0.752-2.309; $P = .316$).

3.3. Tumor-infiltrating lymphocytes

The percentage of TILs >5% versus ;≤5% showed significant advantages for the higher TILs group concerning the parameters pT stage (0.;$P < .001$), pN stage (p<0,001), M stage (0.;$P < .001$), TNM stage (0.;$P < .001$), lymphatic vessel invasion (0.;$P < .001$), and also venous invasion ($P = .021$). In Kaplan-Meier analysis, tumors with >5% TILs had a significantly longer OS (0.;$P < .001$, log rank; medians, 95.0 months [95% CI, 71.8-98.2] for >5% TILs versus 49.0 months [95% CI, 37.9-60.1] for ;≤5% TILs).

In multivarate analyses, no significant correlations were found between grading (grade 1 or 2) and the parameters budding and TILs (Pearson, $P = .198$ and $P = .636$). Also, no correlation was found concerning the parameters budding and TILs (Pearson, $P = .206$).

Also, there was no correlation between budding, gland formation, and TILs with MSI immunohistochemistry, and RAS, EGFR, and BRAF analyses.

3.4. Scoring: gland formation, budding, and TILs

Applying our proposed Bayreuth score to the 576 cases revealed the following results: 9 cases (1.6%) had the minimum possible sum score of 3 points, 192 cases (33.3%) had 4 points, 272 cases (47.2%) had a total score of 5 points, and the remaining 103 cases (17.9%) had the maximum total score of 6 points. Lower sum scores correlated significantly with lower pT stages (0.;$P < .001$), lower pN stages (0.;$P < .001$), lower M stage (0.;$P < .001$), lower TNM stage (0.;$P < .001$), less lymphatic vessel invasion (0.;$P < .001$), and also less venous invasion ($P = .007$). Interestingly, cases with lower sum scores were significantly more often localized in the right-sided hemicolon (Pearson, $P = .021$) and vice versa. Male patients belonged significantly more often to the intermediate- and high-risk groups than did female patients ($P = .018$). In Kaplan-Meier analysis, the difference in OS between all cases (3-6 points) was statistically significant ($P = .002$). Also, we found significantly better OS for cases with 4 points compared with cases with 5 points ($P = .043$, log rank) as well as for cases with 5 points compared with cases with 6 points ($P = .036$, log rank). No significant difference in OS was found between the 3-point group and the 4-point group ($P = .865$; Fig. 4).

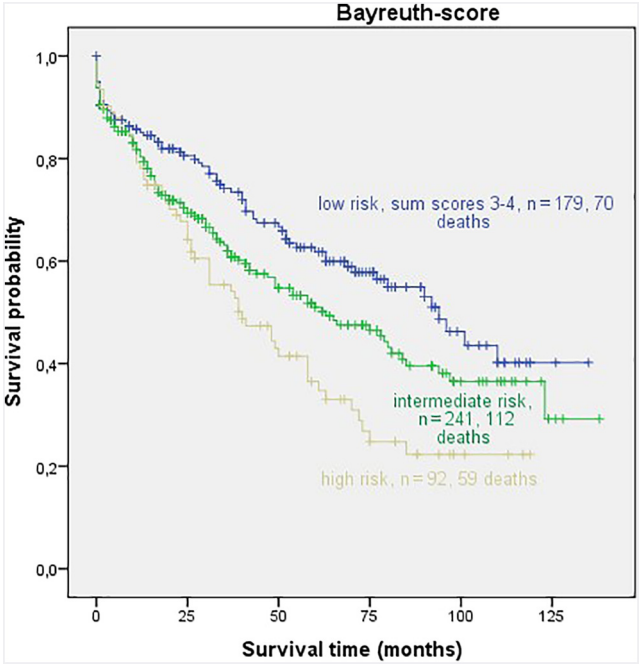

Fig. 5 Kaplan-Meier analysis showing significant differences in OS between each of the 3 subgroups defined by Bayreuth score. Low-risk cases (sum scores 3 and 4, blue) do better than intermediate-risk cases (sum score 5, green) and high-risk cases (sum score 6, gray; 0.;$P < .001$, log rank). The differences between low- and intermediate-risk cases as well as between intermediate- and high-risk cases are also significant ($P = .038$ and $P = .036$, log rank, respectively).

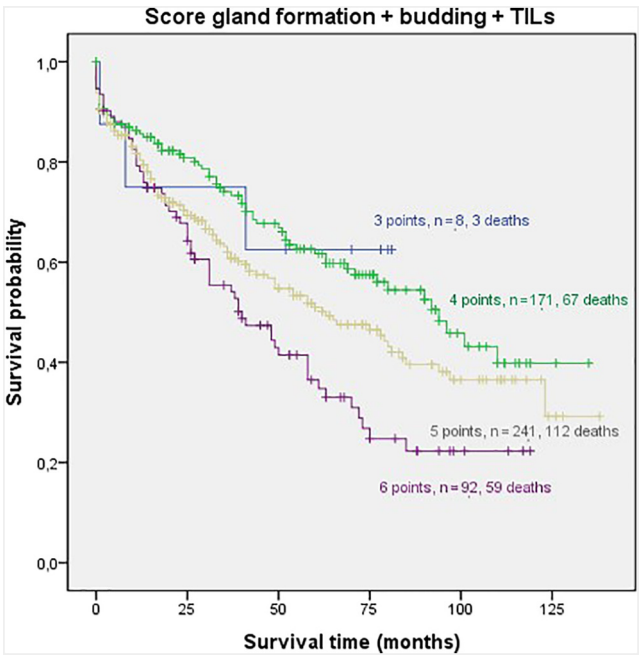

Fig. 4 Kaplan-Meier curve showing differences in OS for sum scores 3 to 6 (score 3, blue; score 4, green; score 5, gray; and score 6, purple [$P = .002$, log rank]).

Summarizing cases with 3 and 4 points to one low-risk group resulted in 201 (34.9%) low-risk cases. The low-risk group differed significantly in OS from the intermediate-risk group ($P = .038$, log rank) and the high-risk group (0.;$P < .001$, log rank). OS for the intermediate-risk group was also significant better than for the high-risk group ($P = .036$, log rank). Median OS for low-risk cases was 94.0 months (95% CI, 74.9-113); for intermediate-risk cases, 63.0 months (95% CI, 44.0-82.0); and for high-risk cases, 40.0 months (95% CI, 23.4-56.7). Kaplan-Meier analysis for our proposed Bayreuth score is shown in Fig. 5.

Correlations between our proposed score and various clinicopathological parameters are shown in Table 2. No correlations were found with the molecular parameters MSI, RAS, EGFR, and BRAF. Multivariable analysis revealed independent prognostic effects for our proposed Bayreuth score with pT, pN, M, and TNM stages; venous invasion; and age. The results of the multivariable analysis with the clinicopathological features are shown in Table 3.

Table 2 Association of proposed Bayreuth score with clinicopathological features

Features	Bayreuth score				*P* value low *vs* intermediate risk	*P* value low *vs* high risk	*P* value intermediate *vs* high risk	*P* value (overall)
	Low risk (n/%)		Intermediate risk (n/%)	High risk (n/%)				
	3 points	4 points	5 points	6 points				
Age (y; n = 576), mean	70.3	73.2	73.0	71.4	0.974	0.171	0.157	0.221
Sex (n = 576)								
Male	7 (2.2)	88 (27.8)	161 (50.8)	61 (19.2)	.010 *	.049 *	.996	.018 *
Female	2 (0.8)	104 (40.2)	111 (42.9)	42 (16.2)				
pT (n = 576)								
pT1	6 (15.8)	23 (60.5)	9 (23.7)	0 (0.0)	<.001 *	<.001 *	.006 *	<.001 *
pT2	1 (1.0)	45 (46.4)	39 (40.2)	12 (12.4)				
pT3	1 (0.3)	105 (30.9)	174 (51.2)	60 (17.6)				
pT4	1 (1.0)	19 (18.8)	50 (49.5)	31 (30.7)				
pN (n = 576)								
pN0	6 (1.7)	151 (43.1)	152 (43.4)	41 (11.7)	<.001 *	<.001 *	.003 *	<.001
pN1	3 (2.1)	31 (21.4)	76 (52.4)	35 (24.1)				
pN2	0 (0.0)	10 (12.3)	44 (54.3)	27 (33.3)				
M (n = 576)								
M0	9 (1.8)	183 (37.4)	220 (45.0)	77 (15.7)	<.001 *	<.001 *	.193	<.001 *
M1	0 (0.0)	9 (10.3)	52 (59.8)	26 (29.9)				
TNM stage (n = 576)								
I	5 (4.3)	60 (51.7)	41 (35.3)	10 (8.6)	<.001 *	<.001 *	.012 *	<.001
II	1 (0.5)	90 (40.7)	101 (45.7)	29 (13.1)				
III	3 (2.0)	32 (21.8)	77 (52.4)	35 (23.8)				
IV	0 (0.0)	10 (10.9)	53 (57.6)	29 (31.5)				
Tumor location (right/left; n = 576)								
Right	8 (2.2)	132 (36.2)	166 (45.5)	59 (16.2)	.053	.032 *	.510	.021 *
Left	1 (0.5)	60 (28.4)	106 (50.2)	44 (20.9)				
Grading (WHO 2010; n = 576)								
1	9 (31.0)	17 (58.6)	3 (10.3)	0 (0.0)	<.001 *	<.001 *	.286	<.001 *
2	0 (0.0)	175 (32.0)	269 (49.2)	103 (18.8)				
Venous invasion (n = 576)								
V0	8 (1.7)	167 (35.6)	218 (46.5)	76 (16.2)	.045 *	.007 *	.298	.007 *
V1	1 (1.0)	24 (23.3)	51 (49.5)	27 (26.2)				
V2	0 (0.0)	1 (25.0)	3 (75.0)	0 (0.0)				
Lymphatic invasion (n = 576)								
L0	8 (2.1)	155 (41.1)	167 (44.3)	47 (12.5)	<.001 *	<.001 *	.006	<.001 *
L1	1 (0.5)	37 (18.6)	105 (52.8)	56 (28.1)				
MMR status (n = 36)								
MMR proficient	0 (0.0)	3 (33.3)	13 (48.1)	5 (18.5)	.931	.874	.249	.896
MMR deficient	0 (0.0)	3 (33.3)	4 (44.4)	2 (22.2)				
RAS (n = 112)								
Wild type	0 (0.0)	11 (19.3))	33 (57.9)	13 (22.8)	.570	.312	.505	.307
Mutated	0 (0.0)	14 (25.5)	32 (58.2)	9 (16.4)				

* Statistically significant values (0.;$P < .05$).

4. Discussion

Today for CRC, estimating the amount of gland formation semiquantitatively serves as a triage to filter out WHO high-grade (grade 3) cases for further molecular grading [12]. Molecular pathology, though, failed to further characterize the much larger group of WHO low-grade cases the way it did in high-grade cases. Concerning classical, mostly H&E-based morphology, budding and TILs have been established as promising morphologic additive prognostic factors during the last years. As a reaction to the defined International Tumor Budding Consensus Conference consensus criteria, budding has recently been included as an additional reporting parameter in the protocol of the College of American

Table 3 Comparison of relative risks of Bayreuth score in multivariable analysis with clinicopathological features

Feature	Bayreuth score	
	HR (95% CI)	*P*
Score category		
Low risk	1.0	.001
Intermediate risk	1.38 (1.02-18.6)	
High risk	1.94 (1.37-2.75)	
Age (years)		
Baseline	1.0	
Increasing by year	1.04 (1.02-1.05)	<.001
Sex		
Male	1.0	.027
Female	0.75 (0.57-0.97)	
pT		
4	1.0	<.001
1	0.33 (0.14-0.65)	
2	0.37 (0.23-0.57)	
3	0.57 (0.42-0.78)	
pN		
0	1.0	<.001
1	1.39 (1.01-1.88)	
2	2.36 (1.67-3.33)	
M		
0	1.0	<.001
1	3.15 (2.32-4.27)	
TNM stage		
IV	1.0	<.001
I	0.22 (0.14-0.34)	
II	0.30 (0.21-0.43)	
III	0.37 (0.26-0.53)	
Tumor location (right/left)		
Right	1.0	.620
Left	1.07 (0.82-1.39)	
Grading (WHO 2010)		
1	1.0	.887
2	1.04 (0.58-1.88)	
Venous invasion		
0	1.0	<.001
1	1.92 (1.43-2.58)	
2	6.11 (2.23-16.80)	
Lymphatic invasion		
0	1.0	.060
1	1.30 (0.99-1.69)	
MMR status		
MSS	1.0	.967
MSI	1.035 (0.21-5.23)	
RAS		
Wild type	1.0	.273
Mutated	1.43 (0.76-2.69)	

Abbreviations: HR, hazard ratio, MSS - microsatelite stable, MSI - microsatelite instable.

Pathologists in pT1 and stage II CRC [19]. As we could show in our previous work, budding is now well reproducible with high interobserver correlation [14]. For the assessment of the amount of inflammation in CRC, many similar studies refer to the H&E-based Klintrup-Mäkinen score [20-23]. In their 4-tiered scale, they took not only TILs into account but also neutrophilic and eosinophilic granulocytes. From our experience, many CRCs present with heavily florid inflammation, especially at the invasive margin with abscess formations, which—in our opinion—are not equal to the host-responsive effect of TILs. Also, there are approaches to quantify the amount of TILs by immunohistochemistry and computer-assisted systems [21,24,25]. Recently, results were presented from a worldwide taskforce that prospectively validated the so-called immunoscore by a combination of immunohistochemistry with CD3 and CD8 as well as digital pathology [26]. As in our previous work, we characterized the amount of TILs as defined in the consensus article by Salgado et alE morphology, excluding polymorphonuclear leukocytes from analysis and focussing on stromal lymphocytes; [18] for breast cancer, based on H&E morphology, excluding polymorphonuclear leukocytes from analysis and focusing on stromal lymphocytes. This method, although H&E based, revealed highly significant differences between the group with >5% ;and ≤5% TILs concerning the clinocopathological parameters as well as OS. Nevertheless, we point out that it will be of interest to see if using immunohistochemistry and digital image analysis for measuring TILs is able to add additional information in a next step.

In 2009, Lugli et al [21] made a similar approach and proposed a CD8+ lymphocytes/tumor budding index with impact on prognosis, but budding was not as well defined then. Since then, there have been several approaches to describe the relationship between budding and tumor microenvironment [13,22,27].

To our knowledge, our study is the first that integrates budding and TILs in a combined scoring system and includes gland formation as a traditional grading parameter.

Compared with the conventional, gland formation only–based WHO grading, our proposed score showed higher significant correlations when tested against the histopathologic parameters pN (P = .013 versus <0.001, Pearson), M (P = .021 versus <0.001), and lymphatic invasion (P = .010 versus <0.001). It even revealed a correlation with venous invasion, which conventional grading did not (P = .097 versus P = .007). Our score also showed better OS for the low-risk group (sum scores 3 to 4), which conventional grading in our collective did not (P = .330 versus P = .001, log rank).

Further interesting results of our study are that male patients belonged significantly more often to the high-risk group than did female patients and that cases with lower sum scores were significantly more often localized in the right-sided hemicolon and vice versa, both phenomena for which we do not have an explanation yet.

The amount of gland formation, budding, and TILs did not correlate with MSI immunohistochemistry, RAS, EGFR, and BRAF in our study, which fits to the fact that molecular grading is not the appropriate method of grading for patients with WHO low-grade CRC.

We show a simple to reproduce, routinely applicable, cost-effective, H&E-based method that integrates the promising new additive prognostic markers budding and TILs into

one grading system. Our proposed score allows to further characterize the large and inhomogeneous group of WHO low-grade CRC with impact on survival. Further studies are necessary to validate the score in prospective series and evaluate, if adding immunohistochemistry or digital pathology for TILs assessment really brings additional information.

References

[1] Ferlay J, Ervik M, Lam F, et al. Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer; 2018 Available from: https://gco.iarc.fr/today Accessed June 12, 2018.

[2] Freedman LS, Macaskill P, Smith AN. Multivariate analysis of prognostic factors for operable rectal cancer. Lancet 1984;29:733-6.

[3] Fisher ER, Sass R, Palekar A, Fisher B, Wolmark N. Dukes´classification revisited. Findings from the National Surgical Adjuvant Breast and Bowel Projects (Protocol R-01). Cancer 1989;64:2354-60.

[4] Hamilton SR, Vogelstein B, Kudo S, et al. In: Hamilton SR, Aaltonen LA, editors. Carcinoma of the Colon and Rectum. Lyon: International Agency for Research on Cancer; 2000. p. 110-1.

[5] Thomas GD, Dixon MF, Smeeton NC, Williams NS. Observer variation in the histological grading of rectal carcinoma. J Clin Pathol 1983;36:385-91.

[6] Chandler I, Houlston RS. Interobserver agreement in grading of colorectal cancer-findings from a nationwide web-based survey of histopathologists. Histopathology 2008;52:494-9.

[7] Hamilton SR, Bosman FT, Boffetta P, et al. In: Bosman FT, Carneiro F, Hruban RH, Theise ND, editors. Carcinoma of the Colon and Rectum. Lyon: International Agency for Research on Cancer; 2010. p. 138.

[8] Wright CM, Dent OF, Barker M, et al. Prognostic significance of extensive microsatellite instability in sporadic clinicopathological stage C colorectal cancer. Br J Surg 2000;87:1197-202. https://doi.org/10.1046/j.1365-2168.2000.01508.

[9] Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005;23:609-18.

[10] Benatti P, Gafà R, Barana D, et al. Microsatellite instability and colorectal cancer prognosis. Clin Cancer Res 2005;11:8332-40. https://doi.org/10.1158/1078-0432.CCR-05-1030.

[11] Lanza G, Gafà R, Matteuzzi M, Santini A. Medullary-type poorly differentiated adenocarcinoma of the large bowel: a distinct clinicopathologic entity characterized by microsatellite instability and improved survival. J Clin Oncol 1999;17:2429-38. https://doi.org/10.1200/JCO.1999.17.8.2429.

[12] Neumann JH, Jung A, Kirchner T. Molecular pathology of colorectal cancer. Der Pathologe 2015;36 (2): 137-144. https://doi.org/10.1007/s00292-015-0005-3

[13] Eriksen AC, Sorensen FB, Lindebjerg J, Hager H. The prognostic value of tumour stroma ratio and tumour budding in stage II colon cancer. A nationwide population-based study. Int J Colorect Dis 2018;33:1115-24. https://doi.org/10.1016/j.tranon.2018.03.008.

[14] Lang-Schwarz C, Melcher B, Haumaier F, et al. Budding and tumor-infiltrating lymphocytes—combination of both parameters predicts survival in colorectal cancer and leads to new prognostic subgroups. Hum Pathol 2018;79:160-7. https://doi.org/10.1016/j.humpath.2018.05.010.

[15] Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:403-10.

[16] Trojani M, Contesso G, Coindre JM, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. Int J Cancer 1984;33:37-42.

[17] Lugli A, Kirsch R, Ajioka Y, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol 2017;30:1299-311. https://doi.org/10.1038/modpathol.2017.46.

[18] Salgado R, Denkert C, Demaria S, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2015. Ann Oncol 2014;26:259-71. https://doi.org/10.1093/annonc/mdu450.

[19] Kakar S; Shi C; Berho ME et al. College of American Pathologists: Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum. Version 4.0.1.0. Available at: https://cap.objects.frb.io/protocols/cp-gilower-colonrectum-17protocol-4010.pdf. Accessed January 5, 2019.

[20] Klintrup K, Mäkinen JM, Kauppila S, et al. Inflammation and prognosis in colorectal cancer. Eur J Cancer 2005;41(17):2645-54. https://doi.org/10.1016/j.ejca.2005.07.017.

[21] Lugli A, Karamitopoulou E, Panayiotides I, et al. CD8+ lymphocytes/tumour-budding indes: an independent prognostic factor representing a 'pro-/anti-tumour' approach to rumour host interaction in colorectal cancer. Br J Cancer 2009;101:1382-92. https://doi.org/10.1038/sj.bjc.6605318.

[22] Park JH, van Wyk H, Roxburgh CSD, Horgan PG, Edwards J, McMillan DC. Tumour invasiveness, the local and systemic environment and the basis of staging systems in colorectal cancer. Br J Cancer 2017;116:1444-50. https://doi.org/10.1038/bjc.2017.

[23] Dawson IH, Galuppini F, Träger RS, et al. Validation of the international tumor budding consensus conference (ITBCC 2016) recommendations on tumor budding in stage I-IV colorectal cancer. J Clin Oncol 2019;85:145-51.

[24] Lee WS, Park S, Lee WY, Yun SH, Chun HK. Clinical impact of tumor-infiltrating lymphocytes for survival in stage II colon cancer. Cancer 2010;Nov 15;116(22):5188-99. https://doi.org/10.1002/cncr.25293.

[25] Eriksen AC, Andersen JB, Kristensson M, dePont Christensen R, Hansen TF, Kjaer-Frifeldt S, Sorensen FB. Computer assisted stereology and automated image analysis for quantification of tumor infiltration lymphocytes in colon cancer. Diagn Pathol 2017;29;12(1):65. https://doi.org/10.1186/s13000-017-0653-0

[26] Pagès F, Mlecnik B, Marliot F et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 2018;26;391:2128-39. https://doi.org/10.1016/S0140-6736(18)30789-X

[27] VanWyk HC, Park JH, Edwards J, Horgan PG, McMillan DC. The relationship between tumour budding, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Br J Cancer 2016;115:156-63. https://doi.org/10.1002/ijc.32045.