



Published in final edited form as:

Am J Surg Pathol. 2018 June ; 42(6): 705–714. doi:10.1097/PAS.0000000000001059.

Poorly Differentiated Clusters Predict Colon Cancer Recurrence: An In-Depth Comparative Analysis of Invasive-Front Prognostic Markers

Tsuyoshi Konishi, MD^{*,†}, Yoshifumi Shimada, MD^{*,‡}, Lik Hang Lee, MD[§], Marcela S. Cavalcanti, MD[§], Meier Hsu, MS[#], J. Joshua Smith, MD, PhD^{*}, Garrett M. Nash, MD^{*}, Larissa K. Temple, MD^{*}, José G. Guillem, MD^{*}, Philip B. Paty, MD^{*}, Julio Garcia-Aguilar, MD, PhD^{*}, Efsevia Vakiani, MD, PhD[§], Mithat Gonen, PhD[#], Jinru Shia, MD[§], and Martin R. Weiser, MD^{*}

^{*}Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY

[§]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY

[#]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

[†]Department of Gastroenterological Surgery, Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo, Japan

[‡]Division of Digestive and General Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan

Abstract

This study aimed to compare common histologic markers at the invasive front of colon adenocarcinoma in terms of prognostic accuracy and interobserver agreement. Consecutive patients who underwent curative resection for stage I–III colon adenocarcinoma at a single institution in 2007–2014 were identified. Poorly differentiated clusters (PDCs), tumor budding (TB), perineural invasion (PN), desmoplastic reaction (DR), and Crohn-like lymphoid reaction (CLR) at the invasive front, as well as the World Health Organization (WHO) grade of the entire tumor, were analyzed. Prognostic accuracies for recurrence-free survival (RFS) were compared, and interobserver agreement among three pathologists was assessed. The study cohort consisted of 851 patients. Although all the histologic markers except WHO grade were significantly associated with RFS (PDCs, TB, PN and DR: $p < .001$; CLR: $p = 0.021$), PDCs (grade 1 [G1]: $n = 581$; G2: $n = 145$; G3: $n = 125$) showed the largest separation of 3-year RFS in the full cohort (G1: 94.1%; G3: 63.7%; hazard ratio [HR] 6.39; 95% confidence interval [CI] 4.11–9.95, $p < 0.001$), stage II patients (G1: 94.0%; G3: 67.3%; HR 4.15; 95% CI 1.96–8.82, $p < 0.001$), and stage III patients (G1: 89.0%; G3: 59.4%; HR 4.50; 95% CI 2.41–8.41, $p < 0.001$). PDCs had the highest prognostic accuracy for RFS with the concordance probability estimate of 0.642, while WHO grade had the lowest.

Correspondence: Jinru Shia, MD, Department of Pathology, Memorial Sloan Kettering Cancer Center, 1275 York Ave., New York, NY 10065. shiaj@mskcc.org. Phone: 212-639-7282. Martin R. Weiser, MD, Colorectal Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065. Weiser1@mskcc.org. Phone: 212-639-6698.

Conflicts of interest: none.

Supplemental Digital Content: Tables S1–S3.

Interobserver agreement was the highest for PDCs, with a weighted kappa of 0.824. The risk of recurrence over time peaked earlier for worse PDCs grade. Our findings indicate that PDCs are the best invasive-front histologic marker in terms of prognostic accuracy and interobserver agreement. PDCs may replace WHO grade as a prognostic indicator.

Keywords

colon cancer; invasive front; prognostic markers; poorly differentiated clusters

Introduction

The TNM staging system of the American Joint Committee on Cancer (AJCC) and the International Union Against Cancer (UICC) remains the most widely used prognostic indicator for patients with colon cancer.(1) However, recurrence and survival vary widely within each stage grouping.(2) The AJCC lists histologic markers to help stratify recurrence risk including histologic type, histologic grade, lymphovascular invasion, perineural invasion (PN), tumor deposits, extramural venous invasion, and tumor configuration.(1) Tumor heterogeneity, including areas of necrosis, however, often limits the standard evaluation of these histologic markers.

In recent years, histologic patterns at the invasive front have been correlated with tumor behavior and oncologic outcome.(3–5) The invasive front represents a dynamic interaction of tumor and host, with cellular dedifferentiation and single and clustered cells advancing and interacting with the host immune system.(4, 5)

Tumor budding (TB), which is defined as single cancer cells or clusters comprising < 5 cancer cells(6–8) and represents the dynamic process of epithelial-mesenchymal transition (EMT), is listed among poor-prognosis indicators by the European Society of Medical Oncology (ESMO)(9) and the Union for International Cancer Control (UICC).(10, 11) Another feature of EMT, poorly differentiated clusters (PDCs) of 5 or more cells (which have been investigated mainly in Japan(3, 4, 12) and Italy(13–15)), are reportedly a more accurate prognostic indicator than TB. Desmoplastic reaction (DR), which is the cancer-associated response of fibroblasts at the invasive front and which is critical to EMT,(16–18) is strongly associated with prognosis.(19) PN, which refers to the spread of cancer around a nerve, is a commonly used prognostic feature listed by ESMO(9) and UICC(10, 11) and included in a widely validated nomogram for predicting colon cancer recurrence.(2) Crohn-like lymphoid reaction (CLR), an enhanced host immune response at the invasive front,(20) is associated with better prognosis regardless of microsatellite instability status or TNM stage.(21)

Although these histologic markers have been studied individually, only limited analyses have compared these markers to identify those with the highest prognostic accuracy in predicting cancer recurrence and those with the highest diagnostic agreement among pathologists. Previous studies on these markers had limitations such as cohort heterogeneity and outdated chemotherapy regimens. Many were conducted in Japan and Italy, where the methods of processing pathological specimens, preoperative staging, follow-up protocols,

and recurrence rates differ from those in the United States. Our study was aimed at determining which markers are the most clinically useful in terms of prognostic accuracy and diagnostic agreement among pathologists. We analyzed data from a large cohort of patients with stage I–III colon adenocarcinoma treated by a multidisciplinary team including dedicated colorectal cancer surgeons at a comprehensive cancer center.

Materials and Methods

Study Design and Patient Cohort

The study was approved by the Institutional Review Board of Memorial Sloan Kettering Cancer Center. Clinicopathologic factors were collected from a prospectively maintained database at Memorial Sloan Kettering and data manually reviewed from an electronic medical record. The study cohort included consecutive patients who underwent curative colon (including rectosigmoid colon) resection for stage I, II, or III adenocarcinoma from January 2007 through December 2014 (n=1596). The exclusion criteria were history of treatment for malignancies other than colon cancer within the last 5 years (n=134), rectal cancer within 12 cm of the anal verge (n=196), appendiceal cancer (n=30), Tis or adenoma (n=25), other benign disease (n=8), preoperative chemotherapy or radiotherapy for colon cancer (n=17), noncurative palliative resection (n=2), stage IV disease at surgery (n=63), nonprimary recurrent or metastatic cancer (n=26), and unavailability of hematoxylin and eosin (H&E) stained slides (n=244).

Histologic Evaluation

All specimens were processed according to an institutional protocol concordant with the recommendations of the College of American Pathologists(22). Specimens were grossed by designated gross room assistants and reviewed by specialized gastrointestinal pathologists. There were no significant changes in how the specimens were processed during the study period. All available H&E-stained slides that included full-thickness sections of the tumor encompassing the deepest portion of the invasive front (mean, 5 tumor slides per patient; range, 1 to 13 slides per patient) were reviewed. This review was conducted by a Japanese pathologist (Y.S.) with expertise on histologic features at the invasive front,(3, 23–26) for consistency with previously described diagnostic criteria. The pathologist was blinded to clinical outcomes. A BX51 microscope (Olympus, Tokyo, Japan) with a standard 22-mm diameter eyepiece (specimen area of 0.950 mm² under an objective lens with a magnification of ×20) was used (27). For assessing interobserver agreement among pathologists on each of the pathologic parameters, a random sample of 50 patients was selected irrespective of specific histologic features as described previously (3) and evaluated independently by the Japanese expert pathologist (Y.S.) and two Memorial Sloan Kettering pathologists with no previous experience in assessing these parameters (L.H.L. and M.S.C). (Y.S. was the only pathologist who reviewed the entire cohort [N=851].)

All histologic markers were assessed as described previously.(3) Briefly, all histologic markers were assessed using longitudinal sections of the whole tumor including its deepest part. PDCs were defined as clusters of 5 cancer cells that lacked a gland-like structure. The whole tumor was first scanned at low-power magnification to identify areas with the greatest

number of PDCs, generally at the invasive front. The number of PDCs in a single field of highest activity was then determined and graded as G1 (<5 clusters), G2 (5–9 clusters), or G3 (10 clusters) under an objective lens with a magnification of $\times 20$ (Fig. 1A). TB was defined as a single cancer cell or a cluster of <5 cancer cells at the invasive front and was graded as G1 (<5 buds), G2 (5–9 buds), or G3 (10 buds) on the basis of the highest number of buds observed under an objective lens with a magnification of $\times 20$ (Fig. 1B).^(6, 28) PN was defined as a histologic finding of tumor cells invading into or spreading along nerve fascicles and was classified as intramural, extramural, or absent⁽²⁹⁾ (Fig. 1C). CLR was defined as nodular lymphoid aggregates of 1 mm or larger at the tumor periphery and was classified as present or absent⁽²⁰⁾ (Fig. 1D). DR was evaluated in the reactive fibrous zone at the invasive front for T3/4 tumors only, as described previously,⁽¹⁹⁾ and was classified as mature (fibrotic stroma composed of fine mature collagen fibers stratified into multilayers without presence of keloid-like or myxoid stroma), intermediate (keloid-like collagens intermingled with mature stroma), or immature (amorphous intermixed myxoid stroma) (Fig. 1E). When a mixture of different types of desmoplastic stroma was present, DR was classified based on the area with the most immature stroma as described previously (19).

The following conventional histologic markers were also investigated: WHO tumor grade (for the entire tumor), lymphovascular invasion, tumor-infiltrating lymphocytes (TILs), and peritumoral lymphocytes. Lymphocytes located within the boundaries of tumor cell nests or glands were categorized as TILs^(21, 30) and were counted in five consecutive high-power microscopic fields within the area of the highest lymphocyte infiltration. Lymphocytes that were outside the boundaries were categorized as peritumoral,⁽³¹⁾ and those that were readily identifiable in the peritumoral tissue at scanning power were counted. Apoptotic bodies were excluded. WHO tumor grade was based on the proportion of gland formation and categorized as grade 1 (well-differentiated, >95%), grade 2 (moderately differentiated, 50–95%) or grade 3 (poorly differentiated, 0–49%) (32). In mucinous carcinoma, WHO grade was assessed in the area outside the mucinous component. In signet ring cell carcinoma, the areas with signet ring cells were categorized as lacking gland formation. For testing mismatch repair (MMR) protein deficiency, immunohistochemistry (IHC) was performed by the Memorial Sloan Kettering Department of Pathology, using the standard streptavidin-biotin-peroxidase procedure. Primary monoclonal antibodies against MLH1 (clone G168–728, diluted 1:250 (PharMingen)), MSH2 (clone FE11, diluted 1:50 (Oncogene Research Products)), MSH6 (clone GRBP.P1/2.D4, diluted 1:200 (Serotec Inc.)), and PMS2 (clone A16–4, diluted 1:200 (BD PharMingen)) were used. Specialized pathologists interpreted all IHC results. MMR testing was performed selectively in this dataset (n=445), and therefore, MMR was not included in prognostic analyses.

Staging and Surveillance Protocol

Preoperative staging included contrast-enhanced computed tomography (CT) of the chest, abdomen, and pelvis and colonoscopy. Adjuvant chemotherapy was recommended for patients with stage III or high-risk stage II disease after histologic evaluation of the surgical specimen as recommended in the NCCN guidelines.^(33, 34) The general practice for postoperative surveillance of stage I–III colon cancer at Memorial Sloan Kettering was in accordance with the NCCN guidelines.^(34, 35)

Radiographic imaging reports were reviewed, and definitive diagnoses of recurrence were established based on the appearance of new lesions on CT, MRI, and/or PET images and/or biopsy-based histologic confirmation of recurrence. Right colon was defined as consisting of the cecum and the ascending and transverse colon. Left colon was defined as consisting of the descending, sigmoid, and rectosigmoid colon.

Statistical Analysis

The primary outcome was recurrence-free survival (RFS) following surgical resection. Patients were monitored from the date of surgery until recurrence, death, or last follow-up. RFS was estimated using the Kaplan-Meier method. Patients who died or had a recurrence during the study period represented events in the analysis. Patients who were alive without recurrence at last follow-up were censored.

Each pathologic parameter was assessed for association with RFS using the log-rank test. Hazard ratios (HR) along with 95% confidence intervals (CI) and Akaike information criterion (AIC) model fit statistics were estimated using Cox proportional hazards regression modeling, with a lower AIC indicating a better fit of the data. The prognostic accuracy of each pathologic parameter in predicting recurrence or death was assessed using the concordance probability estimate (CPE) for proportional hazards regression.(36) The CPE can range from 0 to 1, with 1 indicating that for any two randomly selected patients, the parameter predictions are perfectly concordant with the observed outcomes, 0 indicating that they are perfectly discordant, and 0.5 indicating that the parameter's prognostic accuracy is no better than random chance. The hazard function of recurrence or death was plotted using the kernel-smoothing method.(37, 38) The AIC, CPE, and hazard function of recurrence or death were calculated for the entire cohort including all disease stages and were based on assessment by a single pathologist. For evaluation of agreement between pathologists on each of the pathologic parameters, the overall weighted kappa was calculated as the mean of the pairwise Fleiss-Cohen weighted kappas (39) based on Light's method. (40, 41)

Continuous variables were compared using the Mann-Whitney U test. Categorical variables were compared using the χ^2 test. All statistical analyses were performed using JMP version 10.1.2 software (SAS Institute Inc., Cary, NC), SAS version 9.4 (SAS Institute Inc.), or R version 3.2.4 (www.R-project.org). All tests were two-sided, and p-values < 0.05 were considered significant.

Results

The study cohort consisted of 851 patients with a median age of 66 years (range, 25 to 99 years), and median follow-up was 36 months. Descriptive statistics for the patient cohort are listed in Table S1 in Supplemental Digital Content. A total of 90 patients (10.6%) had recurrence, and 39 patients (4.6%) had died before the last follow-up; there were 106 RFS events in the analysis. The 3-year RFS rate for all patients was 86.6% (95% CI, 83.7% to 89.0%). Table 1 lists the results of univariate analyses of patient and tumor characteristics for each histologic marker. In general, the invasive-front markers were all associated with advanced disease stage and poor clinicopathologic features. The markers were also

associated with higher TILs, peritumoral lymphocytes, and MMR protein deficiency and were strongly correlated with each other (Table S2 in Supplemental Digital Content)

Figure 2 shows the probabilities of 3-year RFS based on marker grade. All the markers examined were significantly associated with RFS. The largest difference in 3-year RFS between the best and worst grades of a marker was in PDCs (30.4% difference; HR 6.39, 95% CI 4.11–9.95, $p < .0001$), and the smallest was in CLR (8.6% difference; HR 1.86, 95% CI 1.09–3.16, $p = .02$). PDCs, TB, and PN were associated with RFS in both stage II and stage III patients, while DR and CLR were prognostic only in stage III (Table 2). Again, PDCs grades were associated with the largest separation of 3-year RFS in both stage II (26.7% difference; HR 4.15, 95% CI 1.96–8.82, $p = .0002$) and stage III (29.6% difference; HR 4.50, 95% CI 2.41–8.41, $p < .0001$) patients. When the patients were grouped according to whether they received adjuvant chemotherapy, PDCs, TB, PN, and DR were prognostic in both treatment groups, while WHO grade was not prognostic in either group (Table S3 in Supplemental Digital Content).

Prognostic accuracy as indicated by the CPE was the highest for PDCs grade (0.642) and the lowest for WHO grade (0.526). Although DR had a low AIC (1139), it was applicable only for a subset of T3/4 tumors and could not be compared with the other markers for overall prognostic accuracy. Among the four markers applicable for the entire cohort, PDCs grade had the lowest AIC (1266), indicating that the PDCs prognostic model had the best fit. The WHO grade model had the worst fit (Table 2).

The smoothed graph of the hazard function over time for PDCs, TB, PN, and DR demonstrates that the risk of recurrence or death was higher for worse-grade tumors (Figure 3). The risk of recurrence or death associated with PDCs peaked earlier for G3 (7.2 months) than for G2 (18.6 months) (G1 had no peak). The risk for G3 TB was higher than for G1 and G2 but peaked close in time to G2 (5.7 and 8.9 months, respectively). The risk associated with extramural PN was higher but peaked later than the risk associated with intramural PN (22.1 and 14.2 months, respectively). Likewise, the risk associated with immature DR was higher but peaked later than the risk associated with intermediate DR (14.3 and 10.6 months, respectively).

Table 3 lists the data on interobserver agreement. The PDCs kappa value for interobserver agreement was 0.824, indicating excellent agreement, while kappa values for the other markers ranged from 0.464 to 0.696, indicating fair to good agreement (42).

Discussion

All invasive-front markers examined were associated with tumor recurrence and had better prognostic accuracy for recurrence than WHO grade (which is a broad measure of the entire tumor excluding the invasive front). The prognostic value of these markers stems from the dynamic nature of the invasive front, with tumor cells undergoing EMT and interacting with the host immune system as they advance.⁽⁴³⁾ PDCs grade had the highest accuracy in predicting RFS, with the highest interobserver agreement among pathologists. Our findings indicate that PDCs are clinically more useful in predicting recurrence than the other

invasive-front histologic markers. In contrast, WHO grade performed poorly as a prognostic marker and had poor interobserver agreement between pathologists. The results of our study suggest that using the grade of PDCs at the invasive front instead of WHO grade of the entire tumor can help identify patients with a high risk of recurrence more accurately. Furthermore, the excellent interobserver agreement on PDCs grade among Japanese and American pathologists suggests that PDCs assessment should be considered for introduction into clinical practice by U.S. pathologists.

Accurate prediction of outcome in colon cancer patients has many advantages, such as individualization of postoperative surveillance. We found that the risk of recurrence not only was higher but peaked earlier for high-grade PDCs: at 7.2 months for G3 and at 18.6 months for G2. Clinical guidelines for stage II–III colon cancer recommend surveillance every 3–6 months after surgery for 2–3 years and then every 6 months for a total of 5 years.(34, 35) Our findings indicate that closer surveillance may be needed in the first year in patients with G3 PDCs.

Current clinical guidelines recommend adjuvant chemotherapy for high-risk stage II colon cancer with histologic features associated with poor prognosis.(33, 34) In the current study, different PDCs, TB, and PN grades were associated with different outcomes in stage II colon cancer. PDCs and TB grades had similar associations with 3-year RFS: 94% RFS for G1 PDCs and TB and 67–68% RFS for G3 PDCs and TB. PN separated 3-year RFS curves in stage II patients to a lesser degree: 95% RFS for patients without PN and 75–77% RFS for patients with intramural or extramural PN. Interestingly, stage II patients with either G3 PDCs or G3 TB had lower RFS than stage III patients with either G1 PDCs or G1 TB. Our findings suggest that PDCs and TB grades can be used to further stratify stage II tumors into high and low risk and may inform the decision to proceed with adjuvant chemotherapy. Host immune response is also an important determinant of prognosis in colon cancer patients. Interestingly, PDCs and TB were inversely associated with immune response at the invasive front: higher grades of PDCs and TB were associated with fewer TILs and peritumoral lymphocytes. For TB, this finding is consistent with those of previous studies,(5, 44, 45) but to our knowledge the association of PDCs with reduced immune response has not been reported previously. A recent study revealed loss of major histocompatibility complex class I in budding cancer cells,(46) which may be the mechanism whereby budding cancer cells or PDCs evade host immune response via EMT.

As for direct markers of immune response, a recent large study found that CLR and TILs were strongly associated with improved prognosis regardless of the status of microsatellite instability or TNM stage.(21) Another study indicated that combining CD8+ lymphocytes with TB improved prognostic accuracy.(47) However, the standard methods for quantifying CLR and TILs are pathologist-dependent and labor-intensive. To minimize the subjectivity of assessment, we applied H. Ueno's validated criteria for assessing CLR,(20) which are based on presence of active lymphoid aggregate 1mm.(20, 48) However, CLR had the poorest prognostic accuracy among the invasive-front histologic markers examined. Although host immune response does have promise as a prognostic marker, further refining of the diagnostic criteria is needed before clinical use.

Although PDCs and TB are distinguished by being defined as ≥ 5 and <5 cancer cells, respectively, the two markers likely belong in the same spectrum of dedifferentiated biological features. Because of their larger size, PDCs likely require less time and effort to identify, and in our study interobserver agreement was higher for PDCs than for TB. We also found that PDC grade and TB grade were strongly correlated (Table S2 in Supplemental Digital Content). This finding suggests that PDCs and TB could potentially be combined as a novel dedifferentiation marker at the invasive front. Further investigation is warranted to determine whether combining PDCs and TB would improve prognostic accuracy and interobserver agreement.

The strengths of our study include the large size of the cohort of patients with stage I–III colon adenocarcinomas resected at a high-volume comprehensive cancer center by dedicated colorectal surgeons using uniform technique. Other strengths were the comprehensive histologic assessment by specialized pathologists, the availability of detailed clinicopathologic information, and the use of the latest, highly effective chemotherapy. However, the study was subject to the potential selection bias inherent in observational retrospective studies and the potential selection bias associated with the use of a random subsample of 50 patients instead of the whole cohort. Because the study was conducted at a single, high-volume center, the generalizability of the results may be restricted to such specialized centers. Our results therefore need to be validated in an external dataset, using the ocular normalization factors of Lugli et al. (27) to ensure reproducibility (which is essential for PDCs to have clinical significance). In addition, CLR grading was two-tiered, and DR was applicable only for T3/4 tumors, making the CPE, AIC, and kappa for these markers not directly comparable to those of the other markers.

Our identification of PDCs as the best prognostic marker in terms of accuracy and interobserver agreement among the invasive-front histologic markers examined does not negate the clinical value of the other histologic markers, particularly in combination with PDCs. As all the markers are relatively new, we expect further improvement in interobserver agreement and prognostic accuracy. Future confirmatory investigations are warranted to determine whether PDCs can replace WHO grade and whether PDCs should be incorporated in staging systems or guidelines as an indicator of high risk in colon cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: National Cancer Institute grant P30 C008748.

This research was funded in part by the NIH/NCI Cancer Center Support Grant P30 CA008748. We gratefully acknowledge the editorial assistance of Arthur Gelmis.

References

1. Ratto C, Ricci R. Potential pitfalls concerning colorectal cancer classification in the seventh edition of the AJCC Cancer Staging Manual. *Dis Colon Rectum*. 2011; 54:e232. [PubMed: 21730773]

2. Weiser MR, Landmann RG, Kattan MW, et al. Individualized prediction of colon cancer recurrence using a nomogram. *J Clin Oncol.* 2008; 26:380–385. [PubMed: 18202413]
3. Ueno H, Hase K, Hashiguchi Y, et al. Site-specific tumor grading system in colorectal cancer: multicenter pathologic review of the value of quantifying poorly differentiated clusters. *Am J Surg Pathol.* 2014; 38:197–204. [PubMed: 24418853]
4. Ueno H, Shinto E, Kajiwara Y, et al. Prognostic impact of histological categorisation of epithelial-mesenchymal transition in colorectal cancer. *Br J Cancer.* 2014; 111:2082–2090. [PubMed: 25247323]
5. Zlobec I, Lugli A. Invasive front of colorectal cancer: dynamic interface of pro-/anti-tumor factors. *World J Gastroenterol.* 2009; 15:5898–5906. [PubMed: 20014453]
6. Kawachi H, Eishi Y, Ueno H, et al. A three-tier classification system based on the depth of submucosal invasion and budding/sprouting can improve the treatment strategy for T1 colorectal cancer: a retrospective multicenter study. *Mod Pathol.* 2015; 28:872–879. [PubMed: 25720321]
7. Prall F. Tumour budding in colorectal carcinoma. *Histopathology.* 2007; 50:151–162. [PubMed: 17204028]
8. Puppa G, Senore C, Sheahan K, et al. Diagnostic reproducibility of tumour budding in colorectal cancer: a multicentre, multinational study using virtual microscopy. *Histopathology.* 2012; 61:562–575. [PubMed: 22765314]
9. Schmoll HJ, Van Cutsem E, Stein A, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. *Ann Oncol.* 2012; 23:2479–2516. [PubMed: 23012255]
10. Compton C, Fenoglio-Preiser CM, Pettigrew N, et al. American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer.* 2000; 88:1739–1757. [PubMed: 10738234]
11. Compton CC, Fielding LP, Burgart LJ, et al. Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med.* 2000; 124:979–994. [PubMed: 10888773]
12. Ueno H, Konishi T, Ishikawa Y, et al. Prognostic value of poorly differentiated clusters in the primary tumor in patients undergoing hepatectomy for colorectal liver metastasis. *Surgery.* 2015; 157:899–908. [PubMed: 25731780]
13. Barresi V, Reggiani Bonetti L, Ieni A, et al. Poorly Differentiated Clusters: Clinical Impact in Colorectal Cancer. *Clin Colorectal Cancer.* 2016
14. Barresi V, Reggiani Bonetti L, Ieni A, et al. Prognostic significance of grading based on the counting of poorly differentiated clusters in colorectal mucinous adenocarcinoma. *Hum Pathol.* 2015; 46:1722–1729. [PubMed: 26344416]
15. Reggiani Bonetti L, Barresi V, Bettelli S, et al. Poorly differentiated clusters (PDC) in colorectal cancer: what is and ought to be known. *Diagn Pathol.* 2016; 11:31. [PubMed: 27004798]
16. Kalluri R, Zeisberg M. Fibroblasts in cancer. *Nat Rev Cancer.* 2006; 6:392–401. [PubMed: 16572188]
17. Karagiannis GS, Poutahidis T, Erdman SE, et al. Cancer-associated fibroblasts drive the progression of metastasis through both paracrine and mechanical pressure on cancer tissue. *Mol Cancer Res.* 2012; 10:1403–1418. [PubMed: 23024188]
18. McAllister SS, Weinberg RA. Tumor-host interactions: a far-reaching relationship. *J Clin Oncol.* 2010; 28:4022–4028. [PubMed: 20644094]
19. Ueno H, Shinto E, Shimazaki H, et al. Histologic categorization of desmoplastic reaction: its relevance to the colorectal cancer microenvironment and prognosis. *Ann Surg Oncol.* 2015; 22:1504–1512. [PubMed: 25395146]
20. Ueno H, Hashiguchi Y, Shimazaki H, et al. Objective criteria for crohn-like lymphoid reaction in colorectal cancer. *Am J Clin Pathol.* 2013; 139:434–441. [PubMed: 23525613]
21. Rozek LS, Schmit SL, Greenson JK, et al. Tumor-Infiltrating Lymphocytes, Crohn's-Like Lymphoid Reaction, and Survival From Colorectal Cancer. *J Natl Cancer Inst.* 2016:108.
22. Washington MK, Berlin J, Branton P, et al. Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum. *Arch Pathol Lab Med.* 2009; 133:1539–1551. [PubMed: 19792043]

23. Okamura T, Shimada Y, Nogami H, et al. Tumor Budding Detection by Immunohistochemical Staining is Not Superior to Hematoxylin and Eosin Staining for Predicting Lymph Node Metastasis in pT1 Colorectal Cancer. *Dis Colon Rectum*. 2016; 59:396–402. [PubMed: 27050601]
24. Shimada Y, Ajioka Y, Nishikura K, et al. Multiple-sectioning Study to Establish a Standardized Immunohistochemical Method for Detecting Isolated Tumor Cells in Lymph Nodes of Patients with Colorectal Cancer. *Acta medica et biologica*. 2008; 56:19–25.
25. Shimada Y, Kido T, Kameyama H, et al. Clinical significance of perineural invasion diagnosed by immunohistochemistry with anti-S100 antibody in Stage I–III colorectal cancer. *Surg Today*. 2015; 45:1493–1500. [PubMed: 25502403]
26. Yagi R, Shimada Y, Kameyama H, et al. Clinical Significance of Extramural Tumor Deposits in the Lateral Pelvic Lymph Node Area in Low Rectal Cancer: A Retrospective Study at Two Institutions. *Ann Surg Oncol*. 2016; 23:552–558.
27. 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; 30:1299–1311. [PubMed: 28548122]
28. Ueno H, Mochizuki H, Hashiguchi Y, et al. Risk factors for an adverse outcome in early invasive colorectal carcinoma. *Gastroenterology*. 2004; 127:385–394. [PubMed: 15300569]
29. Ueno H, Shirouzu K, Eishi Y, et al. Characterization of perineural invasion as a component of colorectal cancer staging. *Am J Surg Pathol*. 2013; 37:1542–1549. [PubMed: 24025524]
30. Shia J, Ellis NA, Paty PB, et al. Value of histopathology in predicting microsatellite instability in hereditary nonpolyposis colorectal cancer and sporadic colorectal cancer. *Am J Surg Pathol*. 2003; 27:1407–1417. [PubMed: 14576473]
31. Shia J, Black D, Hummer AJ, et al. Routinely assessed morphological features correlate with microsatellite instability status in endometrial cancer. *Hum Pathol*. 2008; 39:116–125. [PubMed: 17949789]
32. Hamilton, SR., Bosman, FT., Boffetta, P., et al. Carcinoma of the colon and rectum. In: Bosman, FT., Carneiro, F., Hruban, RH., et al., editors. *WHO Classification of Tumours of the Digestive System*. 4. Lyon: International Agency for Research on Cancer; 2010. p. 134–146.
33. Benson AB 3rd, Schrag D, Somerfield MR, et al. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. *J Clin Oncol*. 2004; 22:3408–3419. [PubMed: 15199089]
34. Network@ NCC. [Accessed 10/19/2016] Clinical Practical Guidelines in Oncology (NCCN Guidelines®) Colon Cancer Version 2.2016. 2016. Available at: https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf
35. Meyerhardt JA, Mangu PB, Flynn PJ, et al. Follow-up care, surveillance protocol, and secondary prevention measures for survivors of colorectal cancer: American Society of Clinical Oncology clinical practice guideline endorsement. *J Clin Oncol*. 2013; 31:4465–4470. [PubMed: 24220554]
36. Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. *Biometrika*. 2005; 92:965–970.
37. Hess KR, Serachitopol DM, Brown BW. Hazard function estimators: a simulation study. *Stat Med*. 1999; 18:3075–3088. [PubMed: 10544307]
38. Muller HG, Wang JL. Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics*. 1994; 50:61–76. [PubMed: 8086616]
39. Fleiss JL, Cohen J. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement*. 1973; 33:613–619.
40. Light R. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol Bull*. 1971; 76:365–377.
41. Warrens M. Inequalities between multi-rater kappas. *Adv Data Anal Classif*. 2010; 4:271–286.
42. Fleiss, JL., Levin, B., Paik, MC. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc; 2004. The Measurement of Interrater Agreement; p. 598–626.
43. Spaderna S, Schmalhofer O, Hlubek F, et al. A transient, EMT-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer. *Gastroenterology*. 2006; 131:830–840. [PubMed: 16952552]

44. Shinto E, Mochizuki H, Ueno H, et al. A novel classification of tumour budding in colorectal cancer based on the presence of cytoplasmic pseudo-fragments around budding foci. *Histopathology*. 2005; 47:25–31. [PubMed: 15982320]
45. Zlobec I, Lugli A, Baker K, et al. Role of APAF-1, E-cadherin and peritumoral lymphocytic infiltration in tumour budding in colorectal cancer. *J Pathol*. 2007; 212:260–268. [PubMed: 17516584]
46. Koelzer VH, Dawson H, Andersson E, et al. Active immunosurveillance in the tumor microenvironment of colorectal cancer is associated with low frequency tumor budding and improved outcome. *Transl Res*. 2015; 166:207–217. [PubMed: 25797890]
47. Lugli A, Karamitopoulou E, Panayiotides I, et al. CD8+ lymphocytes/tumour-budding index: an independent prognostic factor representing a ‘pro-/anti-tumour’ approach to tumour host interaction in colorectal cancer. *Br J Cancer*. 2009; 101:1382–1392. [PubMed: 19755986]
48. Kim JH, Kim KJ, Bae JM, et al. Comparative validation of assessment criteria for Crohn-like lymphoid reaction in colorectal carcinoma. *J Clin Pathol*. 2015; 68:22–28. [PubMed: 25322692]

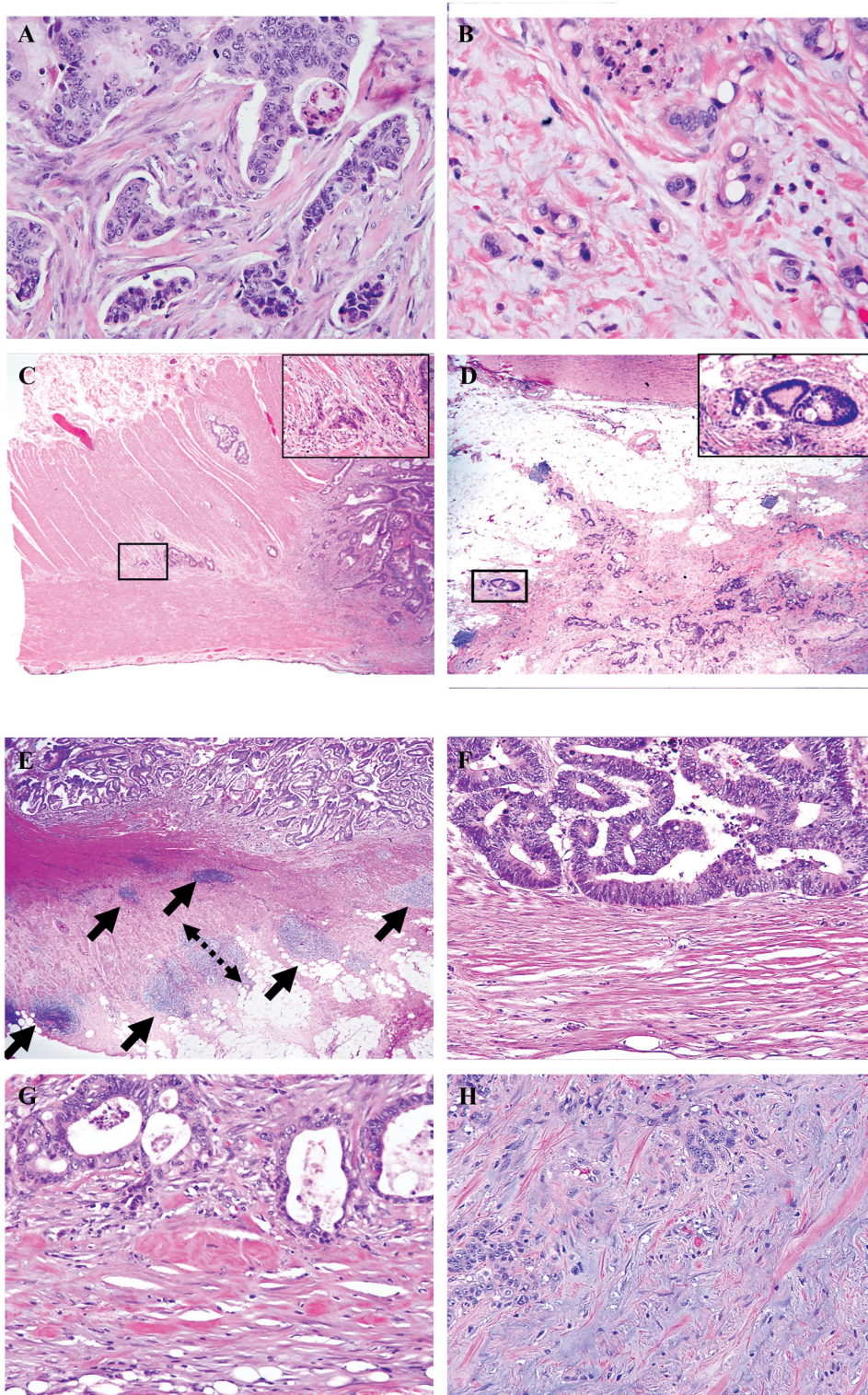


Figure 1.
Histologic features at the invasive front of colon cancer.
A) Poorly differentiated clusters (H&E staining; original magnification $\times 100$).

- B) Tumor budding (H&E staining; original magnification $\times 200$).
- C) Intramural perineural invasion spreading along the Auerbach plexus zone (H&E staining; original magnification $\times 20$, inset $\times 200$).
- D) Extramural perineural invasion (H&E staining; original magnification $\times 20$, inset $\times 200$).
- E) Crohn-like lymphoid reaction. Several nodular lymphoid aggregates of different sizes (arrows) are present. The maximum diameter of the largest lymphoid aggregate was measured (dotted arrow). (H&E staining; original magnification, $\times 20$.)
- F–H) Desmoplastic reaction: mature (F), intermediate (G), and immature (H). (H&E staining; original magnification, $\times 200$).

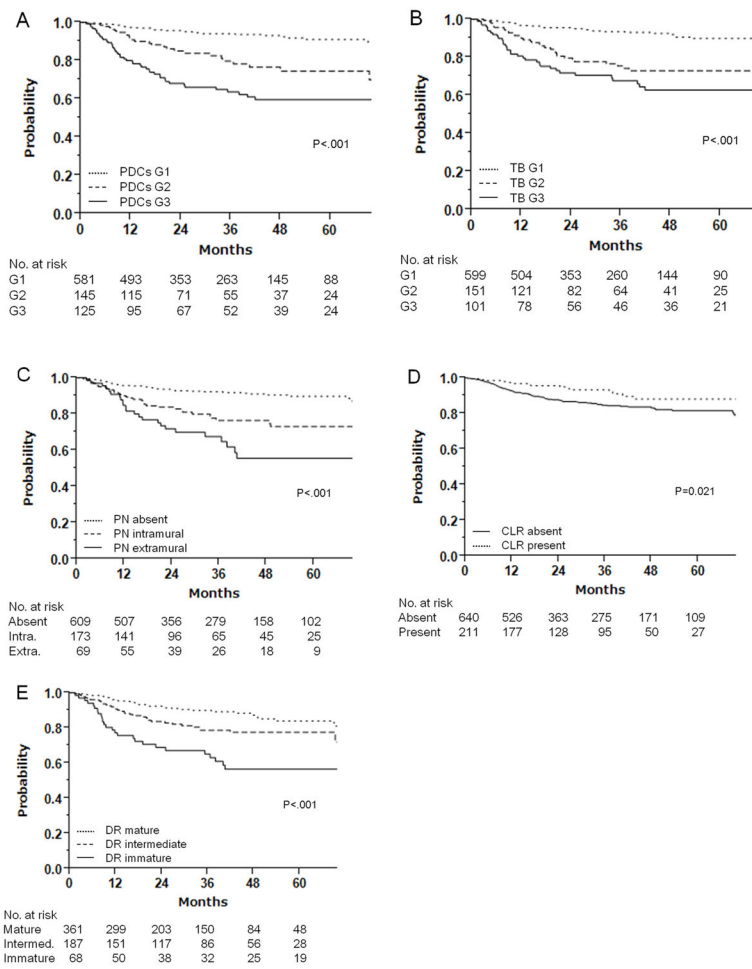


Figure 2. Kaplan-Meier curves for association between recurrence-free survival and invasive-front histologic markers.
 A) Poorly differentiated clusters
 B) Tumor budding
 C) Perineural invasion
 D) Crohn-like lymphoid reaction
 E) Desmoplastic reaction

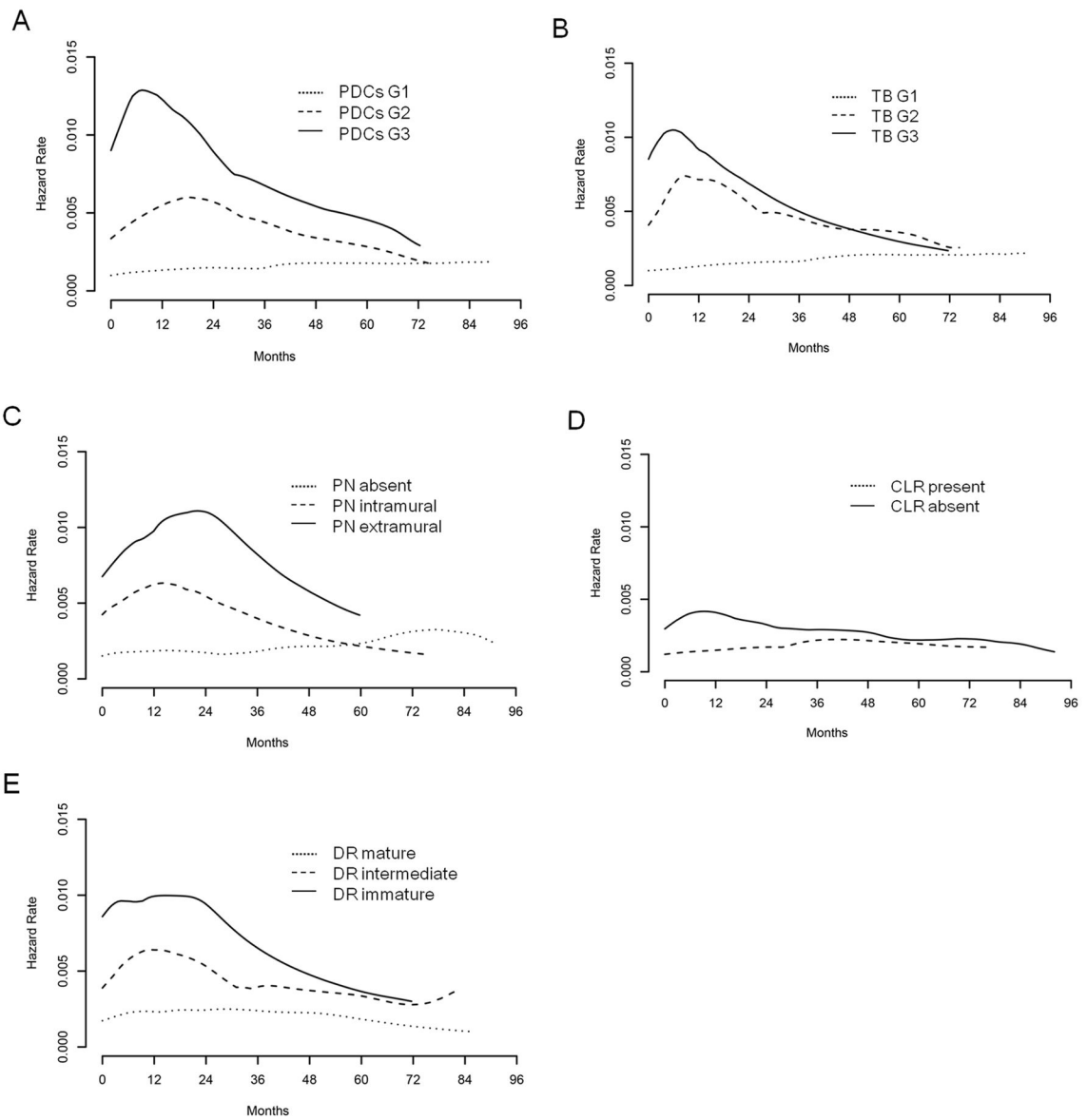


Figure 3.
 Hazard function of recurrence or death for histologic markers at the invasive front.
 A) Poorly differentiated clusters
 B) Tumor budding
 C) Perineural invasion
 D) Crohn-like lymphoid reaction
 E) Desmoplastic reaction

Table 1

Patient and tumor characteristics in relation to histological markers^a

Characteristic	% Patients with the following PDCs:				P	% Patients with the following TB:				P	% Patients with the following PN:				P	% Patients with the following CLR:				P	% Patients with the following DR ^b :				P	
	G1	G2	G3	G3		G1	G2	G3	G3		Absent	Intra	Extra	P		Absent	Present	P	Mature		Intermed	Immature	P			
Total	68.3	17.0	14.7	14.7	0.39	70.4	17.7	11.9	11.9	0.75	71.6	20.3	8.1	8.1	0.30	75.2	24.8	24.8	24.8	0.024	58.6	30.4	11.0	11.0	0.12	
Sex																										
Male	49.1	42.8	48.8	48.8	0.39	48.6	47.7	44.6	44.6	0.75	48.9	48.0	39.1	39.1	0.30	50.2	41.2	41.2	41.2	0.024	48.2	51.3	36.8	36.8	0.12	
Female	50.9	57.2	51.2	51.2	0.39	51.4	52.3	55.4	55.4	0.75	51.1	52.0	60.9	60.9	0.30	49.8	58.8	58.8	58.8	0.024	51.8	48.7	63.2	63.2	0.12	
Age																										
>65 yr	53.2	53.8	43.2	43.2	0.11	53.9	48.3	44.6	44.6	0.14	55.2	46.8	34.8	34.8	0.002	50.6	55.5	55.5	55.5	0.22	51.5	54.6	51.0	51.0	0.72	
65 yr	46.8	46.2	56.8	56.8	0.11	46.1	51.7	55.4	55.4	0.14	44.8	53.2	65.2	65.2	0.002	49.4	44.5	44.5	44.5	0.22	48.5	45.4	49.0	49.0	0.72	
Site of primary tumor																										
Right colon	57.0	44.1	44.8	44.8	0.003	55.4	47.7	46.5	46.5	0.090	56.2	42.8	50.7	50.7	0.007	50.5	36.5	36.5	36.5	<0.0001	50.7	61.0	47.1	47.1	0.040	
Left colon	43.0	55.9	55.2	55.2	0.003	44.6	52.3	53.5	53.5	0.090	43.8	57.2	49.3	49.3	0.007	49.5	63.5	63.5	63.5	<0.0001	49.3	39.0	52.9	52.9	0.040	
CEA																										
>5.0 ng/mL	30.4	35.9	47.4	47.4	0.002	30.1	38.3	48.9	48.9	<0.001	31.3	33.3	57.1	57.1	<0.001	32.8	36.7	36.7	36.7	0.31	34.2	43.1	60.3	60.3	<0.001	
5.0 ng/mL	69.6	64.1	52.6	52.6	0.002	69.9	61.7	51.1	51.1	<0.001	68.7	66.7	42.9	42.9	<0.001	67.2	63.3	63.3	63.3	0.31	65.8	56.9	39.7	39.7	<0.001	
T stage																										
T1	11.4	5.5	2.4	2.4	<0.001	11.9	2.6	2.0	2.0	<0.001	11.5	4.1	0.0	0.0	<0.001	10.2	5.7	5.7	5.7	<0.0001	7.5	21.4	42.7	42.7	<0.001	
T2	21.7	15.2	8.0	8.0	<0.001	21.5	15.9	5.0	5.0	<0.001	22.3	12.7	0.0	0.0	<0.001	20.5	12.8	12.8	12.8	<0.0001	92.5	78.6	57.4	57.4	<0.001	
T3	60.9	62.1	60.8	60.8	<0.001	61.6	62.3	56.4	56.4	<0.001	59.3	64.7	68.1	68.1	<0.001	55.8	77.3	77.3	77.3	<0.0001	92.5	78.6	57.4	57.4	<0.001	
T4	6.0	17.2	28.8	28.8	<0.001	5.0	19.2	36.6	36.6	<0.001	6.9	18.5	31.9	31.9	<0.001	13.6	4.3	4.3	4.3	<0.0001	7.5	21.4	42.7	42.7	<0.001	
N stage																										
N0	74.9	47.6	33.6	33.6	<0.001	75.0	45.0	28.7	28.7	<0.001	72.3	50.3	27.5	27.5	<0.001	62.0	70.6	70.6	70.6	0.075	66.8	48.1	27.9	27.9	<0.001	
N1	18.6	38.6	38.4	38.4	<0.001	18.7	36.4	44.6	44.6	<0.001	21.5	30.6	40.6	40.6	<0.001	26.6	19.9	19.9	19.9	0.075	24.1	34.2	41.2	41.2	<0.001	
N2	6.5	13.8	28.0	28.0	<0.001	6.3	18.5	26.7	26.7	<0.001	6.2	19.1	31.9	31.9	<0.001	11.4	9.5	9.5	9.5	0.075	9.1	17.7	30.9	30.9	<0.001	
AJCC stage (5 th ed.)																										
I	29.1	14.5	4.8	4.8	<0.001	29.2	12.6	2.0	2.0	<0.001	29.1	11.0	0.0	0.0	<0.001	25.5	15.6	15.6	15.6	<0.001	66.8	48.1	27.9	27.9	<0.001	
II	45.8	33.1	28.8	28.8	<0.001	45.7	32.5	26.7	26.7	<0.001	43.2	39.3	27.5	27.5	<0.001	36.6	55.0	55.0	55.0	<0.001	66.8	48.1	27.9	27.9	<0.001	
III	25.1	52.4	66.4	66.4	<0.001	25.0	55.0	71.3	71.3	<0.001	27.8	49.7	72.5	72.5	<0.001	38.0	29.4	29.4	29.4	<0.001	33.2	51.9	72.1	72.1	<0.001	

Characteristic	% Patients with the following PDCs:				% Patients with the following TB:				% Patients with the following PN:				% Patients with the following CLR:				% Patients with the following DR ^b :			
	G1	G2	G3	P	G1	G2	G3	P	Absent	Intra	Extra	P	Absent	Present	P	Mature	Intermed	Immature	P	
WHO grade				<0.001				<0.001				<0.001								0.005
1	3.1	0.7	0.0		3.2	0.0	0.0	3.0	0.6	0.0			2.2	2.4		2.2	0.0	0.0		0.74
2	86.8	87.6	56.0		86.8	82.8	55.4	86.5	74.0	66.7			81.9	83.9		83.9	79.1	72.1		
3	10.2	11.7	44.0		10.0	17.2	44.6	10.5	25.4	33.3			15.9	13.7		13.9	20.9	27.9		
LVI				<0.001																<0.001
Yes	34.4	64.1	81.6		33.9	70.2	85.1	36.5	65.9	85.5			48.0	41.7		41.8	64.2	92.7		
No	65.6	35.9	18.4		66.1	29.8	14.9	63.5	34.1	14.5			52.0	58.3		58.2	35.8	7.4		
Median no. of TILs/5 HPF (IQR)	6 (2-18)	3 (1-7)	1 (0-4)	<0.001	6 (2-19)	3 (1-6)	1 (0-3)	<0.001	5 (1-16)	2 (0-9)	1 (0-5)	<0.001	3 (1-9)	10 (2-29)	<0.001	5 (1-14)	2 (0-7)	1 (0-3)		<0.001
Peritumoral lymphocytes				<0.001				<0.001				<0.001								<0.001
Yes	62.1	45.5	32.8		63.9	40.4	23.8	60.4	47.4	26.1			47.8	76.8		55.1	49.2	13.2		
No	37.9	54.5	67.2		36.1	59.6	76.2	39.6	52.6	73.9			52.2	23.2		44.9	50.8	86.8		
MMR protein ^c				0.005				0.001				<0.001								0.046
Intact	72.5	86.5	86.4		71.7	90.7	91.1	71.6	87.4	91.7			84.4	58.8		73.1	78.2	91.7		
Deficient	27.5	13.5	13.6		28.3	9.3	8.9	28.4	12.6	8.3			15.6	41.2		26.9	21.8	8.3		

^aP values were calculated using the Mann-Whitney U test for continuous variables and the χ^2 test for categorical variables.

Intra, intratumoral; Extra, extratumoral; Intermed, intermediate; CEA, carcinoembryonic antigen; WHO, World Health Organization; LVI, lymphovascular invasion; TILs, tumor infiltrating lymphocytes; HPF, high power field; IQR, interquartile range; MMR, mismatch repair.

^bApplicable for T3/4 tumors only (n=616).

^cApplicable for patients who underwent MMR testing only (n=445)

Table 2
 Three-year RFS rates in relation to histologic marker grade, disease stage, and prognostic accuracy^a

Marker grade	Total (N=851)						Stage II (N=350)						Stage III (N=305)					
	RFS	P	HR (95% CI)	CPE	AIC	RFS	P	HR (95% CI)	RFS	P	HR (95% CI)	RFS	P	HR (95% CI)				
PDCs		<0.001	Reference	0.642	1266	94.0	<0.001	Reference	89.0	<0.001	Reference	89.0	<0.001	Reference				
G1	94.1		Reference			94.0		Reference			Reference			Reference				
G2	79.8		3.15 (1.90–5.24)			83.7		2.30 (0.96–5.5)			2.57 (1.28–5.16)			2.57 (1.28–5.16)				
G3	63.7		6.39 (4.11–9.95)			67.3		4.15 (1.96–8.82)			4.50 (2.41–8.41)			4.50 (2.41–8.41)				
TB		<0.001	Reference	0.629	1281	94.0	0.001	Reference	86.6	<0.001	Reference	86.6	<0.001	Reference				
G1	93.5		Reference			94.0		Reference			Reference			Reference				
G2	75.5		3.52 (2.22–5.58)			79.6		2.38 (1.07–5.27)			2.64 (1.41–4.93)			2.64 (1.41–4.93)				
G3	67.8		4.98 (3.13–7.93)			68.0		3.94 (1.72–9.03)			2.98 (1.61–5.52)			2.98 (1.61–5.52)				
DR ^b		<0.001	Reference	0.604	1139	92.3	0.39	Reference	85	<0.001	Reference	85	<0.001	Reference				
Mature	89.9		Reference			92.3		Reference			Reference			Reference				
Intermediate	78.8		1.91 (1.20–3.04)			86.1		1.44 (0.70–2.95)			1.87 (0.98–3.57)			1.87 (0.98–3.57)				
Immature	65.2		3.79 (2.28–6.31)			75.9		1.89 (0.64–5.59)			3.52 (1.82–6.78)			3.52 (1.82–6.78)				
CLR	0.021		0.556	1326		0.32		0.020			0.020			0.020				
Present	93.3		Reference			93.4		Reference	91.7		Reference	91.7		Reference				
Absent	84.7		1.86 (1.09–3.16)			87.6		1.47 (0.69–3.13)			2.61 (1.13–6.05)			2.61 (1.13–6.05)				
PN		<0.001	Reference	0.607	1295	94.5	0.002	Reference	83.1	0.003	Reference	83.1	0.003	Reference				
Absent	92.3		Reference			94.5		Reference			Reference			Reference				
Intramural	76.7		2.63 (1.69–4.07)			75.4		3.12 (1.55–6.29)			1.67 (0.93–3.01)			1.67 (0.93–3.01)				
Extramural	67.6		4.60 (2.82–7.52)			76.5		3.10 (1.05–9.18)			2.69 (1.49–4.85)			2.69 (1.49–4.85)				
WHO grade ^c		0.26	0.526	1332		0.37												
1	94.4		0.90 (0.22–3.66)															
2	87.3		Reference						77.7		Reference			Reference				
3	83.6		1.47 (0.92–2.36)						73.8		1.28 (0.75–2.16)			1.28 (0.75–2.16)				

^a P-values were calculated using the log-rank test. HR, hazard ratio. Analysis for Stage I was not conducted because the number of events (n=5) was too small.

^b Applicable for T3/4 tumors only (total N=616, stage II N=350, stage III N=266).

^c Not analyzed for stage II due to low numbers of events in some strata.

Table 3

Interobserver agreement for histologic markers

Marker	Kappa			
	Observer 1 vs. 2	Observer 2 vs. 3	Observer 1 vs. 3	Light's Overall Kappa
PDCs	0.850	0.821	0.800	0.824
TB	0.742	0.661	0.683	0.696
DR	0.582	0.655	0.520	0.586
CLR	0.465	0.414	0.649	0.509
PN	0.353	0.385	0.654	0.464
WHO grade	0.555	0.594	0.555	0.568

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript