ORIGINAL ARTICLE

Correlation of tumor microenvironment from biopsy and resection specimens in untreated colorectal cancer patients: a surprising lack of agreement

Phillip M. Kemp Bohan[1] · Robert C. Chick[1] · Annelies T. Hickerson[1] · Lynn M. Messersmith[2] · Grant M. Williams[2] · Jessica L. Cindass[1] · Jamie Lombardo[2] · Ryan Collins[2] · Robert O. Brady[2] · Diane F. Hale[1] · George E. Peoples[3] · Timothy J. Vreeland[1] · Guy T. Clifton[1]

Abstract

Background Colorectal cancer (CRC) tumor microenvironment (TME) characteristics, such as tumor infiltrating lymphocyte (TIL) densities and PD-L1 status, are predictive of recurrence, disease-free survival, and overall survival. In many malignancies, TME characteristics are also predictive of response to immunotherapy. As window of opportunity studies using neoadjuvant immunotherapy become more common and treatment guidelines incorporate TME features, accurate assessment of the pre-treatment TME using the biopsy specimen is critical. However, no study has thoroughly evaluated the correlation between the TMEs of the biopsy and resection specimens.

Methods We conducted a retrospective analysis of patients with stage I–III CRC with matched biopsy and resection specimens. CD3+, CD4+, CD8+, and FoxP3+ lymphocyte populations at the center of tumor (CT) and invasive margin (IM) and tumor PD-L1 status in the biopsy and resection specimens were evaluated. TIL populations were compared using Mann–Whitney U tests or Student's t tests and correlated using Pearson r.

Results CD3+ and CD4+ densities were significantly higher in the CT of the biopsy relative to the resection specimen Comparing biopsy and resection specimens, no TIL population at either the CT or IM had a correlation coefficient > 0.5. Determining PD-L1 status based on biopsy tissue resulted in a sensitivity of 37.1%, specificity of 81.4%, and accuracy of 61.5%.

Conclusions These findings demonstrate significant discordance between the TME of the biopsy and resection specimens. Caution should be used when basing treatment decisions on pre-treatment endoscopic biopsy findings and when interpreting changes in the TME between pre-treatment biopsy and resection specimens after neoadjuvant therapy.

Keywords Tumor microenvironment · Tumor infiltrating lymphocytes · Colorectal cancer · Biopsy

---

✉ Phillip M. Kemp Bohan
phillip.m.kempbohan.mil@mail.mil

1 Department of Surgery, Brooke Army Medical Center, 3551 Roger Brooke Dr., Ft. Sam Houston, TX 78234, USA

2 Department of Pathology, Brooke Army Medical Center, 3551 Roger Brooke Dr., Ft. Sam Houston, TX 78234, USA

3 Cancer Vaccine Development Program, 1422 E. Grayson St, San Antonio, TX 78208, USA

Introduction

Treatment of localized colorectal cancer (CRC) is based upon the tumor, nodes, metastasis (TNM) staging system, with surgery alone recommended for stage I and most stage II disease and adjuvant chemotherapy recommended for high-risk stage II and stage III disease [1]. However, the TNM staging system offers only a rough assessment of tumor biology. The high-risk stage II disease classification

includes patients with clinical findings of obstruction and perforation [1] and excludes patients with microsatellite instability (MSI)-high tumors, a subtle distinction that represents a major shift in cancer care as it acknowledges the prognostic importance of tumor biology [2, 3]. This shift towards incorporating tumor biology into staging and treatment algorithms is further evidenced by the recent FDA approval of the anti-PD-1 antibody, pembrolizumab, for treatment of any MSI-high metastatic or unresectable tumor, regardless of histology or origin [4]. This unprecedented approval for a therapy across all tumors types based on a single biologic factor highlights the importance of accurately understanding tumor biology.

The incorporation of MSI status into staging and treatment recommendations [1] highlight recent interest in the tumor microenvironment (TME), a factor that influences response to immunotherapy and serves as a marker of overall tumor biology and prognosis [5, 6]. In CRC, the immunoscore—a measure of CD3+ and CD8+ tumor infiltrating lymphocyte (TIL) densities at the center of tumor (CT) and invasive margin (IM)—has been shown to predict recurrence independent of, and more accurately than, T or N stage [7, 8]. Additionally, a greater degree of inflammation, higher CD8+ lymphocyte density, and higher FoxP3+ lymphocyte density have also been associated with improved survival [9–11]. Finally, higher levels of programmed cell death-ligand 1 (PD-L1, a component of the PD-1/PD-L1 checkpoint) within the TME have been correlated with higher CD8+ densities, improved disease-free survival, and improved overall survival [12]. Given the predictive value of the TME, it seems likely that future staging and treatment guidelines will continue to incorporate these features.

As TME characterization and MSI status become increasingly important in cancer care, it is vital that providers accurately define these factors in a pre-therapy setting before the final resection specimen is subjected to thorough pathologic analysis. This pre-therapy assessment is used to identify patients who would benefit from neoadjuvant therapy or be eligible for an immunotherapy window of opportunity trial, which involves administering an immunomodulatory agent over a limited period of time between diagnosis on biopsy and final surgical resection. Mischaracterization of the pre-intervention TME could lead to inappropriate enrollment in a trial and incorrect appraisal of the effect of the trial agent on the TME.

Despite the importance of TME evaluation from pre-treatment biopsies, little is known regarding the accuracy of the biopsy TME and the correlation between the TME from biopsy and surgical specimens. The purpose of our study was to thoroughly compare the TME of CRC (stages I–III) sampled on endoscopic biopsy to that of the definitive surgical resection specimen in patients without a history of neoadjuvant chemo- or immunotherapy. We sought to determine the density of CD3+, CD4+, CD8+, and FoxP3+ lymphocyte populations in both the CT and IM and the correlation of these densities between biopsy and resection specimens. Additionally, we sought to assess the accuracy of an endoscopic biopsy at determining the PD-L1 status of the tumor.

Methods

Patients

This study was conducted after approval by our institutional review board. Patients with pathologic stage I–III CRC diagnosed on endoscopic biopsy who underwent resection with curative intent from the years 2006–2016 at our center were included. Cases were identified through the electronic pathology database at our institution (CoPath Plus, Cerner, Kansas City, MO) using the terms "colonic adenocarcinoma" and "colorectal adenocarcinoma" under the category "final diagnosis". To meet inclusion criteria, formalin-fixed paraffin embedded tissue (FFPE) blocks from both the initial biopsy and the resection specimen of the same malignancy had to be on site at our facility and needed to have at least 2 mm^2 of tumor remaining in the block. Cases were excluded if one of the procedures had been performed at an outside institution, blocks were unavailable, or less than 2 mm^2 of tumor remained in the blocks. Patients with stage IV disease at time of diagnosis, in situ disease only on final resection specimen, or who underwent neoadjuvant chemotherapy prior to resection were also excluded. Tumors were staged using the TNM classification from the American Joint Committee on Cancer (AJCC) 7th edition. Demographic, tumor-specific, and treatment data were all collected from retrospective chart review.

Tissue preparation, staining, and analysis

FFPE tissue blocks from the endoscopic biopsies and surgical resections were visually examined for tumor quantity. Blocks that contained the most representative tumor material were examined histologically (hematoxylin and eosin) to ensure adequate tumor was available and selected blocks contained representative samples of the invasive margin and center of tumor. Each block was cut at 4 μ and immunohistochemically stained with FoxP3, CD3, CD4, CD8, and PD-L1 (CAL10, Biocare Medical, Pacheco, CA). The immunohistochemical-stained slides were digitally captured at 400 × magnification, using the Aperio ScanScope AT Turbo (Leica Biosystems, Buffalo Grove, IL) digital imager system. De-identified images were uploaded into eSlide Manager (version 12.3.2.5030) so that all interpreting pathologists were blinded to any patient information corresponding to a given slide. Image

analysis was performed using Aperio ImageScope (version 12.0.1.5012) and an in-house analysis algorithm. The algorithm was tuned to detect nuclear and cytoplasmic positivity while excluding larger tumor nuclei in an attempt to interpret only lymphocytes. To validate this algorithm, images were captured on the ScanScope and representative fields were manually counted by three different pathologists and compared to the result obtained using the developed algorithm. After validation, slides were scanned into the eSlide Manager server and 1mm^2 areas were manually selected from both the tumor center and the invasive margin (Fig. 1). CD3+, CD8+, and FoxP3+ immunostained lymphocyte counts were reported as positive cells/ mm^2.

PD-L1-stained slides were read manually by a pathologist and reported as positive/negative. The PD-L1 stain was interpreted per manufacturer guidelines for interpretation in non-small cell lung cancer (the only FDA-approved indication for this test at the time of this investigation; Fig. 2) [13]. Slides containing tissue from the CT, IM, and tumor surface were all examined for PD-L1 staining. If greater than or equal to 1% of the tumor cells were positive for the stain, the specimen was considered positive. If the stain was dark and continuous around the cell membrane, the specimen was considered high positive. If the stain was faint or patchy around the membrane, it was considered low positive.

All photomicrographs were captured using the Aperio ScanScope AT Turbo (Leica Biosystems, Buffalo Grove, IL) at 400× magnification using a 20X FN 26.5 lens with a 2× doubler. Images were processed using the Aperio ImageScope software, version 12.

Statistical analysis

SPSS v. 24 (IBM, Armonk, NY) was used for all statistical analyses. Continuous variables were assessed for normality using Shapiro–Wilk tests and reported as either mean and standard deviation (SD) or median and interquartile range (IQR) as appropriate. Continuous variables were compared either with Student's t tests (parametric) or Mann–Whitney U tests (nonparametric). Correlations between continuous variables were performed using a Pearson correlation (r). Statistical significance was set at $p < 0.05$.
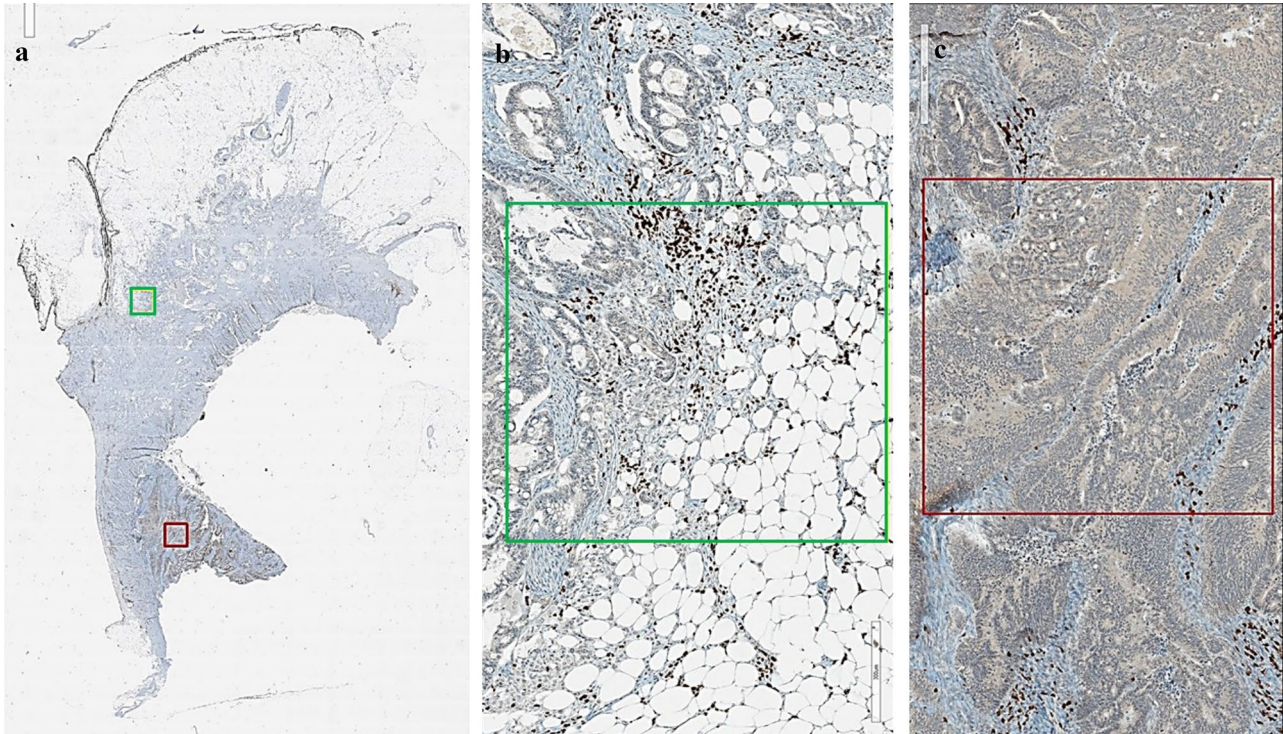

Fig. 1 CD3+ immunohistochemical staining of colorectal adenocarcinoma as seen on an Aperio® scanned slide. *A*: 1 mm^2 of tissue is selected from the invasive margin and the center of tumor at 1×. *B*: CD3+ lymphocytes at the invasive margin at 10×. *C*: CD3+ lymphocytes at the center of tumor at 10×
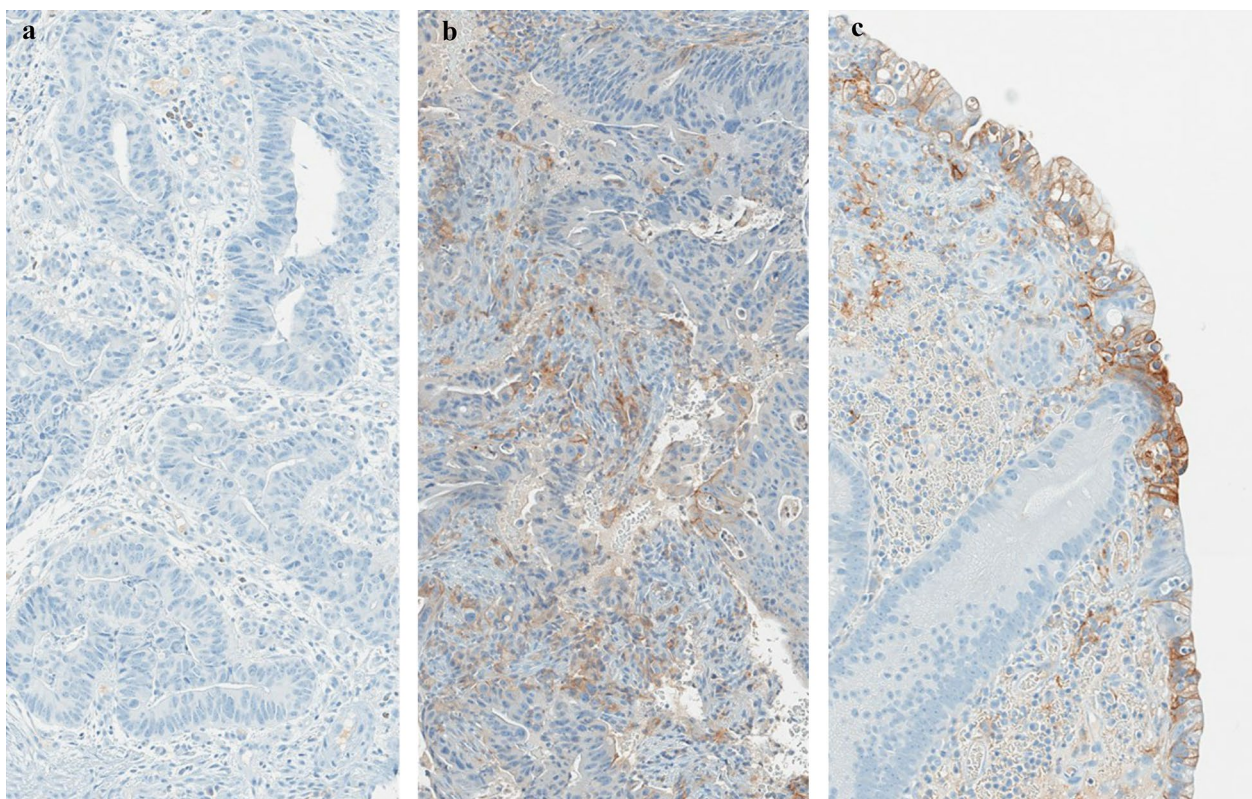
Fig. 2 PD-L1 immunohistochemical staining of colorectal adenocarcinoma. *A*: negative stain. *B*: low positive stain. *C*: high positive stain. Both *B* and *C* are scored as "positive".

Results

Patients

Matched endoscopic biopsy and resection specimens from 78 patients with sufficient tissue remaining were identified and included. Clinical and pathologic descriptors of those patients are described in Table 1. Median age was 61 years, and the majority of patients were male (56.4%). The most frequent tumor and node classifications were T3 (53%) and N0 (54%), respectively. The most frequent overall pathologic stage was stage III (46.2%). Tumors were most commonly in the right colon (consisting of all colon supplied by the superior mesenteric artery). MSI testing was performed in 52.6% of patients, of which 85.4% (35/41) were reported as low probability. Lymphovascular and perineural invasion were present in 30.8% and 11.5% of cases, respectively.

TIL analysis

The TME was assessed at four sites: the CT and IM of the biopsy specimen (CT-B and IM-B), and the CT and IM of the resection specimen (CT-R and IM-R). Comparison of the TME of CT-B and CT-R (Table 2) demonstrated significantly larger populations of CD3+ and CD4+ lymphocytes in the CT-B relative to the CT-R specimen ($p < 0.001$ and $p = 0.004$, respectively). The CD3+ lymphocyte population was larger in the IM-B relative to the IM-R ($p = 0.001$). There was no difference in FoxP3+ or CD8+ lymphocyte populations at either the CT or the IM when comparing biopsy and resection specimens.

Two sets of correlations using TME populations were performed (Table 3). First, CT and IM TME populations were correlated in the biopsy and resection specimens separately (Supplemental Figures S1 and S2). There were moderate correlations between FoxP3+ and CD8+ lymphocyte populations at CT and IM in the biopsy (CT-B to IM-B, $r = 0.700$ and $r = 0.617$, respectively) and resection specimens (CT-R and IM-R, $r = 0.673$ and $r = 0.621$, respectively). Second, biopsy and resection specimen TME populations were correlated at the CT and IM separately (Supplemental Figures S3 and S4). No lymphocyte population in either the CT or IM had a Pearson $r > 0.5$ when comparing the biopsy and resection. CD3+ and CD8+ lymphocyte populations at the IM and CT moderately correlated (r values between 0.394 and 0.444) and CD4+ and FoxP3+ lymphocyte populations at either site weakly correlation (r values all < 0.250) between biopsy and resection specimens.

Table 1 Patient demographics

Variable	
Age, median (interquartile range)	61 (49.8, 77.3)
Sex, *n* (%)	
Male	44 (56.4)
Female	34 (43.6)
Pathologic tumor (T)/nodes (N) stage, *n* (%)	
T1	7 (9.0)
T2	21 (26.9)
T3	41 (52.6)
T4	9 (11.5)
N0	42 (53.8)
N1	25 (32.1)
N2	11 (14.1)
Pathologic stage, *n* (%)	
I	19 (24.4)
II	23 (29.5)
III	36 (46.2)
Location of tumor, *n* (%)	
Right (cecum, ascending/transverse colon)	40 (51.3)
Left (descending/sigmoid colon)	31 (39.7)
Rectum	7 (9.0)
Microsatellite instability (MSI) status, *n* (%)	
High	6 (7.7)
Low	35 (44.9)
Not performed	37 (47.4)
Lymphovascular invasion, *n* (%)	
Present	24 (30.8)
Absent	47 (60.3)
Unknown	7 (9.0)
Perineural invasion, *n* (%)	
Present	9 (11.5)
Absent	51 (65.4)
Unknown	18 (23.1)

PD-L1 analysis

Biopsy and resection specimens were then divided into groups based on PD-L1 status (Table 4). Of 78 patients, 21 (26.9%) had PD-L1+ biopsies, 35 (44.9%) had PD-L1+ resection specimens, and 13 (16.7%) had both biopsy and resection stain as PD-L1+. Patients with a PD-L1+ biopsy specimen had a greater number of CD3+ and CD4+ cells at CT-B and CD3+ cells at IM-B relative to PD-L1- biopsy specimens. Patients with PD-L1+ resection specimens had a greater number of FoxP3+, CD4+ and CD8+ cells at the CT-R, and FoxP3+ and CD3+ cells at IM-R (all $p < 0.05$) relative to PD-L1- resection specimens.

Given the difference in PD-L1 status between biopsy and resection specimens, the accuracy of the biopsy at predicting final specimen PD-L1 status was then assessed (Table 5). Biopsy specimens were considered the test and resection specimens were considered the gold standard. Only 16.7% of specimens were true positives and 45% of specimens were true negative, with a false-negative rate of 28.2%. The overall accuracy of the biopsy at correctly identifying PD-L1 status of the final tumor was 61.5% (95% CI 49.8–72.3%).

Discussion

This study investigated the degree of correlation amongst TMEs as measured at the CT and IM in CRC biopsy and resection specimens. Our data demonstrate that there is concordance between the IM and CT FoxP3+ and CD8+ lymphocyte populations within individual biopsy and resection specimens. However, there were only moderate correlations between TIL populations (all $r < 0.5$) at any location between the biopsy and resection specimens. PD-L1+ specimens consistently had higher TIL populations at CT-R and IM-R,

Table 2 Comparison of tumor microenvironment in biopsy and resection specimens

Cell type	Center of tumor (CT)		
	Biopsy (CT-B)	Resection (CT-R)	*p* value
FoxP3+, median (IQR[a])	309 (187, 483)	288 (159, 469)	0.622
CD3+, median (IQR)	1060 (570, 1464)	468 (231, 932)	< 0.001
CD4+, median (IQR)	1275 (931, 1628)	975 (581, 1439)	0.004
CD8+, median (IQR)	190 (87, 372)	193 (74, 381)	0.782
Cell type	Invasive margin (IM)		
	Biopsy (IM-B)	Resection (IM-R)	*p* value
FoxP3+, median (IQR)	251 (165, 385)	231 (115, 396)	0.347
CD3+, median (IQR)	1186 (829, 1732)	847 (536, 1322)	0.001
CD4+, median (IQR)	1149 (821, 1523)	1174 (824, 1656)	0.734
CD8+, median (IQR)	313 (176, 611)	433 (215, 634)	0.156

[a]Interquartile range

Table 3 Tumor microenvironment correlations

Lymphocyte type	IM[a]/CT[b] correlation (r[c])	p value
Biopsy specimen		
FoxP3+	0.700	<0.001
CD3+	0.497	<0.001
CD4+	0.36	0.001
CD8+	0.617	<0.001
Resection specimen		
FoxP3+	0.673	<0.001
CD3+	0.555	<0.001
CD4+	0.491	<0.001
CD8+	0.621	<0.001
Lymphocyte type	Biopsy/resection correlation (r)	p value
Center of tumor		
FoxP3+	0.250	0.028
CD3+	0.444	<0.001
CD4+	0.241	0.034
CD8+	0.394	<0.001
Invasive margin		
FoxP3+	0.226	0.047
CD3+	0.427	<0.001
CD4+	0.128	0.263
CD8+	0.417	<0.001

[a]Invasive margin

[b]Center of tumor

[c]Pearson r

but the overall accuracy of the biopsy at predicting resection PD-L1 status was only 61.5%, with a false-negative rate of 28.2%.

Only recently has immunotherapy expanded to CRC, and only two studies to date have compared the TME of endoscopic biopsies to those of surgical specimens in CRC. Koelzer et al. [14] compared CD8+ and CD45RO lymphocyte populations in pre-operative biopsies and resection specimens from 130 patients with stage I–III CRC. Higher CD8+ lymphocyte infiltration in the biopsy specimen was found to be independently predictive of improved overall survival ($p < 0.01$). Lower CD8+ lymphocyte densities on biopsy were also predictive of higher T stage and positive nodal status. Yet when matched biopsy and resection specimens were correlated, only a moderate correlation existed for CD8+ lymphocytes ($r = 0.42$) and a weak correlation existed for CD45RO lymphocytes ($r = 0.16$). The authors did not comment on any relationship between TIL populations found in the resection specimen and survival, which would have been useful to determine the clinical significance of this weaker correlation between biopsy and resection specimens and the prognostic significance of the biopsy relative to the resection specimen.

Similarly, Park et al. [15] compared CD3+ lymphocytes, CD8+ lymphocytes, and tumor stroma percentage (TSP, an assessment of the degree of intratumoral stroma at the deepest point of tumor invasion) in matched biopsy and resection specimens of 115 patients with stage I–III CRC. No correlation coefficients were calculated, but incorrect characterization of TIL density occurred at frequencies of between 26.1 and 41.1%, depending on the type and location of the various T cells. The group also noted a difference in high- and low-density TSP when comparing biopsy and resection specimens ($p = 0.001$). The authors concluded that biopsy specimen provides a representative assessment of TME microenvironment, despite the statistically significant difference in lymphocytes populations and tumor stroma between biopsy and resection specimens.

Cumulatively, these data suggest a lack of correlation in certain TIL populations between biopsy and final surgical resection. While comparing Pearson r correlations and sensitivity/specificity calculations across studies is difficult as each of these can be independently influenced by a number of factors inherent to each data set, our study now supports a growing a body of literature that question the reliability of a biopsy at accurately representing the resection

Table 4 PD-L1 status and the tumor microenvironment

Biopsy specimen	PD-L1 negative (n=57)	PD-L1 positive (n=21)	p value
Center of tumor			
FoxP3+, median (IQR[a])	295 (167, 467)	419 (226, 589)	0.133
CD3+, median (IQR)	977 (539, 1398)	1329 (798, 1918)	0.031
CD4+, mean (SD[b])	1247 (537)	1752 (802)	0.013
CD8+, median (IQR)	168 (82, 333)	236 (101, 708)	0.101
Invasive margin			
FoxP3+, median (IQR)	239 (152, 288)	394 (187, 528)	0.055
CD3+, median (IQR)	1038 (813, 1658)	1620 (1038, 2138)	0.028
CD4+, mean (SD)	1228 (563)	1252 (604)	0.874
CD8+, median (IQR)	290 (163, 549)	436 (206, 913)	0.059
Resection specimen	**PD-L1 negative (n=43)**	**PD-L1 positive (n=35)**	**p value**
Center of tumor			
FoxP3+, median (IQR)	229 (143, 415)	386 (238, 586)	0.012
CD3+, median (IQR)	1100 (605, 1686)	545 (261, 1305)	0.062
CD4+, median (IQR)	847 (581, 1366)	1162 (570, 1680)	0.345
CD8+, median (IQR)	162 (18, 266)	254 (144, 513)	0.015
Invasive margin			
FoxP3+, median (IQR)	142 (104, 271)	308 (188, 497)	0.001
CD3+, median (IQR)	757 (523, 1013)	1100 (605, 1686)	0.008
CD4+, mean (SD)	1187 (616)	1372 (555)	0.170
CD8+, median (IQR)	378 (169, 587)	482 (260, 856)	0.116

[a]Programmed cell death-ligand 1

[b]interquartile range

[c]standard deviation

Table 5 PD-L1 accuracy

Biopsy-resection concordance	n (%)	
True positive	13 (16.7)	
False positive	8 (10.3)	
False negative	22 (28.2)	
True negative	35 (44.9)	
Test statistic	Value	95% CI[a]
Sensitivity	37.1%	21.5–55.1%
Specificity	81.4%	66.6–91.6%
Positive predictive value	61.9%	43.2–77.6%
Negative predictive value	61.4%	54.3–68.1%
Accuracy	61.5%	49.8–72.3%

Programmed cell death-ligand 1

[a]95% confidence interval

specimen TME. Our correlation of CD8+ lymphocytes at the IM ($r=0.417$) and CT ($r=0.394$) between biopsy and resection mirrors the correlation of CD8+ lymphocytes seen by Koelzer et al. ($r=0.40$) [14]. Additionally, the accuracy of biopsy intraepithelial CD3+ lymphocytes density at

predicting density in resection specimen reported by Park et al. is similarly lacking (73%) [15]. While our current study corroborates the initial data from previous studies, we have included a more exhaustive examination of the TME of both biopsy and resected specimens, including additional TIL populations (FoxP3+ and CD4+). We found no correlation between biopsy and resection specimens with regards to any of the examined TIL lines. This lack of correlation is concerning as such an assessment is becoming crucial for both deciding patient treatment and determining the efficacy of therapies.

Our study also examined PD-L1 positivity and its correlation between biopsy and resection specimens. We found significant discordance of PD-L1 status between biopsy and resection specimens, with 38.5% of patients misclassified based on the endoscopic biopsy. This inaccuracy is driven by a low sensitivity (37.1%) and thus a high false-negative rate (28.2%), which could clinically result in the withholding of potentially effective immunotherapy from these patients. A possible reason for this low accuracy is a high degree of intratumoral PD-L1 heterogeneity. PD-L1 heterogeneity is well documented in lung adenocarcinoma [16, 17] and breast cancer [18, 19] and it would be

reasonable to hypothesize that such variability might also be present in CRC. This heterogeneity likely applies not just to PD-L1, but to different TIL phenotypes and other aspects of the TME.

Understanding TME heterogeneity is particularly important because individual TIL populations [10, 20–23] as well as PD-L1/PD-1 staining [24, 25] have all been shown to have prognostic importance. However, these conclusions are based on studies that only analyzed the TME of surgical resection specimen. As researchers attempt to apply such analyses of the TME in the neoadjuvant setting, an inadequate analysis on biopsy secondary to TME heterogeneity may result in the misclassification of patients. In addition, comparing biopsy and resection specimens separated by chemo- or immunotherapy and drawing conclusions on the effect of therapy on the TME may lead to false conclusions as no reliable correlation between biopsy and resection TME has been published. We must, therefore, use caution when making important treatment decisions based on any biopsy specimen alone, and must similarly be cautious when judging the efficacy of a therapy based on the relative change in the TME that are calculated by comparing a pre-treatment biopsy to a post-treatment biopsy or resection specimen.

Future studies must focus on the development of an adequate means of comparing biopsy and resection specimens. While no studies to date have addressed this problem, we would advocate for thorough mapping of the TME using multiple samples throughout the tumor, with initial analysis aimed at understanding the implications of differences throughout the TME. More extensive sampling is also necessary both before and after treatment in window of opportunity trials to more definitively characterize the effects of any therapy on the TME. Similarly, additional studies are needed to fully understand the clinical implications of PD-L1 heterogeneity and if reviewing a greater number of slides to better characterize a patient's PD-L1 expression status can improve our ability to predict response to targeted therapies.

There are a number of limitations to our study. First, this was a retrospective study constrained by data accessible in the medical record and remaining tissue available for analysis. Lack of sufficient biopsy tissue remaining after original diagnosis certainly limited the power of this study as patients without available tissue were appropriately excluded. Second, a more thorough analysis would have addressed how individual TIL populations, and particularly those from the biopsy specimen, correlated with clinical outcome. Our data set is limited by a lack of long-term follow-up data for a number of patients, precluding our ability to perform such an analysis accurately. Finally, we did not compare multiple CT or IM TMEs from the same specimen to evaluate for spatial heterogeneity at that particular location of the TME. We would expect that a certain degree of heterogeneity would exist even within the CT or the IM, but determining

if a specific location more consistently represents the TME would be of great clinical relevance. This question should be investigated further in future studies.

In conclusion, our results demonstrate a weak correlation between TME assessed on biopsy with the TME present in the final resection specimen. In particular, the accuracy of endoscopic biopsy of assessing tumor PD-L1 status was only 61%. Until a more accurate means of assessing the TME using biopsy alone is available, we would use caution when excluding patients from specific therapies based solely on the findings of endoscopic biopsy and interpreting results of window of opportunity studies in colon cancer based on comparisons between biopsy and resection specimens.

Compliance with ethical standards

References

1. National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology: Colon Cancer (2019) https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf. Accessed 8 Feb 2020

2. Koenig JL, Toesca DAS, Harris JP, Tsai CJ, Haraldsdottir S, Lin AY, Pollom EL, Chang DT (2019) Microsatellite instability and adjuvant chemotherapy in stage II colon cancer. Am J Clin Oncol 42(7):573–580. https://doi.org/10.1097/coc.0000000000000554

3. Babcock BD, Aljehani MA, Jabo B, Choi AH, Morgan JW, Selleck MJ, Luca F, Raskin E, Reeves ME, Garberoglio CA, Lum SS, Senthil M (2018) High-risk stage II colon cancer: not all risks are created equal. Ann Surg Oncol 25(7):1980–1985. https://doi.org/10.1245/s10434-018-6484-8

4. Marcus L, Lemery SJ, Keegan P, Pazdur R (2019) FDA approval summary: pembrolizumab for the treatment of microsatellite instability-high solid tumors. Clin Cancer Res 25(13):3753–3758. https://doi.org/10.1158/1078-0432.Ccr-18-4070

5. Gibney GT, Weiner LM, Atkins MB (2016) Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Lancet Oncol 17(12):e542–e551. https://doi.org/10.1016/s1470-2045(16)30406-5

6. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, Moreno V, Warren RS, Lothe RA, Guinney J (2019) Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. Ann Oncol 30(10):1622–1629. https://doi.org/10.1093/annonc/mdz287

7. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, Lugli A, Zlobec I, Rau TT, Berger MD, Nagtegaal ID, Vink-Borger E, Hartmann A, Geppert C, Kolwelter J, Merkel S, Grutzmann R, Van den Eynde M, Jouret-Mourin A, Kartheuser A, Leonard D, Remue C, Wang JY, Bavi P, Roehrl MHA, Ohashi PS, Nguyen LT, Han S, MacGregor HL, Hafezi-Bakhtiari S, Wouters BG, Masucci GV, Andersson EK, Zavadova E, Vocka M, Spacek J, Petruzelka L, Konopasek B, Dundr P, Skalova H, Nemejcova K, Botti G, Tatangelo F, Delrio P, Ciliberto G, Maio M, Laghi L, Grizzi F, Fredriksen T, Buttard B, Angelova M, Vasaturo A, Maby P, Church SE, Angell HK, Lafontaine L, Bruni D, El Sissy C, Haicheur N, Kirilovsky A, Berger A, Lagorce C, Meyers JP, Paustian C, Feng Z, Ballesteros-Merino C, Dijkstra J, van de Water C, van Lent-van Vliet S, Knijn N, Musina AM, Scripcariu DV, Popivanova B, Xu M, Fujita T, Hazama S, Suzuki N, Nagano H, Okuno K, Torigoe T, Sato N, Furuhata T, Takemasa I, Itoh K, Patel PS, Vora HH, Shah B, Patel JB, Rajvik KN, Pandya SJ, Shukla SN, Wang Y, Zhang G, Kawakami Y, Marincola FM, Ascierto PA, Sargent DJ, Fox BA, Galon J (2018) International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 391 (10135):2128-2139. https://doi.org/10.1016/s0140-6736(18)30789-x

8. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoue F, Bruneval P, Cugnenc PH, Trajanoski Z, Fridman WH, Pages F (2006) Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 313(5795):1960–1964. https://doi.org/10.1126/science.1129139

9. Hynes SO, Coleman HG, Kelly PJ, Irwin S, O'Neill RF, Gray RT, McGready C, Dunne PD, McQuaid S, James JA, Salto-Tellez M, Loughrey MB (2017) Back to the future: routine morphological assessment of the tumour microenvironment is prognostic in stage II/III colon cancer in a large population-based study. Histopathology 71(1):12–26. https://doi.org/10.1111/his.13181

10. Nosho K, Baba Y, Tanaka N, Shima K, Hayashi M, Meyerhardt JA, Giovannucci E, Dranoff G, Fuchs CS, Ogino S (2010) Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. J Pathol 222(4):350–366. https://doi.org/10.1002/path.2774

11. Shang B, Liu Y, Jiang SJ, Liu Y (2015) Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis. Sci Rep 5:15179. https://doi.org/10.1038/srep15179

12. Huang CY, Chiang SF, Ke TW, Chen TW, You YS, Chen WT, Chao KSC (2018) Clinical significance of programmed death 1 ligand-1 (CD274/PD-L1) and intra-tumoral CD8+ T-cell infiltration in stage II-III colorectal cancer. Sci Rep 8(1):15658. https://doi.org/10.1038/s41598-018-33927-5

13. Oncology M A Pathologists Guide to PD-L1 Testing in NSCLC. https://www.keytruda.com/static/pdf/pathologist-guide-to-pd-l1-testing-in-advanced-nsclc.pdf.

14. Koelzer VH, Lugli A, Dawson H, Hadrich M, Berger MD, Borner M, Mallaev M, Galvan JA, Amsler J, Schnuriger B, Zlobec I, Inderbitzin D (2014) CD8/CD45RO T-cell infiltration in endoscopic biopsies of colorectal cancer predicts nodal metastasis and survival. J Transl Med 12:81. https://doi.org/10.1186/1479-5876-12-81

15. Park JH, van Wyk H, McMillan DC, Edwards J, Orange C, Horgan PG, Roxburgh CS (2019) Preoperative, biopsy-based assessment of the tumour microenvironment in patients with primary operable colorectal cancer. J Pathol Clin Res. https://doi.org/10.1002/cjp2.143

16. Zito Marino F, Rossi G, Montella M, Botti G, De Cecio R, Morabito A, La Manna C, Ronchi A, Micheli M, Salatiello G, Micheli P, Rocco D, Accardo M, Franco R (2019) Heterogeneity of PD-L1 expression in lung mixed adenocarcinomas and adenosquamous carcinomas. Am J Surg Pathol. https://doi.org/10.1097/pas.0000000000001400

17. Haragan A, Field JK, Davies MPA, Escriu C, Gruver A, Gosney JR (2019) Heterogeneity of PD-L1 expression in non-small cell lung cancer: Implications for specimen sampling in predicting treatment response. Lung Cancer 134:79–84. https://doi.org/10.1016/j.lungcan.2019.06.005

18. Dill EA, Gru AA, Atkins KA, Friedman LA, Moore ME, Bullock TN, Cross JV, Dillon PM, Mills AM (2017) PD-L1 expression and intratumoral heterogeneity across breast cancer subtypes and stages: an assessment of 245 primary and 40 metastatic tumors. Am J Surg Pathol 41(3):334–342. https://doi.org/10.1097/pas.0000000000000780

19. Li M, Li A, Zhou S, Xu Y, Xiao Y, Bi R, Yang W (2018) Heterogeneity of PD-L1 expression in primary tumors and paired lymph node metastases of triple negative breast cancer. BMC Cancer 18(1):4. https://doi.org/10.1186/s12885-017-3916-y

20. Yoon HH, Orrock JM, Foster NR, Sargent DJ, Smyrk TC, Sinicrope FA (2012) Prognostic impact of FoxP3+ regulatory T cells in relation to CD8+ T lymphocyte density in human colon carcinomas. PLoS ONE 7(8):e42274. https://doi.org/10.1371/journal.pone.0042274

21. Lee WS, Park S, Lee WY, Yun SH, Chun HK (2010) Clinical impact of tumor-infiltrating lymphocytes for survival in stage II colon cancer. Cancer 116(22):5188–5199. https://doi.org/10.1002/cncr.25293

22. Flaherty DC, Lavotshkin S, Jalas JR, Torisu-Itakura H, Kirchoff DD, Sim MS, Lee DJ, Bilchik AJ (2016) Prognostic utility of immunoprofiling in colon cancer: results from a prospective, multicenter nodal ultrastaging trial. J Am Coll Surg 223(1):134–140. https://doi.org/10.1016/j.jamcollsurg.2016.03.003

23. Sinicrope FA, Rego RL, Ansell SM, Knutson KL, Foster NR, Sargent DJ (2009) Intraepithelial effector (CD3+)/regulatory (FoxP3+) T-cell ratio predicts a clinical outcome of human colon carcinoma. Gastroenterology 137(4):1270–1279. https://doi.org/10.1053/j.gastro.2009.06.053

24. Wyss J, Dislich B, Koelzer VH, Galvan JA, Dawson H, Hadrich M, Inderbitzin D, Lugli A, Zlobec I, Berger MD (2019) Stromal PD-1/PD-L1 expression predicts outcome in colon cancer patients. Clin Colorectal Cancer 18(1):e20–e38. https://doi.org/10.1016/j.clcc.2018.09.007

25. Enkhbat T, Nishi M, Takasu C, Yoshikawa K, Jun H, Tokunaga T, Kashihara H, Ishikawa D, Shimada M (2018) Programmed cell death ligand 1 expression is an independent prognostic factor in colorectal cancer. Anticancer Res 38(6):3367–3373. https://doi.org/10.21873/anticanres.12603