Histopathology

Histopathology 2026, **88**, 868–880. DOI: 10.1111/his.70043

Interobserver variability of histopathological assessment in pT1 colorectal carcinoma

Karmele Saez de Gordoa,[1,2,3] María Daca-Alvarez,[4,5] Maite Rodrigo-Calvo,[1,2,3] Ivan Archilla,[1,2,3] Sandra Lopez-Prades,[1,2,3] Jose Javier Aguirre,[6,7] Lorena Alarcón-Molero,[8] María Cámara Jurado,[9] Adriana Canosa,[10] Francisco Giner,[11,12] Carmen González-Lois,[13] Mireya Jimeno,[14,15] Ismael Jurado,[16] Isidro Machado,[17,18] Carolina Martínez-Ciarpaglini,[12,19,20] Eva Musulen,[21,22] Dolores Naranjo,[23] Natalia Papaleo,[24] Cristina Peña,[25] Òria Rosiñol,[26] Rosario Sánchez-Yuste,[13] Greissy Tibisay Vázquez Benítez,[13] María Pellisé,[4,5,27] Miriam Cuatrecasas[1,2,3,27] & on behalf of the EpiT1 Consortium

[1]*Pathology Department, Hospital Clínic,* [2]*Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS),* [3]*Department of Clinical Foundations,* [4]*Department of Medicine, Universitat de Barcelona,* [5]*Gastroenterology Department, Hospital Clínic, Barcelona,* [6]*Pathology Department, Hospital Universitario Araba,* [7]*Bioaraba HRI, Vitoria-Gasteiz,* [8]*Pathology Department, Hospital General de Valdepeñas, Ciudad Real,* [9]*Pathology Department, Hospital Universitario 12 de Octubre, Madrid,* [10]*Pathology Department, Hospital Universitario de La Ribera, Alzira,* [11]*Pathology Department, Hospital Universitario y Politécnico La Fe,* [12]*University of Valencia, Valencia,* [13]*Pathology Department, Hospital Universitario Puerta de Hierro, Majadahonda,* [14]*Pathology Department, Hospital Universitari Germans Trias i Pujol,* [15]*Institut Germans Trias i Pujol (IGTP), Badalona,* [16]*Pathology Department, Hospital Universitari Consorsi Sanitari de Terrassa, Terrassa,* [17]*Pathology Department, Instituto Valenciano de Oncologia,* [18]*Pathology Laboratory, Hospital QuirónSalud,* [19]*Pathology Department, Hospital Clínico Universitario of Valencia,* [20]*Biomedical Research Institute INCLIVA, Valencia,* [21]*Pathology Department, Hospital Universitari General de Catalunya-Grupo QuirónSalud, Sant Cugat del Vallès,* [22]*Cancer Epigenetics Group, Josep Carreras Leukaemia Research Institute (IJC), Badalona,* [23]*Pathology Department, Hospital del Mar, Barcelona,* [24]*Pathology Department, Corporació Sanitària Parc Taulí, Sabadell,* [25]*Pathology Department, Complejo Hospitalario Universitario, Albacete,* [26]*Pathology Department, Althaia Xarxa Assistencial Universitària, Manresa and* [27]*Centro de Investigación Biomédica en Red en Enfermedades Hepáticas y Digestivas (CIBEREHD), Madrid, Spain*

Date of submission 12 June 2025
Accepted for publication 26 October 2025

de Gordoa K S, Daca-Alvarez M, Rodrigo-Calvo M, Archilla I, Lopez-Prades S, Aguirre J J, Alarcón-Molero L, Jurado M C, Canosa A, Giner F, González-Lois C, Jimeno M, Jurado I, Machado I, Martínez-Ciarpaglini C, Musulen E, Naranjo D, Papaleo N, Peña C, Rosiñol Ò, Sánchez-Yuste R, Benítez G T V, Pellisé M, Cuatrecasas M, the EpiT1 Consortium
(2026) *Histopathology* **88**, 868–880. https://doi.org/10.1111/his.70043

Interobserver variability of histopathological assessment in pT1 colorectal carcinoma

Abbreviations: AC1, Agreement Coefficient 1; CRC, colorectal carcinoma; DSI, deep submucosal invasion; ESMO, European Society of Medical Oncology; H&E, haematoxylin-eosin; ICC, intraclass correlation coefficient; IHC, immunohistochemistry; ITBCC, International Tumour Budding Consensus Conference; JSCCR, Japanese Society for Cancer of the Colon and Rectum; LNM, lymph node metastasis; LVI, lymphovascular invasion; NCCN, National Comprehensive Cancer Network; PDC, poorly differentiated clusters; PNI, perineural invasion; RCPath, Royal College of Pathologists; TB, tumour budding; WHO, World Health Organization.

Address for correspondence: Miriam Cuatrecasas, Department of Clinical Foundations, Faculty of Medicine and Health Sciences, University of Barcelona (UB), Pathology Department, Hospital Clinic – IDIBAPS – UB, Villarroel, 170, 08036, Barcelona, Spain. e-mail: mcuatrec@clinic.cat
María Pellisé and Miriam Cuatrecasas contributed equally to this work

Aims: Pathological evaluation of colorectal carcinoma (CRC) diagnosed at stage pT1 is challenging. Nevertheless, it is crucial for treatment guiding and to determine the patient's prognosis. This study aimed to assess the interobserver variability in the histopathological evaluation of pT1 CRC.

Methods and results: A retrospective multicentre pT1 CRC cohort study was designed (EpiT1 consortium). A task force comprising 20 experienced pathologists conducted the histopathological evaluation using digitalized haematoxylin-eosin (H&E) slides. A pilot study was performed with 10 cases, and afterwards, a consensus meeting was held to assess interobserver variability. Then, a concordance study was performed by assessing 70 new pT1 CRC cases. We used percentage agreement and Gwet's Agreement Coefficient 1 for categorical variables, and intraclass correlation coefficient (ICC) for continuous variables. In the pilot study, histological grade and perineural invasion (PNI) demonstrated 100% agreement, with good concordance for lymphovascular invasion (LVI), tumour budding (TB), poorly differentiated clusters (PDC) and margin assessment. The concordance study showed high agreement (≥90%) on histological grade, PDC, PNI and LVI. Submucosal invasion depth showed excellent reliability in the concordance study (ICC = 0.97). Notably, in both studies, the agreement of PDC was higher than for TB. Lower concordance was observed on stromal lymphocytes and the status of muscularis mucosae.

Conclusions: Our results emphasize the need for standardization in evaluating pT1 CRC to improve the concordance among pathologists, and the precision of digital measurements. Moreover, the addition of PDC assessment in pT1 CRC diagnostic guidelines could help to improve the accuracy of risk stratification and reliably predict prognosis.

Keywords: colorectal carcinoma, concordance, interobserver, poorly differentiated clusters, T1 CRC, tumour budding, variability

Introduction

The high prevalence and mortality of colorectal carcinoma (CRC)[1] has led to the implementation of CRC screening programs in many countries, achieving early detection and reducing mortality rates. Approximately 40% of CRC cases detected in the context of screening programs are diagnosed at pT1 stage.[2]

An accurate pT1 CRC histopathological evaluation, particularly in the identification of risk factors for lymph node metastasis (LNM), has significant implications for patient management.[3–7] Recent studies have established histological grade, tumour budding (TB) and lymphovascular invasion (LVI) as the most significant predictors of LNM. Contrarily, deep submucosal invasion (DSI ≥1000 μm) may not be an independent risk factor.[8] Other histological features, such as poorly differentiated clusters (PDC), perineural invasion (PNI), muscularis mucosae disruption, stromal lymphocytes and submucosal invasion width and area, have also been proposed as additional risk factors.[9–15]

According to the European Society of Medical Oncology (ESMO) guidelines, surgical resection including lymphadenectomy is recommended for patients with pT1 CRC exhibiting at least one of the following histological risk factors: lymphatic or venous invasion, high tumour grade, intermediate-to-high TB (Bd2 or Bd3), or Haggitt level 4 submucosal invasion in pedunculated polyps. In contrast, it recommends endoscopic surveillance for patients without risk factors or with only a positive resection margin.[3] The National Comprehensive Cancer Network (NCCN) Guidelines follow a similar approach, and consider histologically high-grade, angiolymphatic invasion or a positive resection margin as unfavourable histological features. Of note, a definition of positive resection margin is lacking, which has been defined as the presence of tumour <1 mm or <2 mm from the transected margin, or tumour cells present within the diathermy artefact. Additionally, the NCCN states that high TB may be associated with worse outcomes.[4] The Royal College of Pathologists (RCPath) dataset for histopathological reporting of CRC considers LVI, high TB (Bd2-Bd3) and high tumour grade as pT1 CRC risk factors for LNM.[7] The Japanese Society for Cancer of the Colon and Rectum (JSCCR) recommends intestinal resection with lymph node dissection in cases of pT1 CRC with an affected vertical margin, DSI (pT1b, ≥1000 μm), LVI, poor differentiation, signet ring cell or mucinous carcinoma, or high TB (Bd2-Bd3).[5]

As stated, the guidelines among different professional societies are not standardized. In addition, the

significant interobserver variability impacts patient risk classification, which could alter treatment decisions from 9.7% to 30% of patients.[16] Therefore, given the reliance on pathological evaluation for therapeutic decisions, a standardized and reliable histological assessment is of utmost importance for optimal patient management.[17]

This study aimed to assess the interobserver variability in evaluating histopathological features in pT1 CRC, with the goal of improving the reproducibility and clinical utility of these criteria to enhance patient's risk stratification and reduce over- or undertreatment.

Materials and Methods

STUDY DESIGN

This study is part of a population-based multicentre cohort study involving 33 tertiary hospitals across 12 Spanish regions (EpiT1 consortium). All included patients were diagnosed with pT1 CRC between 2007 and 2018, regardless of the treatment received. K.S.G. and M.C. selected a sample of 80 cases for pathological evaluation, based on two criteria: presence of unequivocal submucosal invasion and en bloc resection. Cases with faint staining or marked artefact were excluded. In those cases where the tumour was represented in more than one slide, the most representative section was selected. A task force was created comprising 20 experienced pathologists, with a median of 12 years of practice. Most pathologists (85%) worked in tertiary or university hospitals or were specialized in gastrointestinal pathology (75%). All pathologists were based in hospitals involved in CRC screening programs, diagnosing a mean of 3.78 pT1 CRC cases per month.

The histopathological evaluation was carried out using high-resolution scanned haematoxylin-eosin (H&E) slides (200× magnification; Pannoramic 250 FLASH II, 3DHISTECH). The slides were viewed through a digital platform (Digital Slide Archive - DSA: https://dsa.athenatechai.com), which enabled making exact measurements. Each pathologist had a personal user account which did not allow seeing the measurements made by others. A pilot phase was first conducted, where all 20 pathologists evaluated the same 10 representative cases, four pedunculated and six sessile pT1 CRC. Seven cases were initially endoscopically treated and ultimately underwent surgical resection due to clinical decision, and three patients were only surgically treated. The histological evaluation was performed on the endoscopic resected tumour in seven cases (70%). Then, a consensus meeting was held to discuss the most discordant histological features. Subsequently, an interobserver concordance study was performed on 70 new cases, 34% pedunculated and 66% sessile, including 45 endoscopically treated and 25 surgically resected pT1 CRC. The histological evaluation was performed on the endoscopic resected tumour (45/70; 65%). For all tumours, each pathologist analyzed all histological features, made measurements on the digitized slides and recorded the data in a database. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Hospital Clinic de Barcelona (HCB/2019/0224) on 21/05/2019.

CRITERIA FOR EVALUATING HISTOLOGICAL PARAMETERS

Histological grade was determined as described in the 5th edition of the World Health Organization (WHO) Classification of Digestive System Tumours, based on gland formation on the least differentiated component and divided into low and high grade.[18] LVI was defined as the presence of tumour cells in the lumen of channels coated by endothelial cells and comprised lymphatic or venous invasion. TB was evaluated according to The International Tumour Budding Consensus Conference (ITBCC) definition and was divided into Bd1 (low; 0–4 buds), Bd2 (intermediate; 5–9 buds) and Bd3 (high; ≥10 buds).[19] Two cases from the pilot phase and 15 from the concordance study showed prominent inflammation and glandular rupture, making TB not reliably assessable. These cases were omitted from the analysis of TB, since we had no other areas for TB assessment, as recommended by ITBCC Reporting Consensus[20] and the RCPath, which advise excluding areas with active inflammation.[7] PDC were defined as groups of ≥5 tumour cells, at the infiltrating margin or within the tumour, which were not forming a glandular structure. PDC were counted at the hotspot area of 0.785 mm^2 and classified into G1 (0–4 PDCs), G2 (5–9 PDCs) and G3 (≥10 PDCs).[9] PNI was defined as present when tumour cells surrounded nerve sheaths. Stromal lymphocytes included the mononuclear infiltrate (lymphocytes and plasma cells), and were assessed within the entire invasive tumour area, excluding granulocytes, necrotic areas and non-infiltrating carcinoma. It was scored as: absent or low (0%–10%), intermediate (15%–50%) and high (55%–100%).[21,22] The status of the muscularis mucosae was divided into intact, partially disrupted and completely disrupted.[12]

The depth of submucosal invasion was assessed differently depending on the polyp morphology. On sessile or flat polyps with intact or partially disrupted muscularis mucosae, the invasion was measured from the lower part of the muscularis mucosae to the deepest invading tumour cell (Figure 1A). When the muscularis mucosae was effaced by the tumour, the depth of invasion was measured from the surface of the tumour to the deepest invading tumour cell (Figure 1B).[11] In well-oriented pedunculated polyps, Haggitt's classification was used to evaluate the level of invasion as follows: Haggitt 1 when submucosal invasion was limited to the head of the polyp, Haggitt 2 for tumour invasion that reached the neck of the polyp, Haggitt 3 for invasion into the polyp stalk, and Haggitt 4 for tumour invasion beyond the stalk, extending to the submucosa.[23] In addition, a quantitative submucosal depth invasion was measured in μm from the Haggitt line at the junction between normal and neoplastic mucosa (Figure 1C). For flat, non-pedunculated carcinomas, Kikuchi's classification was also used to assess the level of submucosal invasion only when the muscular propria was identified and the entire submucosal layer could be evaluated. It was divided into sm1, sm2 and sm3, corresponding to invasion of the superficial third, two-thirds, or the deepest third of the submucosa, respectively.[24] According to the JSCCR, the European Society of Gastrointestinal Endoscopy and European Society of Digestive Oncology guidelines,[5,6] a submucosal invasion of ≥1000 μm was considered a risk factor for LNM. Moreover, a submucosal invasion of ≥2000 μm was identified as a risk factor for LNM by Ueno *et al.*[25] Therefore, we divided this feature into three categories (<1000 μm, 1000–1999 μm and ≥2000 μm).

The width of the adenocarcinoma included the maximum horizontal size of the malignant component and was further divided into two categories (<4000 μm and ≥4000 μm), as it has been considered to confer a higher risk for LNM.[25] The polyp width included both the adenoma and the carcinoma components (Figure 1D).[15] The area of infiltrating adenocarcinoma was evaluated by measuring the total area of the submucosal invasive tumour, excluding the superficial dysplastic part.[15]

In this study, the vertical or deep margin was considered affected by carcinoma when the tumour cells presented cautery artefact, were in contact, or at less than 0.1 mm from the inked margin.[26] Free margin was divided into three categories: at 1 mm or less, at

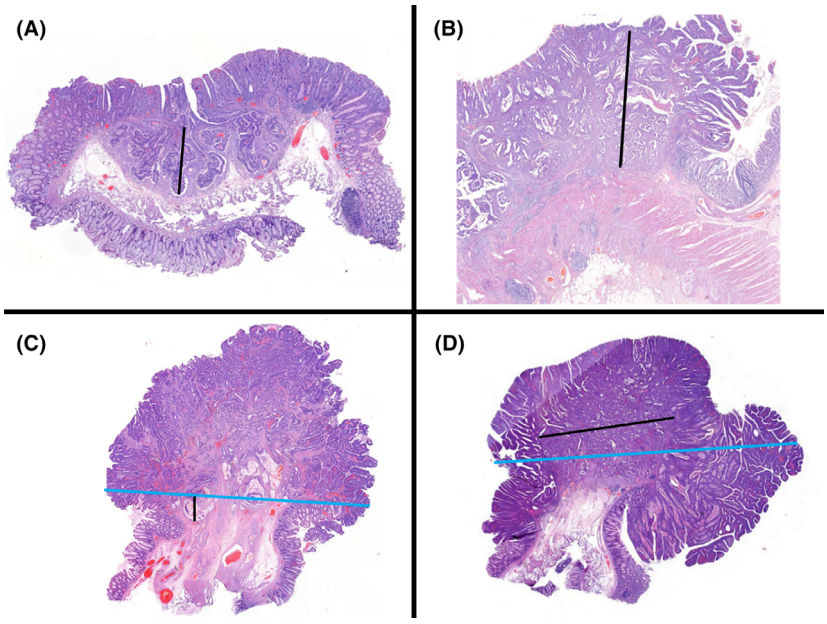

Figure 1. Measurement of pT1 CRC submucosal invasion on sessile and pedunculated lesions. (**A**) Sessile polyp with visible muscularis mucosae. The invasion depth (black line) is measured from the lower part of the muscularis mucosae to the deepest tumour cell (haematoxylin-eosin [H&E] 1×). (**B**) Sessile polyp with complete effacement of the muscularis mucosae. The invasion depth (black line) is measured from the surface of the lesion to the deepest tumour cell (H&E 1×). (**C**) Pedunculated polyp. The Haggitt line (blue line) separates the head from the stalk of the polyp. Submucosal invasion depth (black line) is measured from the Haggitt line to the deepest part of the tumour (H&E; 1×). (**D**) The size of the polyp comprises the whole lesion, including both the carcinoma and adenoma components (blue line). The width of the carcinoma only includes the carcinoma component (black line) (H&E; 1×).

more than 1 mm, and indeterminate when there was fragmentation or marked artefact. The horizontal margin or lateral mucosal margin was defined as free, indeterminate and affected by adenoma with low- or high-grade dysplasia, in situ carcinoma or by infiltrating carcinoma.

We stratified each case according to its risk of LNM, including the unfavourable features included in the ESMO and NCCN guidelines. Hence, cases at risk showed high-grade, LVI, high TB (Bd2 or Bd3) or a positive margin (in contact with the tumour).[3,4]

STATISTICAL ANALYSIS

The category with the highest agreement among pathologists was considered the gold standard diagnosis. All clinicopathological characteristics evaluated in both studies are shown in Table 1. The concordance values of continuous variables were calculated using the average intraclass correlation coefficient (ICC), two-way random-effects model (with 95% confidence intervals). It was interpreted as poor reliability in the case of ICC lower than 0.50; moderate reliability between 0.5 and 0.75; good reliability between 0.75 and 0.9; and excellent reliability with an ICC greater than 0.9.[27] The percentage agreement value (median and IQR – interquartile range) and Gwet's Agreement Coefficient 1 (AC1) were employed for assessing the concordance of categorical variables.

Results

PILOT STUDY

The evaluation of the histological grade and PNI had 100% agreement among pathologists, with an AC1 of 0.88 and 0.99, respectively. The agreement on PDC (90.91%; 0.83) was higher than that for TB (77.27%; 0.63). The evaluation of LVI, horizontal margin and width of adenocarcinoma showed high agreement, whereas the agreement in the assessment of stromal lymphocytes, status of muscularis mucosae, Haggitt's and Kikuchi's classifications for the depth of submucosal invasion and the vertical margin was below 75%. Finally, the percentage of agreement for risk stratification according to the ESMO and NCCN guidelines was 74.57% (AC1 0.36). For continuous variables, a high ICC was reached in the evaluation of the polyp size, width of adenocarcinoma and the area of the submucosal invasion by the adenocarcinoma, with excellent reliability. Submucosal invasion depth had moderate reliability (Table 2).

CONCORDANCE STUDY

High agreement was achieved in the evaluation of histological grade, TB, PDC, LVI, PNI, horizontal margin and the width of the adenocarcinoma. The evaluation of stromal lymphocytes, disruption of muscularis mucosae, vertical margin, depth of submucosal invasion, Haggitt's and Kikuchi's classifications showed poorer agreement. The agreement on the ESMO and NCCN guidelines risk stratification was 80.65% and AC1 0.28. In the case of ICC, submucosal invasion depth had excellent reliability, with ICC of 0.97. The polyp size and the width of the adenocarcinoma had moderate reliability, and the area of the infiltrating adenocarcinoma had poor reliability (Table 2).

COMPARISON OF BOTH STUDIES

In both studies, all measurements were made on digitalized H&E slides; therefore, exact measurements could be performed and were comparable. The concordance of most of the histological features improved in the second assessment. The highest improvement was observed in the measurement of the submucosal invasion depth as a continuous variable, reaching an excellent reliability (ICC = 0.97). Haggitt's and Kikuchi's classifications improved to 94.44% and 73.68% of agreement, respectively. The agreement of PDC was higher than for TB in both studies, with a higher improvement of TB in the second study. The assessment of LVI and horizontal margin had small differences, while histological grade and PNI remained the same, with 100% agreement. Contrarily, the agreement of stromal lymphocytes and disruption of muscularis mucosae was lower in the second assessment. Nevertheless, the vertical margin improved in the second assessment. The ICC of the polyp size, width of adenocarcinoma and area of infiltrating adenocarcinoma was lower in the concordance study (Table 2).

Of note, cautery artefact and inflammation were causes of discordances, especially for the assessment of PDC, TB and margin status. In addition, tumours with mucinous differentiation generated challenges in the evaluation of LVI, TB and PDC (Figure 2). On the contrary, other cases were evaluated more homogeneously regarding grade, TB, submucosal invasion depth or PDC (Figure 3).

PATIENTS' FOLLOW-UP AND ADVERSE OUTCOMES

All the pilot study patients ultimately underwent surgical resection. Nine patients had at least 5 years of

Table 1. Histological features of included cases in both the pilot and the concordance study

	Pilot study (10 cases)		Concordance study (70 cases)	
	N	%	*N*	%
Histological grade				
Low-grade	10	100	67	95.71
High-grade	0	0	3	4.29
Tumour budding (TB)	Non-inflamed cases (8)		Non-inflamed cases (55)	
Bd1 (0–4 buds)	8	100	52	94.55
Bd2 (5–9 buds)	0	0	0	0
Bd3 (>10 buds)	0	0	3	5.45
Poorly differentiated clusters (PDC)				
G1 (0–4 PDC)	10	100	67	95.71
G2 (5–9 PDC)	0	0	2	2.86
G3 (>10 PDC)	0	0	1	1.43
Lymphovascular invasion (LVI)				
Yes	1	10	6	8.57
No	9	90	64	91.43
Perineural invasion (PNI)				
Yes	0	0	0	0
No	10	100	70	100
Stromal lymphocytes				
Absent/Low grade (0%–10%)	4	40	23	32.86
Intermedium (15%–50%)	5	50	43	61.43
High grade (55%–100%)	1	10	4	5.71
Status of muscularis mucosae				
Intact	0	0	0	0
Incompletely disrupted	1	10	21	30
Completely disrupted	9	90	49	70
Vertical margin				
Indeterminate	2	20	0	0
Free at >1 mm	6	60	52	74.29
Free between 0.1 and 1 mm	1	10	13	18.57
Positive	1	10	5	7.14
Horizontal margin				
Indeterminate	2	20	0	0

Continued

Table 1. (*Continued*)

	Pilot study (10 cases)		Concordance study (70 cases)	
	N	%	*N*	%
Free	8	80	68	97.14
Affected by adenoma or in situ carcinoma	0	0	1	1.43
Affected by infiltrating carcinoma	0	0	1	1.43
Haggitt's classification	Endoscopic samples (7 cases)		Endoscopic samples (45 cases)	
NA (not pedunculated polyp)	4	57.14	25	55.56
Haggitt 1	2	28.57	11	24.44
Haggitt 2	1	14.29	4	8.89
Haggitt 3	0	0	5	11.11
Haggitt 4	0	0	0	0
Kikuchi's classification	Surgical samples (3 cases)		Surgical samples (25 cases)	
NA (non-flat or sessile polyp)	0	0	4	16
sm1	1	33.33	4	16
sm2	1	33.33	4	16
sm3	1	33.33	13	52
Submucosal invasion depth				
<1000 μm	2	20	17	24.29
1000–1999 μm	0	0	9	12.86
≥2000 μm	8	80	44	62.86
Width of adenocarcinoma				
<4000 μm	1	10	4	5.71
≥4000 μm	9	90	66	94.29
Risk stratification				
Low-risk	7	70	48	68,57
High-risk	3	30	22	31,43

follow-up, and one was lost to follow up. One case without risk histological features had LNM and received chemotherapy, being alive without disease after 6 years of follow-up. No disease recurrence or death occurred in the rest of the patients.

In the concordance study, 36 patients underwent surgery (25 as the initial treatment and 11 after endoscopic resection). Seven patients (19.4%) had LNM; six of them were classified as high-risk, but none of them recurred. Among the six high-risk patients with LNM, the most frequent features observed were LVI in four patients, followed by TB in two, high grade and a positive vertical margin in one patient, respectively. Two died due to causes unrelated to CRC. Four other patients had distant metastases; two of them had high-risk histopathological features and the two low-risk patients died of disease at 57 and 63 months after diagnosis. Seven additional patients died due to other causes.

Table 2. Concordance values on the histopathological features of the tumours included in the study

Categorical variables	Pilot study (N = 10) Median percentage of agreement (IQR)	Concordance study (N = 70) Median percentage of agreement (IQR)	Pilot study (N = 10) Gwet's AC1 (0.95 CI)	Concordance study (N = 70) Gwet's AC1 (0.95 CI)
Histological grade	100% (88.64 to 100.00)	100% (94.20 to 100.00)	0.88 (0.81 to 0.96)	0.84 (0.76 to 0.91)
Tumour budding (TB)[a]	77.27% (67.05 to 87.01)	89.47% (74.35 to 100)	0.63 (0.46 to 0.81)	0.66 (0.54 to 0.78)
Poorly differentiated clusters (PDC)	90.91% (87.39 to 95.45)	93.93% (78.33 to 100.00)	0.83 (0.74 to 0.92)	0.73 (0.63 to 0.83)
Lymphovascular invasion (LVI)	90.91% (85.88 to 95.45)	94.43% (79.80 to 100.00)	0.71 (0.57 to 0.85)	0.68 (0.57 to 0.79)
Perineural invasion (PNI)	100% (100 to 100)	100% (100 to 100)	0.99 (0.97 to 1)	0.96 (0.95 to 0.98)
Stromal lymphocytes	61.36% (54.55 to 71.5)	52.63% (45.10 to 60.54)	0.23 (0.14 to 0.33)	0.26 (0.16 to 0.38)
Status of muscularis mucosae	72.73% (63.64 to 79.46)	68.42% (59.40 to 78.95)	0.66 (0.53 to 0.79)	0.42 (0.34 to 0.51)
Vertical margin	73.11% (50.00 to 95.45)	78.95% (61.11 to 94.74)	0.24 (0.16 to 0.32)	0.45 (0.40 to 0.49)
Horizontal margin	93.56% (73.86 to 98.86)	90.19% (81.53 to 100.00)	0.57 (0.47 to 0.67)	0.77 (0.68 to 0.81)
Haggitt's classification[b]	72.73% (63.64 to 79.55)	94.44% (68.42 to 100.00)	0.22 (0.15 to 0.29)	0.30 (0.26 to 0.35)
Kikuchi's classification[c]	66.67% (66.67 to 69.70)	73.68% (57.89 to 89.47)	0.07 (−0.16 to 0.29)	0.13 (0.09 to 0.16)
Submucosal invasion depth	70.45% (52.08 to 89.77)	88.85% (75.37 to 100.00)	0.39 (0.23 to 0.55)	0.34 (0.27 to 0.42)
Width of adenocarcinoma	95.45% (77.84 to 100)	94.44% (88.89 to 94.74)	0.69 (0.56 to 0.82)	0.75 (0.69 to 0.82)
Risk stratification	76.84% (69.32 to 85.23)	82.84% (63.54 to 94.49)	0.20 (0.01 to 0.397)	0.17 (0.04 to 0.29)

Continuous variables	Pilot study (N = 10) Average ICC (95% confidence interval)	Concordance study (N = 70) Average ICC (95% confidence interval)
Submucosal invasion depth	0.59 (0.12 to 0.88)	0.97 (0.96 to 0.98)
Polyp size	0.96 (0.91 to 0.99)	0.55 (0.37 to 0.69)
Width of adenocarcinoma	0.94 (0.87 to 0.98)	0.71 (0.60 to 0.80)
Area of infiltrating adenocarcinoma	0.97 (0.93 to 0.99)	0.06 (−0.30 to 0.35)

ICC, intraclass correlation coefficient; IQR, interquartile range.
[a]Evaluated in cases without prominent inflammation (8 in the pilot phase and 55 in the concordance phase).
[b]Haggitt's classification was only evaluated on endoscopic samples (7 cases in the pilot study and 45 cases in the concordance study).
[c]Kikuchi's classification could only be assessed on surgical samples (3 cases in the pilot study and 25 cases in the concordance study).

Focusing on PDC, all patients from the pilot study had G1 PDC. In the concordance study, three patients had G2 or G3 PDC with other high-risk features. One of the latter patients had distant metastasis detected 45 months after surgery. The other two patients did not show a recurrence after 5 years of follow-up.

Discussion

An accurate histopathological assessment is determinant for pT1 CRC patient management. This study underscores the complexity of some histopathological features that need to be assessed in pT1 CRC. This can result in significant variability among pathologists. However, it also demonstrates that high inter-observer agreement can be achieved when clear criteria are defined, and that the use of digital slides improves the precision of measurements.

The three histopathological features most related to prognosis in pT1 CRC are histological grade, LVI and TB; thus, a precise and reproducible assessment is essential. We found perfect agreement (100%) on histological grading using a two-tiered system (low vs. high grade), as recommended by the WHO.[18] The use

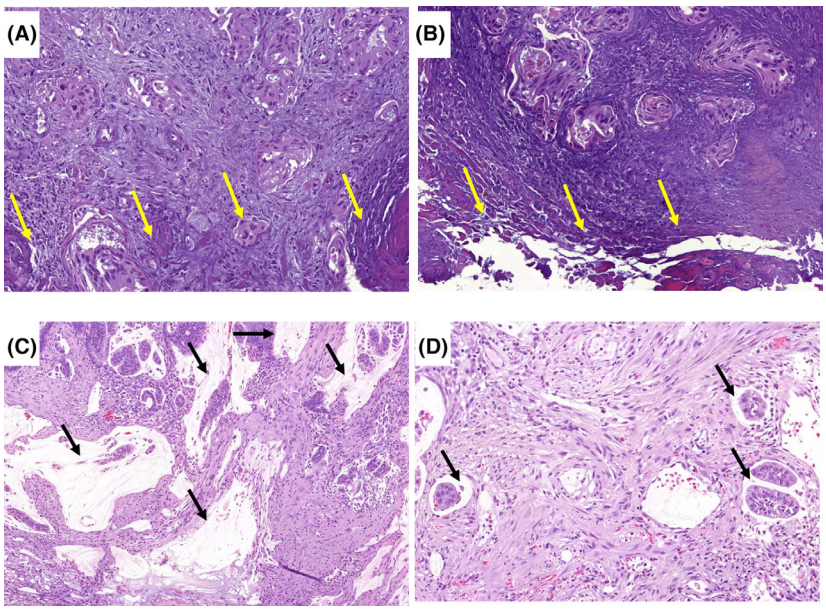

Figure 2. Difficult evaluation of tumour budding (TB) and poorly differentiated clusters (PDC) due to inflammation and cautery artefact. (**A**, **B**) The marked cautery artefact (yellow arrows) present in this case challenged the evaluation of the deep margin and TB at the invading front (haematoxylin-eosin [H&E]; 20×); (**C**) Tumour with mucinous component (black arrows), which made difficult the assessment of PDC and TB (H&E; 10×); (**D**) Tumour groups surrounded by a white space (black arrows), made it difficult with H&E alone to differentiate shrinkage artefact from lymphovascular invasion (LVI) (H&E; 20×).

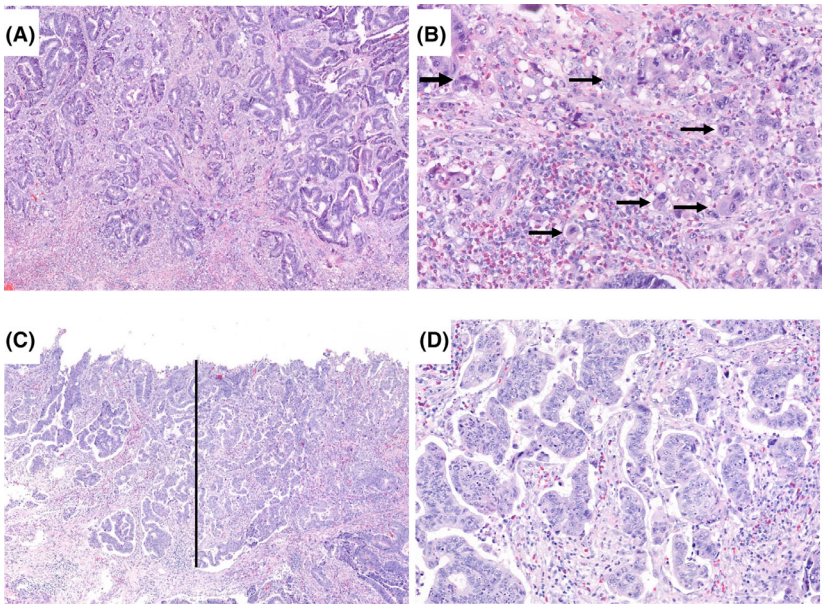

Figure 3. Examples of the assessment of some histological features. (**A**) A tumour showing mostly glandular architecture, therefore classified as low grade by most observers (haematoxylin-eosin [H&E]; 5×); (**B**) Despite the inflammatory component, most observers agreed on the presence of high TB (black arrows) (H&E; 20×). (**C**) This case had surface ulceration, and submucosal invasion depth was measured from the surface (black line) (H&E; 5×); (**D**) Tumour composed mostly of clusters of cells with high agreement on PDC grade (H&E; 20×).

of three or four grading tiers based on gland formation often leads to lower agreement.[28–30] Importantly, the recent WHO, RCPath and International Collaboration on Cancer Reporting have recommended grading pT1 CRC based on the least differentiated component, rather than on the predominant

one.[7,18,31] Although the WHO two-tiered grading system approach has been demonstrated to increase interobserver agreement, it may result in tumour upgrading and impact directly on patients' prognosis, compared to the 8th edition of the American Joint Committee on Cancer.[32] Thus, there is a need to reliably grade CRC, and to demonstrate the WHO system's clinical value.[32] Concerning the assessment of lymphatic and blood vessel invasion (LVI) on H&E stains, we obtained an almost perfect level of agreement (90%), while other authors have reported slight or fair agreement.[29,30] When needed, the incorporation of immunohistochemistry (IHC) markers such as CD31, D2-40, or cytokeratins, has been shown to improve the concordance and accuracy of LVI assessment.[33–35] TB is another important prognostic factor in pT1 CRC, strongly associated with LNM.[36] Although some studies have shown moderate to fair agreement on TB assessment using H&E slides,[30,33] we had an almost perfect agreement after consensus (89.47%). The ITBCC Reporting Consensus and the RCPath dataset recommend excluding areas with active inflammation for TB assessment, or performing cytokeratin IHC for TB visualization.[7,20] Nevertheless, there is still some controversy about adding cytokeratin IHC to improve TB detection, based on the risk of overdiagnosis due to cytoplasmic pseudobuds.[29,33,37,38] As for TB risk classification, the suggestion of a two-tiered classification, i.e., low-grade Bd1 versus high-grade Bd2/3, may further improve interobserver concordance.[29,37,39] We found PDC to be a highly reliable feature, with an almost perfect agreement (93.93%). This is aligned with reports of high interobserver consistency on PDC assessment for being easier to identify on H&E slides.[40,41] In addition, PDC has been proposed as a surrogate or even a substitute for tumour grade, due to its reproducibility and reliable risk stratification.[40–43] Notably, the presence of G2 or G3 PDC is not frequently observed in pT1 CRC. In the concordance study, one of the three patients with G2-3 PDC had distant metastases. Furthermore, the inclusion of PDC in pT1 CRC guidelines could help improve the accuracy of risk stratification, as there is evidence that PDC can better predict LNM in endoscopically resected pT1 CRC.[7,41–43] However, broader multicentre validation and stronger evidence in this field are needed, prior to the introduction of PDC as a prognostic factor into clinical practice.[7]

Although the depth of submucosal invasion and the width of invasion are commonly included in risk stratification models, the assessment of these parameters can be challenging. As in other studies, we obtained excellent agreement after consensus of standardized criteria on the measurement of submucosal invasion depth, with substantial improvement in reliability (ICC = 0.97) compared to the pilot study (ICC = 0.59).[44] The width of invasion also showed good agreement, and the use of the Ueno system for categorizing invasion (with a 4000 μm cut-off) provided robust results in our study.[25] Contrarily, we had lower agreement on the assessment of margin status, particularly the vertical margin. While vertical margin assessment improved after the criteria were clarified in the consensus meeting, it still had lower concordance compared to horizontal margin assessment, probably due to the fact that it was divided into four categories. In previous studies, margin status agreement was reported as moderate to very good, but it was divided into two categories, affected or unaffected (R1 or R0, respectively).[29,33] Finally, the features where we obtained lower agreement were the status of the muscularis mucosae and the assessment of stromal lymphocytes, likely due to the infrequent reporting of these features in clinical practice. Likewise, other studies have found only fair to moderate agreement on the evaluation of muscularis mucosae disruption,[29,30] and stromal lymphocytes, the latter having shown varying levels of concordance depending on the cut-off system used.[45]

Although we achieved a high overall pT1 CRC risk stratification agreement (82.84%), the relatively low AC1 (0.17) suggests that discordances in evaluating some key histological parameters contributed to the variability of patient risk assignment.[3,4] This is consistent with previous reports showing moderate to good agreement in risk assessments using H&E.[33] Therefore, a consistent and reproducible assessment of the most relevant pathologic parameters of pT1 CRC is paramount due to the significant clinical implications. As we demonstrated, this can be achieved through consensus meetings and can be easily incorporated in guidelines. Another important issue is the experience and subspecialty of pathologists, which has been assumed to play a significant role in enhancing agreement, particularly in TB evaluation. Expert pathologists had better concordance on TB assessment with H&E alone, while junior pathologists benefited from IHC staining.[37] Similarly, subspecialty training in gastrointestinal pathology may improve the agreement on TB assessment.[39] In addition, a second opinion by an experienced gastrointestinal pathologist has been reported to significantly affect treatment decisions of pT1 CRC patients, potentially altering management in up to 30% of cases.[16,46] Expert review can also reduce

misclassification of pseudo-invasion or high-grade dysplasia.[47] This highlights the critical need for training programmes and standardized histopathological evaluation to avoid unnecessary over- or undertreatment of pT1 CRC patients.

One of the limitations of our study could be the small number of pT1 CRC assessed by the 20 pathologists' task force, being infrequent features scarcely represented or absent (e.g. perineural invasion, high histological grade and high TB). Due to the limited number of cases with LNM or recurrences, the correlation with histological features could not be thoroughly assessed.[13,48] However, we considered that these 80 cases were a representative sample to achieve our aims, which would reflect the daily practice. Moreover, pathologists were from different institutions, and no instructions for assessment were given for the pilot study, which adds value to the results obtained. Another advantage of the study is the use of digitalized slides, where all measurements and annotations were exact and comparable, although this is not yet applicable in many institutions. As this retrospective cohort included cases from 2007 onward, mismatch repair protein status was not routinely assessed and was unavailable for many cases.

Conclusion

This study highlights the variability in the assessment of some histopathological pT1 CRC features and emphasizes the need for standardized criteria to improve interobserver agreement. While several histological features show high agreement, others, such as the status of muscularis mucosae and stromal lymphocytes, may require further refinement in reporting methods and guidelines. The inclusion of PDC as a new well-defined high-risk feature in future guidelines could enhance the accuracy of risk stratification and improve clinical outcomes for pT1 CRC patients.

Author contributions

K.S.G., M.D., M.P. and M.C. performed study concept and design; M.R.-C., I.A., S.L.-P., J.J.A., L.A.-M., M.C.J., A.C., F.G., C.G.-L., M.J., I.J., I.M., C.M.-C., E.M., D.N., N.P., C.P., O.R., R.S.-Y., G.T.V.B. and M.C. performed development of methodology and acquisition of data; K.S.G. and M.C. provided data and statistical analysis, and writing of the paper; all authors reviewed and revised the paper, and approved the final version of the paper.

Acknowledgements

We would like to thank all investigators of the EPIT1 Consortium who contributed to the database of the study. We are indebted to the HCB-IDIBAPS Biobank, integrated in the Spanish National Biobanks Network, for the biological human samples and data procurement. We acknowledge the support of the Xarxa de Bancs de Tumours de Catalunya (XBTC), sponsored by Pla Director d'Oncologia de Catalunya (PDO). CIBEREHD is funded by the Instituto de Salud Carlos III. We also acknowledge the support of the CERCA Programme/Generalitat de Catalunya.

Funding information

This study was financed by Beca de la Marató de TV3 2020 (201932-30) and Instituto de Salud Carlos III (PI19/01050), co-funded by the European Regional Development Fund/European Social Fund; "A way to make Europe"/ "Investing in your future". The funding institutions did not have any involvement in the study design; data collection, analysis or interpretation; manuscript writing; and in the decision to submit the article for publication.

Conflict of interests

The authors declare that they have no conflicts of interest.

Data availability statement

The datasets generated during the current study are not publicly available but can be made available from the corresponding author upon reasonable request.

References

1. Sung H, Ferlay J, Siegel RL *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 2021; **71**; 209–249.
2. Cubiella J, González A, Almazán R *et al.* pT1 colorectal cancer detected in a colorectal cancer mass screening program: treatment and factors associated with residual and extraluminal disease. *Cancers (Basel)* 2020; **12**; 12.
3. Argilés G, Tabernero J, Labianca R *et al.* Localised colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 2020; **31**; 1291–1305.
4. Benson AB, Venook AP, Adam M *et al.* Colon cancer, version 2.2021, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* 2021; **19**; 329–359.

5. Hashiguchi Y, Muro K, Saito Y *et al.* Japanese Society for Cancer of the colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int. J. Clin. Oncol.* 2020; **25**; 1–42.
6. Hassan C, Wysocki PT, Fuccio L *et al.* Endoscopic surveillance after surgical or endoscopic resection for colorectal cancer: European Society of Gastrointestinal Endoscopy (ESGE) and European Society of Digestive Oncology (ESDO) guideline. *Endoscopy* 2019; **51**; 266–277.
7. Loughrey MB, Quirke P, Shepherd NA. Dataset for histopathological reporting of colorectal cancer. https://www.rcpath.org/static/c8b61ba0-ae3f-43f1-85ffd3ab9f17cfe6/c19a5cd7-3485-44c2-b5e1c87154830582/G049-Dataset-for-histopathological-reporting-of-colorectal-cancer.pdf.
8. Zwager LW, Bastiaansen BAJ, Montazeri NSM *et al.* Deep submucosal invasion is not an independent risk factor for lymph node metastasis in T1 colorectal cancer: a meta-analysis. *Gastroenterology* 2022; **163**; 174–189.
9. Ueno H, Kajiwara Y, Shimazaki H *et al.* New criteria for histologic grading of colorectal cancer. *Am. J. Surg. Pathol.* 2012; **36**; 193–201.
10. Brockmoeller SF, West NP. Predicting systemic spread in early colorectal cancer: can we do better? *World J. Gastroenterol.* 2019; **25**; 2887–2897.
11. Kajiwara Y, Oka S, Tanaka S *et al.* Nomogram as a novel predictive tool for lymph node metastasis in T1 colorectal cancer treated with endoscopic resection: a nationwide, multicenter study. *Gastrointest. Endosc.* 2023; **97**; 1119–1128.e5.
12. Tateishi Y, Nakanishi Y, Taniguchi H, Shimoda T, Umemura S. Pathological prognostic factors predicting lymph node metastasis in submucosal invasive (T1) colorectal carcinoma. *Mod. Pathol.* 2010; **23**; 1068–1072.
13. Backes Y, Elias SG, Groen JN *et al.* Histologic factors associated with need for surgery in patients with pedunculated T1 colorectal carcinomas. *Gastroenterology* 2018; **154**; 1647–1659.
14. Cappellesso R, Nicolè L, Zanco F *et al.* Synchronous nodal metastatic risk in screening detected and endoscopically removed pT1 colorectal cancers. *Pathol. Res. Pract.* 2020; **216**; 152966.
15. Toh E-W, Brown P, Morris E, Botterill I, Quirke P. Area of submucosal invasion and width of invasion predicts lymph node metastasis in pT1 colorectal cancers. *Dis. Colon Rectum* 2015; **58**; 393–400.
16. Rampioni Vinciguerra GL, Antonelli G, Citron F *et al.* Pathologist second opinion significantly alters clinical management of pT1 endoscopically resected colorectal cancer. *Virchows Arch.* 2019; **475**; 665–668.
17. Norton EJ, Bateman AC. Risk assessment in pT1 colorectal cancer. *J. Clin. Pathol.* 2023; **77**; 225–232.
18. WHO Classification of Tumours Editorial Board. *Digestive system tumours*. 5th ed. International Agency for Research on Cancer (IARC), Lyon, France: WHO Classification of Tumours, 2019.
19. 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.
20. Haddad TS, Lugli A, Aherne S *et al.* Improving tumor budding reporting in colorectal cancer: a Delphi consensus study. *Virchows Arch.* 2021; **479**; 459–469.
21. Jakubowska K, Koda M, Kisielewski W, Kańczuga-Koda L, Famulski W. Tumor-infiltrating lymphocytes in primary tumors of colorectal cancer and their metastases. *Exp. Ther. Med.* 2019; **18**; 4904–4912.
22. Fuchs TL, Sioson L, Sheen A *et al.* Assessment of tumor-infiltrating lymphocytes using international TILs working group (ITWG) system is a strong predictor of overall survival in colorectal carcinoma: a study of 1034 patients. *Am. J. Surg. Pathol.* 2020; **44**; 536–544.
23. Haggitt RC, Glotzbach RE, Soffer EE, Wruble LD. Prognostic factors in colorectal carcinomas arising in adenomas: implications for lesions removed by endoscopic polypectomy. *Gastroenterology* 1985; **89**; 328–336.
24. Kikuchi R, Takano M, Takagi K *et al.* Management of early invasive colorectal cancer. Risk of recurrence and clinical guidelines. *Dis. Colon Rectum* 1995; **38**; 1286–1295.
25. Ueno H, Mochizuki H, Hashiguchi Y *et al.* Risk factors for an adverse outcome in early invasive colorectal carcinoma. *Gastroenterology* 2004; **127**; 385–394.
26. Scott N, Cairns A, Prasad P *et al.* Resection margin involvement after endoscopic excision of malignant colorectal polyps: definition of margin involvement and its impact upon tumour recurrence. *Histopathology* 2023; **83**; 80–90.
27. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 2016; **15**; 155–163.
28. Chandler I, Houlston RS. Interobserver agreement in grading of colorectal cancers - findings from a nationwide web-based survey of histopathologists. *Histopathology* 2008; **52**; 494–499.
29. Pontarollo G, Bonjour M, Walter T *et al.* Histopronostic factors in superficial colorectal adenocarcinomas treated by endoscopy: reproducibility and impact of immunohistochemistry and digital pathology. *Virchows Arch.* 2024; **485**; 233–244.
30. Davenport A, Morris J, Pritchard SA, Salmo E, Scott M, Haboubi NY. Interobserver variability amongst gastrointestinal pathologists in assessing prognostic parameters of malignant colorectal polyps: a cause for concern. *Tech. Coloproctol.* 2016; **20**; 647–652.
31. Rosty C, Webster F, Nagtegaal ID *et al.* Pathology reporting of colorectal local excision specimens: recommendations from the international collaboration on cancer reporting (ICCR). *Gastroenterology* 2021; **161**; 382–387.
32. Karamchandani DM, Gonzalez RS, Lee H, Westerhoff M, Cox B, Pai RK. Interobserver agreement and practice patterns for grading of colorectal carcinoma: World Health Organization (WHO) classification of tumours 5th edition versus American joint committee on cancer (AJCC) 8th edition staging manual. *Histopathology* 2025; **86**; 1101–1111.
33. Barel F, Auffret A, Cariou M *et al.* High reproducibility is attainable in assessing histoprognostic parameters of pT1 colorectal cancer using routine histopathology slides and immunohistochemistry analyses. *Pathology* 2019; **51**; 46–54.
34. Harris EI, Lewin DN, Wang HL *et al.* Lymphovascular invasion in colorectal cancer. *Am. J. Surg. Pathol.* 2008; **32**; 1816–1821.
35. Watanabe J, Ichimasa K, Kataoka Y *et al.* Additional staining for lymphovascular invasion is associated with increased estimation of lymph node metastasis in patients with T1 colorectal cancer: systematic review and meta-analysis. *Dig. Endosc.* 2024; **36**; 533–545.
36. Zhang H, Simmer F, Lugli A, Nagtegaal ID. Tumor budding as a risk factor for lymph node metastasis and local recurrence in pT1 colorectal cancer: a systematic review and meta-analysis. *Gastro. Hep. Adv.* 2025; **4**; 100713.
37. Kai K, Aishima S, Aoki S *et al.* Cytokeratin immunohistochemistry improves interobserver variability between unskilled

pathologists in the evaluation of tumor budding in T1 colorectal cancer. *Pathol. Int.* 2016; **66**; 75–82.
38. 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.
39. Hacking S, Angert M, Jin C *et al*. Tumor budding in colorectal carcinoma: an institutional interobserver reliability and prognostic study of colorectal adenocarcinoma cases. *Ann. Diagn. Pathol.* 2019; **43**; 151420.
40. Ji X, Kang M, Zhao X *et al*. Poorly differentiated cluster grade-a vital predictor for lymph node metastasis and oncological outcomes in patients with T1 colorectal cancer: a retrospective study. *BMC Gastroenterol.* 2022; **22**; 409.
41. Barresi V, Bonetti LR, Branca G *et al*. Colorectal carcinoma grading by quantifying poorly differentiated cell clusters is more reproducible and provides more robust prognostic information than conventional grading. *Virchows Arch.* 2012; **461**; 621–628.
42. Barresi V, Reggiani Bonetti L, Ieni A, Domati F, Tuccari G. Prognostic significance of grading based on the counting of poorly differentiated clusters in colorectal mucinous adenocarcinoma. *Hum. Pathol.* 2015; **46**; 1722–1729.
43. Barresi V, Bonetti LR, Ieni A, Branca G, Baron L, Tuccari G. Histologic grading based on counting poorly differentiated clusters in preoperative biopsy predicts nodal involvement and pTNM stage in colorectal cancer patients. *Hum. Pathol.* 2014; **45**; 268–275.
44. Karamchandani DM, Westerhoff M, Arnold CA *et al*. Interobserver variability in assessment of depth of submucosal invasion for colonic endoscopic resections among subspecialized gastrointestinal pathologists. *Arch. Pathol. Lab Med.* 2023; **147**; 534–545.
45. Kazemi A, Gharib M, Mohamadian Roshan N *et al*. Assessment of the tumor–stroma ratio and tumor-infiltrating lymphocytes in colorectal cancer: inter-observer agreement evaluation. *Diagnostics* 2023; **13**; 2339.
46. Smits LJH, Vink-Börger E, van Lijnschoten G *et al*. Diagnostic variability in the histopathological assessment of advanced colorectal adenomas and early colorectal cancer in a screening population. *Histopathology* 2022; **80**; 790–798.
47. Backes Y, Moons LM, Novelli MR *et al*. Diagnosis of T1 colorectal cancer in pedunculated polyps in daily clinical practice: a multicenter study. *Mod. Pathol.* 2017; **30**; 104–112.
48. Haasnoot KJC, Backes Y, Moons LMG *et al*. Associations of non-pedunculated T1 colorectal adenocarcinoma outcome with consensus molecular subtypes, immunoscore, and microsatellite status: a multicenter case-cohort study. *Mod. Pathol.* 2020; **33**; 2626–2636.