ORIGINAL ARTICLE

UNITED EUROPEAN GASTROENTEROLOGY
ueg journal WILEY

Lymph node metastases and recurrence in pT1 colorectal cancer: Prediction with the International Budding Consortium Score—A retrospective, multi-centric study

Heather Dawson[1] | John-Melle Bokhorst[2] | Linda Studer[1,3] | Michael Vieth[4] | Ayse Selcen Oguz Erdogan[2] | Sonay Kus Öztürk[2] | Richard Kirsch[5] | Scarlett Brockmoeller[6] | Gieri Cathomas[7] | Rolf Buslei[8] | David Fink[9] | Marie Roumet[10] | Inti Zlobec[1] | Jeroen van der Laak[2] | Iris D. Nagtegaal[2] | Alessandro Lugli[1]

[1]Institute of Tissue Medicine and Pathology, University of Bern, Bern, Switzerland

[2]Department of Pathology, Radboudumc, Nijmegen, the Netherlands

[3]Institute of Artificial Intelligence and Complex Systems, University of Applied Sciences and Arts Western Switzerland, Fribourg, Switzerland

[4]Institute of Pathology, Friedrich-Alexander-University Erlangen-Nuremberg, Klinikum Bayreuth, Bayreuth, Germany

[5]Pathology and Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

[6]Pathology and Data Analytics, Leeds Institute of Medical Research at St. James's School of Medicine, Leeds, UK

[7]Institute of Pathology, Kantonsspital Baselland, Liestal, Switzerland

[8]Institut und Praxis für Pathologie, Neuropathologie, Molekulare Diagnostik und Zytologie, Sozialstiftung Bamberg, Bamberg, Germany

[9]Department of Pathology and Immunology, Baylor College of Medicine, Houston, Texas, USA

[10]Clinical Trials Unit, University of Bern, Bern, Switzerland

Correspondence

Heather Dawson, Institute of Tissue Medicine and Pathology, University of Bern, Murtenstrasse 31, Bern 3008, Switzerland.
Email: heather.dawson@unibe.ch

Present address

Gieri Cathomas, Institute of Tissue Medicine and Pathology, University of Bern, Bern, Switzerland.

Abstract

Background: The International Collaboration on Cancer Reporting proposes histological tumour type, lymphovascular invasion, tumour grade, perineural invasion, extent, and dimensions of invasion as risk factors for lymph node metastases and tumour progression in completely endoscopically resected pT1 colorectal cancer (CRC).

Objective: The aim of the study was to propose a predictive and reliable score to optimise the clinical management of endoscopically resected pT1 CRC patients.

Methods: This multi-centric, retrospective International Budding Consortium (IBC) study included an international pT1 CRC cohort of 565 patients. All cases were reviewed by eight expert gastrointestinal pathologists. All risk factors were reported according to international guidelines. Tumour budding and immune response (CD8+ T-cells) were assessed with automated models using artificial intelligence. We used the information on risk factors and least absolute shrinkage and selection operator

John-Melle Bokhorst and Linda Studer contributed equally to this paper.

logistic regression to develop a prediction model and generate a score to predict the occurrence of lymph node metastasis or cancer recurrence.

Results: The IBC prediction score included the following parameters: lymphovascular invasion, tumour buds, infiltration depth and tumour grade. The score has an acceptable discrimination power (area under the curve of 0.68 [95% confidence intervals (CI) 0.61–0.75]; 0.64 [95% CI 0.57–0.71] after internal validation). At a cut-off of 6.8 points to discriminate high-and low-risk patients, the score had a sensitivity and specificity of 0.9 [95% CI 0.8–0.95] and 0.26 [95% 0.22, 0.3], respectively.

Conclusion: The IBC score is based on well-established risk factors and is a promising tool with clinical utility to support the management of pT1 CRC patients.

KEYWORDS

colorectal cancer, CRC, early, grade, histology, lymph node metastases, lymphovascular invasion, prediction, pT1, recurrence

INTRODUCTION

In the aera of personalised healthcare, interdisciplinary clinical management supports the optimisation of risk assessment in colorectal cancer (CRC) patients. Indeed, three clinical CRC scenarios highlight this statement: in pT1 CRC, the prediction of local lymph node metastases indicates a potential oncologic resection,[1,2] while in stage II CRC adjuvant therapy[1,2] and in rectal cancer neoadjuvant therapy should be considered.

Risk assessment is usually based on the Tumor Nodes Metastasis classification proposed by the union internationale contre le cancer[3] and the additional evidence-based biomarkers suggested in the World Health Organization classification.[3,4] According to national and international CRC guidelines, the International Collaboration on Cancer Reporting (ICCR) proposes additional 'core' prognostic factors for pT1 CRC, such as histological tumour type, lymphovascular invasion, tumour grade, perineural invasion, extent of invasion, invasive carcinoma dimensions and margin status.[5] In endoscopically resected pT1 CRC, the histological features of lymphovascular invasion, tumour grade and tumour infiltration depth are traditionally used to determine further therapy (observation vs. oncologic resection with en bloc removal of regional lymph nodes).[1,2] The incidence of lymph node metastases among all pT1 CRC is estimated to be about 7%–15% in the literature,[6,7] and lower for pedunculated polyps.[8] However, if histological risk factors are considered, this probability varies from <1% (no risk factors present) to about 35% (several risk factors present).[9] Typically, once a single unfavourable histological risk factor is present, oncologic resection is recommended.[1,2,10] However, it has been suggested that this approach is too stringent, leading to over-treatment and potentially unnecessary clinical complications.[11,12] Therefore, a more refined risk assessment could be of great utility to aid patient management.

In the recent years, several studies have promoted the increasing prognostic evidence of two additional morphologic biomarkers, namely tumour budding and tumour immunology. Tumour budding is

Key summary

Established knowledge on this subject

- Risk factors in endoscopically resected pT1 colorectal cancer (CRC) are often reported descriptively and are used to guide patient management.
- Most treatment guidelines do not take the actual probability of an adverse event (e.g. lymph node metastases) into account and may be too stringent, leading to over-treatment.

What are the significant findings of this study?

- The International Budding Consortium (IBC) score delivers the probability of adverse events in pT1 colorectal cancers as an additional tool for individualised therapeutic management.
- The IBC score includes well-established risk factors in pT1 CRC while minimising inter-observer variability and is a promising tool with clinical utility in daily practice.

a tumour-related factor and is defined as single tumour cells or clusters of up to four cells at the invasive CRC front.[13] Tumour budding is a strong marker of tumour progression and prognosis[14,15] including lymph node metastasis in pT1 CRC and can also be used among the above-mentioned histological risk factors for risk stratification.[16-18] The prognostic role of immune response, especially by T-cell infiltrates, is supported by several studies[19,20] as well as its potential predictive value.[19,21]

A promising approach for the integration of tumour budding and tumour inflammation into a predictive score for local lymph node metastases and cancer recurrence in pT1 CRC is based on the 'attacker-defender model' (referring to tumour buds as the 'attacker'

and immune cells as the 'defender'). This concept was first published in 2009[22] and reflects the pathogenetic dynamic of the tumour microenvironment (TME). Since then, several studies have shown the importance of this model, including the tumour budding grade and the cytotoxic T-lymphocyte count, in prognostic models.[23,24]

The implementation of tumour budding in national and international guidelines was boosted by the standardised scoring system proposed by the International Tumour Budding Consensus Conference (ITBCC) in 2016.[13] The International Budding Consortium (IBC) is a spinoff of the ITBCC and aims to validate the prognostic and predictive role of tumour budding in CRC and to decrease potential inter-observer variability using available artificial intelligence (AI) algorithms.

The aim of the present retrospective, multicentric IBC study is not to replace the already well-established additional risk factors in pT1 CRC but to investigate the potential role of the two pillars of the attacker-defender model, the tumour budding and immune response, in a predictive and reproducible scoring system for the clinical management of pT1 CRC patients.

MATERIALS AND METHODS

Study design

Cases of diagnosed pT1 CRC from 1991 until 2018 were retrieved from the pathology institutes of the University of Bern, Switzerland ($n = 196$), Radboud University Nijmegen, Netherlands ($n = 105$), Klinikum Bayreuth, Germany ($n = 96$), Cantonal Hospital Liestal, Switzerland ($n = 92$), Baylor Scott and White Health, Temple, Texas, USA ($n = 44$) and Medizinisches Versorgungszentrum am Bruderwald, Bamberg, Germany ($n = 34$). All participating centres complied with local ethics committee regulations (as applicable): Bern and Liestal: 2017-01803 (Kantonale Ethikkommission Bern); Nijmegen: 10,602 (Dutch Cancer Society); Bayreuth: 23-109-Br (Friedrich-Alexander-Universität).

Participants

Inclusion criteria were histopathological diagnosis of submucosally invasive (pT1) CRC with either (A) initial resection or completion oncologic resection (no salvage surgery following surveillance of endoscopic removal) and therefore with known lymph node status or (B) endoscopic specimens with at least 36 months follow-up with documented information on local/distant recurrence. Exclusion criteria included a history of inflammatory bowel disease, prior CRC or synchronous CRC and polyposis syndromes (Lynch syndrome only if first CRC). Patients with fragmented/piecemeal resection were included only if completion oncologic resection was performed. Participating centres confirmed compliance to corresponding local ethics regulations and provided a formalin-fixed, paraffin embedded (FFPE) block and corresponding haematoxylin and eosin (H&E) stained

slides. In case of several blocks/case, the FFPE block with the most advanced tumour was selected. If no block with the most advanced tumour could be determined, all tumour blocks were provided.

A total of 773 patients were submitted to the study. During the study, a total of 209 cases were excluded due to reasons listed in Figure 1. In the final cohort consisting of 565 patients, 462 underwent either initial or subsequent resection (nodal status known) and 103 were treated by endoscopic resection alone. Seventeen patients with piecemeal resection included in the study were amenable to complete histologic review of all parameters and underwent oncologic resection.

Test methods

Histology review

A group of 8 gastrointestinal pathologists (AL, IDN, HD, RK, MV, SB, SKÖ, ASOE) were involved in a comprehensive histopathological review of all cases. Details concerning reviewed parameters (including histological risk factors of tumour grade, lymphovascular invasion and infiltration depth among others) and immunohistochemical assays including mismatch repair (MMR) proteins for a subset of patients can be found in Supporting Information S1. Of note, 86 cases (11%) were excluded due to reclassification as adenomas with epithelial misplacement (no invasive carcinoma). Only cases with full data for key predictor factors ($n = 565$) considered for the prediction model were included for further analysis.

Assessment of tumour budding

Tumour budding was assessed using an automated detection algorithm (Figure 2)[25] in three different areas: in a manually selected hotspot, in an automatically detected hotspot and as a mean value among 10 automatically detected hotspots. All hotspots measured 0.785 mm^2 according to the ITBCC guidelines.[13] All manual tumour budding areas were selected using the Automated Slide Analysis Platform tool (https://computationalpathologygroup.github.io/ASAP/) with a square annotation (preprint available: https://www.researchsquare.com/article/rs-2093909/v1). The algorithm is available at https://grand-challenge.org/algorithms/colon-budding-in-ihc/.

Assessment of CD8+ lymphocytes

Colour deconvolution was used to separate the AE1/AE3 and CD8 stains. CD8-positive T-cells were automatically detected in the same areas as tumour buds (manually selected hotspot, automatically detected hotspot and mean value among 10 automatically detected hotspots). The algorithm used for lymphocyte detection has been described previously in detail.[26] Manual quality control of the algorithm output revealed a precision of 89.9% and recall of 93%.[27]
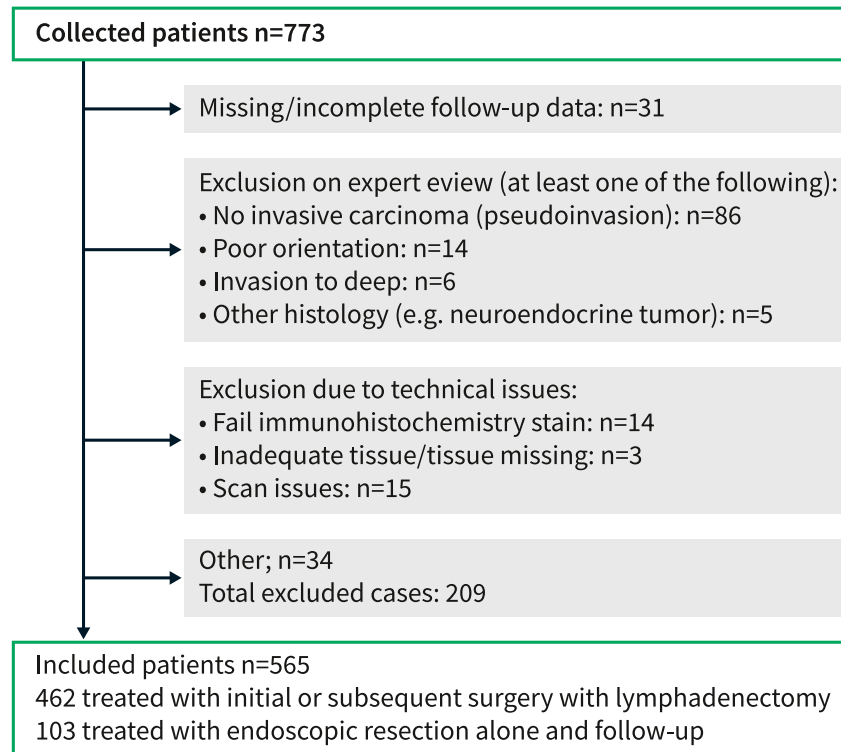
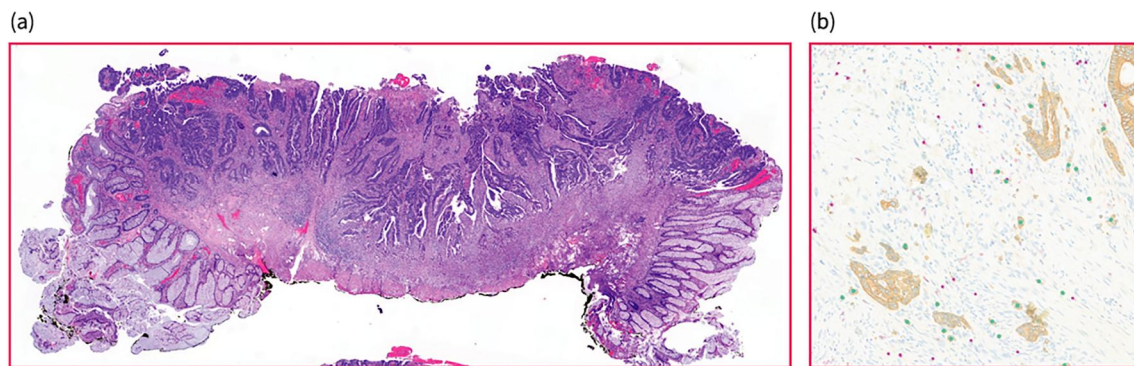**FIGURE 1**   Flowchart of patients through the study.

FIGURE 2   Scoring of cases. Haematoxylin stain of whole tissue slide (a, endoscopic resection specimen) used for the assessment of pathological features by an expert panel. Tumour budding and T-cells were analysed using an automatic algorithm on AE1/AE3 and CD8 double-stained slides (output in b; tumour buds marked green, T-cells marked pink).

Statistical analysis

Outcome and predictor variables

The primary outcome was the occurrence of an adverse event defined as either (1) lymph node metastasis in cases with known nodal status or (2) local cancer recurrence in cases undergoing endoscopic resection alone. To predict this outcome, we used information according to the following well-known risk factors based on guidelines[1,2]: occurrence of lymphovascular invasion, tumour grade, invasion depth and depth classification and tumour buds. CD8+ T-cells were considered as an additional potential risk factor. Tumour grade was assessed as a numerical, a categorical (G1–G3), and a binary (i.e., G1–G2 vs. G3) variable; the invasion depth was measured from the submucosa and from the surface; the detected number of buds was considered in a manually selected hotspot, an automatically detected hotspot and mean value among the top 10 automatically detected hotspots. Due to their distribution, variables related to the number of buds were log10 transformed prior to the analysis, and null values were replaced by 0.1 prior to log10 transformation.

Prediction modelling

As detailed above, some risk factors were assessed using different methods/measuring techniques. In the current study, we decided to only consider the measurement with a stronger association with the outcome. To do so, we used a bootstrapping approach. For this, we built 500 bootstrap samples and in each sample, and for each risk factor, we regressed the outcome on the different measurements and selected the assessment with the best Akaike information criterion. Only the most commonly retained measurement was kept for the analysis (Table 1).

We used least absolute shrinkage and selection operator (LASSO) logistic regression analysis to predict adverse events as a function of the selected risk factors and the interaction between the invasion depth measurement and category. Alpha was set to one throughout. The optimum lambda penalty parameter was chosen using 10-fold cross-validation in a range between 0.01 and 0.37. To assess the performance of the obtained prediction model, we characterised the overall accuracy and the discriminatory power of the model by calculating the Brier score and the area under the receiver operating characteristic (AUROC) curve, respectively. Then, we evaluated the model calibration by using calibration curves.

The internal validity of the model was assessed using 1000 bootstrap samples, where samples were drawn with replacement from the derivation sample. The bootstrap-corrected performance estimates were calculated by subtracting the optimism from the performance of the original model (in-sample estimates are generally too high—too optimistic—so bootstrap samples are used to estimate how overly optimistic the model is and correct the statistics). The 95% confidence intervals for the boot-strapped performance measures were derived using the percentile method. Measurement selection (as described in the paragraph above) was included in the bootstrapping step.

TABLE 1 Model coefficients.

Label	Level	Coef	OR
(Intercept)		−3.116	
Lymphovascular invasion	No	Ref.	Ref.
	Yes	0.239	1.27
Invasion depth, SM or Haggitt classification	Haggitt	Ref.	Ref.
	SM	0	1
log10 Nb of buds (manual hotspot)		0.555	1.742
log10 invasion depth from the submucosa (mm)		0.321	1.379
Tumour grade		0.19	1.209
Invasion depth, SM or Haggitt classification x log10 invasion depth from the submucosa (mm)	Haggitt	−0.011	0.989
	SM	−0.011	0.989

Note: Note that standard errors from penalised estimation methods such as LASSO are biased, generally leading to too small confidence intervals. As such, confidence intervals are not reported.

In a second step, we used the prediction model to create a user-friendly point scoring system using the regression coefficient-based scoring method.[28] The IBC score points were obtained by multiplying each beta-coefficient by 10 and rounding to the closest integer (what Mehta et al. 2016[29] refer to as Beta10/integer). Adding up the converted coefficients yielded the score of increasing risk of an adverse event.

Finally, we built a model predicting the risk of an adverse event as a function of the IBC score value. From this model, we estimated the predicted probabilities and associated 95% predicted interval and assessed the discriminatory power of the IBC score by calculating AUROC (ROC) curve (i.e., area under the curve [AUC]). The validity of the AUC value was assessed by performing internal validation as described above.

In the last step, we illustrated the distribution of the IBC score among all patients and split the patients into two (low and high-risk) categories. The reference point (cut-off) was chosen to maximise specificity while maintaining a minimum sensitivity of 0.9. To determine the accuracy of this method to predict an adverse event, we estimated sensitivity, specificity, positive and negative predictive values and likelihood ratios.

The R 4.2.1 software was used for statistical analysis.

The methods are based on the Standards for Reporting Diagnostic Accuracy Studies and Transparent Reporting of a multivariate Prediction model for Individual Prognosis Or Diagnosis[30] guidelines.

RESULTS

Baseline characteristics

Baseline characteristics of patients included in the final dataset used for the prediction model with clinicopathological information and key variables for the predictive model are listed in Table 2. Of 462 patients who were treated with surgery, positive lymph nodes were found in 58 (13%) and 404 patients were nodal negative (87%). Of all patients treated with surgery, 371 (80%) had at least one of the following risk factors: tumour infiltration depth >1 mm/Haggitt level >2, presence of LVI and/or high tumour grade.

Of the 103 patients treated by endoscopic resection alone, 3 patients had local cancer recurrence (3%).80 patients (76%) had at least one of the risk factors mentioned above. In 57 of the 80 patients with histologic risk factors, tumour infiltration depth was the sole risk factor.

Prediction model

Of the initially considered potential predictor variables (see 'outcome and predictor variables'), four remain in the model after LASSO logistic regression analysis (Table 3): lymphovascular invasion, number of buds (manual hotspot), invasion depth and tumour grade (numerical). In our original model, which is shown in Table S2, CD8+ T-

TABLE 2 Baseline patient characteristics stratified by adverse events (lymph node metastasis or cancer recurrence).

Characteristic	Overall, N = 565	No adverse event, N = 504	Adverse event, N = 61
Demographics			
Sex–female, n (%)	232 (41%)	210 (42%)	22 (37%)
Missing	4	4	0
Age at diagnosis			
Median (interquartile range; IQR)	69 (62, 76)	69 (63, 76)	67 (58, 75)
Missing	4	4	0
Surgery			
Surgery, n (%)	462 (81%)	404 (80%)	58 (94%)
Lymph nodes; LN (assessed on patients with surgery only)			
Total number of LN			
No. obs.	462	404	58
Median (IQR)	16 (12, 23)	16 (11, 23)	16 (12, 21)
Missing	37	34	3
Tumour characteristics			
Location, n (%)			
Left	257 (45%)	228 (45%)	28 (45%)
Rectum	133 (23%)	114 (23%)	18 (31%)
Right	128 (23%)	117 (23%)	11 (18%)
Transverse	27 (5%)	25 (5%)	2 (3%)
Missing	22	20	2
Tumour type, n (%)			
Adenocarcinoma	533 (94%)	479 (95%)	54 (89%)
Mucinous adenocarcinoma	28 (4.9%)	23 (4.6%)	5 (8.1%)
Other	4 (0.8%)	2 (0.4%)	2 (3.2%)
Tumour grade, n (%)			
1	147 (26%)	137 (27%)	10 (16%)
2	371 (66%)	328 (65%)	43 (71%)
3	47 (8.3%)	39 (7.7%)	8 (13%)
Tumour precursor, n (%)			
Conventional colorectal adenoma	446 (79%)	405 (80%)	41 (67%)
Not observed	67 (12%)	57 (11%)	10 (16%)
Sessile serrated lesion	19 (3.4%)	16 (3.2%)	3 (4.8%)
Traditional serrated adenoma	33 (5.8%)	26 (5.1%)	7 (11%)
Other pathological variables			
Lymphovascular invasion, n (%)	96 (17%)	79 (16%)	17 (28%)
Perineural invasion, n (%)	10 (1.8%)	5 (1.0%)	5 (8.2%)
Invasion width (mm)			
Median (IQR)	7.1 (4.5, 9.8)	7.0 (4.5, 9.7)	7.9 (5.4, 10.2)
Missing	1	1	0
Invasion depth from the surface (mm)			
Median (IQR)	4.33 (3.08, 6.40)	4.25 (3.05, 6.27)	5.12 (3.20, 7.48)

TABLE 2 (Continued)

Characteristic	Overall, N = 565	No adverse event, N = 504	Adverse event, N = 61
Invasion depth from the submucosa (mm)			
Median (IQR)	2.83 (1.81, 4.36)	2.79 (1.80, 4.18)	3.85 (2.07, 5.58)
Invasion depth - Haggitt classification, n (%)	154 (27%)	136 (27%)	18 (29%)
Haggitt 1	36 (6%)	36 (7%)	0 (0%)
Haggitt 2	83 (15%)	71 (14%)	12 (20%)
Haggitt ≥3	35 (6%)	29 (6%)	6 (9%)
Invasion depth—SM classification (Kikuchi level of submucosal invaion)			
SM1	39 (7%)	38 (8%)	1 (1.6%)
SM2	112 (20%)	104 (20%)	8 (13.1%)
SM3	78 (14%)	71 (14%)	7 (11.5%)
Measurement not possible	336 (59%)		
No. of buds (manual hotspot)			
Median (IQR)	8 (2, 17)	7 (2, 16)	13 (7, 25)
No. of buds (automatic hotspot, top-1)			
Median (IQR)	16 (10, 27)	15 (9, 26)	21 (15, 34)
No. of buds (automatic hotspot, top-10)			
Median (IQR)	11 (7, 19)	11 (7, 18)	14 (9, 25)
No. of CD8 T-cells (manual hotspot)			
Median (IQR)	159 (81, 286)	158 (80, 280)	181 (88, 330)
Mismatch repair status			
Proficient	179	163	16
Deficient	16	16	0
Missing	373		

TABLE 3 Score derivation weighted according to coefficients.

Label	Level	Score
Lymphovascular invasion	No	
	Yes	2
No. of buds (manual hotspot), per log10-unit		6
Invasion depth from the submucosa (mm), per log10-unit		3
Tumour grade (1, 2, 3)		2

cell counts emerged as a risk factor. This result is surprising and may perhaps not reflect true biology due to the methodology (see discussion, paragraph 6). Considering this and the negligible effect of CD8+ T-cell counts on our model performance, we decided not to include this parameter in the final prediction model.

Each included parameter was assigned a weight by multiplying each coefficient by 10 and rounding to the nearest integer, resulting in a scoring model (Table 3, see Figure 3 for detailed instructions on score calculation). The prediction model showed good accuracy and acceptable discrimination power, with a validated Brier score of 0.1 [0.08; 0.11] and a validated AUROC of 0.64 [0.57; 0.71] (Table 4).

The distribution of the score among all patients is shown in Figure S4. The score, ranging from −6 to 24 points, increases with the probability of an adverse event (1.18%–36.2%; Table S3) and can be used to define high versus low-risk patients. As we wanted to maximise specificity and maintain a minimum of sensitivity of 0.9, we placed the cut-off value at 7 and obtained 141 low risk and 424 high-risk patients. With this cut-off we obtained score sensitivity of 0.9 [0.8, 0.95] and specificity of 0.27 [0.23, 0.31], positive and negative predicted values were estimated at 0.13 [0.1, 0.17] and 0.96 [0.91, 0.98], respectively (Table S5). Table S6 shows a comparison of low versus high-risk patients in terms of characteristics and clinical outcomes.

CD8+ T-cells, tumour budding and mismatch repair status

CD8+ T-cell infiltrates were examined in various tumour budding regions of the main tumour, namely in the manual and automated tumour budding hotspots and as an average among the 10 tumours budding hotspots of the main tumour. A weak positive correlation

Calculation of the IBC score
1 Assess all parameters in the model (Table 3)
2 Assign the score value for categorical variables (e.g. 2 points for lymphovascular invasion if present) Multiply tumor grade by 2 (e.g. 4 points for G2)
3 For continious variables (tumour buds and invasion depth) transform the obtained values to log 10 and multiply with the score value (e.g. patient with 15 tumor buds will have a score of log 10 (15)* 6) = 8.23
4 All values are summed up to obtain the final score
5 The final score can be reported as a category (low - vs. high-risk; cutoff score 6.8) along with the predicted probability of the actual score value (e.g. score 12: high risk, predicted probability 11.58% [95% CI 9.07-14.67]

FIGURE 3 Calculation of the International Budding Consortium score.

TABLE 4 Performance statistics of the prediction model.

	Raw	Corrected (internal validation)[a]
AUC	0.68 [95% CI 0.61–0.75]	0.64 [95% CI 0.57–0.71]
Brier score	0.09	0.1 [95% CI 0.08–0.12]

[a]Corrected by internal validation, 1000 bootstraps.

between increasing tumour budding and CD8 counts in the manual tumour budding hotspot was demonstrated (log transformed variables; $R = 0.099$; $p = 0.018$). Mismatch repair status was available for 195 cases, of which 16 (8.2%) demonstrated a loss of expression of at least one immunohistochemical marker (MLH1, MSH2, PMS2 and/or MSH6). Significantly higher CD8 counts were seen in MMR-deficient cases ($p = 0.002$).

DISCUSSION

This retrospective, international multicentric IBC study included 567 pT1 CRC patients and aimed to investigate established risk factors for lymph node metastases and cancer recurrence in pT1 CRC, including tumour budding and tumour inflammation. The resulting IBC score is a prediction model consisting of four histological parameters: lymphovascular infiltration, tumour grade, infiltration depth and tumour budding visualised by immunohistochemistry and assessed using AI. The score gives an estimate of the probability of an adverse event in pT1 CRC and also stratifies patients into a low- and high-risk groups.

In the last few years, several systematic reviews and meta-analyses have investigated the risk factors for pT1 CRC with partly different conclusions. Cracco et al. proposed in a review including eight studies ($n = 1727$ patients) the width and the area of submucosal invasion as reliable prognostic factors for cancer progression in pT1 CRC, but without agreement on ideal cut-offs.[31] The systematic review and meta-analysis of Dykstra et al. included 60 studies ($n = 10,000$–13,128 patients) and confirmed established factors such as lymphovascular invasion, neural invasion, poor differentiation, and

infiltration depth (1500 μm); additionally, poorly differentiated clusters and tumour budding were proposed.[17] In the same year, Ebbehoj et al. published a further meta-analysis including 16 observational studies ($n = 10,181$ patients) confirming lymphovascular invasion, tumour budding, depth of submucosal invasion (at least 1000 mm), high tumour grade, polypoid growth pattern and rectal tumour location as potential risk factors for lymph node metastases in pT1 CRC.[18] A further meta-analysis[12] entered eight studies ($n = 3621$ patients) out of 67 included studies ($n = 21,238$ patients) into a multi-variable analysis showing a significant association between Lymp node metastases (LNM), poor differentiation, high-grade tumour budding and lymphovascular invasion. This contrasted with the deep submucosal invasion, which did not emerge as a significant independent predictor of LNM. Infiltration depth as a core element of the ICCR guidelines was considered a predictor of our model and contributed to predicting the outcome in our dataset. However, infiltration depth alone will virtually never classify a patient as 'high-risk' (see instructions on calculating the IBC score; Figure 3) Therefore, our results certainly support the recent conclusion that deep submucosal invasion should be reconsidered as a strong indicator of oncologic surgery.[12]

Despite the inclusion of risk factors in national and international guidelines, the question arises why a standardised predictive score system for the clinical management of pT1 CRC patients is still lacking. Therefore, considering the core risk factors proposed by the ICCR such as histological tumour type, lymphovascular invasion, tumour grade, perineural invasion, extent of invasion, invasive carcinoma dimensions and margin status, the following lines of thought need to be discussed. First, 'histopathological' factors such as tumour budding may be subject to high inter-observer variability[32,33]; second, some factors may be assessed by measurements on the slides leading to suboptimal results; third, the description and listing of these factors itself may qualitatively help but does not optimally support the interdisciplinary team during clinic-pathological conferences or tumour boards, respectively. These points lead to the potential impact and limitations of the present study.

The proposed IBC score includes several strong points: (a) cost-effectiveness: Three out of four factors (lymphovascular invasion, invasion depth and tumour grade) can be assessed on H&E slides; (b) a potential inter-observer variability of the recently standardised factor 'tumour budding' can be avoided by the AI application which can be used for free; (c) the score mainly avoids arbitrary cut-offs normally based on single studies with small sample sizes and uses linear parameters; (d) the IBC score takes multiple histological risk factors into account to deliver a probability estimate of adverse events which can be taken as an additional tool for clinical management (e) the cohort is international and multicentric, reviewed by the IBC and the sample size seems to be quite competitive in comparison with other original studies.[17] In contrast to a recently published risk stratification model,[34] our cohort underwent rigorous review by a panel of gastrointestinal pathologists, which excluded 11% of initially identified cases due to reclassification as adenomas with epithelial misplacement ('pseudoinvasion'), which is commonly mistaken for invasive carcinoma.[35]

Nevertheless, some limitations of the IBC score must be clarified: A two-tier risk category (low-/high) was established to facilitate clinical management with the cut-off at 90% sensitivity. This means that the low-risk category will still miss patients who have an adverse event. For this reason, it is important to report the estimated probability alongside the risk category for better identification of the extent of the risk (for instance, the model predicts around 7% probability of an adverse event for a patient with a score of 8%, and 39% respectively for a patient with a score of 24, yet both patients are in the high-risk category).

Other potential drawbacks are that the study is retrospective, hybrid by using partially conventional histology and AI and unexpectedly the impact of CD8+ cytotoxic T-cells was not a protective factor as expected in CRC. Indeed, the role of T-cell infiltrates in pT1 CRC is poorly characterised.[36] Potential reasons may be the often observed focality of the tumour invasion in pT1 CRC and the specific TME in the tunica submucosa. Additionally, the approach of the 'attacker-defender model' used in the present study takes only lymphocytes in the tumour budding hotspot into account and not the lymphocyte number observed in the whole tissue section independently of the tumour budding hotspot. AI-based methods will continue to provide a promising approach for a more in-depth analysis of the role of CD8+ T-cells in submucosally invasive CRC. Also, in this cohort the measured depth of invasion (in mm from the submucosa) had no different effect in pedunculated or sessile polyps and was therefore used as the sole parameter for measuring the extent of tumour infiltration. External validation of the IBC score, which is required prior to widespread clinical use, may consider a separate analysis of pedunculated and sessile lesions. Also, MMR status was only available for a subset of cases and could therefore not be included in the model. As the role of MMR status in predicting adverse events may be better characterised in pT1 CRC in the future, further developments and refinements of the IBC score may also include this and other molecular parameters.

In conclusion, the IBC score is a solid, reproducible tool with potential clinical utility in daily practice and from the prospective point of view reflects an excellent basis for a future fully digitised IBC score.

CONFLICT OF INTEREST STATEMENT

None of the authors have potential conflicts of interest relevant to the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Heather Dawson https://orcid.org/0000-0001-7527-336X

REFERENCES

1. Benson AB, Venook AP, Al-Hawary MM, Azad N, Chen YJ, Ciombor KK, et al. NCCN clinical practice guidelines in oncology (NCCN Guidelines®) colon cancer V2 Oct 27,2022; 2022. [cited Dec 08 2022].
2. Argiles G, Tabernero J, Labianca R, Hochhauser D, Salazar R, Iveson T, et al. Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2020;31(10):1291–305. https://doi.org/10.1016/j.annonc.2020.06.022
3. International Union against Cancer. Committee on TNM Classification. TNM classfication of malignant tumors. 8 ed. 2017.
4. Nagtegaal IOR. Colorectal adenocarcinomas. In: WHO classification of tumours editorial board. Digestive system tumours [internet]. Lyon: International Agency for Research on Cancer; 2019 [2022 Dec 20]. (WHO classification of tumours series, 5th ed.; vol. 1). https://tumourclassification.iarc.who.int/chapters/31
5. Loughrey MBAM, Brown I, et al. Colorectal cancer histopathology reporting guide. Sydney: International Collaboration on Cancer Reporting; 2020. ISBN: 978-1-922324-01-6.
6. Nivatvongs S, Rojanasakul A, Reiman HM, Dozois RR, Wolff BG, Pemberton JH, et al. The risk of lymph node metastasis in colorectal polyps with invasive adenocarcinoma. Dis Colon Rectum. 1991;34(4): 323–8. https://doi.org/10.1007/BF02050592
7. Kikuchi R, Takano M, Takagi K, Fujimoto N, Nozaki R, Fujiyoshi T, et al. Management of early invasive colorectal cancer. Risk of recurrence and clinical guidelines. Dis Colon Rectum. 1995;38(12):1286–95. https://doi.org/10.1007/BF02049154
8. Backes Y, Elias SG, Groen JN, Schwartz MP, Wolfhagen FH, Geesing JM, et al. Histologic factors associated with need for surgery in patients with pedunculated T1 colorectal carcinomas. Gastroenterology. 2018;154(6):1647–59. https://doi.org/10.1053/j.gastro.2018.01.023
9. Ueno H, Mochizuki H, Hashiguchi Y, Shimazaki H, Aida S, Hase K, et al. Risk factors for an adverse outcome in early invasive colorectal carcinoma. Gastroenterology. 2004;127(2):385–94. https://doi.org/10.1053/j.gastro.2004.04.022
10. Pimentel-Nunes P, Libanio D, Bastiaansen BAJ, Bhandari, P, Bisschops, R, Bourke, MJ, et al. Endoscopic submucosal dissection for superficial gastrointestinal lesions: European society of

gastrointestinal endoscopy (ESGE) guideline - update 2022. Endoscopy 2022; 54(06): 591–622. 20220506. https://doi.org/10.1055/a-1811-7025

11. Vermeer NCA, Backes Y, Snijders HS, Bastiaannet E, Liefers GJ, Moons LMG, et al. National cohort study on postoperative risks after surgery for submucosal invasive colorectal cancer. BJS Open. 2019;3(2):210–7. https://doi.org/10.1002/bjs5.50125

12. Zwager LW, Bastiaansen BAJ, Montazeri NSM, Hompes R, Barresi V, Ichimasa K, 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(1):174–89. https://doi.org/10.1053/j.gastro.2022.04.010

13. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol. 2017;30(9):1299–311. https://doi.org/10.1038/modpathol.2017.46

14. Lugli A, Zlobec I, Berger MD, Kirsch R, Nagtegaal ID. Tumour budding in solid cancers. Nat Rev Clin Oncol. 2021;18(2):101–15. https://doi.org/10.1038/s41571-020-0422-y

15. Rogers AC, Winter DC, Heeney A, Gibbons D, Lugli A, Puppa G, et al. Systematic review and meta-analysis of the impact of tumour budding in colorectal cancer. Br J Cancer. 2016;115(7):831–40. https://doi.org/10.1038/bjc.2016.274

16. Cappellesso R, Luchini C, Veronese N, Lo Mele M, Rosa-Rizzotto E, Guido E, et al. Tumor budding as a risk factor for nodal metastasis in pT1 colorectal cancers: a meta-analysis. Hum Pathol. 2017;65:62–70. https://doi.org/10.1016/j.humpath.2017.04.013

17. Dykstra MA, Gimon TI, Ronksley PE, Buie WD, MacLean AR. Classic and novel histopathologic risk factors for lymph node metastasis in T1 colorectal cancer: a systematic review and meta-analysis. Dis Colon Rectum. 2021;64(9):1139–50. https://doi.org/10.1097/DCR.0000000000002164

18. Ebbehoj AL, Jorgensen LN, Krarup PM, Smith HG. Histopathological risk factors for lymph node metastases in T1 colorectal cancer: meta-analysis. Br J Surg. 2021;108(7):769–76. https://doi.org/10.1093/bjs/znab168

19. Pages F, Andre T, Taieb J, Vernerey D, Henriques J, Borg C, et al. Prognostic and predictive value of the Immunoscore in stage III colon cancer patients treated with oxaliplatin in the prospective IDEA France PRODIGE-GERCOR cohort study. Ann Oncol. 2020;31(7):921–9. https://doi.org/10.1016/j.annonc.2020.03.310

20. Idos GE, Kwok J, Bonthala N, Kysh L, Gruber SB, Qu C. The prognostic implications of tumor infiltrating lymphocytes in colorectal cancer: a systematic review and meta-analysis. Sci Rep. 2020;10(1):3360. https://doi.org/10.1038/s41598-020-60255-4

21. Oshi M, Sarkar J, Wu R, et al. Intratumoral density of regulatory T cells is a predictor of host immune response and chemotherapy response in colorectal cancer. Am J Cancer Res. 2022;12:490–503.

22. Lugli A, Karamitopoulou E, Panayiotides I, Karakitsos P, Rallis G, Peros G, 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(8):1382–92. https://doi.org/10.1038/sj.bjc.6605318

23. Dawson H, Christe L, Eichmann M, Reinhard S, Zlobec I, Blank A, et al. Tumour budding/T cell infiltrates in colorectal cancer: proposal of a novel combined score. Histopathology. 2020;76(4):572–80. https://doi.org/10.1111/his.14006

24. Nearchou IP, Lillard K, Gavriel CG, Ueno H, Harrison DJ, Caie PD. Automated analysis of lymphocytic infiltration, tumor budding, and their spatial relationship improves prognostic accuracy in colorectal cancer. Cancer Immunol Res. 2019;7(4):609–20. https://doi.org/10.1158/2326-6066.CIR-18-0377

25. Bokhorst JM, Nagtegaal ID, Zlobec I, Dawson H, Sheahan K, Simmer F, et al. Semi-supervised learning to automate tumor bud detection in cytokeratin-stained whole-slide images of colorectal cancer.

Cancers (Basel). 2023;15(7):2079. https://doi.org/10.3390/cancers15072079

26. Swiderska-Chadaj Z, Pinckaers H, van Rijthoven M, Balkenhol M, Melnikova M, Geessink O, et al. Learning to detect lymphocytes in immunohistochemistry with deep learning. Med Image Anal. 2019;58:101547. ARTN 101547. https://doi.org/10.1016/j.media.2019.101547

27. Studer LBJ, Zlobec I, et al. Validation of computer-assisted tumour-bud and T-cell detection in pT1 colorectal cancer. In: ECP glasgow. Springer; 2020. p. S37. (online).

28. Janssen KJ, Siccama I, Vergouwe Y, Koffijberg H, Debray T, Keijzer M, et al. Development and validation of clinical prediction models: marginal differences between logistic regression, penalized maximum likelihood estimation, and genetic programming. J Clin Epidemiol. 2012;65(4):404–12. https://doi.org/10.1016/j.jclinepi.2011.08.011

29. Mehta HB, Mehta V, Girman CJ, Adhikari D, Johnson ML. Regression coefficient-based scoring system should be used to assign weights to the risk index. J Clin Epidemiol. 2016;79:22–8. https://doi.org/10.1016/j.jclinepi.2016.03.031

30. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Br J Surg. 2015;102(3):148–58. https://doi.org/10.1002/bjs.9736

31. Cracco N, Todaro V, Pedrazzi G, Del Rio P, Haboubi N, Zinicola R. The risk of lymph node metastasis in T1 colorectal cancer: new parameters to assess the degree of submucosal invasion. Int J Colorectal Dis. 2021;36(1):41–5. https://doi.org/10.1007/s00384-020-03738-0

32. Bokhorst JM, Blank A, Lugli A, Zlobec I, Dawson H, Vieth M, et al. Assessment of individual tumor buds using keratin immunohistochemistry: moderate interobserver agreement suggests a role for machine learning. Mod Pathol. 2020;33(5):825–33. https://doi.org/10.1038/s41379-019-0434-2

33. Aizawa D, Sugino T, Oishi T, Hotta K, Imai K, Shiomi A, et al. The essential problem of over-measuring the depth of submucosal invasion in pT1 colorectal cancer. Virchows Arch. 2022;480(2):323–33. https://doi.org/10.1007/s00428-021-03221-3

34. Kajiwara Y, Oka S, Tanaka S, Nakamura T, Saito S, Fukunaga Y, 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(6):1119–28.e5. https://doi.org/10.1016/j.gie.2023.01.022

35. Griggs RK, Novelli MR, Sanders DS, Warren BF, Williams GT, Quirke P, et al. Challenging diagnostic issues in adenomatous polyps with epithelial misplacement in bowel cancer screening: 5 years' experience of the Bowel Cancer Screening Programme Expert Board. Histopathology. 2017;70(3):466–72. https://doi.org/10.1111/his.13092

36. Cappellesso R, Nicole L, Zanco F, Lo Mele M, Fassina A, Ventura L, et al. Synchronous nodal metastatic risk in screening detected and endoscopically removed pT1 colorectal cancers. Pathol Res Pract. 2020;216(6):152966. https://doi.org/10.1016/j.prp.2020.152966

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.