



A machine learning–based transcriptomic signature for predicting tumor recurrence after curative resection in T1 colorectal cancer: a retrospective multicenter cohort study (The Tw1CE trial)

Takayuki Noma, MD^{a,b}, Karmele Saez de Gordo, MD^c, María Daca-Alvarez, MD, PhD^d, Katsuki Miyazaki, MD, PhD^{a,b}, Yuma Wada, MD, PhD^{a,b}, Alessandro Mannucci, MD^{a,e}, Takumi Onoyama, MD, PhD^{a,f}, Mitsuo Shimada, MD, PhD^b, Miriam Cuatrecasas, MD, PhD^c, Luis Bujanda, MD, PhD^g, Maria Pellise, MD, PhD^{d,*}, Ajay Goel, PhD^{a,h,*} on behalf of part of the EpiT1 Consortium

Background: T1 colorectal cancer (T1 CRC) is increasingly treated with curative-intent endoscopic resection, but tumor recurrence remains a critical factor influencing patient prognosis. However there is no validated biomarker exists to reliably predict post-resection recurrence, limiting risk-adapted follow-up and adjuvant therapy decisions.

Materials and methods: In this multicenter retrospective cohort study across academic centers in Spain, 138 patients with T1 CRC (2023–2025; ClinicalTrials.gov NCT06314971) were enrolled. From FFPE endoscopic specimens, expression of five mRNAs and two miRNAs was quantified by RT-qPCR, and an XGBoost-based transcriptomic panel was developed. Patients were assigned to training and independent testing cohorts by treatment type. The primary outcome was 3-year recurrence-free survival (RFS); secondary outcomes included 5-year RFS and overall survival (OS).

Results: The transcriptomic panel demonstrated high predictive performance in both the training (AUROC = 91.7%) and testing (AUROC = 88.2%) cohorts. Patients classified as high-risk by the panel exhibited significantly worse RFS and OS compared with those classified as low-risk (log-rank $P < 0.001$). Furthermore, integrating lymphatic invasion with the transcriptomic panel into a combined risk stratification model further improved predictive accuracy (AUROC = 94.6%), and decision curve analysis confirmed its superior clinical utility compared to conventional criteria.

Conclusion: This study established a validated machine learning–based transcriptomic classifier derived from endoscopic resection specimens that accurately predicts tumor recurrence in patients with T1 CRC. Our findings highlight the potential of this biomarker panel to enable risk-adapted surveillance strategies and guide decisions regarding additional therapy after curative resection.

Keywords: machine learning, T1 CRC, transcriptomic panel, tumor recurrence, XGBoost

Introduction

Colorectal cancer (CRC) is the third most common malignant tumor in the world^[1,2,3], but with widespread screening, it is

increasingly being diagnosed at an earlier and more treatable stage, especially invasive submucosal CRC (T1 CRC) with submucosal invasion^[4,5]. This increased diagnosis at this stage has

^aDepartment of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Biomedical Research Center, Monrovia, CA, USA, ^bDepartment of Surgery, Tokushima University, Tokushima, Japan,

^cPathology Department, Centre for Biomedical Diagnosis, Hospital Clinic Barcelona, Barcelona, Spain, ^dDepartment of Gastroenterology, Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), University of Barcelona, Barcelona, Spain,

^eGastroenterology and Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, IRCCS San Raffaele Hospital, Milan, Italy, ^fDepartment of Gastroenterology and Nephrology, Tottori University, Tottori, Japan, ^gGastroenterology Department, Instituto Biodonostia, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Universidad del País Vasco (UPV/EHU), San Sebastián, Spain and ^hCity of Hope Comprehensive Cancer Center, Duarte, CA, USA

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding authors. Address: Department of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Biomedical Research Center, Monrovia, CA, 91016, USA. Tel.: +1 626-218-3452.

E-mail: ajgoel@coh.org (A. Goel); Department of Gastroenterology, Hospital Clinic de Barcelona. Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD). Universitat de Barcelona, Barcelona, 08036, Spain. E-mail: mpellise@clinic.cat (M. Pellisé).

Copyright © 2026 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

International Journal of Surgery (2026) 00:1–13

Received 9 August 2025; Accepted 15 December 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.ij-surgery.com.

Published online 28 January 2026

<http://dx.doi.org/10.1097/JS9.0000000000004690>

led to the widespread adoption of minimally invasive, organ-sparing endoscopic resection^[6]. Nonetheless, endoscopic resection does not permit complete pathological staging of the regional lymph nodes. In current practice, pathological risk features observed in endoscopically resected T1 CRC specimens – such as vascular invasion, poor differentiation, deep submucosal invasion, and tumor budding – often trigger recommendations for additional surgical resection, even in the absence of direct evidence of lymph node metastasis (LNM)^[7–10]. This approach prioritizes oncologic thoroughness but results in overtreatment for many patients who, despite having high-risk histopathologic features, ultimately have no nodal involvement upon surgical resection. We previously reported that a panel of five miRNAs and eight mRNAs could reliably detect LNM in endoscopic tissue specimens and even blood from patients with T1 colorectal cancer, reducing overtreatment rates^[11–14].

However, the current surveillance strategies outlined in treatment guidelines for T1 CRC primarily focus on assessing LNM risk, which may be insufficient for predicting the overall prognosis of T1 CRC patients. Emerging evidence suggests that distant metastases may often arise independently of nodal involvement, as studies have shown that >65% of distant metastases do not originate from the same subclones as LNM in CRC^[15]. Moreover, in early-stage CRC, recurrence rates after endoscopic resection without lymphadenectomy are comparable to those after surgical resection with lymph node dissection^[16]. Thus, in T1 CRC, tumor recurrence may arise not only via lymphatic pathways but also through hematogenous dissemination^[17], underscoring the need for risk-stratification strategies that extend beyond predicting LNM alone.

Numerous studies have demonstrated that CRC prognosis is influenced by a wide range of biological and clinical determinants, including transcriptomic alterations, inflammatory signaling pathways, postoperative complications, and host-related factors such as sarcopenia or sarcobesity, which have been associated with adverse outcomes in colorectal surgery cohorts^[18–21]. Although many such risk factors have been reported, no clinically established tool currently provides accurate prediction of prognosis or recurrence in early-stage CRC, including T1 lesions. Previous biomarker studies in CRC such as Oncotype DX Colon and ColoPrint have primarily focused on stage II–III disease and relied on surgically resected specimens spanning a heterogeneous spectrum from early pT2 tumors to high-risk stage II^[22–25]. Importantly, these assays were not designed to specifically evaluate early-stage cases such as T1 CRC, and their applicability to this population remains uncertain. Moreover, because these platforms depend on surgically resected tissue, they are inherently unsuitable for the growing proportion of T1 CRC patients who undergo endoscopic treatment alone and therefore lack access to surgical specimens.

In other words, considering that tumor recurrence cannot be accurately assessed by current risk factors, the current surveillance strategies prescribed in treatment guidelines for T1 CRC may be inadequate in predicting the prognosis of these patients. Although histopathological features such as tumor budding, lymphovascular invasion, and poor differentiation have been shown to be associated with recurrence in T1 CRC, their predictive accuracy and consistency remain limited^[26–28]. These factors alone are insufficient to identify patients at risk of distant metastasis, which has the greatest impact on long-term

HIGHLIGHTS

- Tumor recurrence is a significant prognostic factor in T1 colorectal cancer (T1CRC), and current histopathological criteria are insufficient for accurate risk prediction.
- We developed a machine learning–based transcriptomic biomarker panel which accurately predicted tumor recurrence risk after curative resection.
- The panel demonstrated high predictive accuracy in both training (AUROC = 91.7%) and independent testing (AUROC = 88.2%) cohorts.
- Combining the transcriptomic panel with lymphatic invasion further improved recurrence prediction (AUROC = 94.6%) and outperformed conventional criteria in decision curve analysis.
- This clinically applicable tool may guide personalized surveillance and inform decisions about adjuvant treatment after tumor resection.

prognosis. Therefore, there is a critical need to incorporate molecular biomarkers that more accurately reflect the biological drivers of recurrence and allow precise risk stratification in T1 CRC.

Building on prior evidence supporting the biological relevance of specific mRNAs and miRNAs in CRC progression^[29–37], and leveraging recent advances in transcriptomic and machine learning–based prognostic modeling^[23,38–44], we sought to evaluate the clinical utility of transcriptomic biomarkers for predicting recurrence in patients with T1 CRC. Our goal was to develop a robust, biologically informed risk stratification model capable of improving prognostication and guiding postoperative management in this increasingly common early-stage malignancy. This work has been reported in line with the STROCSS criteria (Supplemental Digital Content, available at: <http://links.lww.com/J9/G597>)^[45].

Materials and methods

Study design

The Tw1CE study was a multi-institutional, retrospective, multi-phase biomarker study. It was designed to complete the EDRN phases II and III study (Early Detection Research Network) for a tissue-based prognostic study. In our previous publications, we identified a panel of transcriptomic biomarkers in the EDRN phase I study (biomarker discovery phase). Building upon those findings, the current study aimed to evaluate these biomarkers for predicting tumor recurrence in patients with T1 CRC. Two independent tissue-based cohorts (training and testing cohorts, respectively) were utilized, and we evaluated the performance of the selected biomarkers in both cohorts independently. This phase II/III validation assessed whether these markers could stratify patients based on recurrence risk, thereby supporting their potential use in guiding postoperative management. In our previous transcriptomic analysis, we reported candidates of eight mRNAs and five miRNAs that were significantly associated with high-risk T1 CRC^[11–14]. From these candidates, we selected five mRNAs and two miRNAs that showed the strongest association with postoperative recurrence

and validated their predictive performance in independent cohorts.

Patient cohorts

A total of 3161 patients diagnosed with pT1 CRC between 2007 and 2018 were registered, regardless of treatment received. Exclusion criteria were histological types other than adenocarcinoma, hereditary cancer syndromes, inflammatory bowel disease, synchronous or metachronous CRC within 5 years, or metastatic disease at diagnosis. Among the remaining patients, 138 had sufficient biological material from endoscopically resected specimens and were considered for this study (Supplemental Digital Content sTable 1, available at: <http://links.lww.com/JS9/G597>). From these, patients with lymph node metastasis (LNM) at the time of initial surgery, local tumor recurrence, or no recurrence but with a follow-up period shorter than 3 years were excluded from the Tw1CE analysis. The final study population was divided into a training cohort (patients who underwent surgery, *n* = 57) and a testing cohort (patients managed with endoscopic resection alone, *n* = 58; see Table 1, Supplemental Digital Content sFigure 1, available at: <http://links.lww.com/JS9/G597>). These inclusion and exclusion criteria were uniformly applied across all participating centers to ensure consistency and reproducibility.

Pathological risk classification was evaluated by pathologists at each participating institution, based on features such as deep submucosal (SM) invasion ≥1000 μm or SM2–3 (deep SM invasion), presence of lymphatic or vascular invasion, high-grade tumor budding, and poorly differentiated histology, as well as affected margins^[8]. This study received approval from the ethical review boards of each participating institution. Biological specimen collection and subsequent laboratory analyses were conducted in strict accordance with the principles outlined in the Declaration of Helsinki. All patients provided written informed consent post-treatment. The study was registered and completed on clinicaltrials.gov (NCT06314971). Prior to endoscopic resection, all patients underwent standard pre-procedural preparation, including a thorough review of medical history and current medications. Patients were instructed to remain nil by mouth for at least 6 hours prior to the intervention. All resections were performed in the endoscopy suite of each participating institution under standard clinical protocols.

Definitions and study endpoints

Recurrence-free survival (RFS) was defined as the interval from curative-intent treatment to the first documented recurrence. Overall survival (OS) was defined as the time from curative-intent treatment to death from any cause. Since the majority of CRC recurrences occur within 3 years after surgery, and early recurrence is consistently associated with poor prognosis^[46–48], we identified RFS <3 years as the primary endpoint. As recurrence may still occur beyond 3 years in T1 colorectal cancer^[49,50], 5-year RFS and 5-year OS were also evaluated as secondary endpoints.

Laboratory methodology

RNA extraction from formalin-fixed paraffin-embedded (FFPE) specimens involved manual microdissection to enrich the

Table 1
Clinical cohort characteristics

Characteristics	Surgery cohort (<i>n</i> = 57)	Endoscopy cohort (<i>n</i> = 58)	<i>P</i> -Value
Median (range)	63 (50–86)	60.5 (46–88)	0.30
Biological sex			
Male, <i>N</i> (%)	39 (68.4)	40 (69.0)	1.00
Female, <i>N</i> (%)	18 (31.6)	18 (31.0)	
Differentiation			
Well/Moderate, <i>N</i> (%)	51 (89.5)	56 (96.6)	0.11
Poor, <i>N</i> (%)	5 (8.8)	1 (1.7)	
Unknown, <i>N</i> (%)	1 (1.7)	1 (1.7)	
Tumor location			
Right side, <i>N</i> (%)	8 (14.0)	6 (10.3)	0.78
Left side, <i>N</i> (%)	49 (86.0)	51 (87.9)	
Unknown, <i>N</i> (%)	0 (0)	1 (1.8)	
Tumor size, mm			
Median (range)	15 (2–50)	10 (1.5–80)	0.01
Submucosal invasion, μm			
≥1000, <i>N</i> (%)	46 (80.7)	51 (87.9)	0.15
<1000, <i>N</i> (%)	10 (17.5)	4 (6.9)	
Unknown, <i>N</i> (%)	1 (1.8)	3 (5.17)	
Budding grade			
≥2, <i>N</i> (%)	13 (22.8)	6 (10.3)	0.13
1, <i>N</i> (%)	43 (75.4)	48 (82.8)	
Unknown, <i>N</i> (%)	1 (1.8)	4 (6.9)	
Lymphatic invasion			
Present, <i>N</i> (%)	7 (12.3)	3 (5.2)	0.32
Absent, <i>N</i> (%)	49 (86.0)	51 (87.9)	
Unknown, <i>N</i> (%)	1 (1.7)	4 (6.9)	
Vascular invasion			
Present, <i>N</i> (%)	8 (14.0)	1 (1.7)	0.03
Absent, <i>N</i> (%)	48 (84.2)	53 (91.4)	
Unknown, <i>N</i> (%)	1 (1.8)	4 (6.9)	
Recurrence			
RFS < 3 years, <i>N</i> (%)	15 (26.3)	11 (19.0)	0.65
RFS ≥ 3 years, <i>N</i> (%)	42 (73.7)	47 (81.0)	

representation of neoplastic cells, with more than 75% of the micro-dissected tissue comprising tumor cells. Subsequently, RNA isolation was performed using the AllPrep DNA/RNA FFPE Kit manufactured by Qiagen. RNA purity and concentration were assessed spectrophotometrically, and only samples with an A260/A280 ratio ≥1.8 were included. Following the extraction of RNA from the FFPE specimens, complementary DNA (cDNA) was generated before proceeding to the polymerase chain reaction (PCR) assays. For miRNA expression analyses, cDNA synthesis from total RNA was conducted using the TaqMan microRNA Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA). For mRNA profiling, the conversion of RNA into complementary DNA (cDNA) was achieved using the High-Capacity cDNA Reverse Transcription Kit, provided by Thermo Fisher Scientific. The RT-qPCR analysis was performed using the SensiFAST™ probe Lo-ROX Kit (Bioline, London, UK) on the QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems, Foster City, CA). No-template controls were included on every plate, and inter-plate variability was minimized using an inter-plate calibrator. Samples in which the internal normalizer (U6 for miRNA or β-actin for mRNA) had a Ct value ≥35 were excluded or repeated. PCR data were normalized using the ΔCt method

with U6 (for miRNA) or β -actin (for mRNA) as internal controls, and log10 transformed prior to statistical analysis^[51].

All miRNA primers employed in this study were procured from Thermo Fisher Scientific, sharing a catalog number (4427975). Individual miRNAs used in this study were identified by their respective assay IDs, including U6 (001973), hsa-miR-32-5p (002109), hsa-miR-181b-5p (001098), hsa-miR-193b-3p (002367), hsa-miR-195-5p (000494), and hsa-miR-411-5p (001610) (Supplemental Digital Content sTable 2, available at: <http://links.lww.com/JS9/G597>). The primer sequences and Assay ID for the targets investigated in this study are presented in Supplemental Digital Content sTable 3, available at: <http://links.lww.com/JS9/G597>.

Statistical analysis

From the 13 previously reported candidate biomarkers, stepwise regression was applied in the training cohort to identify the most informative and stable predictors of recurrence, yielding five mRNAs (RCC1, FOXA1, AMT, MMP9, and MMP1) and two miRNAs (miR-411-5p and miR-193b-3p). These selected markers were then used for model development and validation. We employed the eXtreme Gradient Boosting (XGBoost) algorithm, an ensemble learning method that sequentially combines multiple weak decision tree learners to construct a robust predictive model. XGBoost is particularly well suited for high-dimensional molecular datasets, providing efficient computation, extensive hyperparameter optimization, and robust handling of missing data. To minimize the risk of overfitting in the context of a relatively small dataset, several safeguards were implemented. XGBoost has built-in cross-validation and early stopping functions, which were used to monitor model performance during training and confirmed stable performance within the training cohort. In addition, hyperparameters were tuned with an explicit focus on regularization. Model complexity was constrained by limiting the maximum tree depth to 5 and penalizing overly complex splits ($\gamma = 3.5$). A conservative learning rate ($\eta = 0.001$) was applied to ensure gradual learning and to prevent abrupt overfitting, and the total number of boosting rounds was capped at 1500. These adjustments were specifically designed to enhance regularization, reduce variance, and avoid overfitting. After training cohort, the model was locked to prevent further modification and subsequently applied unchanged to the independent validation cohort, thereby ensuring robust external validation and assessment of generalizability. Cohorts were stratified into “high-risk” and “low-risk” groups based on the cut-off value determined in the training cohort (targeting sensitivity $\geq 85\%$). Predictive performance was evaluated using receiver operating characteristic (ROC) curves and the corresponding area under the ROC curve (AUROC), with 95% confidence intervals estimated by the DeLong method (pROC R package, v1.18.5). Sensitivity and specificity were calculated with 95% confidence intervals using the Wilson method. Group comparisons of continuous variables between cohorts were assessed using the unpaired Mann–Whitney U test. Categorical variables were compared using Fisher’s exact test or χ^2 test, as appropriate. RFS and OS were analyzed with Kaplan–Meier survival curves, and differences between groups were assessed using the log-rank test at 3 and 5 years. The independence of the prediction panel from conventional clinicopathological variables was examined using both univariate and multivariate logistic

regression analyses for recurrence and Cox proportional hazards regression analyses for the time-to-event outcome of RFS and OS. Considering the small and imbalanced strata, subgroup analyses were not pre-specified and were therefore not performed. All statistical tests were two-sided, and statistical significance was defined as $P < 0.05$. All statistical analyses and machine-learning procedures were performed using R (version 4.3.1). The following packages were used: xgboost (1.7.7.1), SHAPforxgboost (0.1.3), pROC (1.18.5), survival (3.6-4), survminer (0.4.9), swimmerplot (1.2.0), and ggplot2 (3.5.1).

Results

XGBoost-based machine training established a biomarker panel that can predict tumor recurrence in T1 CRC patients

To identify transcriptomic biomarkers associated with tumor recurrence in T1 CRC, we performed RT-qPCR assays to examine the associations between the 13 biomarkers we previously identified and their ability to predict tumor recurrence in the training cohort. The predictive performance of each candidate is summarized in Supplemental Digital Content sTable 4, available at: <http://links.lww.com/JS9/G597>. Stepwise regression was used to prioritize a narrower list of 5 mRNA (RCC1, FOXA1, AMT, MMP9, and MMP1) and 2 miRNA biomarkers (miR-411-5p and miR-193b-3p), which were the most significant predictors of tumor recurrence (Fig. 1A). Next, we used an XGBoost-based machine learning approach in the training cohort to establish individual mRNA and miRNA panels (mRNA panel: AUROC 90.2 (81.2–99.1)%; miRNA panel: AUROC 76.7 (63.6–89.8)%; Fig. 1B and C). These panels showed significantly higher scores in recurrence cases (mRNA: $P < 0.001$, miRNA: $P = 0.002$; Fig. 1D and E). Global feature importance analysis consistently highlighted RCC1 and MMP9 as the strongest mRNA contributors and miR-411-5p as the main miRNA, across all XGBoost gain/cover metrics, decision tree architecture, and SHAP summaries (Supplemental Digital Content sFigures 2A and B, 3A and B, 4A and B, 5A and B, available at: <http://links.lww.com/JS9/G597>). These results demonstrate that the combination of stepwise regression and XGBoost effectively identified key transcriptomic biomarkers linked to tumor recurrence.

A combined mRNA/miRNA transcriptomic panel robustly predicted tumor recurrence in T1 CRC patients

We then combined the mRNA and miRNA panels developed in the training cohort using logistic regression analysis to establish an mRNA/miRNA panel, which exhibited even superior predictive accuracy for tumor recurrence [AUROC = 91.7 (83.7–99.8)%, sensitivity = 86.7 (62.1–96.3)%, specificity = 85.7 (72.2–93.3)%; Fig. 2A]. The cutoff values were selected to ensure sensitivity of at least 85%. Furthermore, in the testing cohort, using the same XGBoost model and parameters as the training cohort, this mRNA/miRNA panel was able to predict tumor recurrence with high accuracy, demonstrating the robustness of this panel [AUROC = 88.2 (78.7–97.7)%, sensitivity = 81.8 (52.3–94.9)%, specificity = 76.6 (62.8–86.4)%; Fig. 2B]. To further validate the superiority of our XGBoost-based approach, we compared it with a logistic regression model constructed from the same seven candidate biomarkers and

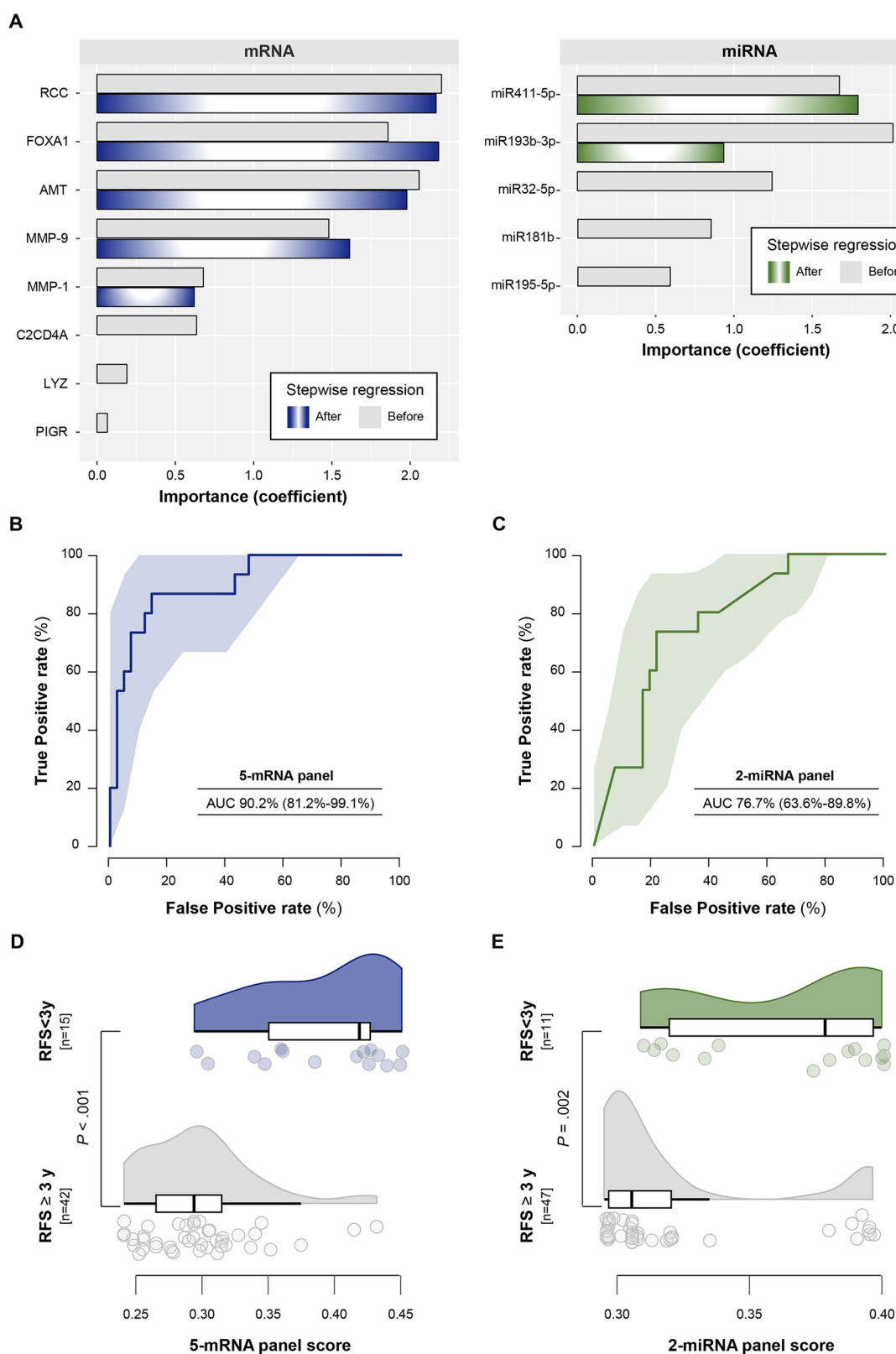


Figure 1. The training phase of transcriptomic panel predicting tumor recurrence in a training cohort. (A) Bar graph showing the weight of each marker before and after stepwise regression analysis. (B) ROC of the 5-mRNA panel (AUROC = 90.2%). (C) ROC of the 2-miRNA panel: (AUROC = 76.7%). (D) Rain cloud plots for the 5-mRNA panel ($P < 0.001$). (E) Rain cloud plots by the 2-miRNA panel ($P = 0.002$). AUROC, area under the receiver operating characteristic curve; mRNA, messenger RNA; miRNA, microRNA; RFS, recurrence-free survival.

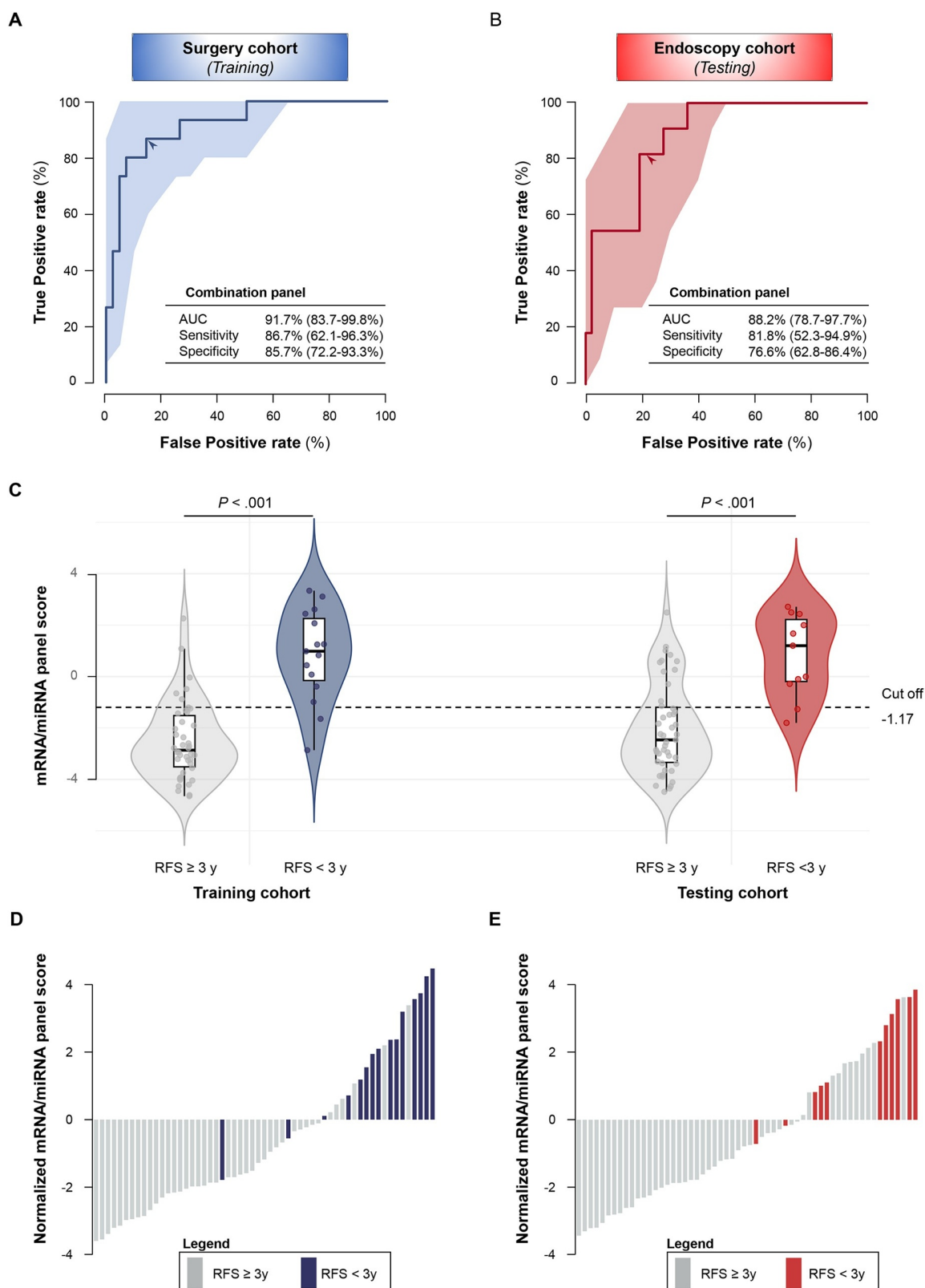


Figure 2. Performance of the transcriptomic panel. (A) AUROC and performance index of the combination panel in the training cohort. Arrowhead indicates the positivity threshold for high-risk transcriptomic risk scores. (B) AUROC and performance index of the combination panel in the testing cohort. Arrowhead indicates the positivity threshold for high-risk transcriptomic risk scores, which was fixed from the training cohort. (C) Boxplot and violin plot in the training and testing cohorts. Violin plot showing the distribution of panel score values for patients with RFS < 3 years and RFS ≥ 3 years (cutoff value -1.17). (D and E) Waterfall plot representative of the transcriptomic panel values observed in the training and testing cohorts. AUROC, area under the receiver operating characteristic curve; mRNA, messenger RNA; miRNA, microRNA; RFS, recurrence-free survival.

found that XGBoost consistently outperformed logistic regression in both the training and testing cohorts (Supplemental Digital Content sFigure 6, available at: <http://links.lww.com/JS9/G597>).

In addition, violin plots and boxplots reveal that patients within the high-risk group who had less than 3 years of RFS scored significantly higher compared to those in the low-risk group in this independent cohort, even when using the same cutoff value (−1.17) that was determined in the training cohort (Fig. 2C). This panel also clearly separated high- and low-risk patients in the waterfall plots (Fig. 2D and E). In conclusion, our mRNA/miRNA panel developed using XGBoost demonstrated robust predictive accuracy for tumor recurrence across both cohorts and further confirmed that XGBoost outperforms linear logistic regression. These findings underscore the potential of this model as a reliable tool for clinical risk stratification.

The mRNA/miRNA panel also predicted prognostic outcomes beyond 3 years in patients with T1 CRC

In the training cohort, survival curves evaluated by log-rank test revealed that patients with higher scores for this mRNA/miRNA panel had a significantly higher risk of tumor recurrence at all observed time points [log-rank test, $P < 0.001$; 3-year RFS: 31.6 (16.3–61.2)% vs 94.7 (87.9–100)%; 5-year RFS: 31.6 (16.3–61.2)% vs 92.0 (83.8–100)%; Fig. 3A]. Similarly, patients with higher scores consistently exhibited a significantly higher risk of death [log-rank test, $P = 0.02$; 3-year OS: 86.2 (70.0–100)% vs 97.3 (92.3–100)%; 5-year OS: 67.0 (44.6–100)% vs 97.3 (92.3–100)%; Supplemental Digital Content sFigure 7A, available at: <http://links.lww.com/JS9/G597>]. Moreover, in the testing cohort as well, high-risk patients deciphered by this mRNA/miRNA panel were associated with a significantly higher risk of tumor recurrence and mortality [log-rank test, $P < 0.001$; 3-year RFS: 55.0 (37.0–81.8)% vs 94.7 (87.9–100)%; 5-year RFS: 28.1 (13.6–58.0)% vs 83.7 (72.5–96.5)%; and log-rank test, $P = 0.003$; 3-year OS: 94.1 (83.6–100)% vs 100 (100–100)%; 5-year OS: 53.6 (31.8–90.4)% vs 96.9 (91.0–100)%; Figure 3B, Supplemental Digital Content sFigure 7B, available at: <http://links.lww.com/JS9/G597>]. The median follow-up time was 5.1 years (range, 0.6–6.5 years) in the surgery cohort and 5.1 years (range, 0.2–7.3 years) in the endoscopy cohort, thereby minimizing the potential for time-related bias in survival analysis. In conclusion, we have highlighted the value of the mRNA/miRNA panel as a highly reliable predictive tool for stratifying patient prognoses, including long-term prognostic outcomes.

Establishment of a risk stratification model that can more accurately predict tumor recurrence in patients with T1 CRC

To further confirm the clinical significance of our mRNA/miRNA panels, we performed a univariate analysis for all key clinicopathologic factors and their association with tumor recurrence after surgical resection with lymph node dissection (RFS: <3 years). The results revealed that specific clinical factors (tumor diameter ≥ 20 mm, lymphatic invasion, and our mRNA/miRNA panel) were significantly associated with tumor recurrence. Furthermore, multivariate analysis confirmed that our mRNA/miRNA panel and lymphatic invasion were independent predictors of tumor recurrence in patients with T1 CRC. Interestingly, our mRNA/miRNA panel could predict tumor recurrence with the highest OR (Table 2). We also performed

Cox proportional hazards regression analyses to further evaluate the independence of our panel in predicting RFS (Supplemental Digital Content sTable 5, available at: <http://links.lww.com/JS9/G597>). Because lymphatic invasion was an essential variable for subsequent analyses, one patient in the training cohort with missing data for this factor was excluded, leaving 56 patients for analysis. We stratified patients into three prognostic groups: (1) those with lymphatic invasion and a high-risk score, (2) those without lymphatic invasion but with a high-risk score, and (3) those with a low-risk score regardless of lymphatic invasion status (Fig. 4A). Most patients with high-risk scores experienced recurrence, and those with both high-risk scores and lymphatic invasion had particularly poor outcomes. Based on this, we hypothesized that combining lymphatic invasion with our mRNA/miRNA panel would improve prognostic accuracy. Indeed, the combined risk stratification model significantly enhanced predictive ability [AUROC = 94.6 (88.7–100) %; sensitivity = 92.9 (78.6–100) %, specificity = 88.1 (78.6–97.6) %; Fig. 4B].

Decision curve analysis (Fig. 4C) further demonstrated the clinical utility of the model. The diagonal dotted line denotes the “treat-all” strategy (systemic therapy for all patients), and the horizontal grey line denotes the “treat-none” strategy (tumor resection alone). The red line is the combined risk-stratification model, the blue line is the mRNA/miRNA panel, and the green line is lymphatic invasion alone. To align with the ROC analysis, we overlaid a vertical dashed line indicating the cutoff value used in the study. At this threshold, the combined model showed the highest net benefit compared with both treat-all and treat-none strategies, as well as compared with either the transcriptomic panel or lymphatic invasion alone. Finally, as shown in the swimmer plot (Fig. 4D), most high-risk patients identified by this model experienced recurrence and poor survival, whereas many low-risk patients remained recurrence-free for ≥ 5 years. Taken together, these findings indicate that our risk stratification model is a powerful tool to accurately identify high-risk patients unlikely to benefit from surgery alone, thereby guiding selection for systemic therapy while sparing low-risk patients from overtreatment.

Discussion

This is the first study to develop and validate a clinically relevant assay that predicts prognosis using a transcriptomic panel of endoscopically resected specimens in patients with T1 CRC. The primary endpoint of this study was to predict 3-year RFS; however, the use of state-of-the-art machine learning approaches made it robust for OS and recurrence-free survival at longer timepoints, with high sensitivity and specificity.

From a clinical standpoint, this study identified lymphatic invasion as an independent predictor of tumor recurrence after curative resection in T1 CRC (Table 2). While other reports have highlighted LNM positivity, lymphatic invasion, tumor size, histological grade, and lesion site (rectal) as independent risk factors for tumor recurrence after curative resection of T1 CRC^[16,17,49,52–54]. However, there are no studies that distinguish between distant tumor recurrence and local tumor recurrence, considering the risks. The tumor recurrence rate after radical resection of T1 CRC is low, making it difficult to accurately predict the risk of tumor recurrence in these

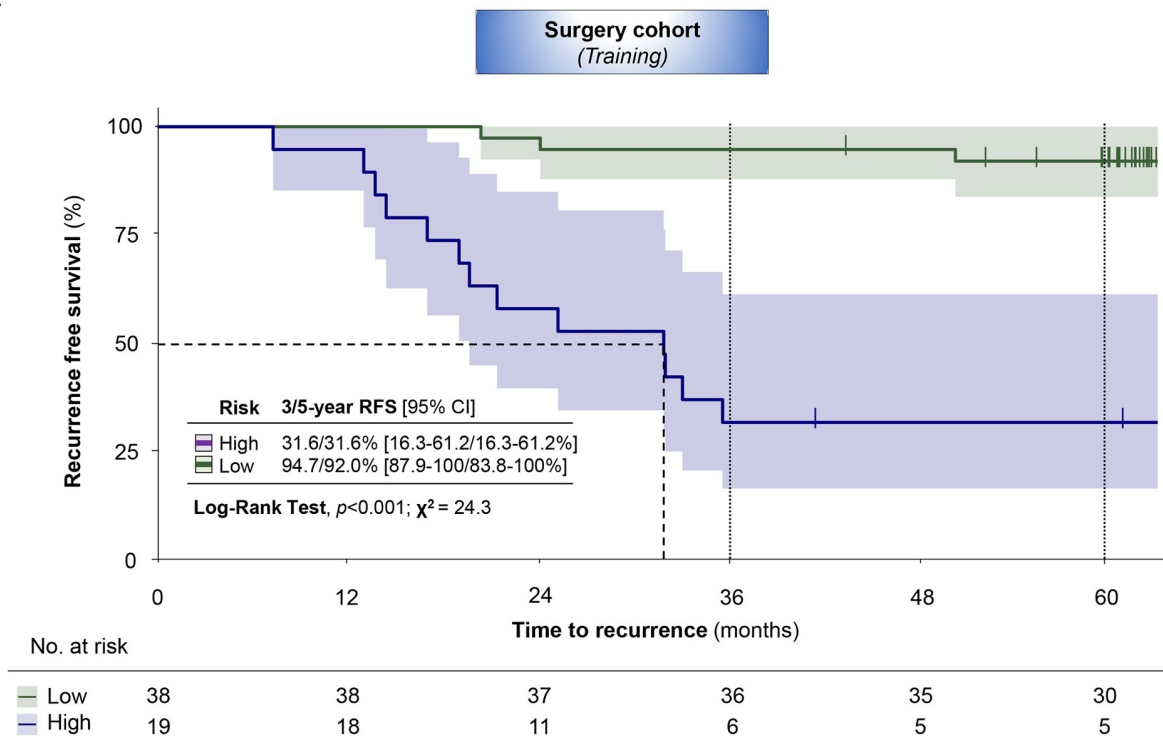
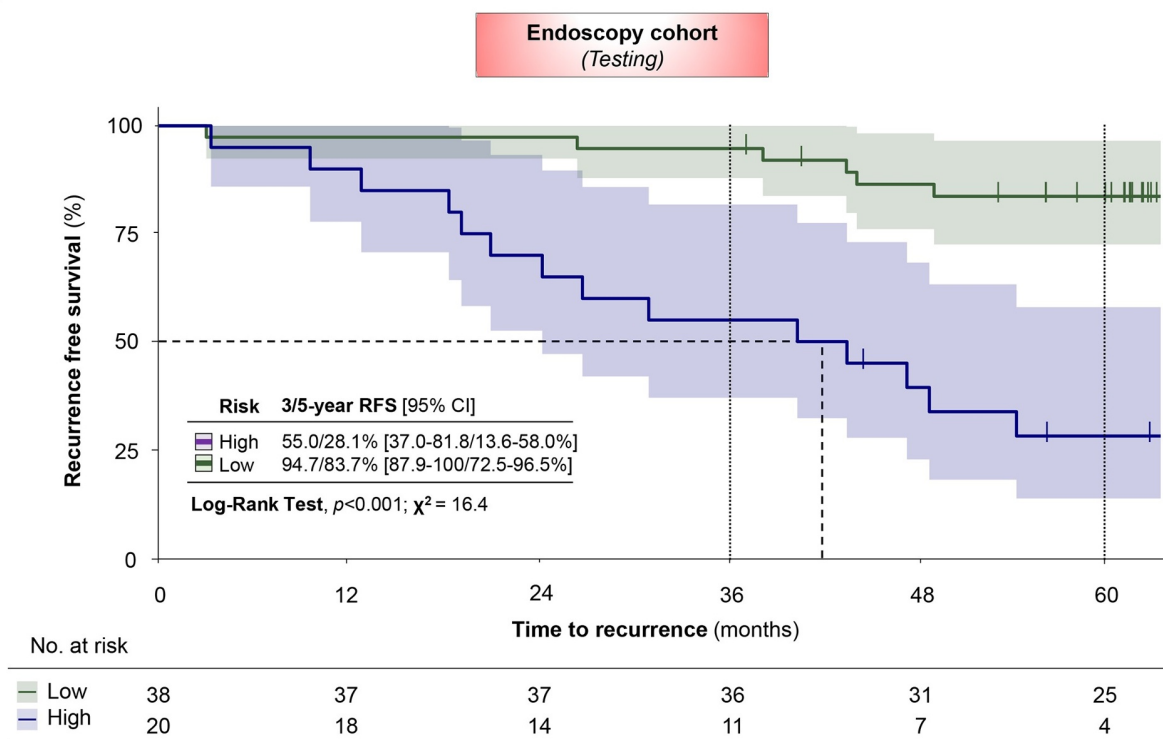
A**B**

Figure 3. Analysis of recurrence-free survival. (A) Kaplan-Meier curve showing recurrence-free survival based on stratification by transcriptomic panel scores in the training cohort ($P < 0.001$). (B) Kaplan-Meier curve showing recurrence-free survival based on stratification by transcriptomic panel scores in the testing cohort ($P < 0.001$). CI, confidence intervals; RFS, recurrence-free survival.

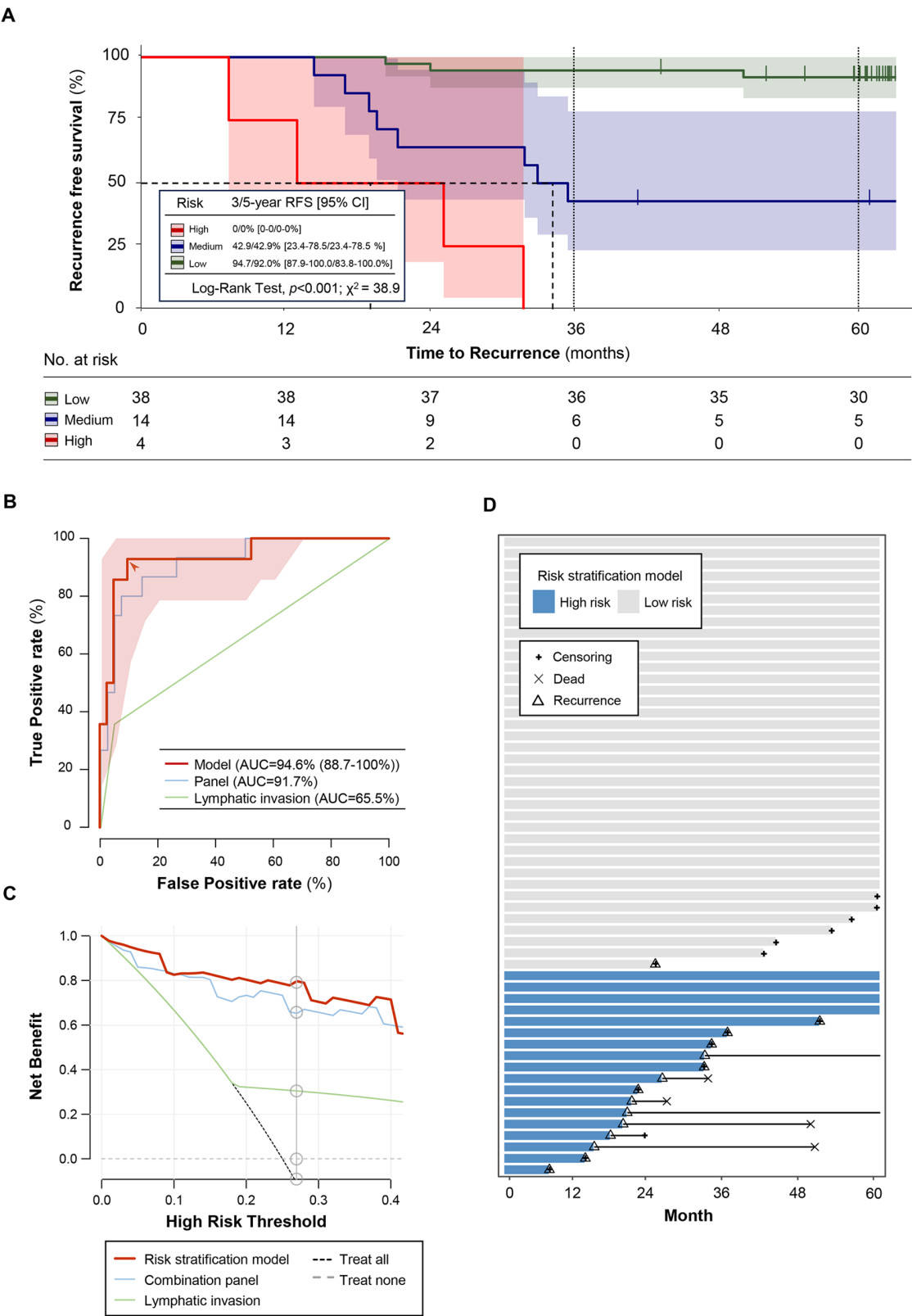


Figure 4. Establishment of a risk-stratification model. (A) Kaplan-Meier curve analysis constructed in the surgical cohort divided into three groups: (1) high transcriptomic panel score and positive lymphatic invasion, (2) high transcriptomic panel score and negative lymphatic invasion, and (3) low transcriptomic panel score. (B) AUROC of the risk stratification model (AUROC = 94.6%). Arrowhead indicated the positivity threshold for indicating a high risk of recurrence (based on which performance values are reported for sensitivity, 92.9%, and specificity, 88.1%). (C) DCA is used for the risk-stratification model, transcriptomic panel, and lymphatic invasion in predicting tumor recurrence. (D) Swimmer plot of survival stratified by risk-stratification model. AUROC, area under the receiver operating characteristic curve; CI, confidence intervals; DCA, decision curve analysis; mRNA, messenger RNA; miRNA, microRNA; RFS, recurrence-free survival.

Table 2
Univariate and multivariate analysis for 3-year recurrence-free survival in the surgery cohort

Characteristics	OR	Univariate analysis		OR	Multivariate analysis	
		95% CI	P-Value		95% CI	P-Value
Age						
(≥60 y/ < 60 y)	2.22	0.59–10.88	0.27			
Biological sex						
(Male/Female)	1.38	0.39–5.67	0.63			
Tumor size, mm						
(≥20 mm/ < 20 mm)	3.66	1.07–13.08	0.04	3.75	0.54–34.54	0.20
Submucosal invasion, μm						
(≥1000 μm/ < 1000 μm)	0.24	0.06–1.04	0.052			
Budding grade						
(≥2/1)	2.36	0.59–9.01	0.21			
Lymphatic invasion						
(Present/Absent)	11.11	2.05–86.95	0.01	35.63	2.34–1272.43	0.02
Vascular invasion						
(Present/Absent)	3.80	0.78–18.85	0.09			
Differentiation						
(Well-Moderate/Poor)	2.17	0.26–14.61	0.43			
mRNA/miRNA panel score						
(High/Low)	39.00	8.31–295.86	<0.01	50.74	7.53–1058.26	<0.01

Reference categories: Age <60 y; female; Tumor size <20 mm; Submucosal invasion <1000 μm; Budding grade = 1; Lymphatic invasion absent; Vascular invasion absent; Differentiation well–moderate; mRNA/miRNA panel score low; P < 0.05 are highlighted in bold in the table.

patients. Considering that there is no significant difference in the rate of distant tumor recurrence between cases involving only endoscopic resection and those involving surgical resection with removal of regional lymph nodes^[16,17], predicting the risk of distant tumor recurrence through hematogenous metastases is crucial. Our training cohort included cases after surgical resection, including regional lymph node dissection, and we were able to accurately evaluate the risk of distant tumor recurrence due to hematogenous metastasis. In addition, the testing cohort included cases that underwent endoscopic treatment only, and in these cases, only the primary tumor was resected without regional lymph node dissection. In other words, the testing cohort may reflect the overall picture of tumor recurrence, including recurrence via both lymphatic and hematogenous routes, and our established panel can comprehensively predict both types of tumor recurrence. According to current guidelines, patients without LNM are not given chemotherapy, but systemic therapy may be appropriate for those with a high risk of tumor recurrence. Our transcriptomic panel and risk stratification model accurately identify such high-risk patients, as demonstrated by the DCA curve, swimmer plot, and Kaplan–Meier analysis. Importantly, this positions the panel as a potential tool for risk-adapted surveillance and adjuvant therapy decision-making. High-risk patients could benefit from intensified follow-up or even consideration of systemic therapies, while low-risk patients might safely avoid overtreatment and undergo less intensive surveillance. This ability to stratify patients based on molecular risk fills a critical gap left by conventional histopathological factors, which often lack sufficient prognostic precision. In addition, the biomarkers incorporated into our panel – five mRNAs (RCC1, FOXA1, AMT, MMP9, and MMP1) and two miRNAs (miR-411-5p and miR-193b-3p) – have each been reported in prior studies to be associated with aggressive tumor behavior, metastasis, or poor prognosis in CRC and

other malignancies (Supplemental Digital Content sTable 6, available at: <http://links.lww.com/JS9/G597>)^[55–60]. In addition, our network and pathway analyses (Supplemental Digital Content sFigure 8, available at: <http://links.lww.com/JS9/G597>) showed that the candidate mRNAs and miRNAs form interconnected regulatory clusters and are enriched in biological processes central to CRC progression, including cell-cycle regulation, focal adhesion, p53/mTOR signaling, and extracellular-matrix remodeling. These findings provide mechanistic support for the biological coherence of the transcriptomic panel and strengthen its relevance as a predictor of aggressive tumor behavior and recurrence risk in T1 CRC.

Notably, the relatively high recurrence rate observed in our cohort – much higher than the 2–8% described in population-based T1 CRC studies^[16,17,49,52,53] – reflects intentional enrichment for high-risk cases rather than unintended sampling bias. Both the surgical training cohort and the endoscopically treated testing cohort consisted largely of patients with high-risk pathological features or long-term intensive surveillance. Therefore, our model is particularly suited for recurrence prediction in high-risk T1 CRC populations, where the need for accurate molecular risk stratification is greatest. Furthermore, this study is specifically focused on high-risk patients; therefore, we did not apply methods such as standardization of recurrence rates, propensity score matching, or inverse probability weighting. However, we recognize that further insights could be gained using these methods if a larger, multi-center dataset with a wider distribution of risk becomes available, and we consider this an important direction for future research.

From an implementation perspective, several challenges remain. Cost and turnaround time of transcriptomic testing may currently limit routine clinical adoption, particularly in community settings. In addition, standardization of RNA extraction and qPCR-based assays across pathology laboratories is critical to

ensure reproducibility and comparability of results. These issues highlight the need for multi-center standardization studies, as well as the potential utility of automated, clinically certified platforms to streamline workflow. Nonetheless, the assay itself can be readily implemented using routinely processed and highly stable FFPE specimens. Only 2–4 unstained 5–10- μ m sections after manual microdissection are required, making it compatible with standard pathology workflows. RNA can be extracted using widely available FFPE extraction kits, and no sequencing infrastructure is needed; all seven biomarkers can be quantified using a conventional real-time PCR. The entire workflow can typically be completed within 3–5 days, supporting feasible adoption across community and academic laboratories. These features suggest that, with forthcoming multi-center standardization and external quality-assurance frameworks, the assay is well positioned for future clinical implementation. In addition, in line with our prior work evaluating the cost-effectiveness of molecular biomarker assays^[41,42,61], the present study also incorporated DCA to provide an initial assessment of cost-effectiveness. The DCA demonstrated that our combined risk stratification model consistently yielded greater net clinical benefit than conventional guideline-based criteria, suggesting that the added implementation costs may be justified by improved patient outcomes.

This study has limitations. First, it was retrospective, and the sample size for biomarker training and testing was limited. Second, most patients in the endoscopic group had deep sub-mucosal invasion, which, although a surgical indication, has been associated with a low rate of LNM and recurrence^[54,62]. Third, tumor size and vascular invasion differed between the training and testing cohorts; however, neither factor was independently associated with recurrence in multivariate analyses, suggesting that their impact on the predictive performance of our panel was minimal. Moreover, the robustness of the panel across these distinct clinical subgroups highlights its potential to detect recurrence – including distant recurrence – beyond local differences between cohorts.

Finally, while this external validation study used a multi-center cohort from Spain, future research will require more extensive international, multi-ethnic, and multi-center validation. Moreover, given the minimally invasive nature of endoscopic resection and the growing interest in non-invasive monitoring, liquid biopsy-based adaptations of the transcriptomic panel represent an attractive avenue for further development. Such approaches could extend the utility of the assay to longitudinal monitoring and early detection of recurrence, thereby enhancing clinical applicability. Indeed, in our previous studies, we successfully translated tissue-based transcriptomic and microRNA biomarkers into blood-based liquid biopsy assays^[13,14], supporting the feasibility of applying the present transcriptomic panel in a liquid biopsy setting.

In conclusion, this study represents the first attempt to adapt surveillance strategies for risk stratification in patients with T1 CRC. Using biological data and machine learning, we independently and externally validated our findings by investigating tumor recurrence, which has a significant impact on prognosis. This machine learning-based signature has excellent potential to provide a personalized clinical approach to identify high-risk patients who require systemic therapies after radical resection of T1 CRC.

Ethical approval

This study received ethical approval from the institutional review boards of all participating institutions. The approval reference numbers are IRB/23228 and HCB/2019/0224, PI2019199.

Consent

Written informed consent was obtained from the patient for publication of this case report and accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal on request.

Sources of funding

This work was supported by CA72851, CA181572, CA184792, CA187956, CA202797, CA214254, and CA271443 grants from the NIH. Instituto de Salud Carlos III (PI19/01050) and Beca de la Marató de TV3 2020 (Beca la Marató – 201932-30). Co-funded by the European Regional Development Fund/ European Social Fund; “A way to make Europe”/ “Investing in your future.” CIBERehd is funded by the Instituto de Salud Carlos III.

Author contributions

T.N., K.S.G., K.M., Y.W., A.M., T.O., M.P., and A.G.: conceptualization; T.N., K.S.G., M.D.A., M.C., and L.B.: data curation; T.N., K.M., A.M., and T.O.: formal analysis; M.P. and A.G.: funding acquisition; T.N., K.S.G., M.D.A., M.C., and L.B.: investigation; T.N., K.M., Y.W., A.M., and T.O.: methodology; M.P. and A.G.: project administration; M.P. and A.G.: resources; T.N., A.M., and T.O.: software; M.P. and A.G.: supervision; T.N., K.M., A.M., and T.O.: validation; T.N., A.M., and T.O.: visualization; T.N., A.M., and A.G.: writing – original draft; K.S.G., M.D.A., K.M., Y.W., A.M., M.S., M.C., L.B., M.P., and A.G.: writing – review & editing.

Conflicts of interest disclosure

The authors have no conflicts of interest to disclose that are relevant to the present study.

Research registration unique identifying number (UIN)

NCT06314971 (ClinicalTrials.gov ID).

Guarantor

Takayuki Noma and Ajay Goel.

Provenance and peer review

Not commissioned; externally peer-reviewed

Data availability statement

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request. Takayuki Noma and Ajay Goel had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Co-authors from the EpiT1 consortium who have collaborated on this project

Fernando Martinez de Juan, Gastroenterology Department, Instituto Valenciano de Oncología, Valencia, Spain; Isidro Machado, Pathology Department, Instituto Valenciano de Oncología, Valencia, Spain; Alberto Perez Pedrosa, Pathology Department, Hospital Ourense, Galicia, Spain; Luis Bujanda Piedrola, Gastroenterology Department, Instituto Bionostia, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Universidad del País Vasco (UPV/EHU), San Sebastián, Basque Country, Spain; Marta Ponce, Gastroenterology Department, Hospital Clínico de Valencia, Valencia, Spain; Carolina Martinez Pathology Department, Hospital Clínico de Valencia, Valencia, Spain; Marco Aurelio Alburquerque, Gastroenterology Department, Hospital de Palamós, Catalonia, Spain; Angeles pizarro moreno Gastroenterology Department, Hospital Virgen del Rocio de Sevilla, Andalusia, Spain; Carolina Castilla Ramirez, Pathology Department, Hospital Virgen del Rocio de Sevilla, Andalusia, Spain; Ignasi Puig, Gastroenterology Department, Althaia Xarxa Assistencial Universitària de Manresa, Barcelona, Catalonia, Spain.; Oria Rosiñol, Pathology Department, Althaia Xarxa Assistencial Universitària de Manresa, Barcelona, Catalonia, Spain.; Alexandra Gene Pathology Department, Hospital Inca, Mallorca, Balearic Islands, Spain; Marita Trelles, Gastroenterology Department, Hospital Inca, Mallorca, Balearic Islands, Spain; Salvador Machlab Gastroenterology Department, Hospital Parc Tauli, Sabadell, Catalonia, Spain; Lluïsa Silva Pathology Department, Hospital Parc Tauli, Sabadell, Catalonia, Spain; Xavier Sanjuan, Pathology Department, Hospital Bellvitge, Barcelona, Catalonia, Spain; Gemma Ibañez; Gastroenterology Department, Hospital Bellvitge, Barcelona, Catalonia, Spain; Ángel Ferrández Gastroenterology Department, Hospital Clínico de Zaragoza, Zaragoza, Aragon, Spain; Teresa Ramirez Pathology Department, Hospital Clínico de Zaragoza, Zaragoza, Aragon, Spain; Eva Musulén, Pathology Department, Hospital Sagrat Cor, Barcelona, Catalonia, Spain; Diana Zaffalon, Gastroenterology Department, Hospital Terrassa, Catalonia, Spain; Ismael Jurado, Pathology Department, Hospital Terrassa, Catalonia, Spain; Rodrigo Jover, Gastroenterology Department, Hospital General Universitario de Alicante, Alicante, Valencia, Spain; Cristina Alenda, Pathology Department, Hospital General Universitario de Alicante, Alicante, Valencia, Spain; and Pilar Diez Redondo, Gastroenterology Department, Hospital Universitario Río Hortega, Valladolid, Castile and León, Spain.

Acknowledgements

Not applicable.

References

- [1] Morgan E, Arnold M, Gini A, *et al.* Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut* 2023;72:338–44.
- [2] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin* 2023;73:17–48.
- [3] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- [4] Chiu HM, Chen SL, Yen AM, *et al.* Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. *Cancer* 2015;121:3221–29.
- [5] Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet* 2019;394:1467–80.
- [6] Jayanna M, Burgess NG, Singh R, *et al.* Cost analysis of endoscopic mucosal resection vs surgery for large laterally spreading colorectal lesions. *Clin Gastroenterol Hepatol* 2016;14:271–78.
- [7] Weinberg RA. Mechanisms of malignant progression. *Carcinogenesis* 2008;29:1092–95.
- [8] 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.
- [9] Pimentel-Nunes P, Dinis-Ribeiro M, Ponchon T, *et al.* Endoscopic submucosal dissection: European Society of Gastrointestinal Endoscopy (ESGE) guideline. *Endoscopy* 2015;47:829–54.
- [10] Kaltenbach T, Anderson JC, Burke CA, *et al.* Endoscopic removal of colorectal lesions-recommendations by the US multi-society task force on colorectal cancer. *Gastroenterology* 2020;158:1095–129.
- [11] Kandimalla R, Ozawa T, Gao F, Wang X, Goel A, Group TCCS. Gene expression signature in surgical tissues and endoscopic biopsies identifies high-risk T1 colorectal cancers. *Gastroenterology* 2019;156:2338–2341e2333.
- [12] Ozawa T, Kandimalla R, Gao F, *et al.* A MicroRNA signature associated with metastasis of t1 colorectal cancers to lymph nodes. *Gastroenterology* 2018;154:844–848e847.
- [13] Wada Y, Shimada M, Murano T, *et al.* A liquid biopsy assay for non-invasive identification of lymph node metastases in T1 colorectal cancer. *Gastroenterology* 2021;161:151–162e151.
- [14] Miyazaki K, Wada Y, Okuno K, *et al.* An exosome-based liquid biopsy signature for pre-operative identification of lymph node metastasis in patients with pathological high-risk T1 colorectal cancer. *Mol Cancer* 2023;22:2.
- [15] Naxerova K, Reiter JG, Brachtel E, *et al.* Origins of lymphatic and distant metastases in human colorectal cancer. *Science* 2017;357:55–60.
- [16] Yeh JH, Tseng CH, Huang RY, *et al.* Long-term outcomes of primary endoscopic resection vs surgery for T1 colorectal cancer: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2020;18:2813–2823e2815.
- [17] Martinez Vila C, Oliveres Montero de Novoa H, Martinez-Bauer E, *et al.* A real world analysis of recurrence risk factors for early colorectal cancer T1 treated with standard endoscopic resection. *Int J Colorectal Dis* 2020;35:921–27.
- [18] Kirby E, Tam W, Gilham I, *et al.* Effect of sarcobesity index and body adipose tissue variables on cardiopulmonary exercise testing performance in colorectal surgery setting: a retrospective cohort study. *Br J Hosp Med (Lond)* 2024;85:1–17.
- [19] Mungan I, Bostancı EB, Türksal E, Tezcan B, Turan S. An early detector of anastomosis leakage following colorectal cancer surgery in intensive care. *Cancer Plus* 2023;5:2374.
- [20] Tong G, Peng T, Chen Y, *et al.* Effects of GLP-1 receptor agonists on biological behavior of colorectal cancer cells by regulating PI3K/AKT/mTOR signaling pathway. *Front Pharmacol* 2022;13:901559.
- [21] Huang B, An H, Gui M, *et al.* Qingjie Fuzheng Granule prevents colitis-associated colorectal cancer by inhibiting abnormal activation of NOD2/NF-kappaB signaling pathway mediated by gut microbiota disorder. *Chin Herb Med* 2025;17:500–12.
- [22] Kopetz S, Tabernero J, Rosenberg R, *et al.* Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors. *Oncologist* 2015;20:127–33.
- [23] Metzger DA, Greenspun BC, Brouwer J, *et al.* Genomic signatures of recurrence after resection of early-stage node-negative colon cancer. *Ann Surg Oncol* 2025;32:3774–80.

- [24] Srivastava G, Renfro LA, Behrens RJ, *et al.* Prospective multicenter study of the impact of oncotype DX colon cancer assay results on treatment recommendations in stage II colon cancer patients. *Oncologist* 2014;19:492–97.
- [25] Wood PC, Zammit AP, Ferris B, *et al.* Establishing high risk factors for recurrence in stage I/II colorectal cancer: a metropolitan perspective. *World J Surg Oncol* 2025;23:241.
- [26] van der Schee L, Verbeek A, Deckers IAG, *et al.* Variation in the detection of lymphovascular invasion in T1 colorectal cancer and its impact on treatment: a nationwide Dutch study. *United Eur Gastroenterol J* 2024;12: 1429–39.
- [27] Busan Ulsan Gyeongnam Intestinal Study Group S. Lee JY, Roh MS, Lee JH, *et al.* Comparison of the prognosis and lymph node metastasis between no tumor budding and low-grade tumor budding in T1 and T2 colorectal cancer. *Sci Rep* 2025;15:212.
- [28] Chen PC, Kao YK, Yang PW, Chen CH, Chen CI. Long-term outcomes and lymph node metastasis following endoscopic resection with additional surgery or primary surgery for T1 colorectal cancer. *Sci Rep* 2025;15:2573.
- [29] Abdelmaksoud-Dammak R, Chamtour N, Triki M, *et al.* Overexpression of miR-10b in colorectal cancer patients: correlation with TWIST-1 and E-cadherin expression. *Tumour Biol* 2017;39:1010428317695916.
- [30] Hao H, Xia G, Wang C, Zhong F, Liu L, Zhang D. miR-106a suppresses tumor cells death in colorectal cancer through targeting ATG7. *Med Mol Morphol* 2017;50:76–85.
- [31] Hur K, Toiyama Y, Okugawa Y, *et al.* Circulating microRNA-203 predicts prognosis and metastasis in human colorectal cancer. *Gut* 2017;66:654–65.
- [32] Li Y, Chen Z. The oncogenic role of microRNA-500a in colorectal cancer. *Oncol Lett* 2020;19:1799–805.
- [33] Mohammadi P, Saidijam M, Kaki A, Etemadi K, Shabab N, Yadegarazari R. A pilot study of CK19, CK20 and GCC mRNA in the peripheral blood as a colorectal cancer biomarker panel. *Int J Mol Cell Med* 2016;5:30–36.
- [34] Tang M, Zhou J, You L, Cui Z, Zhang H. LIN28B/IRS1 axis is targeted by miR-30a-5p and promotes tumor growth in colorectal cancer. *J Cell Biochem* 2020;121:3720–29.
- [35] Toiyama Y, Hur K, Tanaka K, *et al.* Serum miR-200c is a novel prognostic and metastasis-predictive biomarker in patients with colorectal cancer. *Ann Surg* 2014;259:735–43.
- [36] Toiyama Y, Takahashi M, Hur K, *et al.* Serum miR-21 as a diagnostic and prognostic biomarker in colorectal cancer. *J Natl Cancer Inst* 2013;105:849–59.
- [37] Wang H, Zhang X, Wang L, *et al.* Investigation of cell free BIRC5 mRNA as a serum diagnostic and prognostic biomarker for colorectal cancer. *J Surg Oncol* 2014;109:574–79.
- [38] Okumura T, Hotta K, Imai K, *et al.* Comparison of survival and recurrence between colonic and rectal tumors after resection in T1 colorectal cancer: long-term cohort study. *Gastrointest Endosc* 2025. doi:10.1016/j.gie.2025.06.009
- [39] Su J, Liu Z, Li H, *et al.* Artificial intelligence-based model to predict recurrence after local excision in T1 rectal cancer. *Eur J Surg Oncol* 2025;51:109717.
- [40] Li Z, Aihemaiti Y, Yang Q, *et al.* Survival machine learning model of T1 colorectal postoperative recurrence after endoscopic resection and surgical operation: a retrospective cohort study. *BMC Cancer* 2025;25:262.
- [41] Mannucci A, Hernandez G, Uetake H, *et al.* Prediction of long-term recurrence-free and overall survival in early-onset colorectal cancer: the ENCORE multi-centre study. *NPJ Precis Oncol* 2025;9:202.
- [42] Miyoshi J, Mannucci A, Scarpa M, *et al.* Liquid biopsy to identify Barrett's oesophagus, dysplasia and oesophageal adenocarcinoma: the EMERALD multicentre study. *Gut* 2025;74:169–81.
- [43] He B, Wang L, Zhou W, *et al.* A fusion model to predict the survival of colorectal cancer based on histopathological image and gene mutation. *Sci Rep* 2025;15:9677.
- [44] Li H, Wang Z, Guan Z, *et al.* UCFNNet: ulcerative colitis evaluation based on fine-grained lesion learner and noise suppression gating. *Comput Methods Programs Biomed* 2024;247:108080.
- [45] STROCSS Group. Agha RA, Mathew G, Rashid R, *et al.* Revised strengthening the reporting of cohort, cross-sectional and case-control studies in surgery (STROCSS) guideline: an update for the age of artificial intelligence. *Prem J Sci* 2025;10:100081.
- [46] Nors J, Iversen LH, Erichsen R, Gotschalck KA, Andersen CL. Incidence of recurrence and time to recurrence in stage I to III colorectal cancer: a nationwide danish cohort study. *JAMA Oncol* 2024;10: 54–62.
- [47] Lan YT, Chang SC, Lin PC, *et al.* Clinicopathological and molecular features of patients with early and late recurrence after curative surgery for colorectal cancer. *Cancers (Basel)* 2021;13:1883.
- [48] Shahabi F, Mehri A, Abdollahi A, *et al.* Post recurrence survival in early versus late period and its prognostic factors in rectal cancer patients. *Sci Rep* 2024;14:17661.
- [49] Dang H, Dekkers N, le Cessie S, *et al.* Risk and time pattern of recurrences after local endoscopic resection of T1 colorectal cancer: a meta-analysis. *Clin Gastroenterol Hepatol* 2022;20:e298–e314.
- [50] Arthursson V, Medic S, Syk I, Ronnow CF, Thorlacius H. Risk of recurrence after endoscopic resection of nonpedunculated T1 colorectal cancer. *Endoscopy* 2022;54:1071–77.
- [51] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods* 2001;25:402–08.
- [52] Chang LC, Shun CT, Lin BR, *et al.* Recurrence outcomes less favorable in T1 rectal cancer than in T1 colon cancer. *Oncologist* 2021;26: e1548–e1554.
- [53] Kobayashi H, Mochizuki H, Morita T, *et al.* Characteristics of recurrence after curative resection for T1 colorectal cancer: japanese multicenter study. *J Gastroenterol* 2011;46:203–11.
- [54] Yoshii S, Nojima M, Noshio K, *et al.* Factors associated with risk for colorectal cancer recurrence after endoscopic resection of T1 tumors. *Clin Gastroenterol Hepatol* 2014;12:292–302e293.
- [55] Pezeshkian Z, Nobili S, Peyravian N, *et al.* Insights into the role of matrix metalloproteinases in precancerous conditions and in colorectal cancer. *Cancers (Basel)* 2021;13:6226.
- [56] Lazar SB, Pongor L, Li XL, *et al.* Genome-wide analysis of the FOXA1 transcriptional network identifies novel protein-coding and long noncoding rna targets in colorectal cancer cells. *Mol Cell Biol* 2020;40:e00224–20.
- [57] Niu C, Li X, Lei Luo X, *et al.* Identification of novel prognostic biomarkers for colorectal cancer by bioinformatics analysis. *Turk J Gastroenterol* 2024;35:61–72.
- [58] Deng Y, Yu L, Zhao Y, *et al.* RCC1 expression as a prognostic marker in colorectal liver oligometastases. *Pathol Oncol Res* 2021; 27:1610077.
- [59] Dinami R, Petti E, Porru M, *et al.* TRF2 cooperates with CTCF for controlling the oncomiR-193b-3p in colorectal cancer. *Cancer Lett* 2022;533:215607.
- [60] Zhang B, Cao W, Liu Y, Zhao Y, Liu C, Sun B. Circ_0056618 enhances PRRG4 expression by competitively binding to miR-411-5p to promote the malignant progression of colorectal cancer. *Mol Cell Biochem* 2023;478:503–16.
- [61] Xu C, Mannucci A, Esposito F, *et al.* An exosome-based liquid biopsy predicts depth of response and survival outcomes to cetuximab and panitumumab in metastatic colorectal cancer: the EXONERATE study. *Clin Cancer Res* 2025;31:1002–15.
- [62] 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: 174–89.