International Journal of Colorectal Disease

https://doi.org/10.1007/s00384-026-05181-z

Article in Press

Development and internal validation of a nomogram to predict colorectal cancer risk in patients with diverticulitis

Received: 26 Oct 2025

Accepted: 10 Jun 2026

Published online: 23 June 2026

Cite this article as: Usman, M., Correa, E., Chung, W. *et al.* Development and internal validation of a nomogram to predict colorectal cancer risk in patients with diverticulitis. *Int J Colorectal Dis* (2026). https://doi.org/10.1007/s00384-026-05181-z

Muhammad Usman, Elon Correa, Wenyuan Chung, Sayyada Farazi, Konstantinos Baronos, Lasitha Samarakoon & Sudarshan Kadri

Title: Development and Internal Validation of a Nomogram to Predict Colorectal Cancer Risk in Patients with Diverticulitis

Running Title: CRC nomogram in diverticulitis

Authors: Dr Muhammad Usman, MBBS, Department of Gastroenterology, University Hospitals of Leicester, UK; muhammad.usman5@nhs.net

Dr Elon Correa , PhD, Liverpool John Moores University, UK; e.s.correa@ljmu.ac.uk

Mr. Wenyuan Chung, PhD, Department of Surgery, University Hospitals of Leicester; wenyuan.chung@nhs.net

Dr Sayyada Nuha Farazi, MBBS, Department of Gastroenterology, University Hospitals of Leicester; s.nuha18@gmail.com

Dr Konstantinos Baronos, MBBS, Department of Surgery, University Hospitals of Leicester; konbaronos@gmail.com

Mr. Lasitha Samarakoon, MBBS, Department of Surgery, University Hospitals of Leicester; l.samarakoon@nhs.net

Sudarshan Kadri, MD, Department of Gastroenterology, University Hospitals of Leicester; skadri@nhs.net

Corresponding author(s): Name: Dr Muhammad Usman; Degree: MBBS, Department of Gastroenterology, University Hospitals of Leicester, UK; Email address: Muhammad.usman5@nhs.net; Telephone No: 00447448205772

Dr Elon Correa , PhD, Liverpool John Moores University, UK; e.s.correa@ljmu.ac.uk

Statement and Declarations

Data availability declaration: The data supporting the findings of this study are not openly available due to the factors of privacy and sensitivity. However, the data is available from the corresponding author upon request.

Competing interest declaration: None to declare

Funding/Support: Nil

Financial disclosure: Nil

Ethics approval: Not applicable

Poster presentation: Being presented in Midlands Society of Gastroenterology (MGS) conference on 14th November 2025, Loughborough, United Kingdom

Author contribution: **MU** contributed in conception, design, acquisition of data, analysis and interpretation of data and writing and review of manuscript; **EL** contributed equally alongside MU in conception, design, acquisition of data, analysis and interpretation of data and writing and review of manuscript; **WC** contributed in conception, design, acquisition of data, analysis, and review of manuscript; **LS** conception, design, acquisition of data, analysis, and review of manuscript; **SNF** contributed in acquisition of data; **KB** contributed in acquisition of data; **SK** supervised the study and contributed in conception, design, analysis and interpretation of data and review of manuscript

Category: Colorectal neoplasia

Abstract

Aim: Diverticulitis, considered a benign condition, has emerging evidence linking it to colorectal cancer (CRC). Current guidelines lack standardised risk-stratification tools to guide post-diverticulitis surveillance. This study aims to create and validate a nomogram to predict CRC risk in patients with diverticulitis.

Method: This retrospective cohort study included1,546 patients diagnosed with diverticulitis at a UK tertiary hospital between January 2021 and December 2023**.** Patients aged ≥18 years who underwent endoscopic assessment following a diverticulitis episode were included

A logistic regression model with L2 regularisation and balanced class weights (C=0.1) was developed to predict a binary outcome: high CRC risk (histologically confirmed cancer) versus low/moderate risk (no cancer), designed to guide referral for flexible sigmoidoscopy (FS).

Results: The mean age of the cohort was 69.2 years, with near-equal sex distribution. CRC was identified in 42 patients (2.7%). Significant predictors of high CRC risk included older age, male sex, and CT-detected colonic wall thickening, whereas abdominal pain, PR bleeding, and diverticulosis were associated with lower risk. The model achieved a cross-validated mean AUC of 0.736 (95% CI 0.661–0.811) and full-sample AUC of 0.784. At the Youden-optimal threshold (0.523), sensitivity was 72.1% and specificity 73.8%.

Conclusion: We developed and internally validated a nomogram to predict CRC risk in patients with diverticulitis, using a binary classification to guide FS referral. The tool demonstrated robust discriminative ability and could reduce unnecessary colonoscopies while prioritising high-risk patients for timely evaluation. Unlike existing CRC risk models targeting asymptomatic individuals, our nomogram is tailored to diverticulitis patients. External validation in multi-centre cohorts is required before adoption into routine clinical practice.

Introduction

Diverticular disease, of which diverticulitis is a potentially serious inflammatory complication, affects millions worldwide, with a prevalence of up to 70% in individuals over 80 years [1]. Traditionally, it has been considered a benign gastrointestinal condition. However, recent epidemiological studies have suggested its potential association with colorectal cancer (CRC) [2.3]. This presents a complex diagnostic challenge for clinicians as diverticulitis and CRC can often have overlapping radiological and clinical features, which can delay diagnosis of potential colorectal cancer (CRC) [4].

A nationwide Danish cohort study has provided compelling evidence for this association. The study demonstrated a heightened risk of CRC in patients with diverticulitis, especially during the first six months of a diverticulitis episode (hazard ratio 1.7; 95% CI 1.5-1.8) [5]. Similarly, a meta-analysis revealed a CRC pooled prevalence of 1.9% in patients with acute diverticulitis, with prevalence even higher among patients with complicated vs. uncomplicated disease (7.9% (95% CI 3.9-15.3%) vs. 1.3% (95% CI 0.1-2%) [6]. The DIVERT-Ca multicentre retrospective study identified radiological features, particularly Hinchey grade Ib diverticulitis, as significantly associated with CRC diagnosis (OR 4.51, P = 0.028) [7].

Genetic studies have also contributed to this evolving narrative. A bidirectional Mendelian randomization analysis revealed a positive association between genetically predicted diverticular disease and CRC risk in the FinnGen population ($\beta = 0.489$, P = 0.041) [8]. However, these results were not replicated in the UK Biobank cohort [9].These findings suggest population-specific genetic predispositions and biological mechanisms beyond clinical overlap.

Despite the evolving evidence, current clinical guidelines fall short in providing standardized post-diverticulitis endoscopy protocols. The DIVERT-Ca study highlighted this gap suggesting that only 53.7% of patients received appropriate endoscopic assessments following a diverticulitis episode [7]. The inconsistencies in clinical practice and absence of validated predictive tools represent a significant unmet clinical need in this area.

CRC is the third most prevalent cancer and second leading cause of cancer-related deaths worldwide [10]. Current population-based screening strategies may not adequately account for increased risk profile of diverticulitis patients. Endoscopic assessment after diverticulitis, while beneficial for early diagnosis in high-risk patients, can result in

unnecessary procedures for low-risk individuals. This is particularly relevant as it can lead to increased patient anxiety, healthcare costs, and procedure-related complications [11].

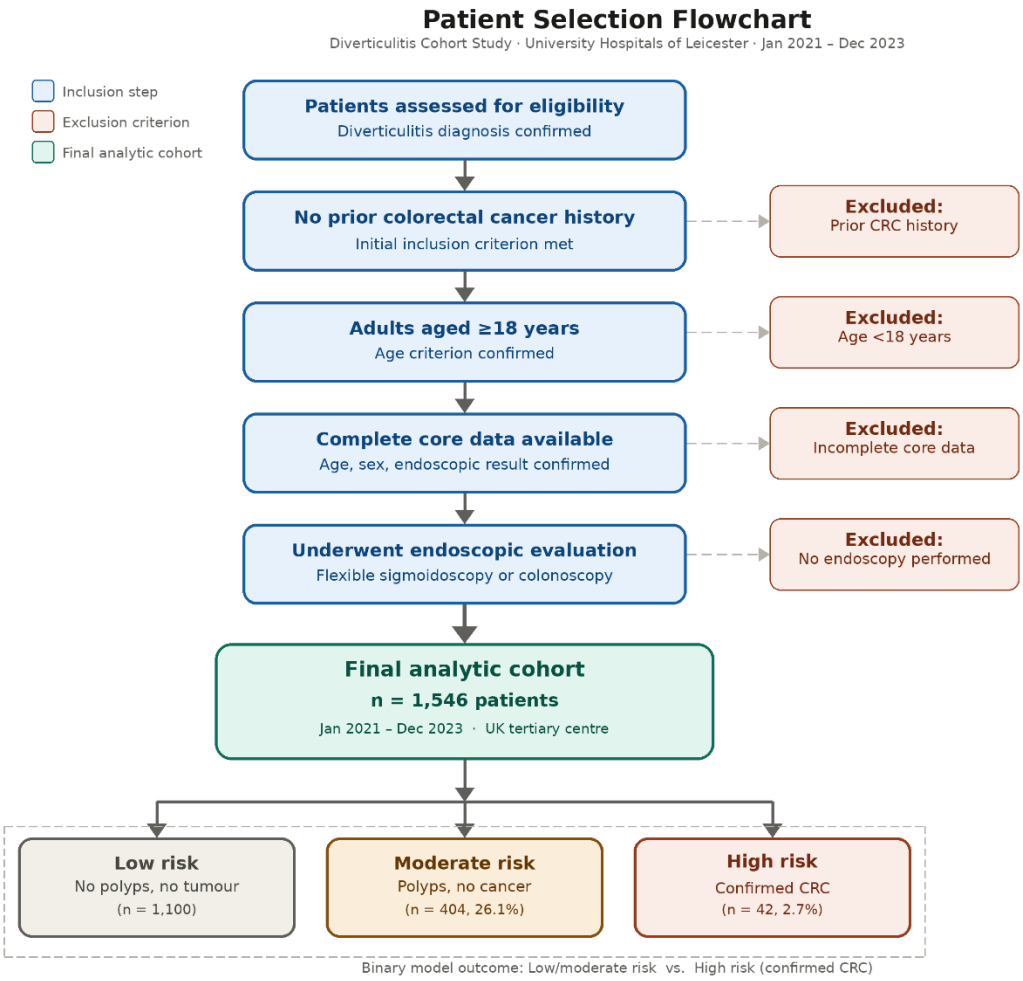
Nomograms, which integrate multiple clinical variables into a visual risk assessment tool, have shown promise in oncology for personalized cancer risk prediction [12]. However, their application in diverticulitis-associated CRC risk remains unexplored. This gap presents an opportunity to develop a nomogram tailored to diverticulitis patients, potentially enhancing early detection and guiding surveillance strategies. By synthesizing demographic, clinical, and radiological data, such a model could offer a pragmatic solution to a complex diagnostic challenge. We aimed to develop and validate a nomogram to predict CRC risk in patients with diverticulitis, using demographic, clinical, and imaging data.

Methods

Study Design and Population

This retrospective cohort study analyzed data from 1,546 patients diagnosed with diverticulitis between January 2021 and December 2023 in a tertiary care hospital in the UK. The study data was derived from the electronic patient records and included patients with confirmed diagnosis of diverticulitis.

Figure I. Flowchart of patient selection.



Prepared in accordance with STROBE reporting guidelines. CRC = colorectal cancer.

Inclusion and Exclusion Criteria

Inclusion Criteria:

- Adults aged ≥18 years.
- Diagnosis of diverticulitis either made clinically or based on CT-findings.
- Patients who underwent endoscopic assessment (flexible sigmoidoscopy or colonoscopy) following the diverticulitis episode.

Exclusion Criteria:

- Previous history of colorectal cancer.
- Incomplete clinical data.
- Patients aged <18.

Variables and Data Collection

All patients underwent flexible sigmoidoscopy (FS). Risk categories were defined as follows: patients with neither polyps nor tumors were classified as low risk, those with polyps but no tumors as moderate risk, and those with histologically confirmed CRC as high risk. For the binary outcome used in model development, the low- and moderate-risk groups were combined, while patients with confirmed cancer (high risk) were retained as a separate category. CT imaging was performed at clinician discretion as part of routine clinical management and was not mandated as an inclusion criterion; patients who did not undergo CT were not excluded. CT-based variables (wall thickening, fat stranding, perforation, abscess) were coded as absent when imaging had not been performed or when the relevant finding was not documented in the electronic patient record, which is standard practice in retrospective cohort studies using routinely collected data. The exclusion criterion of incomplete clinical data applied specifically to core demographic and outcome variables (age, sex, endoscopic result) rather than to imaging variables.

Clinical variables collected included:

- **Demographic factors:** Age, sex.
- **Clinical symptoms:** Abdominal pain, per-rectal (PR) bleeding, change in bowel habits.
- **Imaging findings:** CT wall thickening, fat stranding, perforation, abscess.
- **Endoscopy findings:** Diverticulosis, diverticulitis, polyps, colorectal cancer.

- **Outcome:** The primary outcome was a binary classification designed to guide referral for flexible sigmoidoscopy (FS). Patients were first categorized into three ordered groups based on endoscopic and histological findings:
    - **Low risk:** Neither polyps nor tumors.
    - **Moderate risk:** Presence of polyps but no tumor.
    - **High risk:** Histologically confirmed colorectal cancer.

For the purpose of developing a clinically actionable prediction model, these categories were condensed into a binary outcome: **High risk** (confirmed cancer) versus **Combined Low/Moderate risk** (no cancer). This approach focuses the model on identifying patients with confirmed malignancy, who require urgent endoscopic evaluation.

Feature Engineering

To optimise model performance, several engineered features were created: age-squared (Age^2) to capture non-linear age effects, an Age × Sex interaction term, and an Age × Wall thickening interaction term.

Statistical Analysis

Model Development: The study employed a balanced L2-regularized logistic regression model (C=0.1) as the optimal approach after comparison with tree-based alternatives. This selection was based on superior performance metrics and enhanced clinical interpretability. This study was conducted and reported in accordance with the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) statement for prediction model development and validation. All predictor variables are coded as binary (0/1) unless otherwise specified; age was entered as a continuous variable (years) and also as an engineered squared term (Age^2). Model coefficients and the intercept are available from the corresponding author on request.

Validation Strategy: Internal validation was performed using 5-fold stratified cross-validation to ensure robust performance estimates. This approach provided reliable estimates of model performance on unseen data while maintaining the original distribution of outcomes across folds.

Performance Metrics: Model discrimination was assessed using the area under the receiver operating characteristic curve (AUC). Sensitivity and specificity were evaluated at the Youden-optimal threshold, and calibration was assessed through visual inspection of the bootstrap-corrected calibration curve.

Clinical Tools Development

Excel Calculator: A user-friendly Excel-based calculator was developed implementing the validated logistic regression model, providing real-time risk assessment and clinical recommendations (supplementary material 1).

Bedside Nomogram: A visual nomogram was constructed using positive-only point scales for ease of clinical use, allowing manual calculation of risk scores at the bedside (Fig. II).

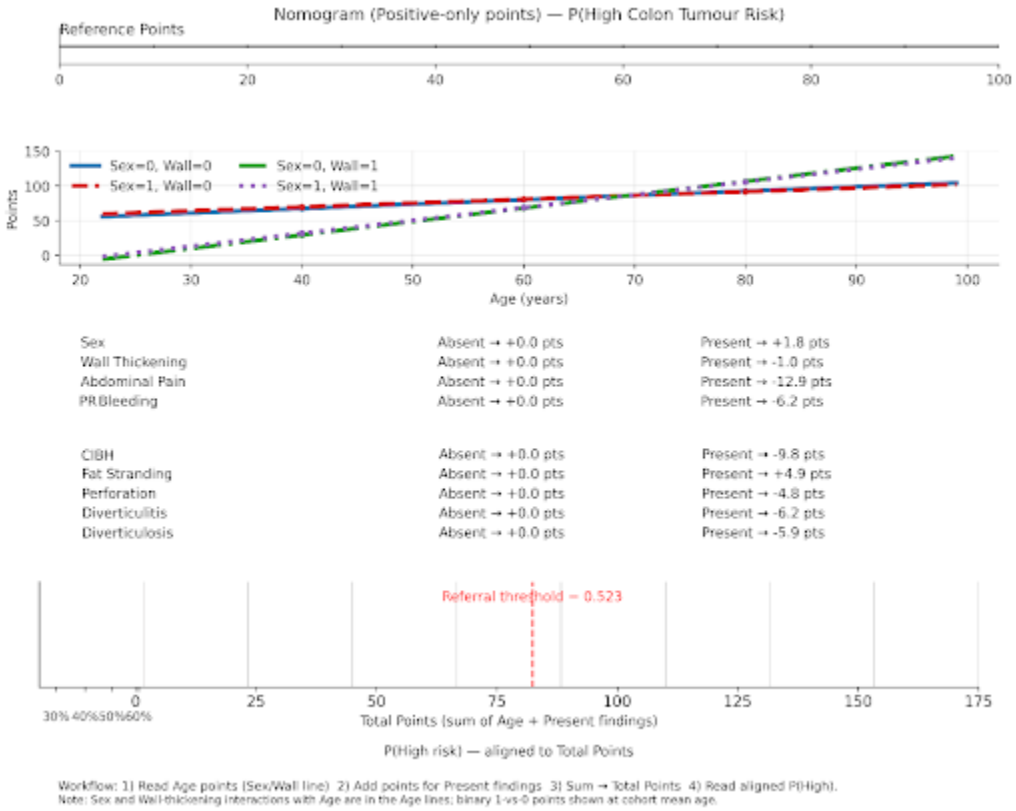


Figure II: Nomogram for estimating probability of high CRC risk.

Statistical Software

All analyses were performed using R Statistical Software (version 4.5.0; R Core Team, 2025) [13] with robust validation procedures and reproducible methodology.

Results

Patient Characteristics and main findings

Baseline characteristics of the study cohort, stratified by endoscopic risk group (low risk, moderate risk, and high risk / confirmed CRC), are presented in Table 1.

Characteristic	Low Risk (n=1,100)	Moderate Risk (n=404)	High Risk / CRC (n=42)	p-value
Demographics				
Age, mean (SD)	67.9 (13.4)	70.7 (11.6)	78.3 (10.2)	<0.001
Male sex, n (%)	500 (44.4%)	236 (59.3%)	21 (50.0%)	<0.001
Clinical symptoms, n (%)				
Abdominal pain	330 (29.3%)	78 (19.6%)	4 (9.5%)	<0.001
Per-rectal bleeding	489 (43.4%)	149 (37.4%)	15 (35.7%)	0.083
Change in bowel habits	271 (24.1%)	95 (23.9%)	7 (16.7%)	0.543
CT imaging findings, n (%)				
Colonic wall thickening	268 (23.8%)	74 (18.6%)	14 (33.3%)	0.026
Fat stranding	142 (12.6%)	34 (8.5%)	7 (16.7%)	0.056
Abscess formation	3 (0.3%)	0 (0.0%)	0 (0.0%)	0.556
Perforation	30 (2.7%)	7 (1.8%)	1 (2.4%)	0.601
Endoscopic findings, n (%)				
Normal	30 (2.7%)	0 (0.0%)	0 (0.0%)	0.003
Diverticulitis	75 (6.7%)	26 (6.5%)	1 (2.4%)	0.544
Diverticulosis	1,020 (90.6%)	352 (88.4%)	34 (81.0%)	0.076

Table 1. Baseline characteristics of the study cohort stratified by endoscopic risk group. SD, standard deviation. Chi-square test for categorical variables; one-way ANOVA for age. CIBH, change in bowel habits; CT, computed tomography; CRC, colorectal cancer; SD, standard deviation.

The study analyzed 1,546 patients with diverticulitis with a mean age of 69.2 years. Based on endoscopic assessment, the cohort was stratified by cancer risk: 1,100 patients (71.2%) were classified as low risk, 404 (26.1%)

as moderate risk, and 42 (2.7%) as high risk (histologically confirmed CRC). For the binary prediction model, the low- and moderate-risk groups were combined (n=1,504) versus the high-risk group (n=42).

The cohort consisted of 796 females (51.5%) and 750 males (48.5%). The most common presenting complaint was PR bleeding, reported in 647 patients (41.8%), followed by abdominal pain in 411 patients (26.6%) and change in bowel habits (CIBH) in 370 patients (23.9%).

Imaging findings (from CT or MRI) revealed colonic wall thickening in 391 patients (25.3%), fat stranding in 201 (13.0%), perforation in 39 (2.5%), and abscess formation in 3 patients (0.2%). Diverticulitis was identified in 101 patients (6.5%), while diverticulosis was highly prevalent, noted in 1,391 patients (90.0%). A total of 42 tumors (2.7%) were identified on endoscopic assessment, including 36 rectal tumors, 4 colonic tumors outside the rectum, and 2 anal tumors. The majority of these 42 confirmed malignancies were adenocarcinomas, consistent with the expected histological distribution of colorectal cancer; granular data on tumour differentiation grade (well, moderate, or poor) were not uniformly available across all cases owing to the retrospective nature of record extraction, which is acknowledged as a limitation. Polyps were detected in 404 patients (26.1%), with tubular adenomas being the most common (262 cases), followed by sessile serrated lesions (63), hyperplastic polyps (45), tubulovillous adenomas (31), leiomyomas (2), and a single inflammatory polyp.

Model Performance

The developed nomogram demonstrated strong discriminative performance. The cross-validated AUC was 0.736 (95% CI 0.661–0.811), with a full-sample AUC of 0.784. At the Youden-optimal threshold of 0.523, sensitivity was 72.1% and specificity was 73.8%.

At this threshold, for the present cohort with a colorectal cancer prevalence of 2.7%, the model yielded a negative predictive value (NPV) of 99.0% and a positive predictive value (PPV) of 7.1%.

Calibration of the model is shown in Figure III. The bootstrap-corrected calibration curve and the distribution of predicted probabilities across the cohort are displayed; interpretation is provided in the Discussion.

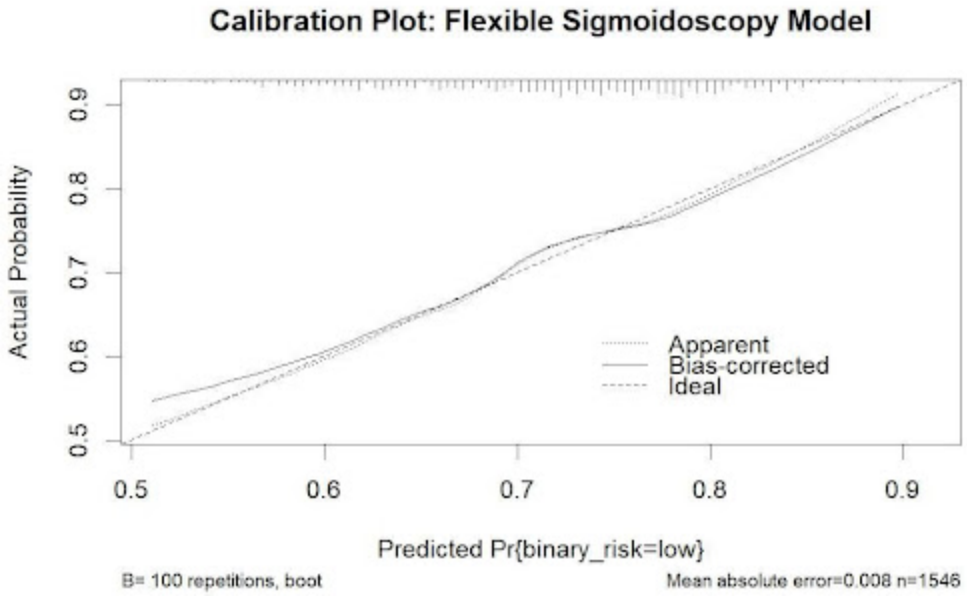


Figure III: Calibration curve of the nomogram for predicting colorectal cancer risk. The bias-corrected curve (black), derived from 100 bootstrap samples, closely follows the ideal line (red), demonstrating that the model's predicted probabilities are accurate and reliable. The histogram shows the distribution of predicted probabilities in the cohort. MAE, mean absolute error.

The ROC analysis showed consistent performance across all validation folds, with the mean ROC curve demonstrating robust discriminative ability and narrow confidence bands around the mean estimate (Fig. IV).

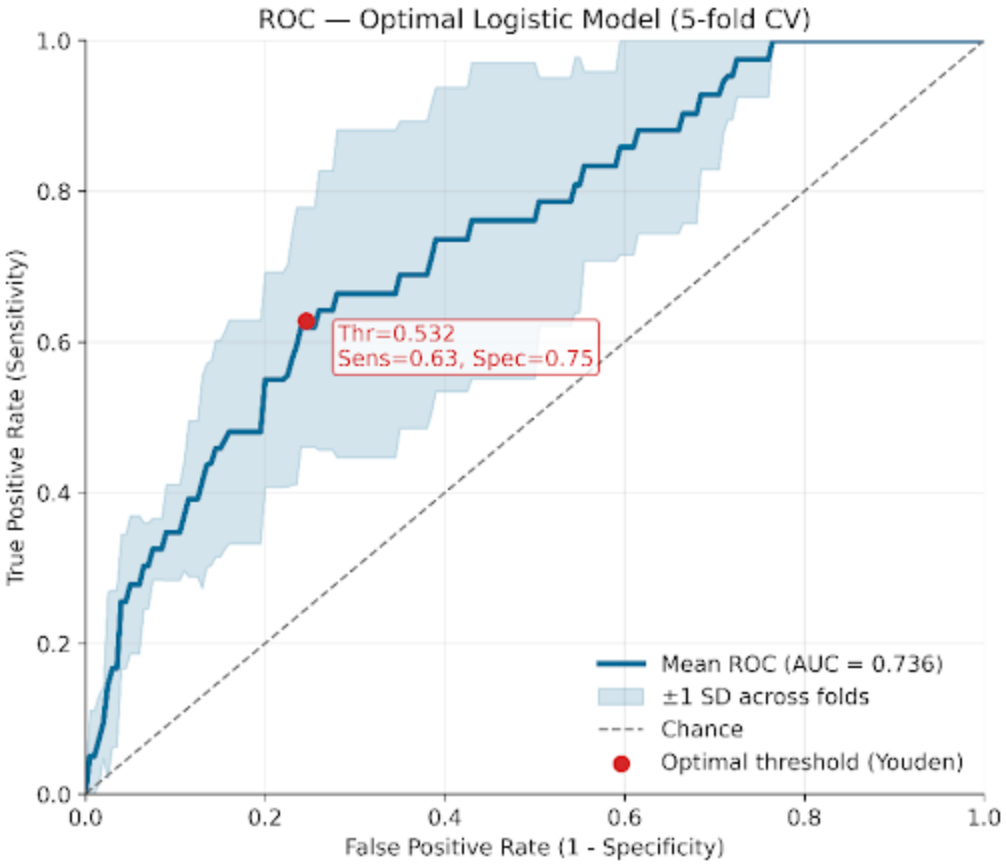


Figure IV: ROC curve for the optimal logistic model (balanced L2 regularization, C=0.1). Mean AUC = 0.736.

The ROC analysis demonstrates the model's strong capacity to distinguish between patients with and without colorectal cancer following an episode of diverticulitis. An AUC of 0.736 reflects a good level of predictive accuracy, offering clinical value as a risk-stratification tool in this challenging diagnostic context, where symptoms and imaging findings frequently overlap between benign inflammation and malignancy.

Multivariable predictor associations with CRC risk, derived from ordered logistic regression across the three risk groups, are presented in Table 2.

Predictor	OR	95% CI	p-value	Direction
Positive predictors				
Age (per SD increase)	1.28	1.14 – 1.45	<0.001	Significant positive predictor
Male sex (vs. female)	1.75	1.39 – 2.21	<0.001	Significant positive predictor
CT wall thickening	2.00	0.98 – 4.12	0.055	Positive trend

Non-significant positive predictors				
Perforation	1.25	0.59 – 2.65	0.564	Not significant
Fat stranding	1.11	0.75 – 1.66	0.598	Not significant
Inverse predictors				
Abdominal pain	0.60	0.44 – 0.80	0.001	Significant inverse predictor
Per-rectal bleeding	0.71	0.56 – 0.89	0.004	Significant inverse predictor
Diverticulosis	0.68	0.47 – 0.97	0.033	Significant inverse predictor
Non-significant inverse predictors				
Change in bowel habits	0.90	0.69 – 1.18	0.458	Not significant
Diverticulitis (endoscopic)	0.89	0.55 – 1.43	0.632	Not significant

Table 2. Multivariable predictor associations with CRC risk (ordered logistic regression). OR, odds ratio; CI, confidence interval; NS, not significant; SD, standard deviation. Ordered logistic regression model fitted to three ordered risk groups (low, moderate, high risk). The OR for age reflects the effect per one standard deviation increase in age. CT wall thickening demonstrated a clinically meaningful OR of 2.00 but did not reach conventional statistical significance (p=0.055), with the 95% CI marginally crossing 1.0.

Predictive Variables

On multivariable ordered logistic regression (Table 2), two variables were identified as significant independent positive predictors of higher CRC risk. Older age was the strongest predictor (OR 1.28 per SD increase, 95% CI 1.14–1.45, p<0.001), with risk increasing consistently across risk categories. Male sex was independently associated with higher CRC risk (OR 1.75, 95% CI 1.39–2.21, p<0.001), consistent with established epidemiological patterns of colorectal cancer [13]. CT-detected colonic wall thickening demonstrated a clinically meaningful association (OR 2.00, 95% CI 0.98–4.12), though this did not reach conventional statistical significance (p=0.055), with the confidence interval marginally crossing 1.0; this finding is consistent with the binary model, in which wall thickening demonstrated approximately two-fold elevated odds for confirmed CRC, and its clinical utility in identifying patients for prioritised endoscopic evaluation is therefore retained. Three variables showed significant inverse associations with CRC risk. Abdominal pain (OR 0.60, 95% CI 0.44–0.80, p=0.001) and per-rectal bleeding (OR 0.71, 95% CI 0.56–0.89, p=0.004) both demonstrated counterintuitive inverse associations; these most likely

reflect selection and referral-pathway bias within this endoscopically-evaluated cohort — patients presenting with overt symptoms are more likely to represent inflammatory rather than malignant pathology — and should not be interpreted causally. Diverticulosis similarly showed a significant inverse association (OR 0.68, 95% CI 0.47–0.97, p=0.033), consistent with a benign phenotype and declining prevalence across higher risk groups (Table 1). Change in bowel habits, perforation, fat stranding, and endoscopic diverticulitis were not significant predictors (all p>0.05, Table 2).

Clinical Decision Tools

The Excel-based calculator implements the validated model, providing automatic standardisation of input variables, real-time probability calculation, and clinical decision recommendations based on the ≥0.523 threshold to guide endoscopic referral. The bedside nomogram enables manual risk calculation using age-specific point assignments stratified by sex and wall thickening status, binary predictor point values for rapid summation, direct probability reading from total point scores, and a clear threshold marker for clinical decision-making. Successful integration of these tools into routine practice will depend on embedding them into electronic health records and clinical pathways. Pilot testing is needed to evaluate usability, clinician uptake, and impact on endoscopy referrals; if proven effective, the nomogram and calculator could be incorporated into decision support dashboards to enhance accessibility and clinical adoption.

Discussion

Our study, to the best of our knowledge, is the first of its kind to present a validated nomogram specifically designed to predict CRC risk in patients with diverticulitis. Our model demonstrates practical utility and robust discriminative performance. This makes it a useful clinical tool that could assist physicians in decision-making regarding the suitability and urgency of endoscopic referral following an episode of diverticulitis.

The model results provide several clinically relevant insights. Consistent with the available literature, age and male sex are fundamental to baseline risk assessment: older male patients with diverticulitis represent the highest-risk group and should be prioritised for endoscopic evaluation [14]. CT wall thickening emerged as a critical discriminating imaging feature; patients with this finding warrant prioritised endoscopic assessment relative to those

with moderate or low radiological burden. The assessment of clinical symptoms remains important for patient evaluation and safety-netting; however, the model demonstrates that the independent predictive value of individual symptoms is limited once baseline demographic and imaging findings are accounted for. This underscores the value of an integrated, multi-variable approach to post-diverticulitis risk stratification rather than reliance on any single clinical feature.

While numerous CRC risk prediction models have been developed for asymptomatic individuals in population-based screening contexts, few, if any, specifically address the post-diverticulitis population. Existing models often rely on demographic and lifestyle factors such as age, sex, BMI, and family history, with some incorporating genetic biomarkers [15]. In contrast, our model integrates diverticulitis-specific clinical and imaging features, such as CT wall thickening and symptom profiles, to stratify CRC risk in a high-risk subgroup. This tailored approach fills a critical gap in current predictive tools. It is important to note that, unlike inflammatory bowel disease, diverticulitis is not considered a pre-malignant condition, and there is no current evidence that diverticular inflammation directly promotes colorectal carcinogenesis. The observed association between diverticulitis and CRC diagnoses most plausibly reflects clinical and radiological overlap — whereby sigmoid malignancy may mimic acute diverticulitis — shared risk factors, and ascertainment effects from endoscopic surveillance, rather than a causal biological relationship.

Our model integrates multiple clinical and radiological variables providing a tool for individualized risk assessment that goes beyond current guideline recommendations. Our model demonstrated a specificity of 73.8%, suggesting the potential to substantially reduce unnecessary endoscopic procedures in appropriately selected low-risk patients. This is strongly supported by the high negative predictive value (NPV) of 99.0%, which suggests the model can reliably identify patients at very low risk of CRC who could potentially forgo endoscopic surveillance. However, these estimates are based on internal validation only and should be interpreted with caution until confirmed in external cohorts. At the same time, the model's high sensitivity of 72.1% ensures high-risk patients are identified and referred for endoscopic assessment. Conversely, the positive predictive value (PPV) was low (7.1%), reflecting the low underlying prevalence of cancer in the cohort; a high-risk prediction from the model should therefore be interpreted as an indicator for prioritization of endoscopic investigation rather than a definitive cancer diagnosis.

The model demonstrated excellent calibration. The bootstrap-corrected calibration curve aligned closely with the ideal diagonal line, indicating minimal overfitting and highly accurate predicted probabilities, particularly in the clinically critical range of 0.6–0.9. The mean absolute error between predicted and observed outcomes was 0.008, confirming the model's reliability for generating trustworthy individual-patient risk estimates. This supports the use of the nomogram for providing meaningful individual risk estimates and enhances its utility in clinical decision-making for post-diverticulitis surveillance.

More importantly, our model represents a novel tool with objective criteria guiding management following diverticulitis. This could help address the inconsistencies in clinical practice suggested by variable post-diverticulitis colonoscopy rates reported in different studies [7, 16].

Our findings, validated by our model, are aligned with the available epidemiological evidence. The population-based studies have suggested an increased risk of CRC in elderly male individuals, consistent with the findings of our study [14]. Similarly, the presence of CT wall thickening, suggesting complicated diverticulitis, and increased risk of CRC incidence has been consistently reported in the literature [16, 17]. CT wall thickening likely reflects more severe inflammatory processes that may either mask underlying malignancy or represent pro-carcinogenic inflammatory environments. The protective associations of certain symptoms may indicate that overtly symptomatic presentations are more likely related to inflammation rather than malignancy [18].

The achieved AUC of 0.736 (cross-validated) represents good discriminative performance for a clinical prediction model, particularly given the complexity of differentiating malignant from inflammatory conditions. This performance compares favorably with other established clinical prediction models in gastroenterology and oncology, while providing superior discriminative ability to simple demographic or symptom-based approaches [19, 20]. The selected threshold of 0.523, while seemingly low, reflects the model's calibration for this specific population and outcome definition. The threshold optimization considered the clinical consequences of both false positives (unnecessary procedures) and false negatives (missed cancers), with the balance favoring sensitivity given the serious implications of delayed cancer diagnosis.

At the Youden-optimal threshold of 0.523, in a theoretical cohort of 1,000 patients with a CRC prevalence of 2.7%, the model would correctly identify approximately 19 of 27 cancers while allowing approximately 715 of 973 non-

cancer patients to forgo immediate endoscopic investigation. Approximately 8 cancers per 1,000 patients would be missed at this threshold. Clinicians may wish to apply a lower decision threshold in settings where the consequences of missed cancer are judged to outweigh procedure burden. Formal decision-curve analysis in an external validation cohort is recommended before clinical deployment.

Furthermore, we acknowledge that flexible sigmoidoscopy is limited in its ability to visualise the proximal colon, where diverticulitis-associated CRC can equally arise. In settings where colonoscopy is the standard post-diverticulitis evaluation — as recommended in many international guidelines — the nomogram should be used to risk-stratify patients for prioritised colonoscopy referral. CT colonography represents an appropriate alternative where full colonoscopy is not feasible.

Limitations and Future Directions

Our study had certain limitations that should be considered. These limitations also highlight opportunities for future research. The number of cancer cases in our cohort was relatively small (n=42), which limited the events-per-variable (EPV) ratio. Although we mitigated the risk of overfitting through regularization, cross-validation, and careful feature selection, the modest event count may still restrict model stability. All patients included in this study underwent endoscopic evaluation as per our institutional surveillance policy, which introduces verification bias: the outcome (CRC) can only be identified in patients who received endoscopy. This reflects a selected population and the model should be interpreted within the context of a post-diverticulitis endoscopic pathway. As a single tertiary centre, the case mix may differ from patients managed in primary and secondary care settings, representing a further source of selection bias. The low proportion of CT-confirmed diverticulitis and the relatively low prevalence of classical acute diverticulitis symptoms suggest this cohort may represent a less severe or more heterogeneous population than typically encountered in acute surgical settings, further limiting generalisability. The distribution of tumour locations, with a predominance of rectal lesions, may reflect coding imprecision in the electronic patient records and should be interpreted accordingly. Data on prior endoscopic history, number of previous diverticulitis episodes, and surgical interventions were not systematically available. The decision threshold of 0.523 is statistically derived and should be considered provisional pending formal decision-curve analysis. Prospective, multi-centre external validation in diverse patient populations is the most critical next step before clinical adoption of this

nomogram. Future studies with larger datasets could also productively explore machine learning-based approaches, which may offer superior discriminative performance in settings with greater event counts than were available in this study.

Conclusion

This study presents the first validated nomogram for predicting CRC risk in patients with diverticulitis, demonstrating robust discriminative performance and practical clinical utility. The tool integrates readily available clinical and imaging variables to provide individualized risk assessment, potentially optimizing resource allocation while ensuring appropriate surveillance for high-risk patients. Beyond risk stratification for flexible sigmoidoscopy, this nomogram has potential utility as a triage instrument to prioritise colonoscopy appointments for patients identified as high risk, thereby optimising endoscopy resource allocation and reducing diagnostic delay in those most likely to harbour malignancy. External validation in independent, multi-centre cohorts is required before this tool can be recommended for routine clinical use. Future research should focus on external validation and implementation strategies to optimize the clinical impact of this predictive tool. The goal remains improving patient outcomes through more precise, individualized approaches to post-diverticulitis cancer surveillance.

References

1. Munie ST, Nalamati SPM. (2018). Epidemiology and pathophysiology of diverticular disease. Clin Colon Rectal Surg.31(4):209–13.
2. Tursi A, Scarpignato C, Strate LL, et al. (2020). Colonic diverticular disease. Nat Rev Dis Primers. 6(1):20. doi:10.1038/s41572-020-0153-5
3. Mohamedahmed AY, Zaman S, Das N, et al. (2024). Systematic review and meta-analysis of the management of acute uncomplicated diverticulitis. Int J Colorectal Dis. 39(1):47. doi:10.1007/s00384-024-04618-7
4. Ou G, Rosenfeld G, Brown J, et al. (2015). Colonoscopy after CT-diagnosed acute diverticulitis: is it really necessary? Can J Surg. 58(4):226–31. doi:10.1503/cjs.014514
5. Mortensen LQ, Andresen K, Thygesen L, et al. (2024). Diverticulitis is associated with increased risk of colon cancer. J Clin Med. 13(9):2503. doi:10.3390/jcm13092503

6. Meyer J, Orci LA, Combescure C, et al. (2019). Risk of colorectal cancer in patients with acute diverticulitis. Clin Gastroenterol Hepatol. 17(8):1448–56.e17. doi:10.1016/j.cgh.2018.07.031
7. Issa MT, Sultana E, Hamid M, et al. (2025). DIVERT-Ca: unveiling the hidden link between acute diverticulitis and colorectal cancer risk. Int J Colorectal Dis. 40(1):68. doi:10.1007/s00384-025-04858-1
8. Zhang Y, Zhang H, Zhu J, et al. (2023). Association between diverticular disease and colorectal cancer. BMC Cancer. 23(1):137. doi:10.1186/s12885-023-10606-x
9. Schafmayer C, Harrison JW, Buch S, et al. (2019). Genome-wide association analysis of diverticular disease. Gut. 2019;68(5):854–65. doi:10.1136/gutjnl-2018-317619
10. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer. Colorectal cancer. Accessed October 5, 2025
11. Dean HF, Britton E, Farrow E, et al. (2023). Can endoscopic follow-up after acute diverticulitis be rationalised? Surg Endosc. 37(7):5114–20. doi:10.1007/s00464-023-09997-6
12. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. (2015). Nomograms in oncology. Lancet Oncol. 16(4):e173–80. doi:10.1016/S1470-2045(14)71116-7
13. R Core Team. R. https://www.R-project.org/. A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2025. Accessed October 5, 2025
14. Cooper GS, Xu F, Schluchter MD, et al. (2014). Diverticulosis and the risk of interval colorectal cancer. Dig Dis Sci. 59(11):2765–72. doi:10.1007/s10620-014-3246-8
15. Usher-Smith JA, Walter FM, Emery JD, et al. (2016). Risk prediction models for colorectal cancer: a systematic review. Cancer Prev Res (Phila). 9(1):13–26. doi:10.1158/1940-6207.CAPR-15-0274
split
16. Cao AMY, Lam VW, Rickard MJFX. (2023). Endoscopic findings after CT proven acute diverticulitis. ANZ J Surg. 93(5):1150–8. doi:10.1111/ans.18190
17. Mangat S, Kozoriz MG, Bicknell S, et al. (2018). Accuracy of colorectal cancer detection by CT in the unprepared bowel. Can Assoc Radiol J. 69(1):92–6. doi: 10.1016/j.carj.2017.09.005
18. Pothoulakis I, Wikholm C, Verma V, et al. (2022). Predictors of colorectal carcinoma and IBD in patients with colonic wall thickening. JGH Open. 6(3):159–65. doi:10.1002/jgh3.12708

19. Hueting TA, van Maaren MC, Hendriks MP, et al. (2023). External validation of 87 clinical prediction models. Breast. 69:382–91. doi:10.1016/j.breast.2023.04.003

20. Collins GS, Dhiman P, Ma J, et al. (2024)/ Evaluation of clinical prediction models (part 1). BMJ. 384:e074819. doi:10.1136/bmj-2023-074819

Supplementary material 1

User-friendly Excel calculator (S1) implementing the validated logistic regression model to provide real-time risk assessment and clinical guidance. This calculator is designed to provide clinicians with a rapid, user-friendly tool for assessing CRC risk in patients following an episode of diverticulitis. To use the calculator, the clinician inputs the patient's age, sex, and binary values for relevant clinical and radiological features (e.g., presence of abdominal pain, per-rectal bleeding, CT wall thickening, diverticulosis). Based on the validated logistic regression model coefficients, the calculator automatically computes the patient's probability of having CRC. If the calculated risk exceeds the predefined threshold (≥0.523), the tool generates a recommendation for flexible sigmoidoscopy (FS) referral, thereby supporting timely and evidence-based clinical decision-making.