Ther Adv Gastroenterol
2021, Vol. 14: 1–12
DOI: 10.1177/17562848211030050

Bowel ultrasound enhances predictive value based on clinical indicators: a scoring system for moderate-to-severe endoscopic activities in patients with ulcerative colitis

Mengmeng Zhang, Huimin Zhang, Qingli Zhu, Xiaoyin Bai, Qingyang Zhou, Gechong Ruan, Wenbo Li, Li Ma, Mengsu Xiao, Hong Yang and Jiaming Qian

Abstract

Background and Aim: The aim was to assess non-invasive factors among clinical features, laboratory, and bowel ultrasound (BUS) characteristics and to develop a scoring system to predict endoscopic activities for ulcerative colitis (UC) patients.

Methods: We performed a retrospective study collecting UC patients between January 2015 to September 2020. Logistic regression was performed to predict moderate-to-severe endoscopic activities, defined as endoscopic Mayo score ⩾2. Model performance was described with discrimination and calibration ability and validated by internal and external methods.

Results: A total of 103 and 29 patients were enrolled in the modeling and validation groups, respectively. Stool frequency ⩾5 times/day, hematochezia, erythrocyte sedimentation rate (ESR), and colonic wall flow in BUS were included into two predictive models for endoscopic activities, both with good discrimination ability [Area under curve (AUC) 0.879 and 0.882, $p < 0.001$] and a sensitivity of 76.7% and specificity of 92.3%, which showed an adequate calibration ability by using the Hosmer–Lemeshow test ($p = 0.14$ and 0.07). The external validation displayed consistent results with the above mentioned. Nomograms were also established for these models.

Conclusion: We developed predictive models for endoscopic disease activities by using noninvasive factors based on stool frequency, hematochezia, ESR, and colonic wall blood flow in BUS. These models performed well in the internal and external validation.

Keywords: bowel ultrasound, endoscopic activities, non-invasive technique, ulcerative colitis

Received: 8 March 2021; revised manuscript accepted: 14 June 2021.

Correspondence to:
Hong Yang
Department of Gastroenterology, Peking Union Medical College Hospital, No.1 Shuaifuyuan, Wangfujing Street, Dongcheng District, Beijing, 100730, China
hongy72@163.com

Jiaming Qian
Department of Gastroenterology, Peking Union Medical College Hospital, No.1 Shuaifuyuan, Wangfujing Street, Dongcheng District, Beijing, 100730, China
qianjiaming1957@126.com

Qingli Zhu
Department of Ultrasound, Peking Union Medical College Hospital, No.1 Shuaifuyuan, Wangfujing Street, Dongcheng District, Beijing, 100730, China
zqlpumch@126.com

Mengmeng Zhang
Huimin Zhang
Xiaoyin Bai
Qingyang Zhou
Gechong Ruan
Department of Gastroenterology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Wenbo Li
Li Ma
Mengsu Xiao
Department of Ultrasound, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Introduction

Ulcerative colitis (UC) is a chronic, idiopathic inflammatory disease, characterized by a relapsing and remitting course that mainly causes colonic continuous mucosal inflammation.[1] Endoscopy is always regarded as the most direct and reliable measurement for disease evaluation. Due to the relatively high risk of complications, low practicality, and compliance of patients, the application of colonoscopy or sigmoidoscopy in severe inflammatory bowel disease (IBD) patients, or during long-term follow-up, is limited.[2] Previous studies have been conducted to improve the reliability and utility of non-invasive assessment based on disease-specific symptoms or biochemical indicators. However, some studies suggested that these clinical indicators may not be consistent with actual mucosal conditions, sometimes resulting in an underestimation of disease severity, treatment delay, or early relapse.[3] The *post hoc* analysis of the ULTRA trials reported discrepancies between rectal bleeding and stool frequency scores and

endoscopic activity in UC and less predictive power of mucosal healing status.[4] Falvey *et al.*[5] revealed that combining simple clinical colitis activity index (SCCAI) data with C reactive protein (CRP) or fecal calprotectin (FC) levels in UC did not significantly improve the accuracy of any single indicator.

Bowel ultrasound (BUS), as a relatively new, applied procedure in clinical work, has certain priority due to its non-invasive properties, its high tolerance, and its cost-effectiveness compared with colonoscopy in the disease assessment for IBD. Although several studies have reported a correlation between BUS and the endoscopic severity of UC, relying on BUS alone to predict mucosal pathologies is inevitably inaccurate and has relatively low sensitivity (71%).[6,7] As a result, coupled with the advancement of intestinal ultrasound, the integration of clinical, biochemical, and BUS measurement, offers a better alternative for disease evaluation and surveillance and is worth further exploration.

Therefore, the aim of our study was to establish a scoring model based on these clinical, biochemical, and BUS indicators to predict moderately to severely active endoscopic activities, through analysis of the clinical data of UC patients with different degrees of endoscopic mucosal severity, to ultimately increase the utilization of BUS and reduce any potential risk of repeated, invasive examinations in UC patients irrespective of whether they are clinically active.

Materials and methods

Patients

The flow chart of patient inclusion is shown in Supplemental Figure 1. Consecutive patients with UC hospitalized in Peking Union Medical College Hospital (PUMCH) between January 2015 to December 2019, defined by the third European Crohn's and Colitis Organization (ECCO) consensus guideline,[1] were retrospectively collected. UC patients underwent colonoscopy for evaluation of disease activity or surveillance, according to the statement of endoscopic assessment or monitoring in the ECCO and the European Society of Gastrointestinal and Abdominal Radiology (ECCO-ESGAR) guidelines,[8] and BUS for further evaluation of disease extent and activity. Patients who had taken BUS, colonoscopy, and biochemical measurement within 1 month were finally included in our study. The endoscopists and radiologists were not blinded for clinical manifestations, but did not know BUS results when performing endoscopy. Patients aged below 16 years or above 75 years, with incomplete medical data, or who had a change in symptoms or therapy between BUS and colonoscopy, were excluded.

In addition, cases hospitalized in PUMCH from January 2020 to September 2020, who complied with the above criteria for diagnosis and examinations, were included as the validation group. The study was approved by the Ethical Committee of PUMCH.

Data collection

The following data were collected, mainly from medical records during patient hospitalization:

Clinical indicators. Age at admission, age of onset, disease duration, Montreal classification, and body mass index (BMI) at the same period as colonoscopy was performed. Symptoms including stool frequency and blood in stool were obtained from medical records. Stool frequency was recorded as defecation frequency per day. The degree of hematochezia was evaluated as three levels: no blood, bloody stool below 50%, and bloody stool over 50%. Symptoms between BUS and colonoscopy examinations were checked to make sure no significant changes occurred (stool frequency changed $\leqslant$2 times/day and the same degree of hematochezia). Symptoms just before BUS and colonoscopy were eventually collected.

Laboratory parameters. Laboratory results including hemoglobin (Hb), platelet (PLT), albumin (Alb), high-sensitivity C reactive protein (hsCRP), and erythrocyte sedimentation rate (ESR) were collected from the closest records before colonoscopy, with a maximum window of 1 week.

Bowel ultrasound examinations and parameters. BUS examinations were performed by two experienced radiologists (23 and 11 years of experience, respectively) using Philips iU22 (Philips, Bothell, WA, USA) with convex (C5–2) and linear (L9–3) transducers following the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines.[9] Patients

fasted for at least 8 hours before BUS examination. The colon was scanned continuously with bowel wall thickness measured at the ileocecal area, ascending colon, transverse colon, descending colon, and sigmoid colon. When the abnormal bowel section was identified, the radiologist further assessed features including maximum bowel wall thickness and blood flow at the diseased location. The following two parameters were collected: bowel wall thickness and bowel wall blood flow. The thickest measurement data of all colonic segments were taken. Colonic wall was evaluated by Limberg classification[10] based on BUS reports: Limberg 0: normal bowel wall; Limberg 1: bowel wall thickening (>0.3 cm); Limberg 2: bowel wall thickening and short vessels; Limberg 3: bowel wall thickening and long vessels; Limberg 4: bowel wall thickening and long vessels compromising the mesenterium. In addition, the worst Limberg classification of all segments was taken into account.

Endoscopic measurement. Colonoscopies were performed according to standard endoscopic procedures in ECCO-ESGAR guidelines.[11] Endoscopic disease activity was evaluated using Mayo endoscopic score by endoscopists immediately after finishing colonoscopy. A Mayo endoscopic subscore of 1, 2, or 3 was referred as mild, moderate, or severe disease activity, respectively. The worst Mayo endoscopic score of all segments and disease extent were collected.

Statistical analysis

Continuous variables are expressed as mean ± standard deviation (SD) if they obey normal distribution; otherwise, the median and interquartile ranges (IQRs) are expressed. These values were compared with the bilateral *t*-test or Wilcoxon–Mann–Whitney test. Categorical data were presented as frequencies (percentages) and compared with chi-square test or Fisher exact tests. Logistic regression analysis was performed to assess the association between non-invasive indicators with the possibly of moderate-to-severe endoscopic activity, defined as an endoscopic Mayo score ⩾2. Variables with $p < 0.10$ at univariate analysis were included in multivariate analysis using a 'backwards elimination procedure'. To make a significant and reasonable model, we first divided the cutoff value for some numerical variables to achieve the best distinction between moderate-to-severe endoscopic activities and quiescent or mild disease. Through each step displayed in multivariate regression analysis, we regarded these variables in three parts, which were symptoms (stool frequency and hematochezia), hematological or biochemical indicators (Hb, PLT, Alb, hsCRP, and ESR), and BUS indicators (colonic wall thickness and blood flow). In each aspect, the most significant indicators were selected (both stool frequency and hematochezia were similarly significant and included) and recombined to two scoring models (all included variables satisfied $p < 0.10$).

The mucosal active predictive score was built according to the coefficient estimates in the multivariate regression analysis. The evaluation of predictive performance was featured by the discrimination and calibration ability: discrimination was judged by the *c*-statistic, which is the area under the receiver operating characteristic (ROC) curve; and calibration was assessed by Hosmer and Lemeshow tests and presented with a calibration curve to reflect the agreement between predicted and observed probabilities. The model validation contained internal validation in the modeling group and external validation in the validation group. Internal validation was conducted through the bootstrap method, with 1000 replicates to calculate the discrimination and calibration performance. In the validation group, the endoscopic activities were assessed according to the developed model. The *p-values* < 0.05 were considered statistically significant. Statistical analysis was performed by using SPSS (version 25.0, IBM Corporation, Chicago, USA) and R (version 3.6.2, R Foundation for Statistical Computing, Vienna, Austria). The reporting of this study adheres to the TRIPOD statement for reporting of prediction models.

Results

Demographic characteristics and clinical outcomes

According to the inclusion criteria, 103 UC patients were included in the modeling group and another 29 patients in the validation group. Demographic information and clinical characteristics are presented in Table 1. There was no significant difference between the modeling group and the validation group in Montreal classification,

Table 1. Demographic and clinical characteristics of the enrolled UC patients.

Variables	Modeling group ($n = 103$)	Validation group ($n = 29$)	p-value
Sex (Male)	55 (53.4%)	12 (41.4%)	0.253
Age (years)	41.39 ± 14.70	38.55 ± 12.80	0.347
Age at onset (years)	32 (25–45)	33.34 (22.0–43.5)	0.654
Duration (months)	48 (12–96)	25.0 (9.5–78.0)	0.310
Clinical type (%)			
Primary	13 (12.6)	3 (10.3)	0.992
Chronic recurrent	90 (87.4)	27 (89.7)	
Disease Extent (%)			
Left-sided	10 (9.7)	4 (13.8)	0.772
Extensive	92 (89.3)	25 (86.2)	
BMI (kg/m^2)	20.52 ± 3.48	20.22 ± 3.52	0.717
Stool frequency (⩾5 times/day) (%)	71 (68.9)	15 (51.7)	0.086
Hematochezia (%)			
None	22 (21.4)	10 (34.5)	0.100
<50%	30 (29.1)	11 (37.9)	
⩾50%	51 (49.5)	8 (27.6)	
Hb (g/l)	108.96 ± 27.82	111.96 ± 27.30	0.612
PLT ($\times 10^9$/l)	341.37 ± 131.67	327.75 ± 87.96	0.607
Alb (g/l)	33.46 ± 7.83	36.79 ± 7.89	0.048
hsCRP (mg/l)	16.5 (4.15–55.18)	5.62 (0.96–29.40)	0.027
ESR (mm/h)	27 (13–49)	18.0 (8.5–39.0)	0.160
Endoscopic Mayo (%)			
Score 0–1	13 (12.6)	5 (17.2)	0.456
Score 2–3	90 (87.4)	24 (82.8)	
Colonic wall thickness in BUS (cm)	0.70 (0.60–0.90)	0.65 (0.50–0.85)	0.107
Colonic wall flow in BUS (%)			
Limberg level 0	9 (8.7)	3 (10.3)	0.891
Limberg level 1	4 (3.9)	1 (3.4)	
Limberg level 2	14 (13.6)	2 (6.9)	
Limberg level 3	47 (45.6)	14 (48.3)	
Limberg level 4	29 (28.2)	9 (31.0)	
Adverse outcomes (%)	24 (23.3)	5 (17.2)	0.486

Continuous variables were expressed as mean ± standard deviation or median [interquartile range (IQR)].
Categorical variables were expressed as frequencies (percentages).
Alb, albumin; BMI, body mass index; BUS, bowel ultrasound; ESR, erythrocyte sedimentation rate; Hb, hemoglobin; hsCRP, high-sensitivity C-reactive protein; PLT, platelet; UC, ulcerative colitis.

Table 2. Univariate analysis on variables associated with possibility of moderate-to-severe endoscopic activities (Mayo endoscopic score ⩾2) in patients with UC.

Variables	*p*-value	OR	95% CI
Age (years)	0.685	1.008	0.969–1.050
Duration (months)	0.126	0.997	0.992–1.001
BMI (kg/m^2)	0.869	1.016	0.845–1.220
Stool frequency (⩾5 times/day)	<0.001	13.327	3.573–49.703
Hematochezia	0.003	3.324	1.501–7.363
Hb (g/l)	0.003	0.957	0.930–0.985
PLT ($\times 10^9$/l)	0.009	1.009	1.002–1.006
Alb (g/l)	<0.001	0.825	0.741–0.919
hsCRP (mg/l)	0.029	1.064	1.006–1.126
ESR (mm/h)	0.007	1.086	1.023–1.153
Colonic wall thickness in BUS (cm)	0.077	11.640	0.765–177.036
Colonic wall flow in BUS*	<0.001	2.492	1.538–4.039

*Defined by Limberg classification.
Alb, albumin; BMI, body mass index; BUS, bowel ultrasound; CI, confidence interval; ESR, erythrocyte sedimentation rate; Hb, hemoglobin; hsCRP, high-sensitivity C-reactive protein; OR, odd's ratio; PLT, platelet; UC, ulcerative colitis.

clinical manifestations, biochemical indicators, endoscopic Mayo score, and BUS results. In each group, patients with high disease activity accounted for a larger proportion than those with inactive UC from the percentages of diarrhea, hematochezia, and endoscopic Mayo 2–3, possibly because patients with active UC were more likely to take examinations before therapy adjustment. The most severe segments assessed by colonoscopy and BUS were consistent, except for one case who presented with mucosal healing with a previous history of proctitis in colonoscopy and a maximum bowel wall thickness (0.7 cm) in the descending colon without blood flow in BUS.

Analysis for common-used parameters predicting endoscopic activities

We chose four indicators widely applied in clinical practice: stool frequency, hematochezia, hsCRP, and ESR. For stool frequency, hsCRP, and ESR, we set the cut-off value under the maximum of sensitivity and specificity in ROC analysis. As shown in Supplemental Table 1 and Supplemental Figure 2, clinical symptoms had a sensitivity of 85.6% and a specificity of 69.2%; in contrast, the inflammatory parameters had a slightly low sensitivity of around 70% and a specificity of 84.6%.

Establishment of a scoring system predicting endoscopic activities

By univariate logistic regression analysis, stool frequency, hematochezia, Hb, PLT, Alb, hsCRP, ESR, and colonic wall flow in BUS were significantly associated with endoscopic activities (Table 2). The above-mentioned variables were evaluated in a multivariable logistic regression, turning out two different combinations involving stool frequency, hematochezia, ESR, and colonic wall flow in BUS [Tables 3(a) and (b)]. According to the estimates of regression coefficients, corresponding values were assigned to these 4 variables (Supplemental Table 2). Afterwards, the predictive models for endoscopic activities were established. The Model A total score = 0.8 × hematochezia score + 1.6 × ESR score + 0.6 × Limberg score. The Model B total score = 1.6 × stool frequency score + 1.5 × ESR score + 0.5 × Limberg score. The nomograms were developed to easily obtain

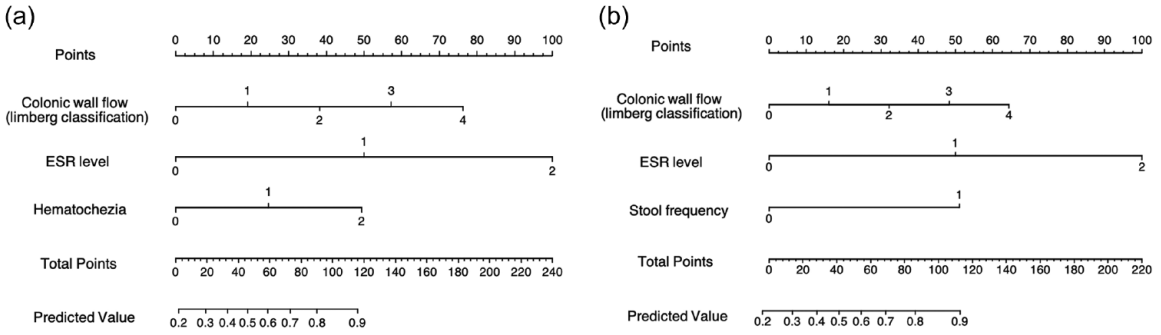

Figure 1. Nomograms of established models for the prediction of moderate-to-severe mucosal activity.
(a) The nomogram based on hematochezia, ESR, and colonic wall flow.
Hematochezia: score 0-none; score 1-bloody stool accounts for less than 50%; score 2- bloody stool accounts for more than 50%.
ESR: score 0-less than 15 millimeters per hour; score 1-range from 15 to 30 millimeters per hour; score 2-above 30 millimeters per hour.
Colonic wall flow: score 0–4 represent Limberg classification level 0 to 4, respectively.
(b) The nomogram based on stool frequency, ESR, and colonic wall flow.
Stool frequency: score 0-less than five times per day; score 1-five times or more per day.
ESR: score 0-less than 15 millimeters per hour; score 1-range from 15 to 30 millimeters per hour; score 2-above 30 millimeters per hour.
Colonic wall flow: score 0–4 represent Limberg classification level 0 to 4, respectively.

Table 3(a) and 3(b). Multivariable analysis on the possibility of moderate-to-severe endoscopic activities (Mayo endoscopic score ⩾2) in patients with UC.

(a) Variables	***p*-value**	**OR**	**95% CI**
Hematochezia	0.087	2.172	0.895–5.273
ESR (mm/h)	0.007	4.819	1.534–15.133
Colonic wall flow in BUS*	0.041	1.822	1.024–3.244
(b) Variables	***p*-value**	**OR**	**95% CI**
Stool frequency (⩾5 times/day)	0.055	4.818	0.969–23.949
ESR (mm/h)	0.011	4.661	1.433–15.159
Colonic wall flow in BUS*	0.084	1.641	0.935–2.879

*Defined by Limberg classification.
BUS, bowel ultrasound; CI, confidence interval; ESR, erythrocyte sedimentation rate; OR, odd's ratio; UC, ulcerative colitis.

the predictive probability of mucosal severity based on model A and B respectively (Figure 1).

Model evaluation and internal validation

For both models predicting endoscopic activities, ROC curves confirmed a good discrimination ability for endoscopic activities, with an area under curve (AUC) of 0.879 [95% confidence interval (CI) 0.758–1.000, $p < 0.001$] and 0.882 (95% CI 0.773–0.992, $p < 0.001$), respectively [Figure 2(a) and (b)] and AUC 0.953 (95% CI 0.901–1.000) and 0.847 (95% CI 0.712–0.983) from the bootstrap method. In Supplemental Table 3, evaluation for accuracy of model A and B with different cut-offs is presented. In model A, a total score ⩾3.70 was set as a threshold that distinguishs patients with moderate-to-severe endoscopic activities with those in mucosal remission under which sensitivity and specificity were 76.7% and 92.3%. In addition, a total score ⩾3.35 was regarded as a threshold in model B under which sensitivity and specificity were 76.7% and 92.3% as well. According to our

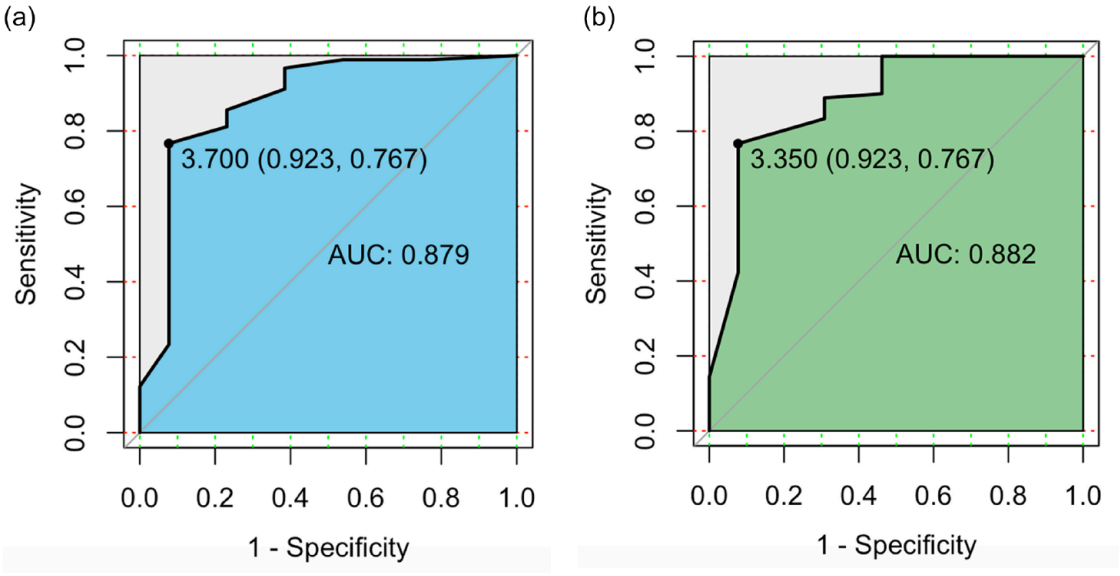

Figure 2. Receiver operating characteristic (ROC) curves for two predicted models for possibility of moderate-to-severe mucosal activity in patients with UC.
(a) ROC curve of model A based on hematochezia, ESR, and colonic wall flow.
Model A showed a good discrimination with area under curve (AUC) of 0.879 [95% confidence interval (CI) 0.758–1.000, $p < 0.001$]. The sensitivity and specificity under the cut-off value of 3.70 performed were 76.7% and 92.3%, respectively.
(b) ROC curve of model B based on stool frequency, ESR, and colonic wall flow.
Model B displayed a similarly good discrimination with an AUC of 0.882 (95% CI 0.773–0.992, $p < 0.001$). The sensitivity and specificity under the cut-off value of 3.35 performed were 76.7% and 92.3%, respectively.

established two models, the scores and endoscopic Mayo scores of 103 UC patients are displayed in Supplemental Figure 3.

With regards to calibration, these two predictive models showed a good fit using the Hosmer–Lemeshow test ($p = 0.14$ in model A, $p = 0.07$ in model B). The calibration curves manifested an acceptable level of agreement between predicted possibility and actual proportions of moderate-to-severe endoscopic activities [Figure 3(a) and (b)]. The difference in discrimination slope based on two predictive models also validated the fine calibration power (shown in Supplemental Figure 4).

External validation of predictive models for endoscopic activities

We applied the developed models to the validation group, mainly to confirm the probability of predicting endoscopic activities for external validation. With model A, ROC analysis displayed with a c-statistic of 0.979; in contrast, the discrimination ability of model B was relatively excellent at 1.0. From calibration analysis, both models had a relatively consistent tendency with the ideal condition and showed a slight possible underestimation at high risk of endoscopic activities (Supplemental Figure 5).

Discussion

This study established predictive models for endoscopic disease activities through the combination of clinical, biochemical, and BUS indicators. Four predictors of mucosal inflammation for UC have been identified: stool frequency above 5 times per day, hematochezia, ESR level, and colonic wall flow based on Limberg classification. Two models including these variables were developed. These models can make up for the deviation caused by a single index to evaluate mucosal severity and achieve non-invasive feature and high practicability in clinical work.

Based on wide application in clinical practice, we first collected data regarding stool frequency, hematochezia, hsCRP, and ESR levels in 103 UC patients and confirmed the inconsistency between these and endoscopic Mayo score. Irrespective of clinical or biochemical parameters, there are not any ideal predictive indicators for mucosal condition. Many

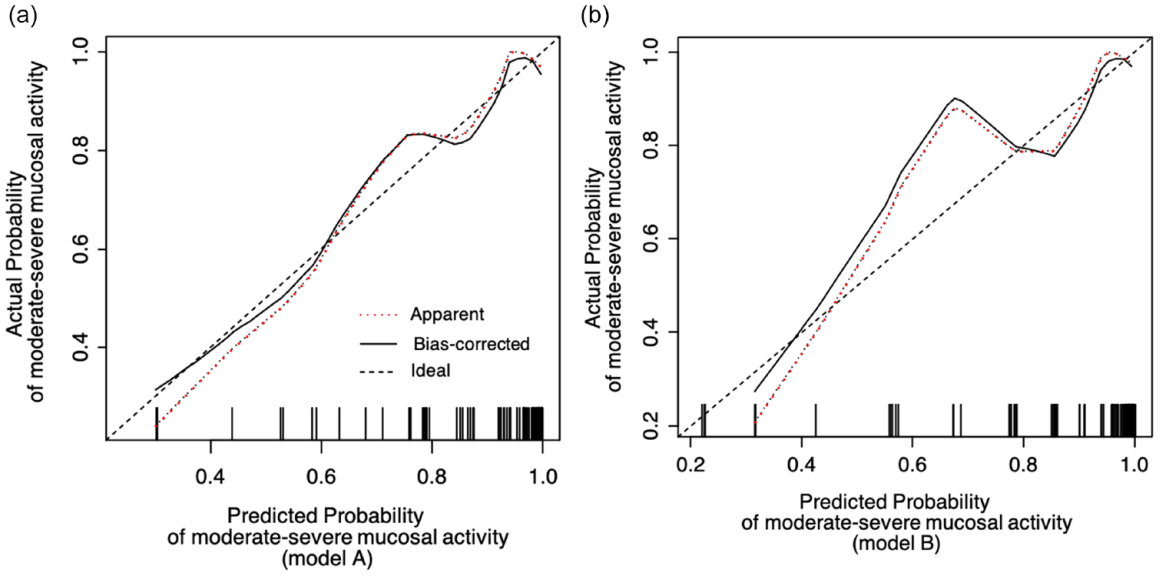

Figure 3. Calibration curves of two models for the prediction of moderate-to-severe mucosal activity in the modeling group.
(a) Calibration curve of model A based on hematochezia, ESR, and colonic wall flow.
(b) Calibration curve of model B based on stool frequency, ESR, and colonic wall flow.
The black dot line on the diagonal of the figure indicated a complete fitting between predictive model and actual data. The red dot line and black solid line illustrated the degree of fitness between model prediction and actual probability of moderate-to-severe mucosal activity.

studies have confirmed the inaccuracy of disease assessment simply through clinical symptoms.[3] Falvey *et al.*[5] found poor accuracy of the Harvey–Bradshaw index (containing general well-being, abdominal pain, abdominal mass, diarrhea, and complications) for the identification of endoscopically active Crohn's disease (CD) and limited value of SCCAI to detect endoscopic active UC. Approximately 50% of patients in clinical remission still had endoscopic active IBD,[12] which was consistent with the low specificity in our study. As a result, endoscopic remission is currently regarded as the treatment goal, rather than clinical remission. In our model, stool frequency and hematochezia were respectively included. Although the discordance between patient's subjective symptoms and endoscopic activities was reported,[13] it is obvious that stool abnormality is still the most intuitive and simple index representing disease activity for UC characterized by diffuse colonic mucosal lesions. Based on clinical applications, we believe that it is an acceptable choice to separate stool frequency and hematochezia into the model respectively with the proper *p*-value. The cut-off value of these two items were appropriately simplified in our model; however, our results are broadly consistent with previous conclusions.

Regarding inflammatory parameters, there were also limitations when we judged disease activity by inflammatory indicators alone. Our results suggested low sensitivity, reflective of similar results in a meta-analysis, with a poor sensitivity [0.49 (95% CI 0.34–0.64)] and a relatively high specificity [0.92 (95% CI 0.72–0.98)].[14] Moreover, there were quite a few patients with persistent active disease and CRP or ESR normalization. In our regression analysis, ESR had significant association with endoscopic activities instead of CRP. There was some agreement on the low correlation between CRP and endoscopic activity.[15–18] A *post hoc* analysis of a prospective clinical trial from Germany showed that the AUC for detecting mucosal healing through CRP was 0.65, with a sensitivity of 45.5%.[19] In addition, there is no optimal threshold of CRP to accurately discriminate active and quiescent conditions.[5] Shin *et al.*[20] found that a normal CRP level (<0.3 mg/dl) also appeared in active UC patients (sensitivity 27.3%). From our results, we speculate that active UC is often accompanied by hematochezia or invisible blood loss from the digestive tract, causing anemia and further aggravation of ESR changes. Furthermore, despite no significant collinearity in our regression analysis, CRP and ESR usually share the roughly same changes,

which leads to the abandonment of the relatively insignificant variable in the process of establishing the composite score, just like Truelove and Witts criteria with simply ESR, Disease Activity Score 28 (DAS-28) based on ESR or CRP respectively in assessments for rheumatoid arthritis.[21,22]

In terms of BUS, bowel wall thickness and blood flow were most frequently applied criteria for evaluating disease location and activity.[23,24] Until now, many studies have proved that BUS has a better advantage for assessing the transmural inflammation and stenotic or penetrated lesion of CD.[25,26] As for UC, a prospective study by Allocca[6] reported that the presence of colonic wall flow and thickening above 3 mm, or just colonic wall thickening above 4.43 mm alone, showed high accuracy for detecting disease activity with great sensitivity (0.71) and specificity (1.00). In addition, there was a significant correlation between endoscopic disease activity and increased bowel wall thickness from the TRUST&UC study[27] and moderate correlation with clinical manifestations and histological grade.[7,28] In our study, we found a significant difference in colonic wall thickness between endoscopic quiescent and active groups [0.60 (0.5–0.75) *versus* 0.80 (0.6–0.93) cm, $p = 0.02$], but this failed to be involved in the final model. This suggests that bowel wall thickness can indeed reflect lesion activity in UC patients. In addition to different inflammatory degree, there may be some other factors influencing bowel wall thickness, including disease duration, previous inflammatory burden, bowel edema, and fibrotic degree. Our enrolled UC patients showed more percentages of extensive colitis, severe phenotype, and longer disease duration, perhaps leading to the higher baseline of colonic wall thickness [0.70 (0.60–0.90) cm]. In addition, our study selected the worst segment when BUS was performed, which may have an effect on the results of bowel wall thickness as well. Therefore, the analysis including all bowel segments by BUS is required to finish in the future to improve the scoring model.

From this study, we hope these models are of great practical value for UC assessment in clinical work. First of all, although it is feasible to predict UC mucosal ability through non-invasive indicators, colonoscopy monitoring is inevitable. We believe that a combination of these non-invasive models and colonoscopy can be achieved. It is recommended to complete all-round and systematic assessment including BUS and colonoscopy at the initial diagnostic period or some key points in order to comparison in entire follow-up course, whereas BUS can be regard as more common routine measurement. Meanwhile, the interval between colonoscopy examinations can be gradually lengthened, especially for long-term quiescent patients with UC. For severely active patients, the risk and tolerance for colonoscopy should be carefully assessed. We can consider BUS first as a preliminary means for evaluation. Then, after the overall condition stabilizes and is improved, colonoscopy can be arranged to accurately assess the mucosal condition and histological examination. Secondly, models of the prediction for endoscopic activities in this study have relatively high specificity instead of sensitivity, and are slightly underestimated, approximately 10%, when predictive value is above 60%. Therefore, actual mucosal damage of the patient with a high predictive possibility may be more severe than expected that need more intervention. As for patients with low predictive value, there is still the possibility of active disease and treatment should not be delayed because of this, and other evaluating indicators or close follow-up observation are required. Thirdly, as utilization of biologics becomes popular, we believe our model has great prospects and developmental potential for patients treated with biological agents. Its good ability of discrimination and calibration can effectively reflect changes in mucosal lesions before and after treatment. Considering the controversial risk of histological fibrosis after long-term anti-tumor necrotic factor therapy,[29] BUS has unique advantages in evaluating bowel wall structure of the entire layers such as edema, fibrotic degree, and blood flow, etc.

This study has several strengths. Firstly, unlike the currently available assessing scores, we establish a multifaceted model including two easily accessible items in BUS innovatively besides clinical and biochemical aspects. In addition, we use a combination of internal and external validation to verify the accuracy and reliability of the developed models, which showed satisfied consistency between the predictive possibility and actual results. Furthermore, the scoring system developed in this study is easy to use because all identified items are commonly used, uncomplicated and also easily available in clinical work. The development of the nomogram also has strong

utility to allow for incorporation of our results into clinical practice.

There were still some limitations in our study. Firstly, the limited sample size of this study may have an unforeseeable effect on the establishment and validation of the model. We hope that with the popularity of BUS taken in UC patients, more cases can be included to improve the model and validation. Secondly, this was a single-center study, which only included patients in another time period as the external validation group, which actually only meets the "time validation". It will be more convincing if validated in data from different medical centers. Thirdly, as a retrospective study, we included UC inpatients, most of whom were moderately to severely active. As a result, our models have a relative advantage on the prediction of moderate-to-severe endoscopic activities. As for identification of mucosal healing, more quiescent subjects or prospective studies are required in the future. Furthermore, the lack of evaluation for consistency of different radiologists or endoscopists and non-blind to clinical symptoms may cause some biases, which requires more prospective studies to validate.

In summary, we developed and validated predictive models for moderate-to-severe endoscopic activities by using noninvasive factors based on stool frequency, hematochezia, ESR level, and colonic wall flow in BUS, which provide a novel approach for disease activity assessment and management in the entire course of UC in the future. In addition, further studies are required to explore more potential biomarkers or measurements to predict mucosal and histologic healing or prognostic outcomes to precisely guide therapeutic intervention in clinical practice.

Authors' contributions

MMZ: Study design, patient recruitment, data collection and analysis, drafting the manuscript.

HMZ: Data collection.

QLZ: Data collection and interpretation, critical revision of manuscript.

XYB, GCR: Data analysis support.

QYZ: Data check.

WBL, LM, MSX: Data collection and interpretation.

HY, JMQ: Study concept and design, data collection and analysis support, critical revision of manuscript.

Conflict of interest statement

The authors declare that there is no conflict of interest.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by National Natural Science Foundation of China (No.81570505 and No.81970495), Health Research & Special Projects Grant of China (No.201002020 and No.201502005), Natural Science Foundation of Beijing, China (No.7202161), and CAMS Innovation Fund for Medical Sciences (No. 2016-I2M-3-001 and No.2019-I2M-2-007).

Guarantor of the article

Hong Yang, MD

ORCID iDs

Hong Yang https://orcid.org/0000-0002-2986-7324

Jiaming Qian https://orcid.org/0000-0001-6611-9475

Data availability statement

The data underlying this study are available from corresponding author upon reasonable request.

Supplemental material

Supplemental material for this article is available online.

References

1. Magro F, Gionchetti P, Eliakim R, *et al*. Third European evidence-based consensus on diagnosis and management of ulcerative colitis. Part 1: definitions, diagnosis, extra-intestinal manifestations, pregnancy, cancer surveillance, surgery, and Ileo-anal pouch disorders. *J Crohns Colitis* 2017; 11: 649–670.
2. Garg R, Singh A, Ahuja KR, *et al*. Risks, time trends, and mortality of colonoscopy-induced perforation in hospitalized patients. *J Gastroenterol Hepatol* 2020; 35: 1381–1386.

3. Bryant RV, Costello SP, Schoeman S, *et al*. Limited uptake of ulcerative colitis "treat-to-target" recommendations in real-world practice. *J Gastroenterol Hepatol* 2018; 33: 599–607.
4. Jharap B, Sandborn WJ, Reinisch W, *et al*. Randomised clinical study: discrepancies between patient-reported outcomes and endoscopic appearance in moderate to severe ulcerative colitis. *Aliment Pharmacol Ther* 2015; 42: 1082–1092.
5. Falvey JD, Hoskin T, Meijer B, *et al*. Disease activity assessment in IBD: clinical indices and biomarkers fail to predict endoscopic remission. *Inflamm Bowel Dis* 2015; 21: 824–831.
6. Allocca M, Fiorino G, Bonovas S, *et al*. Accuracy of humanitas ultrasound criteria in assessing disease activity and severity in ulcerative colitis: a prospective study. *J Crohns Colitis* 2018; 12: 1385–1391.
7. Kinoshita K, Katsurada T, Nishida M, *et al*. Usefulness of transabdominal ultrasonography for assessing ulcerative colitis: a prospective, multicenter study. *J Gastroenterol* 2019; 54: 521–529.
8. Maaser C, Sturm A, Vavricka SR, *et al*. ECCO-ESGAR guideline for diagnostic assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 2019; 13: 144–164.
9. Maconi G, Nylund K, Ripolles T, *et al*. EFSUMB recommendations and clinical Guidelines for Intestinal Ultrasound (GIUS) in inflammatory bowel diseases. *Ultraschall Med* 2018; 39: 304–17.
10. Limberg B. Diagnosis of chronic inflammatory bowel disease by ultrasonography. *Zeitschrift fur Gastroenterologie* 1999; 37: 495–508.
11. Sturm A, Maaser C, Calabrese E, *et al*. ECCO-ESGAR guideline for diagnostic assessment in IBD Part 2: IBD scores and general principles and technical aspects. *J Crohns Colitis* 2019; 13: 273–284.
12. Peyrin-Biroulet L, Reinisch W, Colombel JF, *et al*. Clinical disease activity, C-reactive protein normalisation and mucosal healing in Crohn's disease in the SONIC trial. *Gut* 2014; 63: 88–95.
13. Ma C, Sandborn WJ, D'Haens GR, *et al*. Discordance between patient-reported outcomes and mucosal inflammation in patients with mild to moderate ulcerative colitis. *Clin Gastroenterol Hepatol* 2020; 18: 1760–1768 e1.
14. Mosli MH, Zou G, Garg SK, *et al*. C-reactive protein, fecal calprotectin, and stool lactoferrin for detection of endoscopic activity in symptomatic inflammatory bowel disease patients: a systematic review and meta-analysis. *Am J Gastroenterol* 2015; 110: 802–819; quiz 20.
15. Koido S, Ohkusa T and Takakura K. Clinical significance of serum procalcitonin in patients with ulcerative colitis. *World J Gastroenterol* 2013; 19: 8335–8341.
16. Yoon JY, Park SJ, Hong SP, *et al*. Correlations of C-reactive protein levels and erythrocyte sedimentation rates with endoscopic activity indices in patients with ulcerative colitis. *Dig Dis Sci* 2014; 59: 829–837.
17. Lewis JD. The utility of biomarkers in the diagnosis and therapy of inflammatory bowel disease. *Gastroenterology* 2011; 140: 1817–1826.e2.
18. Florin THJ, Paterson EWJ, Fowler EV, *et al*. Clinically active Crohn's disease in the presence of a low C-reactive protein. *Scand J Gastroenterol* 2006; 41: 306–311.
19. Langhorst J, Boone J, Lauche R, *et al*. Faecal lactoferrin, calprotectin, PMN-elastase, CRP, and white blood cell count as indicators for mucosal healing and clinical course of disease in patients with mild to moderate ulcerative Colitis: post hoc analysis of a prospective clinical trial. *J Crohns Colitis* 2016; 10: 786–794.
20. Shin J, Kong SM, Seong G, *et al*. What is the appropriate cut-off value of CRP to predict endoscopic remission in patients with ulcerative colitis in clinical remission? *Int J Colorectal Dis* 2020; 35: 2249–2255.
21. Truelove S and Witts L. Cortisone in ulcerative colitis final report on a therapeutic trial. *Br Med J* 1955; 2: 1041–1048.
22. Aletaha D, Nell VPK, Stamm T, *et al*. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005; 7: R796–R806.
23. Smith RL, Taylor KM, Friedman AB, *et al*. Systematic review: clinical utility of gastrointestinal ultrasound in the diagnosis, assessment and management of patients with ulcerative Colitis. *J Crohns Colitis* 2020; 14: 465–479.
24. Sagami S, Kobayashi T, Miyatani Y, *et al*. Accuracy of ultrasound for evaluation of colorectal segments in patients with inflammatory bowel diseases: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. Epub ahead of print 7 August 2020. DOI: 10.1016/j.cgh.2020.07.067.

25. Ripollés T, Poza J, Suarez Ferrer C, *et al*. Evaluation of Crohn's disease activity: development of an ultrasound score in a multicenter study. *Inflamm Bowel Dis* 2021; 27: 145–154.

26. Sævik F, Eriksen R, Eide GE, *et al*. Development and validation of a simple ultrasound activity score for Crohn's disease. *J Crohns Colitis* 2021; 15: 115–124.

27. Maaser C, Petersen F, Helwig U, *et al*. Intestinal ultrasound for monitoring therapeutic response in patients with ulcerative colitis: results from the TRUST&UC study. *Gut* 2020; 69: 1629–1636.

28. Omotehara S, Nishida M, Kinoshita K, *et al*. Validation of US evaluation of ulcerative colitis activity. *Ultrasound Med Biol* 2019; 45: 1537–1544.

29. Schaeffer DF, Walsh JC, Kirsch R, *et al*. Distinctive histopathologic phenotype in resection specimens from patients with Crohn's disease receiving anti-TNF-alpha therapy. *Hum Pathol* 2014; 45: 1928–1935.