UNITED EUROPEAN GASTROENTEROLOGY
ueg journal

Original Article

Milan Ultrasound Criteria are accurate in assessing disease activity in ulcerative colitis: External validation

United European Gastroenterology Journal
0(0) 1–5

DOI: 10.1177/2050640620980203
journals.sagepub.com/home/ueg
SAGE

Mariangela Allocca[1,2], Elisabetta Filippi[3], Andrea Costantino[3], Stefanos Bonovas[1,2], Gionata Fiorino[1,2], Federica Furfaro[1], Laurent Peyrin-Biroulet[4], Mirella Fraquelli[3,5], Flavio Caprioli[3,5] and Silvio Danese[1,2]

Abstract

Introduction: The aim of this study was to provide an external validation of bowel ultrasound (US) predictors of activity in ulcerative colitis (UC) and quantitative Milan Ultrasound Criteria (MUC).

Methods: Forty-three consecutive patients with UC (16 in endoscopic remission and 27 with endoscopic activity) underwent bowel US and colonoscopy in a tertiary referral inflammatory bowel disease unit.

Results: A MUC score >6.2 discriminated patients with active versus non-active UC with a sensitivity of 0.85 (95% confidence interval (CI) 0.66–0.96), specificity of 0.94 (95% CI 0.70–0.99) and an area under the curve of 0.902 (95% CI 0.772–0.971) in complete agreement with the derivation study.

Conclusion: The external validation of MUC confirms that it is an accurate tool for assessing disease activity in patients with UC.

Keywords

Inflammatory bowel disease, bowel ultrasound, ulcerative colitis, Milan Ultrasound Criteria, gastrointestinal ultrasound, mucosal healing

Received: 24 September 2020; accepted: 15 November 2020

Introduction

Ulcerative colitis (UC) is a chronic, relapsing and destructive inflammatory disorder of the colon which can lead to organ damage and impair quality of life.[1] A treat-to target approach with tight monitoring of intestinal inflammation is recommended.[2] Hence, we need simple, practical and non-invasive tools for monitoring UC patients in order to reduce the burden related to colonoscopy (CS). Bowel ultrasound (US) is a cost-effective, well-tolerated, non-invasive and readily available tool for the management of patients with inflammatory bowel disease (IBD), including UC, in clinical practice.[3] Several data on the accuracy of bowel US in monitoring disease activity of UC have been published.[4–7] However, clear US-based parameters to assess and grade disease activity and severity in UC patients are lacking. Recently, we determined the bowel US parameters that best identified endoscopic activity, defined by a Mayo score of ≥2, using multivariable analysis, and developed non-invasive

[1]Department of Gastroenterology, Humanitas Clinical and Research Center – IRCCS, Milan, Italy
[2]Department of Biomedical Sciences, Humanitas University, Milan, Italy
[3]Gastroenterology and Endoscopy Unit, La Fondazione IRCCS Ca' Granda Ospedale Maggiore di Milano Policlinico, Milan, Italy
[4]Department of Gastroenterology and Inserm U954, University Hospital of Nancy, Lorraine University, Nancy, France
[5]Department of Pathophysiology and Transplantation, University of Milan, Milan, Italy

Corresponding author:
Mariangela Allocca, Department of Biomedical Sciences, Humanitas University, Via Manzoni 113, 20089, Rozzano, Milan, Italy.
Email: mariangela.allocca@humanitas.it

ultrasonography based criteria (Humanitas Ultrasound Criteria (HUC) now described as Milan Ultrasound Criteria (MUC)) to assess and grade disease activity in UC.[8] The external validity and reproducibility of these criteria remain to be investigated in an independent cohort of UC patients and assessed by an independent IBD team.[9] The aim of this study was to provide external validation of the MUC.

Methods

Study population and examinations

Consecutive adult patients with an established diagnosis of UC (of at least six months), seen in a tertiary referral centre (Fondazione IRCCS Ca' Granda – Ospedale Maggiore Policlinico, Milan, Italy) between May 2019 and May 2020 and requiring routine investigations by CS, were prospectively assessed by CS and bowel US within one week, irrespective of disease activity. Inclusion and exclusion criteria and design of the study were the same as those reported in detail in the derivation study.[8] The endoscopic activity was evaluated by CS according to the Mayo endoscopic sub-score, and mucosal healing (MH) was defined by an absolute Mayo endoscopic sub-score of 0 or 1.[10] Bowel US was performed by two gastroenterologists who were experts in bowel US (6–20 years of experience) using a Philips iU22 apparatus (Philips Ultrasound; Philips Healthcare, Bothell, WA) with a multi-frequency convex (C5-2, 5–2 MHz) and a linear array transducer (L12-5, 12–5 MHz). The following parameters were evaluated: colonic wall thickening (CWT; the average of three measurements, normal values $\leq$3 mm); colonic wall pattern (CWP; 0 = normal, multilayered, 1 = prevalently hypoechogenic, 2 = prevalently hyperechogenic, 3 = lost); colonic wall flow (CWF; 0 = absence, 1 = presence of blood signals at colour Doppler); enlarged mesenteric lymph nodes (short axis >5 mm); and mesenteric hypertrophy (defined as the presence of a hyperechoic area surrounding the pathologic intestinal tract).

Statistical analysis

Descriptive statistics of the baseline data are presented as medians (interquartile range) or as percentages when appropriate. Data comparison was performed using Fisher's exact test for categorical variables and the non-parametric two-sample Wilcoxon rank-sum (Mann–Whitney) test for continuous variables. The coefficients of CWT and CWF (i.e. 1.4 and 2.0, respectively), derived from the multivariable analysis in the derivation study,[8] were used to calculate the MUC in the validation cohort. We performed receiver operating characteristic (ROC) analysis to calculate the area under the curve (AUC) in the validation cohort, and assessed the performance of the model in this new population in terms of sensitivity and specificity.

Ethical considerations

The study was performed according to Good Clinical Practice guidelines and the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by our Institutional Review Board on 5 November 2014 (ICH1330). Written informed consent was obtained from each patient included in the study.

Results

A total of 43 consecutive UC patients were included in the study: 16 (37%) patients were in endoscopic remission (Mayo score 0–1), while 27 (63%) patients displayed endoscopic activity (Mayo score $\geq$2). Nineteen (44%) patients had left-sided UC and 14 patients (33%) had an extensive UC, as evaluated by CS.

Table 1. Characteristics of patients at inclusion in the study (N = 43).

Female	16 (37)
Age at diagnosis (years)	39.01 (28.25–52.20)
Age at inclusion (years)	53.81 (35.06–70.83)
Disease duration (years)	8.77 (1.81–19.05)
Disease extent at diagnosis	
■ E2 left sided	29 (67)
■ E3 extensive	14 (33)
Concomitant treatments[a]	
■ Steroids	16 (37)
■ Immunosuppressants	4 (9.0)
■ Biologic therapy[b]	9 (21)
Smoking	
■ Past	4 (9.0)
■ Active	5 (12)
Partial Mayo score (PMS)	5 (0–8)
■ PMS >2	26 (60)
C-reactive protein (mg/L)	8.5 (3.00–39.60)
Calprotectin (μg/g)	111.0 (48.5–188.0)
Mayo endoscopic sub-score	
■ 0	10 (23)
■ 1	6 (14)
■ 2	9 (21)
■ 3	18 (42)
Disease extent at colonoscopy at inclusion	
■ E2 left sided	19 (44)
■ E3 extensive	14 (33)

Data are presented as medians (interquartile range) or percentages as appropriate.
[a]All patients took mesalazine.
[b]Eight patients were given infliximab, one vedolizumab.

Baseline characteristics and clinical data of the study population are presented in Table 1.

Bowel US findings comparing to CS

Median values of CWT in patients in endoscopic remission (Mayo score 0–1) were 3.0 mm (IQR 3.0–3.5) compared to 6.0 mm (IQR 4.6–6.6) in patients with endoscopic active disease (Mayo score 2–3; $p < 0.0001$). CWF was present in 12% (2/16) of patients in endoscopic remission compared to 78% (21/27) of patients who displayed endoscopic activity ($p = 0.0001$). Hypoechogenic or lost CWP was present in 12% (2/16) of patients in endoscopic remission compared to 52% (14/27) of patients displaying endoscopic activity ($p = 0.02$). Enlarged mesenteric lymph nodes were present in 0% (0/16) of patients in endoscopic remission compared to 15% (4/27) of patients with endoscopic activity ($p = 0.2$).

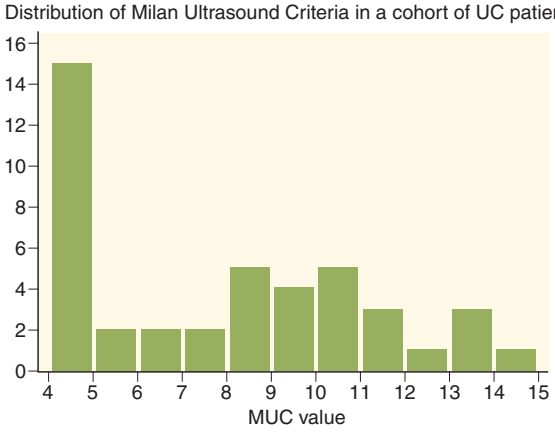

Figure 1. Distribution of the Milan Ultrasound Criteria (MUC) in a cohort of patients with ulcerative colitis.

Table 2. Milan Ultrasound Criteria (MUC): observed risk of activity for each MUC range.

MUC[a] range	Observed risk of endoscopic activity[a]
<6.2	4/19 (21%)
6.3–8.1	1/2 (50%)
8.2–10.6	11/11 (100%)
>10.6	11/11 (100%)

[a]Endoscopic activity defined by a Mayo endoscopic sub-score of 2–3.

Validation of the MUC

The coefficients of CWT and CWF (i.e. 1.4 and 2.0, respectively), derived from the multivariable analysis in the derivation study,[8] were used to calculate the MUC score in the validation cohort (i.e. $MUC = 1.4 \times CWT + 2.0 \times CWF$). The distribution of the MUC in this cohort is presented in Figure 1. The MUC significantly correlated with the Mayo endoscopic sub-score ($r = 0.76$; 95% confidence interval (CI) 0.60–0.86; $p < 0.0001$). A MUC score >6.2 was the best cut-off for discriminating patients with active versus non-active UC at the ROC analysis, with a sensitivity of 0.85 (95% CI 0.66–0.96), specificity of 0.94 (95% CI 0.70–0.99) and an AUC of 0.902 (95% CI 0.772–0.971). The observed risk of activity for each MUC range is shown in Table 2. The diagnostic accuracy of the MUC in the validation cohort was comparable with that in the derivation study[8] (Table 3).

MUC and faecal calprotectin in combination

We performed ROC analysis to investigate the performance of faecal calprotectin (FC), in terms of sensitivity and specificity, for assessing endoscopic activity. A FC value >50 µg/g was the best cut-off for discriminating patients with active disease from patients with non-active disease, with a sensitivity of 0.94, specificity of 0.55 and an AUC of 0.727 (95% CI 0.527–0.877). The comparison of the ROC curves regarding 'MUC alone' and 'MUC+FC' (i.e. a combined diagnostic strategy of MUC and FC, considered positive if having a MUC score >6.2 and/or FC >50 µg/g) showed no difference between the two approaches ($p = 0.101$; Figure 2). Furthermore, we performed a sensitivity analysis using different cut-offs for FC in the combined diagnostic strategy 'MUC+FC'. In the first case, where 'MUC+FC' was considered positive when having a MUC score >6.2 and/or FC >150 µg/g, sensitivity was 0.88 and specificity was 0.63, while in the second case, where 'MUC plus FC' was considered positive when having a MUC score >6.2 and/or FC >250 µg/g, sensitivity was 0.82 and specificity was 0.72.

Table 3. Diagnostic accuracy of Milan Ultrasound Criteria (MUC) in derivation and validation study.

		MUC in derivation study			MUC in validation study		
	Cut-off	ROC	Sens	Spec	ROC	Sens	Spec
Active disease (Mayo endoscopic sub-score ≥2)	>6.2	0.891	0.71	1.00	0.902	0.85	0.94

ROC: receiver operating characteristic; Sens: sensitivity; Spec: specificity.

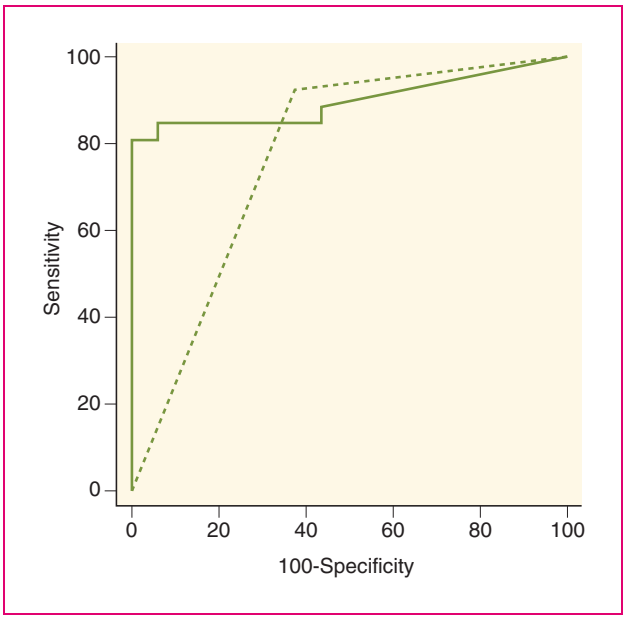

Figure 2. Comparison of receiver operating characteristic curves: 'MUC alone' (continuous line) versus 'MUC+FC' (dashed line). No statistically significant difference was observed ($p = 0.101$). FC: faecal calprotectin.

Discussion

This prospective study provided external validation of US-based criteria (MUC) for assessing disease activity in UC. It confirmed the role of CWT and CWF as predictors of UC activity in an external, independent cohort. The good performance of the MUC was shown by similar ROC areas in the derivation and validation cohorts (AUC: 0.891 and 0.902, respectively). We also confirmed that a MUC score >6.2 is a valid cut-off to discriminate active from non-active UC. We have demonstrated that (a) the presence of a CWF and a CWT >3 mm or (b) the absence of a CWF and a CWT >4.43 mm are both indicative of active disease in UC, and this may be the basis for further investigation about the role of the MUC in monitoring treatment response and driving therapeutic decisions. Finally, we demonstrated that these criteria are reproducible and valid when used by independent operators, confirming they are a simple and valid measure of disease activity that could be widely used.

The role of US in UC patients is still limited. CS remains the standard procedure to assess and confirm disease activity and severity. The main limitations of US are thought to be the poor sensitivity to look at mucosal ulcers, and the poor reproducibility among independent operators. We clearly showed that US is accurate in assessing and scoring disease activity in the derivation and validation cohorts, as the MUC were highly predictive compared to the Mayo endoscopic sub-score. Therefore, these criteria can easily and non-invasively distinguish moderate to severe active disease from inactive disease. Moreover, we clearly demonstrated the high reproducibility of US assessment, as we found similar AUC and cut-off values for our criteria in both derivation and validation cohorts. This will definitely be an interesting starting point to investigate the potential role of US in replacing endoscopy at least in patients who need close monitoring and assessment of response to treatment.

The limitations of this study are that we validated these criteria in a single centre with a rather small sample size, although this was consistent with the previously published derivation study.[8]

In conclusion, we validated the MUC as an accurate tool to assess UC activity. The sensitivity to change of these criteria and their validity to assess treatment response and outcomes both in clinical practice and in clinical trials require further investigation.

Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: M.A. received consulting fees from Nikkiso Europe, Mundipharma, Janssen, Abbvie and Pfizer; G.F. received consultancy fees from Ferring, MSD, AbbVie, Takeda, Janssen, Amgen, Sandoz, Samsung Bioepis and Celltrion; F.F. received consulting fees from Amgen and Abbvie and lecture fees from Janssen and Pfizer; L.P.-B. reports personal fees from Merck, Abbvie, Janssen, Genentech, Mitsubishi, Ferring, Norgine, Tillots, Vifor, Hospira/Pfizer, Celltrion, Takeda, Biogaran, Boerhinger-Ingelheim, Lilly, HAC-Pharma, Index Pharmaceuticals, Amgen, Sandoz, Forward Pharma GmbH, Celgene, Biogen, Lycera and Samsung Biosepsis; F.C. served as a consultant to Mundipharma, Abbvie, MSD, Takeda, Janssen, Roche and Celgene, and received lecture fees from Abbvie, Ferring, Takeda, Allergy Therapeutics and Janssen and unrestricted research grants from Giuliani, Sofar, MS&D, Takeda and Abbvie; S.D. served as a speaker, consultant and advisory board member for Schering-Plough, Abbott (AbbVie) Laboratories, Merck, UCB Pharma, Ferring, Cellerix, Millenium Takeda, Nycomed, Pharmacosmos, Actelion, Alfa Wasserman, Genentech, Grunenthal, Pfizer, AstraZeneca, Novo Nordisk, Cosmo Pharmaceuticals, Vifor and Johnson and Johnson. E.F., A. C., S.B. and M.F. have no conflicts of interest to declare.

Ethics approval

The study was performed according to Good Clinical Practice guidelines and the ethical guidelines of the 1975 Declaration of Helsinki, and was approved by our Institutional Review Board on 5 November 2014 (ICH1330).

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

Informed consent

Written informed consent was obtained from each patient included in the study.

References

1. Gordon IO, Agrawal N, Willis E, et al. Fibrosis in ulcerative colitis is directly linked to severity and chronicity of mucosal inflammation. *Aliment Pharmacol Ther* 2018; 47: 922–939.
2. Peyrin-Biroulet L, Sandborn W, Sands BE, et al. Selecting Therapeutic Targets in Inflammatory Bowel isease (STRIDE): determining therapeutic goals for treat-to-target. *Am J Gastroenterol* 2015; 110: 1324–1338.
3. Bryant RV, Friedman AB, Wright EK, et al. Gastrointestinal ultrasound in inflammatory bowel disease: an underused resource with potential paradigm-changing application. *Gut* 2018; 67: 973–985.
4. Maconi G, Ardizzone S, Parente F, et al. Ultrasonography in the evaluation of extension, activity, and follow-up of ulcerative colitis. *Scand J Gastroenterol* 1999; 34: 1103–1107.
5. Antonelli E, Giuliano V, Casella G, et al. Ultrasonographic assessment of colonic wall in moderate-severe ulcerative colitis: comparison with endoscopic findings. *Dig Liver Dis* 2011; 43: 703–706.
6. Parente F, Molteni M, Marino B, et al. Bowel ultrasound and mucosal healing in ulcerative colitis. *Dig Dis* 2009; 27: 285–290.
7. 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.
8. 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.
9. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003; 56: 826–832.
10. Schroeder KW, Tremaine WJ and Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *New Engl J Med* 1987; 317: 1625–1629.