**Original article**

Thieme

# Reliability of the Endoscopic Ultrasound Ulcerative Colitis (EUS-UC) score for assessment of inflammation in patients with ulcerative colitis ▶

OPEN ACCESS

**Authors**
Brian M. Yan[1], Michael S.L. Sey[1], Paul Belletrutti[2], Gary Brahm[3], Leonardo Guizzetti[4], Vipul Jairath[1,4]

**Institutions**
1 Division of Gastroenterology, Department of Medicine, Western University, London, ON, Canada
2 Division of Gastroenterology, Department of Medicine, University of Calgary, Calgary, AB, Canada
3 Department of Radiology, Western University, London, ON, Canada
4 Alimentiv Inc. (formally Robarts Clinical Trials Inc.), London, ON, Canada

**Corresponding author**
Brian Yan, MD, FRCPC, CAGF, Associate Professor of Medicine, Division of Gastroenterology, Department of Medicine, Western University, Rm E6-319a, LHSC Victoria Campus, 800 Commissioners Road East, London, Ontario, Canada
Fax: +519-667-6820
brian.yan@lhsc.on.ca

**ABSTRACT**

**Background and study aims** Endoscopic ultrasound (EUS) may be a useful modality for disease assessment and risk stratification in ulcerative colitis. We assessed the reliability of a newly developed EUS index of inflammation called the EUS-Ulcerative Colitis (EUS-UC) score.

**Patients and methods** The EUS-UC score components include total wall thickness, hyperemia, and depth of inflammation (DOI). Three blinded expert endosonographers assessed EUS videos of 58 patients with UC in triplicate. Intra- and inter-rater reliability of the hyperemia and DOI component scores were estimated using intra-class correlation coefficients (ICCs). Total wall thickness reliability estimates could not be assessed in this study. The ICCs were compared to the original indices from which they were derived.

**Results** For hyperemia, the inter-class ICC was "moderate" at 0.556 (95% CI=0.434–0.651) and the intra class ICC was "almost perfect" at 0.884 (95% CI=0.835–0.920). The newly defined hyperemia score performed better than the original index from which is was derived. The DOI inter-class ICC was "fair" at 0.335 (95% CI=0.201–0.464), and the intra-class ICC was "substantial" at 0.732 (95% CI=0.642–0.802). The DOI reliability estimates were similar to the original index from which it was derived.

**Conclusions** The hyperemia component of the EUS-UC score performed significantly better than the original index from which it was derived, but the reliability of the DOI component was suboptimal. Intra-class correlation was excellent for both components. The EUS-UC score is a promising instrument for assessment of UC and further validation is required.

## Introduction

Ulcerative colitis (UC) is a chronic inflammatory bowel disease (IBD) of unknown etiology characterized by diarrhea, rectal bleeding, and impaired quality of life. Goals of therapy for UC have evolved from control of bleeding and diarrhea to improvement in more objective measures of inflammation [1]. Endoscopic mucosal healing is associated with favorable outcomes including reduced risk of relapse, need for steroids, hospitalization, and colectomy [2–5]. Although biomarkers such as fecal calprotectin (FCP) and serum C-reactive protein (CRP) are widely used as noninvasive markers of disease activity, endoscopy remains the gold standard for evaluation of inflammation [6–11]. Furthermore, there is increasing interest in resolution of histologic inflammation as a treatment target, which may be

associated with improved outcomes compared to patients with endoscopic mucosal healing alone [12–17].

Although UC is generally considered to be a superficial process restricted to the mucosa, in more severe cases, inflammation extends to deeper layers of the bowel wall with associated submucosal fibrosis that cannot be evaluated by endoscopy or endoscopically procured biopsies. In contrast, endoscopic ultrasound (EUS) can examine all the layers of the bowel wall. Accordingly, EUS is a highly accurate diagnostic and prognostic modality for assessment of diseases of the rectum. Experience with EUS in UC is limited and its potential role as a prognostic tool in UC remains undefined. Interpretation of existing studies is limited because of weaknesses in methodology including lack of blinding, small sample sizes, variations regarding which EUS findings were compared, and differences in the definitions of what constituted normal and abnormal [18–24]. Nevertheless, several studies have demonstrated that patients with active UC may have increased wall thickness, inflammatory changes through the deeper layers of the bowel wall, and/or increased vascularity. In patients with quiescent disease, "deep" disease activity, as specified by increased thickness of the first three layers of the bowel wall, may have prognostic value [18]. Finally, EUS may help evaluate and predict response to therapy [23].

Before its potential can be realized, the operating properties of EUS in UC must be rigorously evaluated. We have previously assessed the intra and inter-rater reliability of EUS indices that included bowel wall thickness, the Tsuga score, and the hyperemia score [25]. The major finding was that although the intra-rater reliability was excellent, the inter-rater was only fair. As a result, a modified Research and Development (RAND) process was completed and the EUS Ulcerative Colitis (EUS-UC) score was developed. The EUS-UC score has the potential to have better inter-rater reliability, as the components of the score are simple, more objective, and possibly reproducible.

This study was conducted to assess the reliability of the novel the properties of the EUS-UC score in patients with ulcerative colitis and compare this to existing indices.

## Patients and methods

This is a retrospective evaluation of a cohort of 58 patients with UC from London Health Sciences Centre and St. Joseph's Healthcare London, tertiary care centers affiliated with Western University (Canada), who were previously enrolled in a reliability study of EUS indices [25]. The archived EUS videos were re-read for the present reliability study. Three expert, central endosonographers not involved in the reading of the videos from the original study (BY, MS, PB) reviewed and evaluated (n = 58) videos from these patients. Each video was rated in triplicate, 2 weeks apart, by each central reader in order to estimate intra- and inter-rater reliability. Identical to the original study, the central readers were given seven training videos prior to assessing the study videos to demonstrate abnormal findings on EUS. Central readers were blinded to the clinical history, endoscopic scores, and histologic findings from the patients. The EUS-UC score is composed of bowel wall thickness, depth of inflammation, and hyperemia, ranging from a score of 0 to 9 points (▶ Table 1).

### Statistical analysis

The primary objective of this study was to evaluate the inter- and intra-rater reliability of the EUS-UC score component items of depth of inflammation and hyperemia. The inter-rater reliability of bowel wall thickness could not be assessed in this study because it must be measured in real time. Reliability is a measure of agreement between measurements. The inter-rater reliability measures the agreement between different raters about measurements on the same subject, whereas intra-rater reliability measures agreement between measurements made by the same rater. This study design used three blinded reviewers measuring each subject three times, with each repeated measure spaced 2 weeks apart. The original sample size was conservatively determined using the one-way random effects model as described by Zou [26]. Assuming a true ICC of 0.75, evaluation of 58 videos by three central readers would yield an 83% chance of obtaining the one-sided 95% confidence interval

▶ Table 1 Endoscopic Ultrasound Ulcerative Colitis (EUS-UC) Score [25].

| | Component score | | | |
|---|---|---|---|---|
| Component | 0 | 1 | 2 | 3 |
| Total wall thickening | Normal (≤3.0 mm) | Mild (3.1 – 4.0 mm) | Moderate (4.1 – 6.0 mm) | Severe (≥6.1 mm) |
| Depth of inflammation | Superficial (No disruption of the 5-layer echo pattern) | Subepithelial (Disruption of the first 3 layers to the submucosa but not beyond) | Deep (Disruption beyond the submucosa to the muscularis propria) | Transmural (Disruption beyond the muscularis propria to the serosa or beyond) |
| Hyperemia | Normal (Absence of intramural vascular signal) | Mild (Intermittent signal) | Moderate (Continuous signal) | Severe (Presence of intramural anechoic vessel seen without power Doppler, with immediate continuous signal on power Doppler) |
| Total score (sum of three items) | | | | 0 to 9 |

lower bound to be >0.60, which is the upper limit of the Landis and Koch benchmark for "moderate" reliability (0.41–0.60).

To assess the reliability of EUS across a range of disease activity, the original sample size was divided into 20 patients with quiescent UC, 20 patients with mild UC, and 20 patients with moderate to severe UC based on the Mayo endoscopic score. In this retrospective analysis, the same videos were assessed from the original cohort of 58 patients. Two patients were excluded as they did not have a true diagnosis of ulcerative colitis.

The reliability of EUSIS components were estimated using the intra-class correlation statistic (ICC), which is equivalent to the weighted Kappa statistic for ordinal data with quadratic weights, and used to quantify inter-rater and intra-rater reliability [27]. The ICC point estimates were obtained using a two-way random-effects analysis of variance model, allowing for an interaction effect between subjects and readers, and where subjects and readers were both treated as random effects [28]. The associated 95% confidence intervals were obtained using a non-parametric percentile bootstrap method with 2000 replicates, in which the data were resampled on the level of the subject to respect the structure of the data. This approach is commonly called the cluster bootstrap method [29]. The magnitude of reliability estimates was interpreted according to the well-known benchmarks of Landis and Koch, where ICCs of <0.00, 0.00 to 0.20, 0.21 to 0.40, 0.41 to 0.60, 0.61 to 0.80, and 0.81 to 1.00 indicate "poor," "slight," "fair," "moderate," "substantial," and "almost perfect" reliability, respectively [30].

## Results

Demographics of the patient population have previously been described. [25] A summary is shown in ▶ **Table 2**.

The ICC estimates of reliability are shown in ▶ **Table 3**. The hyperemia component showed moderate inter-rater reliability (ICC = 0.556) and almost perfect intra-rater reliability (ICC = 0.884). The depth of inflammation component showed fair inter-rater reliability (ICC = 0.335) and substantial intra-rater reliability (ICC = 0.732). In comparison, the original EUS indices of inflammation ICC sores described by Yan et al are shown in ▶ **Table 3**.
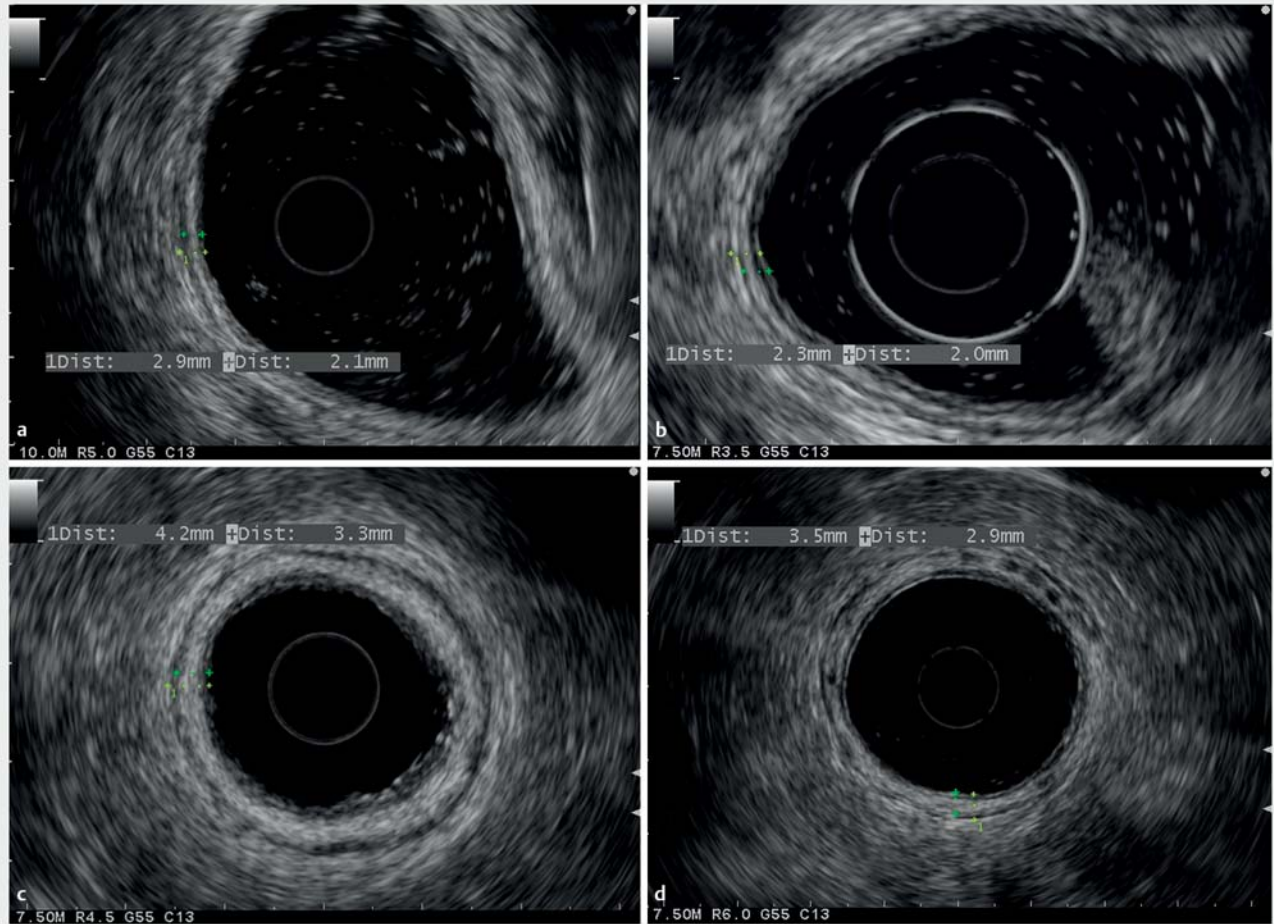
▶ **Table 2** Patient demographics.

| | Demographics (n = 58) |
|---|---|
| Mean age, years (range) | 43 (19–84) |
| Sex (M:F) | 30:28 |
| Mean disease duration (range) | 8.9 y (6 months-44 y) |
| Disease location | |
| ▪ Pancolitis | 31 (53%) |
| ▪ Left-sided colitis | 20 (34%) |
| ▪ Proctitis | 7 (12%) |
| Endoscopic disease activity | |
| ▪ Remission (MES 0) | 16 (28%) |
| ▪ Mild (MES 1) | 22 (38%) |
| ▪ Moderate/severe (MES 2–3) | 20 (34%) |

Modified from Yan et al. [25]
MES, Mayo Endoscopic Score.

▶ **Fig. 1** shows examples of EUS changes in those with UC in remission (▶ **Fig. 1a**), mild (▶ **Fig. 1b**), moderate/severe disease (▶ **Fig. 1c**) based on Mayo endoscopic score. An example of dilated intramucosal vessels (hyperemia score 3) is shown in ▶ **Fig. 1d**. Examples of EUS imaging in ulcerative colitis are shown in ▶ **Video 1** and ▶ **Video 2** to contrast differences in the components of the EUS-UC score. In ▶ **Video 1**, there was transmural involvement with a depth of inflammation score of 3. There was only limited intermittent vascular signal for a hyperemia score of 1. In contrast, ▶ **Video 2** demonstrates relatively well-preserved echo-layering with a depth of inflammation score of 1, but an obvious immediate intramural vascular signal with very small but visible vessels within the submucosal layer to give a hyperemia score of 3. ▶ **Video 2** also shows significant peri-rectal vascular signals, but extramural vascularity is not a component of the EUS-UC score.

▶ **Table 3** Reliability of the EUS-UC score and component items.

| | Inter-rater ICC (95% CI) | Intra-rater ICC (95% CI) |
|---|---|---|
| EUS-UC components | | |
| ▪ Hyperemia | 0.556 (0.434, 0.651) | 0.884 (0.835, 0.920) |
| ▪ Depth of inflammation | 0.335 (0.201, 0.464) | 0.732 (0.642, 0.802) |
| Original EUS indices variables [25] | | |
| ▪ Original hyperemia score | 0.34 (0.25, 0.42) | 0.76 (0.71, 0.80) |
| ▪ Tsuga score | 0.36 (0.24, 0.46) | 0.85 (0.79, 0.89) |

EUS-UC, Endoscopic Ultrasound Ulcerative Colitis.

▶ **Fig. 1 a** EUS-UC of patient in remission. Mayo Endoscopic Score 0. EUS-UC components: total wall thickness score 0, depth score 1, vascularity score (not shown): 1, total EUS-UC score 2. **b** EUS-UC of patient with mild disease. **c** EUS-UC of moderately active Ulcerative colitis. **d** EUS-UC showing large intramural vessels.
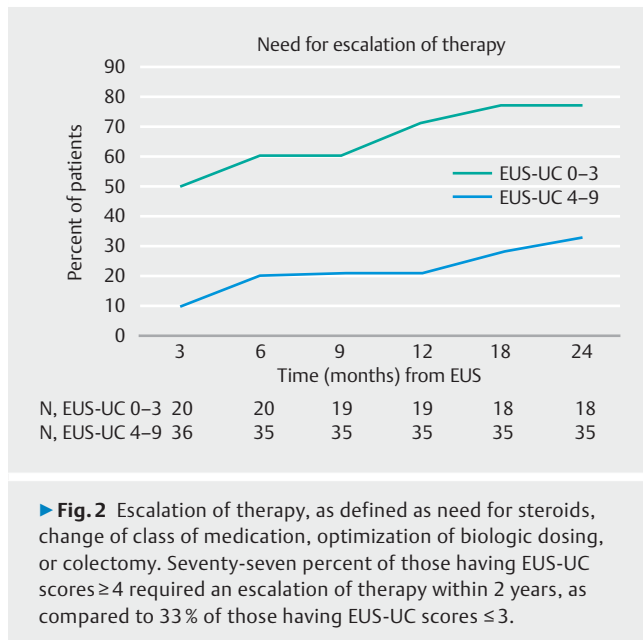
▶ **VIDEO**



▶ **Video 1** Patient with a Mayo Endoscopic score of 1. The video demonstrates transmural inflammation for a Depth of Inflammation Score of 3. Vascular signal is intermittent for a Hyperemia Score of 1.

▶ **VIDEO**
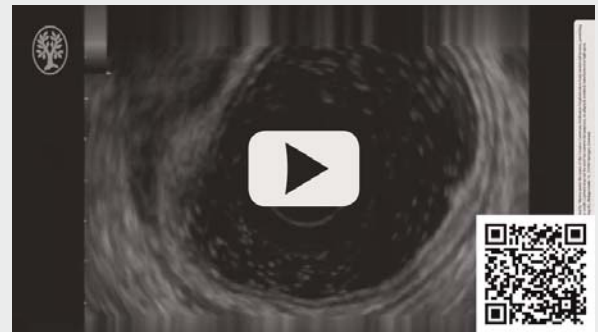


▶ **Video 2** This patient had a Mayo Endoscopic Score of 2. The 5 layer echopattern is relatively well preserved with a Depth of Inflammation Score of 1, but a strong vascular signal is present within visible dilated intramural vessels for a Hyperemia Score of 3.

**Fig. 2** Escalation of therapy, as defined as need for steroids, change of class of medication, optimization of biologic dosing, or colectomy. Seventy-seven percent of those having EUS-UC scores ≥ 4 required an escalation of therapy within 2 years, as compared to 33 % of those having EUS-UC scores ≤ 3.

## Patient outcomes

A retrospective assessment of patient follow-up suggests a substantial difference in the need of escalation of therapy, as defined as flares requiring steroids, change of medication class, escalation of biologic dosing, or colectomy (▶ Fig. 2). In those with EUS-UC scores ≥ 4, 50 % required escalation within 3 months of the procedure and 77 % required escalation by 2 years. In contrast, in those with EUS-UC score of ≤ 3, only 10 % required escalation at 3 months, and 33 % by 2 years. Five patients required colectomy in the 2-year follow-up. Two had resections within 6 months, with EUS-UC scores of 4 and 6. Three had resections after 1 year, with EUS-UC scores of 2, 3, and 5.

## Discussion

We found that the revised definitions of hyperemia in the EUS-UC score resulted in substantial improvements in the inter- and intra-rater reliability compared to the original indices from which it is was derived. However, the assessment of depth of inflammation was unchanged compared to the original Tsuga score, which was the scoring system we initially used to identify layer of bowel wall involvement [22]. Similar to the original study, the intra-rater reliability for depth of inflammation and hyperemia is "substantial" to "almost perfect" respectively.

The hyperemia score performed substantially better (inter-rater ICC of 0.56) with our new definitions compared to the original vague descriptions of "no vascular signal", "slight," "moderate," and "marked" vascularity described by Yan et al. [25] An ICC of 0.56, however, still only falls into a "moderate" reliability category, indicating the potential for improvement. This could be achieved by using sonographic contrast agents and assessment of time intensity curves [31]. This would, however, increase the complexity of the assessment, procedure time, cost, and risk. It is unknown if the enhanced detail in vascularity assessment provided by contrast would yield any clinical bene-

fit beyond the simple use of power Doppler. The hyperemia definitions used in this study, without the use of contrast agents, is more widely applicable to the general endosonographer with any standard ultrasound processor thus increasing uptake and acceptability.

The original study assessed reliability of the Tsuga score, which includes the presence of wall thickening and characterizes inflammatory changes between layers of the bowel wall [22, 25]. The inter-rater ICC was only "fair" at 0.36. Central readers felt this component of subtle abnormalities between the submucosa and muscularis propria to be subjective, prone to imaging errors, and very difficult to accurately assess. Despite our efforts to simplify the definition of depth of inflammation to any disruption of a given layer, central readers were still somewhat uncertain of the accuracy of the statement and scoring. Imaging artifacts, tangential imaging, rectal motility, use of balloon vs water for distension, and degree of rectal distension all can affect the endosonographers interpretation of disruption between layers. This uncertainty was again demonstrated in this study with a suboptimal inter-rater ICC of 0.335, which is unfortunately no better than the previous ICC for the Tsuga score.

For the purpose of index development, this poses an issue given the variability between observers. However, similar to the Physician Global Assessment in the full Mayo score, the component provides the practitioner some freedom to globally assess abnormal subepithelial inflammatory changes which may still have prognostic value. Prior studies have demonstrated that deeper inflammation suggests more severe disease and potentially predicts the need for colectomy [22, 24]. Furthermore, this component may be predictive of response to therapy. Potentially, one may be able to choose a specific therapy based on the predominant "phenotype" of inflammatory change (local wall changes vs increased vascularity). Purely assessing wall thickness as a surrogate for degree of layer involvement would be more objective, but may be erroneous given that longstanding disease with submucosal fibrosis can increase wall thickness without having active inflammation. This has not been studied in a prospective manner to determine if wall thickness is responsive to therapy. For these reasons, we feel that depth of inflammation should still be included in the development of the index in future studies.

Differences in interpretation of the videos for depth of inflammation is potentially related to lack of formalized training about what constitutes abnormal. Although endosonographers are trained to assess depth of cancer invasion through the layers of the bowel wall (T staging), they are not trained to assess subtle hypoechogenicity or irregular borders from inflammation. The readers in both this study and the original study were provided only seven training videos with no instruction accompanying them on what specific abnormalities should be recognized or where specifically on the video to direct their eyes. It was assumed the reader would recognize where the abnormalities were, but this may not have been the case. A more rigorous central reader training program to standardize what observers see as abnormal may improve the inter-rater reliability of both components. This training program has been developed by our

group for use in future studies to validate the responsiveness of the EUS-UC score. EUS assessment of luminal inflammation is not a standard component of EUS training, nor is it used in routine in clinical practice. Practicing endosonographers need to learn how to assess sometimes subtle inflammatory features of the bowel wall.

Other potential limitations may be the quality and length of the video. The recorded videos are not as sharp as the real-time imaging. The reduction in quality of video, inability to zoom in on specific regions or lack of ability to optimize the video to the reader's eye may impact the expert endosonographer's interpretation. The entire EUS exam was provided to the readers, some of which were > 15 minutes long. In this study, readers provided a single score for each component, summarizing the entire video based on what they thought was the worst diseased portion seen sonographically. Readers were not told which portions of the videos depicted the most severe disease. Therefore, different readers may have provided scores for different time periods of a given video, which may result in discordant observations. Shorter, focused videos of the most involved region of inflammation and directing the readers to interpret the same video segments may result in more reliable inter-rater interpretation.

In the EUS-UC score, wall thickness was separated out as an independent component from depth of inflammation and given equal weight to the other two components. Wall thickness is the most objective variable of the three components but can still vary between practitioners depending on degree of distension, peristalsis, where measurements are taken, and errors in tangential imaging. We previously attempted to minimize these errors by using an average of four separate measurements within the area of greatest inflammation during a period of no rectal contractions. In this study, because wall thickness was only measured by one practitioner during the original procedures it was not included in the calculation of an overall EUS-UC inter/intra-rater reliability score. To truly assess inter-observer variability for wall thickness, a given patient would require multiple assessments on the same day by different practitioners, which is unfeasible. However, a standard procedural protocol would help minimize errors between observers. Future prospective studies validating EUS-UC Score by our group will incorporate a standardized rectal EUS technique.

Although the study was not designed to assess patient outcomes, a post hoc assessment indicated that EUS-UC scores may be associated with the need for escalation of therapy. Nearly 80 % of patients with a total EUS-UC score ≥ 4 required some form of escalation within the ensuing 2 years. However, this observation should be considered exploratory since patient management was not controlled and based on the clinical assessment of the primary gastroenterologist, patients had different baseline medications and disease states, and management strategies including treatment targets evolved over the period of assessment. Nevertheless, the data are encouraging that the EUS-UC score may be further evaluated as a prognostic tool and to assess response to therapy.

The strengths of study include the use of experienced endosonographers for central reading who were blinded to the clinical and endoscopic presentation of patients and therefore were completely objective in their scoring. All the endosonographers were trained in formal EUS fellowships and in practice for > 5 years. Providing full videos is superior than simple still images; however, as previously stated, long videos can present some inter-rater interpretation variability. A wide range of disease activity was included allowing the ability to assess the reliability of EUS-UC component scores across all grades of severity of UC inflammation. A limitation of the study is that inter-rater bowel wall thickening assessment could not be measured resulting in an inability to provide an ICC for the entire EUS-UC score. Finally, another limitation is that we used previously recorded videos from which the score was derived rather than a new set of patients and videos. Therefore, there may be bias towards more favorable results.

## Conclusions

In summary, the EUS-UC component scores show improvement in assessment of vascularity, but not in depth of inflammation, as compared to the original EUS indices from which they were derived. A more rigorous reader training program may help improve the inter-rater reliability scores prior to future validation studies using a new set of patients.

## Competing interests

The authors declare that they have no conflict of interest.

## References

[1] Peyrin-Biroulet L, Sandborn W, Sands BE et al. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): determining therapeutic goals for treat-to-target. Am J Gastroenterol 2015; 110: 1324–1338

[2] Ardizzone S, Cassinotti A, Duca P et al. Mucosal healing predicts late outcomes after the first course of corticosteroids for newly diagnosed ulcerative colitis. Clin Gastroenterol Hepatol 2011; 9: 483–489 e3

[3] Bressler B, Marshall JK, Berstein CN et al. Clinical practice guidelines for the medical management of nonhospitalized ulcerative colitis: the Toronto consensus. Gastroenterology 2015; 148: 1035–1058 e3

[4] Laharie D, Filippi J, Roblin X et al. Impact of mucosal healing on long-term outcomes in ulcerative colitis treated with infliximab: a multicenter experience. Aliment Pharmacol Ther 2013; 37: 998–1004

[5] Neurath MF, Travis SP. Mucosal healing in inflammatory bowel diseases: a systematic review. Gut 2012; 61: 1619–1635

[6] Boon GJ, Day AS, Mulder CJ et al. Are faecal markers good indicators of mucosal healing in inflammatory bowel disease? World J Gastroenterol 2015; 21: 11469–11480

[7] Henriksen M, Jahnsen J, Lygren I et al. C-reactive protein: a predictive factor and marker of inflammation in inflammatory bowel disease. Results from a prospective population-based study. Gut 2008; 57: 1518–1523

[8] Ho GT, Lee HM, Brydon G et al. Fecal calprotectin predicts the clinical course of acute severe ulcerative colitis. Am J Gastroenterol 2009; 104: 673–678

[9] Karoui S, Laz S, Serghini M et al. Correlation of C-reactive protein with clinical and endoscopic activity in patients with ulcerative colitis. Dig Dis Sci 2011; 56: 1801–1805

[10] Langhorst J, Elsenbruch S, Koelzer J et al. Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel diseases: performance of fecal lactoferrin, calprotectin, and PMN-elastase, CRP, and clinical indices. Am J Gastroenterol 2008; 103: 162–169

[11] Lewis JD. The utility of biomarkers in the diagnosis and therapy of inflammatory bowel disease. Gastroenterology 2011; 140: 1817–1826 e2

[12] Bessissow T, Lemmens B, Ferrante M et al. Prognostic value of serologic and histologic markers on clinical relapse in ulcerative colitis patients with mucosal healing. Am J Gastroenterol 2012; 107: 1684–1692

[13] Bryant RV, Burger DC, Delo J et al. Beyond endoscopic mucosal healing in UC: histological remission better predicts corticosteroid use and hospitalisation over 6 years of follow-up. Gut 2016; 65: 408–414

[14] Pai RK, Jairath V, Casteele NV et al. The emerging role of histologic disease activity assessment in ulcerative colitis. Gastrointest Endosc 2018; 88: 887–898

[15] Park S, Abdi T, Gentry M et al. Histological disease activity as a predictor of clinical relapse among patients with ulcerative colitis: systematic review and meta-analysis. Am J Gastroenterol 2016; 111: 1692–1701

[16] Rosenberg L, Nanda KS, Zenlea T et al. Histologic markers of inflammation in patients with ulcerative colitis in clinical remission. Clin Gastroenterol Hepatol 2013; 11: 991–996

[17] Zenlea T, Yee EU, Rosenberg L et al. Histology grade is independently associated with relapse risk in patients with ulcerative colitis in clinical remission: a prospective study. Am J Gastroenterol 2016; 111: 685–690

[18] Higaki S, Nohara H, Saitoh Y et al. Increased rectal wall thickness may predict relapse in ulcerative colitis: a pilot follow-up study by ultrasonographic colonoscopy. Endoscopy 2002; 34: 212–219

[19] Hurlstone DP, Sanders DS, Lobo AJ et al. Prospective evaluation of high-frequency mini-probe ultrasound colonoscopic imaging in ulcerative colitis: a valid tool for predicting clinical severity. Eur J Gastroenterol Hepatol 2005; 17: 1325–1331

[20] Rasmussen SN, Riis P. Rectal wall thickness measured by ultrasound in chronic inflammatory diseases of the colon. Scand J Gastroenterol 1985; 20: 109–114

[21] Shimizu S, Tada M, Kawai K. Endoscopic ultrasonography in inflammatory bowel diseases. Gastrointest Endosc Clin N Am 1995; 5: 851–859

[22] Tsuga K, Haruma K, Fujimura J et al. Evaluation of the colorectal wall in normal subjects and patients with ulcerative colitis using an ultrasonic catheter probe. Gastrointest Endosc 1998; 48: 477–484

[23] Watanabe O, Ando T, El-Omar EM et al. Role of endoscopic ultrasonography in predicting the response to cyclosporin A in ulcerative colitis refractory to steroids. Dig Liver Dis 2009; 41: 735–739

[24] Yoshizawa S, Kobayashi K, Katsumata T et al. Clinical usefulness of EUS for active ulcerative colitis. Gastrointest Endosc 2007; 65: 253–260

[25] Yan B, Feagan B, Teriaky A et al. Reliability of EUS indices to detect inflammation in ulcerative colitis. Gastrointest Endosc 2017; 86: 1079–1087

[26] Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat Med 2012; 31: 3972–3981

[27] Fleiss J, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Ed Psycholog Meas 1973; 33: 613–619

[28] Eliasziw M, Young SL, Woodbury MG et al. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Phys Ther 1994; 74: 777–788

[29] Field C, Welsh A. Bootstrapping clustered data. J Royal Stat Soc Series B: Statistical Methodology 2007; 69: 369–390

[30] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174

[31] Girlich C, Schacherer D, Jung EM et al. Comparison between quantitative assessment of bowel wall vascularization by contrast-enhanced ultrasound and results of histopathological scoring in ulcerative colitis. Int J Colorectal Dis 2012; 27: 193–198