Journal Pre-proof

Intestinal ultrasound is accurate to determine endoscopic response and remission in patients with moderate to severe ulcerative colitis: a longitudinal prospective cohort study

Floris de Voogd, MD, Elsa A. van Wassenaer, MD, Aart Mookhoek, MD PhD, Steven Bots, MD, Sara van Gennep, MD, Mark Löwenberg, MD PhD, Geert R. D'Haens, MD PhD, Krisztina B. Gecse, MD PhD

PII: S0016-5085(22)00966-0
DOI: https://doi.org/10.1053/j.gastro.2022.08.038
Reference: YGAST 65301

To appear in: *Gastroenterology*
Accepted Date: 19 August 2022

Please cite this article as: de Voogd F, van Wassenaer EA, Mookhoek A, Bots S, van Gennep S, Löwenberg M, D'Haens GR, Gecse KB, Intestinal ultrasound is accurate to determine endoscopic response and remission in patients with moderate to severe ulcerative colitis: a longitudinal prospective cohort study, *Gastroenterology* (2022), doi: https://doi.org/10.1053/j.gastro.2022.08.038.

Intestinal ultrasound is accurate to determine endoscopic response and remission in patients with moderate to severe ulcerative colitis

Cohort and design

n=30 starting tofacitinib

endoscopic Mayo score (EMS) ≥ 2

Baseline *Week 8*

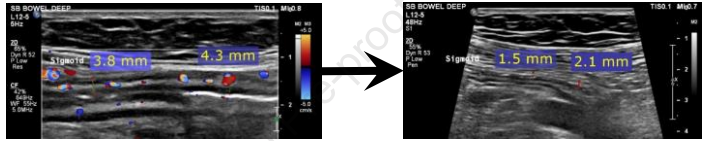
Bowel wall thickness BWT

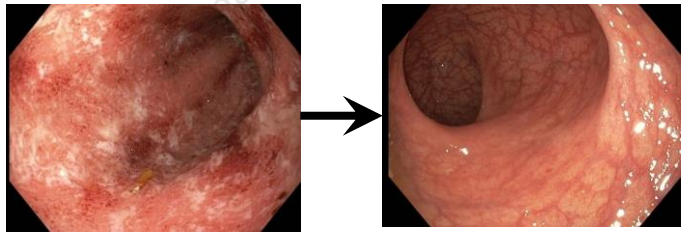
EMS

Robarts Histopathology Index (RHI)

Results

Baseline *Week 8*





score = 26

score = 1

Remission EMS=0 } **BWT** 2.8 mm

Improvement EMS≤1 } 3.9 mm

Response EMS≥1 decrease } 32% decrease

ρ=0.49: RHI vs BWT

Gastroenterology

Intestinal ultrasound is accurate to determine endoscopic response and remission in patients with moderate to severe ulcerative colitis: a longitudinal prospective cohort study

Floris de Voogd, MD[1,2], Elsa A. van Wassenaer, MD[1,3], Aart Mookhoek[4] MD PhD, Steven Bots, MD[1], Sara van Gennep, MD[1], Mark Löwenberg, MD PhD[1], Geert R. D'Haens, MD PhD,[1] Krisztina B. Gecse, MD PhD[1]

1. Amsterdam UMC, University of Amsterdam, Location AMC, Department of Gastroenterology and Hepatology, Amsterdam, The Netherlands
2. Amsterdam Gastroenterology Endocrinology Metabolism Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
3. Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Department of Pediatric Gastroenterology
4. Institute of Pathology, University of Bern, Bern, Switzerland

Short title: Intestinal ultrasound in monitoring patients with ulcerative colitis

Corresponding author:

Krisztina Gecse

Meibergdreef 9

1105 AZ Amsterdam, The Netherlands

E-mail: k.b.gecse@amsterdamumc.nl

Authors contact details:

Elsa van Wassenaer: e.a.vanwassenaer@amsterdamumc.nl

Aart Mookhoek: aart.mookhoek@pathology.unibe.ch

Steven Bots: s.j.bots@amsterdamumc.nl

Sara van Gennep: s.vangennep@amsterdamumc.nl

Mark Löwenberg: m.lowenberg@amsterdamumc.nl

Geert D'Haens: g.dhaens@amsterdamumc.nl

Krisztina B. Gecse: k.b.gecse@amsterdamumc.nl

Conflicts of interest:

Floris de Voogd: received speaker or honoraria fees from AbbVie and Janssen

Elsa van Wassenaer reports no conflicts of interest

Aart Mookhoek reports no conflicts of interest

Sara van Gennep reports no conflicts of interest

Steven Bots: has served as speaker for Abbvie, Merck, Sharp & Dome, Takeda, Jansen Cilag, Pfizer and Tillotts.

Mark Löwenberg: has served as speaker and/or principal investigator for: Abbvie, Celgene, Covidien, Dr. Falk, Ferring Pharmaceuticals, Gilead, GlaxoSmithKline, Janssen-Cilag, Merck Sharp & Dohme, Pfizer, Protagonist therapeutics, Receptos, Robarts Clinical Trials, Takeda, Tillotts, Tramedico. He has received research grants from AbbVie, Merck Sharp & Dohme, Dr. Falk, Achmea healthcare and ZonMW.

Geert D'Haens: has served as advisor for Abbvie, Ablynx, Amakem, AM Pharma, Avaxia, Biogen, Bristol Meiers Squibb, Boerhinger Ingelheim, Celgene, Celltrion, Cosmo, Covidien, Ferring, DrFALK Pharma, Engene, Galapagos, Gilead, Glaxo Smith Kline, Hospira, Immunic, Johnson and Johnson, Lycera, Medimetrics, Millenium/Takeda, Mitsubishi Pharma, Merck Sharp Dome, Mundipharma, Novonordisk, Pfizer, Prometheus laboratories/Nestle, Protagonist, Receptos, Robarts Clinical Trials, Salix, Sandoz, Setpoint, Shire, Teva, Tigenix, Tillotts, Topivert, Versant and Vifor and received speaker fees from Abbvie, Ferring, Johnson and Johnson, Merck Sharp Dome, Mundipharma, Norgine, Pfizer, Shire, Millenium/Takeda, Tillotts and Vifor.

Krisztina B. Gecse: has received grants from Pfizer Inc and Celltrion; consultancy fees from AbbVie, Arena Pharmaceuticals, Galapagos, Gilead, Immunic Therapeutics, Janssen Pharmaceuticals, Novartis, Pfizer Inc., Samsung Bioepis and Takeda and speaker's honoraria from Celltrion, Ferring, Janssen Pharmaceuticals, Novartis, Pfizer Inc, Samsung Bioepis, Takeda and Tillotts.

Authors' contributions:

F.V.: Study design, patient selection, data acquisition, data interpretation, writing first draft of the manuscript and final approval of the manuscript.

E.W.: Data acquisition, central reading, final approval of the manuscript

A.M.: Data acquisition, central reading, final approval of the manuscript

S.G.: Study design, patient selection, data acquisition, final approval of the manuscript

S.B.: Study design, patient selection, data acquisition, final approval of the manuscript

M.L: Study design, final approval of the manuscript

G.D.: Study design, patient selection, central reading, final approval of the manuscript.

K.G.: Study design, patient selection, data acquisition, critical revision and final approval of the manuscript.

Data Availability Statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Word Count: 6352/7000

Abstract

Background and aims

Intestinal ultrasound (IUS) is non-invasive, cost-effective and accurate to determine disease activity in ulcerative colitis (UC). In this study we prospectively evaluated IUS for treatment response in a longitudinal cohort by using endoscopy and histology as gold standards.

Methods

Consecutive patients with moderate-to-severe UC (endoscopic Mayo score (EMS)≥2) starting tofacitinib treatment were included. Patients were evaluated at baseline and after 8 weeks of

tofacitinib induction by means of clinical, biochemical, endoscopic (EMS and ulcerative colitis endoscopic index for severity (UCEIS)), histological (Robarts Histopathologic Index (RHI)) and IUS assessments. Readers of IUS, endoscopy and histology were blinded for all other outcomes. The primary outcome was difference in bowel wall thickness (BWT) for endoscopic improvement versus no endoscopic improvement. Endoscopic remission was defined as EMS=0, improvement as EMS≤1 and response as a decrease of EMS≥1.

Results

Thirty patients were included with 27 patients completing follow-up. BWT correlated with EMS (ρ=0.68, p<0.0001), UCEIS (ρ=0.73, p<0.0001) and RHI (ρ=0.49, p=0.002) at both time-points. BWT in the sigmoid was lower in patients with endoscopic remission (1.4mm vs 4.0mm, p=0.016), endoscopic improvement (1.8mm vs 4.5mm, p<0.0001) and decrease in BWT was more pronounced in patients with endoscopic response (-58.1% vs -13.4%, p=0.018). The most accurate cut-off values for BWT were 2.8 mm (AUC:0.87) for endoscopic remission, 3.9 mm (AUC:0.92) for improvement and decrease of 32% (AUC:0.87) for response. The submucosa was the most responsive wall layer.

Conclusion

IUS, importantly BWT as the single most important parameter, is highly accurate to detect treatment response when evaluated against endoscopic outcomes.

Key Words: intestinal ultrasound; ulcerative colitis; monitoring; treatment response; tofacitinib

Abstract word count=255/260

Introduction

Ulcerative colitis (UC) is chronic inflammatory bowel disease and is characterized by a relapsing-remitting pattern. Clinical symptoms, blood tests, fecal markers, cross-sectional imaging, endoscopy and histopathology are most commonly used to determine disease activity and severity. [1, 2] During follow-up, close monitoring is advocated to assess treatment efficacy and detect early relapse[2, 3].

Endoscopy is generally considered as the gold standard for the diagnosis and follow-up of patients with UC. However, endoscopy is an invasive and costly modality and therefore less attractive to perform frequently during the disease course. Non-invasive fecal biomarkers, most commonly fecal calprotectin (FCP), are frequently used to determine presence of inflammation and treatment response[4]. However, FCP does not supply information on disease extent and was shown to be only a fair surrogate marker for early endoscopic response[4, 5].

Intestinal ultrasound (IUS) is a non-invasive, easily accessible and low-cost alternative to visualize the colon and determine disease activity, extent and treatment response, which does not require bowel preparation [6, 7]. This makes it an ideal tool to use in the point-of-care setting to monitor disease activity and at the same time enable prompt decision making[8, 9]. Recent studies demonstrated that IUS can detect treatment response at two weeks after initiating anti-inflammatory treatment when compared to clinical response[10]. Previous studies in UC focused on the accuracy of IUS to determine endoscopic disease activity in cross-sectional cohorts[11, 12]. However, data is scarce on the responsiveness of IUS as evaluated in a longitudinal cohort using endoscopy (or histology) as gold standard, and cut-off values for IUS parameters corresponding with endoscopic outcomes are lacking.

In this current study, we prospectively evaluated IUS for treatment response after tofacitinib (Xeljanz, Pfizer) treatment in a longitudinal cohort of patients with moderate-to-severe UC by using endoscopy and histology as gold standards.

Methods

In this longitudinal prospective study consecutive patients ≥18 years of age with moderate-severe UC and an endoscopic Mayo score ≥2 in at least one colonic segment starting tofacitinib treatment within days after baseline endoscopy were eligible for inclusion. Exclusion criteria were pregnancy, confirmed infectious gastroenteritis, history of colectomy and imminent need for colectomy. This study was approved by the medical ethical committee of the Amsterdam University Medical Center. All patients gave written informed consent.

Procedures

Patient visits were at baseline and after eight weeks of treatment (tofacitinib, 10 mg bid). Medical history, demographics and disease phenotype (Montreal Classification) were recorded in the electronic patient record. At both visits, the Simple Clinical Colitis Activity Index (SCCAI) and the Lichtiger score were assessed[13] and at baseline, height, weight and Body Mass Index (BMI) were noted. Serum concentrations of C-reactive protein (CRP), albumin, leukocyte count, thrombocyte count, erythrocyte count and haemoglobin and fecal calprotectin (FCP) were measured at baseline and after 8 weeks of tofacitinib treatment. All data was entered in an electronic data management system (Castor EDC).

Intestinal ultrasound

IUS was performed at baseline and at week 8, on a different day than endoscopy by one ultrasonographer (F.V., three years of experience) with a Philips EPIQ 5G machine with a convex 5-1 probe and a linear 5-12 probe. The convex probe was used for screening all segments and visualization of the rectum. The linear probe was used to perform all measurements for the colonic segments. Patients were not fasting and did not receive bowel preparation.

First, we attempted to visualize the rectum and when visualized a cine-loop was recorded. Subsequently bowel wall thickness (BWT) was measured both in longitudinal and in cross-sectional

plane. Subsequently, all colonic bowel segments (sigmoid (SC), descending (DC), transverse (TC) and ascending colon (AC)) were visualized and cine-loops were recorded in longitudinal and cross-sectional plane for both B-mode and Colour Doppler. Settings were optimized for gain, focus and depth. With regard to Colour Doppler, velocity scale was set to 5 cm/s to optimize detection of small vessels within the bowel wall.

All cine-loops were anonymized, numbered and randomized by F.V.. Subsequently one ultrasonographer (E.W., three years of experience) was blinded to all other outcomes, read all cine-loops, and scored per segment for BWT, Colour Doppler Signal (CDS), loss of wall layer stratification (WLS), loss of haustrations, fatty wrapping and presence of lymph nodes as previously described (Supplementary Table 1)[12, 14-16]. A second reader (F.V., three years of experience) scored anonymized cine-loops again in random order (>3 months after initial IUS examination) for all IUS parameters, was blinded to the first reader and all other outcomes. Prior to scoring, readers agreed on the method for IUS measurements based on recent literature[14, 15]. The results for the first reader (E.W.) were used for further analysis. The results for the second reader (F.V.) were used for inter-observer agreement analysis. In addition, individual wall layers were measured in the anterior wall in each segment by the second reader. The mucosa was identified as the black layer between lumen and submucosa. Submucosa was identified as the grey layer between mucosa and muscularis propria. Muscularis propria was identified as the black layer between submucosa and the serosa (Figure 1). Individual wall layers were measured at the site of the BWT measurement.

Endoscopy

Patients underwent a complete colonoscopy or flexible sigmoidoscopy at the discretion of the treating physician at baseline and after eight weeks of treatment. Patients were prepared for the endoscopy as per routine care and during withdrawal the procedure was videotaped. During endoscopy two biopsies were taken between 15 cm and 30 cm from the anal verge in the most severely affected area

and when there was ulceration present from the edge of the ulcer. After eight weeks of treatment the repeated biopsy was acquired in the same segment as baseline at the same distance from the anus. When disease activity at baseline was EMS=0 in the SC, no biopsies were taken and we excluded the patient for histological analysis.

All videotaped loops were centrally read by one expert gastroenterologist (G.D.) for both the endoscopic Mayo score (EMS) and Ulcerative Colitis Endoscopic Index of Severity (UCEIS) for all available colonic segments and per colonic segment. For EMS, endoscopic remission was defined as EMS=0, endoscopic improvement as EMS≤1 and endoscopic response as a decrease of EMS≥1 point. For UCEIS, endoscopic remission and endoscopic response were defined as UCEIS=0 and a decrease of UCEIS ≥ 2 points. If all segments met EMS=0 this was defined as complete endoscopic remission [17, 18].

Pathology

Biopsies were collected, fixed in formalin and thereafter embedded in paraffin. Slides were centrally read by one gastrointestinal pathologist (A.M.) and scored with the Robarts Histopathology Index (RHI)[19]. Histological remission was defined as a RHI≤3 with absence of neutrophils in the lamina propria and epithelium and without ulcerations or erosions. Histological response was defined as a decrease in RHI≥7[19, 20]. Combined histo-endoscopical remission was defined as RHI≤3 and EMS=0.

Outcomes

The primary outcome of this study was BWT in the SC for patients reaching segmental endoscopic improvement versus patients without segmental endoscopic improvement in the SC after eight weeks of treatment with tofacitinib. Secondary outcomes included changes in IUS parameters in the SC and DC with and without endoscopic remission or endoscopic response. Exploratory outcomes were changes in IUS parameters for histological remission and response. Correlation between changes in individual IUS parameters with EMS and UCEIS per segment is investigated.

Sample size

We performed a sample-size calculation to detect a difference in BWT after eight weeks of treatment for patients reaching endoscopic improvement versus no improvement in the SC. In the phase 3 OCTAVE study, tofacitinib induced endoscopic remission and endoscopic improvement in 18.5% and 31% of patients, respectively, after eight weeks of treatment[21]. With regard to IUS, a recent study has shown a BWT difference in the SC of 2.08 mm with a SD of 1.42 mm in patients reaching clinical remission after 6 weeks[10]. With a level of significance of α=0.05 and 80% power we needed 6 patients to reach endoscopic improvement in the SC after eight weeks of treatment. Aiming at 25% endoscopic improvement after eight weeks we needed a total of 24 patients with EMS≥2 in the SC and BWT > 4 mm in the corresponding segment. To overcome poor acquisition of IUS due to disease activity limited to the distal sigmoid or rectum we aimed to include a total number of 30 patients starting tofacitinib.

Statistical analysis

All data were reported as median with inter-quartile range (IQR), mean with standard deviation (SD) or percentages of the total cohort when appropriate. Accuracy was reported as sensitivity, specificity and overall accuracy percentages. Paired dichotomous data were compared using a McNemar test. Paired continuous data were compared using a paired t-test and Wilcoxon-rank test for equally and not equally distributed data, respectively. Area under the ROC curve was used to determine cut-off values for BWT. A logistic regression with forward selection was used to determine odds ratios and create the most optimal IUS prediction model. Correlation was analysed with a Spearman's correlation coefficient (negligible: 0.00–0.09, weak: 0.10–0.39, moderate: 0.40–0.69, strong: 0.70–0.89, very strong: 0.90–1.00)[22] and inter-observer agreement per IUS parameter was assessed using Cohen's kappa statistics, Fleiss' kappa statistics (slight: 0.0-0.20, fair: 0.21-0.40, moderate: 0.41-0.60, substantial: 0.61-0.80 and perfect: 0.81-1.00)[23] or intra-class correlation coefficient (ICC) (poor: <0.50 poor, moderate: 0.50-0.75, substantial: 0.75-0.90 and strong agreement: 0.90-1.00) for dichotomous,

ordinal or continuous data, respectively[24]. A p-value ≤0.05 was considered statistically significant. All data was analysed using SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, N.Y., USA).

Results

A total of 30 consecutive patients started on tofacitinib and were included between October 2018 and August 2020 (Table 1). All but one patient had a SCCAI≥4 and all patients had a Lichtiger score ≥6 (Table 1). Three patients were excluded from further analysis as one patient (BMI = 43) had poor quality IUS cine-loops, one patient was lost to follow-up and one patient had proctitis only that could not be visualized with IUS. All other patients (n=27) completed follow-up and were included in the statistical analysis. At baseline, 24 (86%) and 17 (61%) patients had EMS≥2 in the SC and DC, respectively. In the three patients without EMS≥2 in the SC (EMS=1: n=2, EMS=0: n=1), the DC was scored as EMS≥2. With regards to IUS, at baseline, 25 patients (89%) had a pathological BWT in the SC (≥4.0 mm) and 18 patients (64%) in the DC (≥3.0 mm), respectively. Time between IUS and endoscopy at baseline was mean 5.6 ± 4.0 days.

After eight weeks, in 100%, 97% and 86% rectum, SC and DC were visualized by endoscopy, respectively. A total of three out of 27 (11%) patients had complete endoscopic remission according to EMS and UCEIS in all visualized segments. Endoscopic outcomes per segment after eight weeks is shown in Figure 2. In 22 of 27 patients, biopsies were available for analysis. Of these, five patients (23%) and 12 patients (55%) had histological remission and response in the sigmoid, respectively. Time between IUS and endoscopy was mean 3.6 ± 2.6 days.

Correlation between BWT and endoscopy or RHI

In a pooled analysis for baseline and after eight weeks, BWT in SC and DC showed moderate to strong correlation with EMS per segment (SC: ρ=0.68, p<0.0001, DC: ρ=0.76, p<0.0001) and UCEIS per segment (SC: ρ=0.73, p<0.0001, DC: ρ=0.74, p<0.0001). After eight weeks of treatment, decrease in

BWT showed moderate correlation with decrease in EMS (SC: ρ=0.50, p=0.009, DC: ρ=0.67, p=0.001) and UCEIS (SC: ρ=0.68, p<0.0001, DC: ρ=0.50, p=0.02) (Figure 3). There was a moderate correlation between RHI and BWT (ρ=0.49, p=0.002) in the SC.

Change in BWT in the sigmoid colon

Median BWT in the SC was significantly lower in patients with endoscopic improvement compared to patients without endoscopic improvement after eight weeks of treatment according to the EMS (1.8 mm [1.1-2.5] vs 4.5 mm [4.0-4.8], p<0.0001), Figure 4. For endoscopic remission, we found a similar pattern at eight weeks (EMS: 1.4 mm [1.1-2.4] vs 4.0 mm [2.0-4.6], p=0.016; UCEIS: 1.6 mm [1.1-2.5] vs 4.0 mm [2.0-4.6], p=0.016] according to the UCEIS (Figure 4).

Change in BWT was significantly more pronounced in patients with endoscopic response compared to no endoscopic response according to the EMS (p=0.018) and UCEIS (p=0.046), respectively (Figure 5). Data for BWT in the DC according to endoscopic outcomes is shown in Supplementary Table 2.

Cut-off values for endoscopic remission, endoscopic improvement and endoscopic response according to the EMS in the sigmoid

To determine endoscopic remission, a BWT of 2.8 mm (AUROC: 0.87, 95% CI: 0.74-1.000, p=0.006) had a sensitivity of 73% and a specificity of 100% in the SC. For endoscopic improvement we found 3.9 mm (AUROC: 0.92, 95% CI: 0.82-1.00, p<0.0001) with 81% sensitivity and 100% specificity to be the best cut-off. In addition, a decrease of 32% (AUROC: 0.87, 95% CI: 0.74-1.00, p=0.002) detected endoscopic response with 71% and 90% sensitivity and specificity, respectively. The data for UCEIS in the SC and DC for both endoscopic scores are demonstrated in Supplementary Table 3 and 4, respectively.

Individual wall layer thickness in the sigmoid

At baseline, the submucosa was the most thickened wall layer of the three bowel wall layers (Figure 6). Assessment of individual wall layer thickness was not possible in two patients due to extensive loss

of WLS. After eight weeks, wall layers were significantly different between patients with and without endoscopic improvement (mucosa: 0.6 mm [0.30-0.90] vs 1.0 mm [0.7-1.7], p=0.006; submucosa: 0.9 mm [0.6-1.1] vs 1.9 mm [1.4-2.8], p=0.005) and muscularis propria: 0.6 mm [0.3-0.8] vs 0.9 mm [0.7-1.4], p=0.04) (Supplementary Figure 1a) and endoscopic remission (mucosa: 0.7 mm [0.3-0.9] vs 0.9 mm [0.6-1.5], p=0.11; submucosa: 0.8 mm [0.5-1.0] vs 1.6 mm [0.9-2.7], p=0.03; muscularis propria: 0.4 mm [0.2-0.7] vs 0.8 mm [0.6-1.5], p=0.02) (Supplementary Figure 1b). The submucosa (1.6 mm [0.9-2.7]) was significantly thicker than the mucosa (0.9 mm [0.6-1.5], p=0.023) or muscularis propria (0.8 mm [0.6-1.5], p=0.027) in patients without endoscopic remission (Supplementary Figure 1c) and without endoscopic improvement (Supplementary Figure 1d) (submucosa=1.9 mm [1.4-2.8] vs mucosa=1.0 mm [0.7-1.7], p=0.026; submucosa vs muscularis propria=0.9 mm [0.7-1.4], p=0.016) at eight weeks, respectively-. The submucosa (0.80 mm [0.50-0.95] was not significantly thicker than the mucosa (0.70 mm [0.30-0.80], p=0.21) or muscularis propria (0.40 mm [0.20-0.70], p=0.11) in patients with endoscopic remission or endoscopic improvement (mucosa=0.60 mm [0.30-0.90] vs submucosa=0.90 mm [0.60-1.10], p=0.06; submucosa vs muscularis propria=0.60 mm [0.30-0.80], p=0.13). In patients with endoscopic response the decrease in wall layer thickness was most pronounced for the submucosa and significantly more pronounced compared to the mucosa (p=0.03)(Supplementary Figure 1e).

BWT and histology

There was a moderate correlation between RHI and BWT (ρ=0.49, p=0.002). For histological remission we found lower median BWT (1.9 mm [1.2-3.9] vs 3.9 mm [1.7-4.9], p=0.35) compared to no histological remission but this was not significant. Moreover, BWT was lower when, in addition to histological remission, endoscopic improvement (23%) (1.6 mm [1.1-2.6] vs 4.0 [1.7-6.0], p=0.10) or endoscopic remission (14%) (1.2 mm [1.1-1.2] vs 3.8 mm [1.8-5.3], p=0.05] was reached. In patients with endoscopic improvement nine (64%) patients were not in histological remission. No IUS parameter could detect histological remission in this subset of patients (data not shown). In a per wall

layer analysis, the median thickness of the mucosa was significantly lower when there was histological remission (0.6 mm [0.4-0.7] vs 1.0 mm [0.6-1.7], p=0.029) and we did not find this for the other wall layers.

In the assessment of histological response median BWT decreased significantly in patients with histological response (5.1 mm [2.6-5.5] to 3.2 mm [1.2-4.4], p=0.003) and no significant change was seen for patients without histological response (3.7 mm [1.8-5.1] to 1.9 mm [1.6-4.5], p=0.31).

BWT and clinical and biochemical parameters

After eight weeks, there was good correlation for SCCAI and BWT in the SC (rho=0.65, p<0.0001) and DC (ρ=0.59, p=0.002) and good correlation for the Lichtiger score and BWT in the SC (ρ=0.65, p=0.001) and DC (ρ=0.63, p=0.001), respectively. In addition, correlation for FCP and BWT in the SC (ρ=0.46, p=0.015) and DC (ρ=0.40, p=0.05) was moderate, respectively. Correlation for the segmental endoscopic scores and SCCAI, Lichtiger and FCP is demonstrated in Supplementary Table 5.

For the SC, a FCP level of 142 µg/g was most accurate (AUROC: 0.96, 95% CI: 0.88-1.00, p=0.001) to determine endoscopic remission with 91% sensitivity and 100% specificity. A level of 225 µg/g was accurate to determine endoscopic improvement (AUROC: 0.89, 95% CI: 0.78-1.00, p<0.0001) with 75% sensitivity and 87% specificity and endoscopic response could not be determined with a decrease in FCP (AUROC: 0.68, 95% CI: 0.46-0.90, p=0.13). A FCP level of 114 µg/g was most accurate (AUROC: 0.76, 95% CI: 0.57-0.94, p=0.03) to determine histological remission with 90% sensitivity and 67% specificity. FCP was not accurate to determine endoscopic (AUROC: 0.68, 95% CI: 0.46-0.90, p=0.13) and histological response (AUROC: 0.61, 95% CI: 0.44-0.92, p=0.44).

Correlation and univariable logistic regression for colour Doppler Signal and endoscopic outcomes

The distribution for the four CDS categories according to the endoscopic disease activity scores is demonstrated in Supplementary Figure 2. For CDS we found moderate to good correlation with EMS and UCEIS (Supplementary Table 7). After eight weeks of treatment there was weak to moderate

correlation with change in CDS and change in EMS (SC: ρ=0.47, p=0.025, DC: ρ=0.38, p=0.10) and change in UCEIS (ρ=0.60, p=0.003, DC: ρ=0.58, p=0.009). Normal CDS (CDS≤1) was associated with endoscopic remission (OR: 11.20, 95% CI: 1.23-101.89, p=0.017) and endoscopic improvement (OR: 11.56, 95% CI: 2.64-50.50, p=0.001). In addition, a decrease of one CDS category after eight weeks of treatment was associated with endoscopic response (OR: 2.50, 95%CI: 1.08-5.78, p=0.03) (Supplementary Table 7).

Correlation and univariable logistic regression for other IUS parameters and endoscopic outcomes

Distribution for the loss of WLS, loss of haustrations, presence of lymph nodes and presence of fatty wrapping according to the EMS and UCEIS is demonstrated in Supplementary figures 3-6. Correlation was best between endoscopic scores and loss of haustrations in the SC and DC, respectively (Supplementary Table 6). Loss of WLS (OR: 0.60 [0.46-0.79], p=0.003) and loss of haustrations (OR: 0.24 [0.11-0.51], p<0.0001) were inversely associated with endoscopic improvement in the SC (Supplementary Table 7). Univariate analysis for the other IUS parameters in SC and DC is also reported in Supplementary Table 7.

Combining IUS parameters in a multivariable logistic regression analysis

All patients with BWT ≤2.8 mm and ≤3.9 mm in the SC had endoscopic remission and endoscopic improvement, respectively. In addition, BWT> 2.8 mm was associated with no endoscopic remission (OR: 2.0, 95% CI: 1.14-3.52, p=0.002) and BWT >3.9 mm was associated with no endoscopic improvement (OR: 5.00, 95% CI: 1.82-13,76, p<0.0001) after eight weeks. A BWT decrease of 32% in the SC highly indicated endoscopic response (OR 21.60, 95% CI: 2.14-218.58, p=0.009). Normalization of haustrations further improved the association model (OR: 19.17, 95% CI: 1.49-247.00, p=0.024) whereas other IUS parameters were of no additional value to determine any of the endoscopic endpoints. The addition of FCP, decrease of FCP or cut-off values for FCP did not improve the model to detect endoscopic remission, improvement or response, respectively.

Data on BWT cut-off values in the DC are demonstrated in Supplementary Table 8. In addition to a BWT cut-off of 3.6 mm, decrease in 1-point CDS (OR: 11.11, 95% CI: 1.03-125, p=0.05) improved the model to determine endoscopic improvement. Other IUS parameters or FCP were of no additional value in a multivariable model to determine any endoscopic remission, endoscopic improvement or endoscopic response in the DC, respectively.

Disease extent

For all endoscopic procedures (n=54) the proximal extent of the disease was reached in 39 (72%) cases with 11 (20%) endoscopies showing moderate to severe disease activity (EMS≥2) in the AC and in an additional 3 patients in the TC (n=14, 26%), respectively. For the presence of disease activity (EMS≥2), disease extent at endoscopy showed good correlation (ρ=0.75, p<0.0001) with IUS.

Inter-observer agreement for IUS parameters

Inter-observer agreement for BWT was almost perfect for the SC (ICC: 0.92, 95% CI: 0.86-0.96, p<0.0001) and DC (ICC: 0.89 (95% CI: 0.78-0.94, p<0.0001) (Supplementary Figure 7). Inter-observer agreement for the other IUS parameters is demonstrated in Supplementary Table 9.

Discussion

In this study, we demonstrated that IUS, with BWT as the single most important parameter, is accurate to determine treatment response to tofacitinib in patients with moderate-severe UC when compared against endoscopy. Furthermore, we demonstrated that the submucosa was the most responsive layer for endoscopic response and remains significantly thickened when endoscopic remission or endoscopic improvement are not reached. On the other hand, mucosal change was significantly associated with histological remission. Accurate cut-off values for endoscopic remission (≤2.8 mm),

endoscopic improvement (≤3.9 mm) and endoscopic response (decrease of 32%) have been determined.

A recent systematic review by Smith *et al.* concluded that BWT is the predominant parameter in diagnosing disease flares in UC patients. However, the cut-off for disease activity of 3.0 or 4.0 mm is often debated[7]. This can be due to the fact that definitions of outcomes in previous studies were not consistent and studies lacked assessment of treatment response. In previous cross-sectional studies, Allocca *et al*[11] and Bots *et al*[12] suggested 3.0 mm as cut-off value in the most severely affected segment independent of segment to determine endoscopic improvement (EMS≤1). In our longitudinal cohort, we found higher cut-off values to reflect endoscopic improvement, which might be due to the segmental analysis. The segmental assessment might further explain the slight discrepancy with previous studies and several segments might show different response rates[25, 26]. In addition, the early evaluation of BWT after eight weeks of treatment might also explain the higher cut-off values and it is likely that the BWT decreases further in the subsequent weeks.

Besides endoscopic improvement, we demonstrated an accurate BWT cut-off value for endoscopic remission. Bots *et al* has demonstrated BWT to distinguish EMS=0 from EMS=1[12]. Although the cut-off value of 2.8 mm is higher than the 2.1 mm previously reported by Bots *et al*, it could demonstrate that IUS has potential to determine endoscopic remission from endoscopic improvement. Distinction between these two outcomes by IUS is of significant value, as endoscopic remission is associated with improved long-term outcomes[27].

With regard to endoscopic response, we found change in BWT to correlate with change in EMS and change in UCEIS. Recently, Ilvemark *et al* reported a decrease of 25% in BWT reflecting treatment response, which was predominantly based on an expert opinion rather than data supporting this statement. In agreement, in the current study, we found a 20% to 32% decrease in BWT to accurately reflect segmental endoscopic response according to frequently utilized endoscopic response definitions[17, 18]. Change in BWT was more pronounced in the sigmoid colon than in the descending

colon which might demonstrate a segmental difference in response, which is also supported by recent studies[18, 25]. Furthermore, we included three patients with EMS<2 in the sigmoid and EMS≥2 in the descending colon at baseline demonstrating a segmental difference in disease activity in some patients. On the other hand, the descending colon was in general less thickened at baseline as we included all patients, regardless of EMS, in the BWT analysis for this segment. This resulted in a lower BWT at baseline, consequently with a lower change in BWT when there was response.

In addition to IUS, FCP is a non-invasive alternative for treatment evaluation and several cut-off values have been proposed to determine endoscopic endpoints[5, 28, 29]. However, FCP or decrease of FCP was not accurate to determine segmental endoscopic remission, improvement or response in our study. In patients with a low FCP and absence of clinical symptoms, endoscopic remission or improvement is very likely to be present. However, when treatment does not lead to complete remission, IUS allows evaluation of (segmental) treatment response and could guide further treatment decision making. In clinical practice, the combination of clinical indices, FCP and IUS are likely sufficient to evaluate treatment response for most patients, thereby allowing a more non-invasive but objective approach[9, 30].

In a per wall layer analysis, interestingly the submucosa was most thickened in a predominantly mucosal disease. Furthermore, the submucosa showed most pronounced decrease when there was endoscopic response. A recent study suggested edema within the submucosa resulting in a thickened submucosa[10]. Most vessels are present in the submucosa and hyperaemia at IUS is predominantly present in this wall layer. This might support the role of edema leading to an increased submucosal layer thickness but future studies should correlate IUS findings to histopathological data from resection specimens to confirm this. Moreover, Gordon *et al* found most fibrosis in colectomy resection specimens in the submucosa of UC patients[31]. These results, together with our data, suggest that disease activity is not limited to the mucosa but reach further into the bowel wall. Gordon *et al* also

found fibrosis in the submucosa to contribute substantially to wall layer thickness[31, 32]. In our cohort, we included patients that were difficult to treat (almost half of the patients had been previously treated with at least two biologicals) with likely substantial chronic inflammation which might have led to an increased submucosal wall thickness. However, since the submucosa also decreases when there is endoscopic treatment response not only patients with end-stage UC have an affected submucosa but also in patients with a disease flare.

In addition to endoscopic response, we also found a moderate correlation for BWT with the RHI. There are no previous studies assessing IUS parameters according to histological disease activity. We found the mucosa to be significantly less thickened when there was histological remission, whereas this was not found for the other bowel wall layers. Biopsies are limited to the mucosa whereas, as we have demonstrated, treatment response is not limited to the mucosa in UC. Although assessment of histological disease activity is increasingly studied in UC[33], response might not be limited to the mucosa and transmural evaluation with IUS might also be of merit in monitoring response. In this study, biopsies from 22 patients were available at both time-points and therefore larger studies evaluating both IUS and histology are warranted to confirm our findings.

Recently developed IUS scoring indices for UC incorporating BWT, CDS, loss of haustrations and fatty wrapping, found a strong to moderate correlation with endoscopic scoring indices and accurately detected endoscopic remission and endoscopic improvement[8, 11, 12, 34]. However, the definition of endoscopic reference standards was heterogeneous among studies. In our cohort, evaluation of CDS and haustration pattern, in addition to BWT, contributed to assessment of endoscopic response whereas for endoscopic improvement and endoscopic remission only BWT was accurate. Maaser *et al* found decrease in BWT to precede normalization of CDS[10]. In this study, we performed a second endoscopy at an early time-point and while BWT might already have decreased, CDS was yet to normalize. In addition, another recent score incorporating BWT, CDS and loss of haustrations was

highly accurate to detect endoscopic improvement[12]. Consequently, BWT in combination with CDS and haustration pattern might indicate subtle endoscopic changes in an early phase.

Although IUS showed high accuracy to detect response to treatment, it is traditionally considered operator dependent. On the contrary, a recent study found almost perfect inter-observer agreement for BWT and moderate agreement for CDS[14]. In our study, we found similar inter-observer agreements indicating that BWT and CDS are reliable parameters to score. Another limitation of IUS is the visualization of the rectum. Although FCP could be a non-invasive surrogate marker for endoscopic outcomes in the rectum, we had a low number of patients to analyze this potential role. Alternatively, perineal ultrasound could be of interest to determine treatment response specifically for the rectum[35].

Our study has its limitations. Firstly, our cohort included only patients treated with one mode of action, namely tofacitinib and our results might therefore not reflect treatment response to other anti-inflammatory drugs. However, Maaser *et al* found similar results in a cohort of patients treated with different anti-inflammatory therapies suggesting that our results can be extrapolated to other anti-inflammatory treatments[10]. Secondly, eight weeks is an early time-point to assess endoscopic treatment response and consequently we had a limited number of patients with complete endoscopic remission. However, IUS has the potential to determine treatment response in an early phase and our primary aim was to investigate the accuracy of IUS to objectify response, remission and endoscopic improvement against endoscopy as gold standard. Thirdly, there were limited number of patients to test multiple IUS parameters according to multiple endoscopic and histopathological outcomes which could have caused type one errors.

The current study also had several strengths. Firstly, the complete spectrum of endpoints (clinical, biochemical, mucosal, transmural and histological) was evaluated in a longitudinal real-life cohort with stringent follow-up. Notably, as drugs of different mechanism of action potentially exhibit differences in time to response[3], our cohort represents a homogenous population. Importantly, IUS, endoscopy

and histology were evaluated by readers blinded to other outcomes and we analyzed individual bowel wall layers in assessing treatment response.

For the near future, larger cohorts are warranted to confirm our findings. The reliability of individual wall layer measurements should be further investigated and correlated to histopathologic findings in resection specimens. Further studies should also address the most ideal time-point for IUS to evaluate and to predict treatment response, for drugs of different mechanism of action.

In conclusion, in a prospective, longitudinal cohort we have demonstrated that IUS, and in particular BWT as a single most important parameter, is highly accurate in identifying segmental endoscopic remission, endoscopic improvement and response when compared to globally utilized endoscopic scoring indices and it is even able to detect microscopic treatment response in patients UC. We have also shown that not only the mucosa, but also the submucosa is affected in acute inflammation in UC. Our findings could be paradigm changing in shifting towards less invasive monitoring during follow-up in the treat-to-target setting.

References

1. 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. *Journal of Crohn's and Colitis*. 2019;13(2):144-164K.
2. 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. *Journal of Crohn's and Colitis*. 2017;11(6):649-670.
3. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. *Gastroenterology*. 2021;160(5):1570-1583.
4. Stevens TW, Gecse K, Turner JR, de Hertogh G, Rubin DT, D'Haens GR. Diagnostic accuracy of fecal calprotectin concentration in evaluating therapeutic outcomes of patients with ulcerative colitis. *Clinical Gastroenterology and Hepatology*. 2020;
5. Reinisch W, Bressler B, Curtis R, et al. Fecal calprotectin responses following induction therapy with vedolizumab in moderate to severe ulcerative colitis: a post hoc analysis of GEMINI 1. *Inflammatory bowel diseases*. 2019;25(4):803-810.
6. Maconi G. Ultrasonography in the evaluation of extension, activity, and follow-up of ulcerative colitis. *Scandinavian journal of gastroenterology*. 1999;34(11):1103-1107.
7. Smith RL, Taylor KM, Friedman AB, Gibson RN, Gibson PR. Systematic review: Clinical utility of gastrointestinal ultrasound in the diagnosis, assessment and management of patients with ulcerative colitis. *Journal of Crohn's and Colitis*. 2019;
8. Smith RL, Taylor KM, Friedman AB, Gibson RN, Gibson PR. Systematic review: Clinical utility of gastrointestinal ultrasound in the diagnosis, assessment and management of patients with ulcerative colitis. *Journal of Crohn's and Colitis*. 2020;14(4):465-479.
9. Bots S, De Voogd F, De Jong M, et al. Point-of-care intestinal ultrasound in IBD patients: disease management and diagnostic yield in a real-world cohort and proposal of a point-of-care algorithm. *Journal of Crohn's and Colitis*. 2021;
10. 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(9):1629-1636.
11. 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. *Journal of Crohn's and Colitis*. 2018;12(12):1385-1391.
12. Bots S, Nylund K, Löwenberg M, Gecse K, D'Haens G. Intestinal Ultrasound to Assess Disease Activity in Ulcerative Colitis: Development of a novel UC-Ultrasound index. *Journal of Crohn's and Colitis*. 2021;
13. Turner D, Seow CH, Greenberg GR, Griffiths AM, Silverberg MS, Steinhart AH. A systematic prospective comparison of noninvasive disease activity indices in ulcerative colitis. *Clinical Gastroenterology and Hepatology*. 2009;7(10):1081-1088.
14. De Voogd F, Wilkens R, Gecse K, et al. A reliability study-strong inter-observer agreement of an expert panel for intestinal ultrasound in ulcerative colitis. *Journal of Crohn's and Colitis*. 2021;
15. Ilvemark JF, Hansen T, Goodsall TM, et al. Defining transabdominal Intestinal Ultrasound treatment response and remission in Inflammatory Bowel Disease: Systematic review and expert consensus statement. 2021;
16. Novak KL, Nylund K, Maaser C, et al. Expert consensus on optimal acquisition and development of the International Bowel Ultrasound Segmental Activity Score [IBUS-SAS]: a reliability

and inter-rater variability study on intestinal ultrasonography in Crohn's disease. *Journal of Crohn's and Colitis*. 2021;15(4):609-616.
17. Vuitton L, Peyrin-Biroulet L, Colombel J, et al. Defining endoscopic response and remission in ulcerative colitis clinical trials: an international consensus. *Alimentary pharmacology & therapeutics*. 2017;45(6):801-813.
18. Lobatón T, Bessissow T, De Hertogh G, et al. The Modified Mayo Endoscopic Score (MMES): a new index for the assessment of extension and severity of endoscopic activity in ulcerative colitis patients. *Journal of Crohn's and Colitis*. 2015;9(10):846-852.
19. Pai RK, Khanna R, D'Haens GR, et al. Definitions of response and remission for the Robarts Histopathology Index. *Gut*. 2019;68(11):2101-2102.
20. Magro F, Lopes J, Borralho P, et al. Comparing the continuous Geboes score with the Robarts histopathology index: definitions of histological remission and response and their relation to faecal calprotectin levels. *Journal of Crohn's and Colitis*. 2020;14(2):169-175.
21. Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as induction and maintenance therapy for ulcerative colitis. *New England Journal of Medicine*. 2017;376(18):1723-1736.
22. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*. 2018;126(5):1763-1768.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *biometrics*. 1977:159-174.
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*. 2016;15(2):155-163.
25. Christensen B, Hanauer SB, Gibson PR, Turner JR, Hart J, Rubin DT. Segmental Histological Normalisation Occurs in Ulcerative Colitis but Does Not Improve Clinical Outcomes. *Journal of Crohn's and Colitis*. 2020;14(10):1345-1353.
26. 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. *Clinical Gastroenterology and Hepatology*. 2021;19(5):908-921. e6.
27. Boal Carvalho P, Dias de Castro F, Rosa B, Moreira MJ, Cotter J. Mucosal healing in ulcerative colitis–when zero is better. *Journal of Crohn's and Colitis*. 2016;10(1):20-25.
28. Sandborn WJ, Panés J, Zhang H, Yu D, Niezychowski W, Su C. Correlation between concentrations of fecal calprotectin and outcomes of patients with ulcerative colitis in a phase 2 trial. *Gastroenterology*. 2016;150(1):96-102.
29. Stevens TW, Gecse K, Turner JR, de Hertogh G, Rubin DT, D'Haens GR. Diagnostic accuracy of fecal calprotectin concentration in evaluating therapeutic outcomes of patients with ulcerative colitis. *Clinical Gastroenterology and Hepatology*. 2021;19(11):2333-2342.
30. Wilkens R, Dolinger M, Burisch J, Maaser C. Point-of-care testing and home testing: pragmatic considerations for widespread incorporation: stool tests, serum tests, intestinal ultrasound. *Gastroenterology*. 2022;
31. Gordon IO, Agrawal N, Willis E, et al. Fibrosis in ulcerative colitis is directly linked to severity and chronicity of mucosal inflammation. *Alimentary pharmacology & therapeutics*. 2018;47(7):922-939.
32. Gordon IO, Agrawal N, Goldblum JR, Fiocchi C, Rieder F. Fibrosis in ulcerative colitis: mechanisms, features, and consequences of a neglected problem. *Inflammatory bowel diseases*. 2014;20(11):2198-2206.
33. Chateau T, Feakins R, Marchal-Bressenot A, Magro F, Danese S, Peyrin-Biroulet L. Histological remission in ulcerative colitis: under the microscope is the cure. *American Journal of Gastroenterology*. 2020;115(2):179-189.
34. Parente F, Molteni M, Marino B, et al. Bowel ultrasound and mucosal healing in ulcerative colitis. *Digestive Diseases*. 2009;27(3):285-290.
35. Sagami S, Kobayashi T, Aihara K, et al. Transperineal ultrasound predicts endoscopic and histological healing in ulcerative colitis. *Alimentary Pharmacology & Therapeutics*. 2020;

Figure and Table Legend

Table 1: Baseline characteristics. [SD=standard deviation, SCCAI=Simple Clinical Colitis Activity Index, CRP=C-reactive protein, EMS=Endoscopic Mayo Score, UCEIS=Ulcerative Colitis Endoscopic Index or Severity]

Figure 1: individual wall layers in the sigmoid colon

Figure 2: Percentage of patients that reached segmental endoscopic response (decrease ≥1 EMS), endoscopic improvement (EMS≤1) and endoscopic remission (EMS=0) after 8 weeks of treatment with tofacitinib when baseline EMS≥2 in the corresponding segment. [Rectum: 46%, 33%, 20%; Sigmoid colon: 40%, 30%, 9%; Descending colon: 60%, 53%, 38%].

Figure 3 (A-B): Correlation for decrease in BWT and decrease in EMS and UCEIS. A: decrease in BWT versus decrease in EMS (ρ=0.50, p=0.009) or UCEIS (ρ=0.68, p<0.0001) in the SC. B: decrease in BWT versus decrease in EMS (ρ=0.67, p=0.001) or UCEIS (ρ=0.50, p=0.02) in the DC. [SC=sigmoid colon, DC=descending colon, EMS=Endoscopic Mayo Score, UCEIS: Ulcerative Colitis Endoscopic Index or Severity)

Figure 4: Difference for BWT between different endoscopic outcomes. (UCEIS: ulcerative colitis endoscopic index of severity). A: Difference between patients with and without endoscopic improvement after eight weeks in the sigmoid colon at baseline B: Difference between patients with and without endoscopic remission after 8 weeks in the sigmoid colon at baseline and week 8 according to the endoscopic Mayo score.; C: Difference between patients with and without endoscopic remission after eight weeks in the sigmoid colon at baseline and week eight according to the UCEIS.

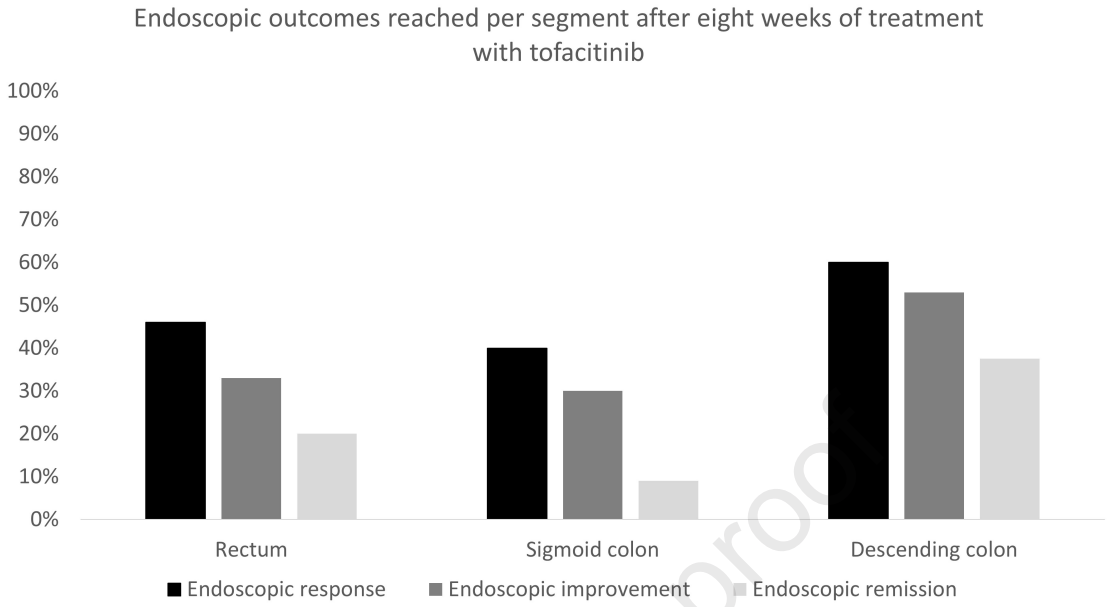
Figure 5: Change in bowel wall thickness (%) in patients with and without endoscopic response. A: change in bowel wall thickness (%) in patients with and without endoscopic response according to the endoscopic Mayo score (median ΔBWT: -58.1% [-67.6- -35.5] vs -13.4% [-40.2- -0.4], p=0.018); B: change in bowel wall thickness (%) in patients with and without endoscopic response according to the UCEIS (median ΔBWT: -49.6% [-68.1- -32.3] vs -18.1% [-55.8- -1.8], p=0.046). [UCEIS: ulcerative colitis endoscopic index of severity]

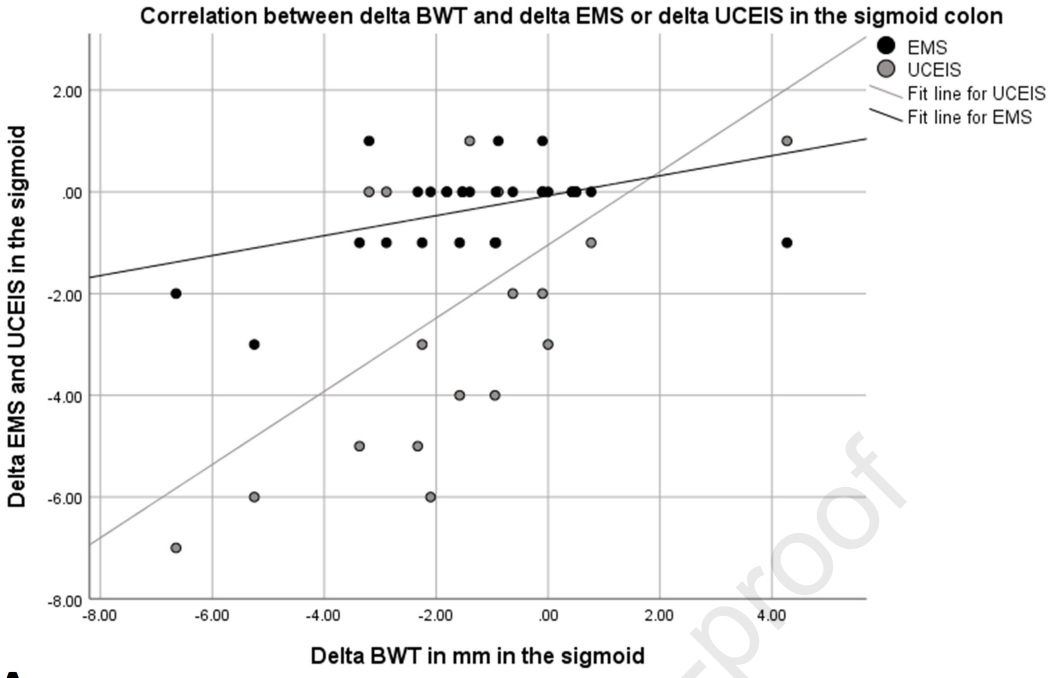
Figure 6: Wall layer thickness in the sigmoid colon for mucosa (median: 1.1 mm [0.7-1.5], submucosa (median: 2.0 mm [1.2-2.6]) and muscularis propria (median: 1.2 mm [0.9-1.8]) when there was active endoscopic disease (EMS≥2) in the sigmoid colon. (EMS=endoscopic Mayo score)

	n=30
Age in years (median)[IQR]	35.5 [27-53]
Gender (female)	18 (60%)
Body Mass Index (median)[IQR]	23.36 [17.27-43.47]
Disease duration in months (median)[IQR]	78 [56-135]
Montreal classification • Proctitis (E1) • Left-sided colitis (E2) • Pancolitis (E3)	 1 (3%) 17 (57%) 12 (40%)
Previous medication history • Aminosalicates • Corticosteroids • Thiopurine • Methotrexate • Biologicals ○ None ○ One ○ Two ○ Three • Previous anti-TNF exposure	 27 (90%) 27 (90%) 21 (70%) 3 (10%) 9 (30%) 7 (23%) 7 (23%) 7 (23%) 21 (70%)
Concomitant medication at the start of tofactinib • Aminosalicates • Corticosteroids • None	 10 (33%) 15 (50%) 10 (33%)
Clinical disease activity • SCCAI (median)[IQR] • Lichtiger score (median)[IQR]	 9.50 [6.50-11] 11 [7.50-14]
Biochemical parameters • CRP in mg/L (mean±SD) • Haemoglobin mmol/L (mean±SD) • Leukocytes 10^9/L (mean±SD) • Thrombocytes 10^9/L (mean±SD) • Erythrocyte count 10^{12}/L (mean±SD) • Albumin g/L (mean±SD) • Faecal calprotectin mg/kg (mean±SD)	 19.48±42.55 7.94±1.11 8.71±3.41 316±100 4.32±0.61 40.62±5.32 2097±2056
Endoscopy at baseline • Rectum ○ EMS=0 ○ EMS=1 ○ EMS=2 ○ EMS=3 • Sigmoid colon ○ EMS=0 ○ EMS=1 ○ EMS=2 ○ EMS=3	 3 (10%) 2 (7%) 11 (37%) 14 (46%) 2 (7%) 3 (10%) 12 (40%) 13 (43%)

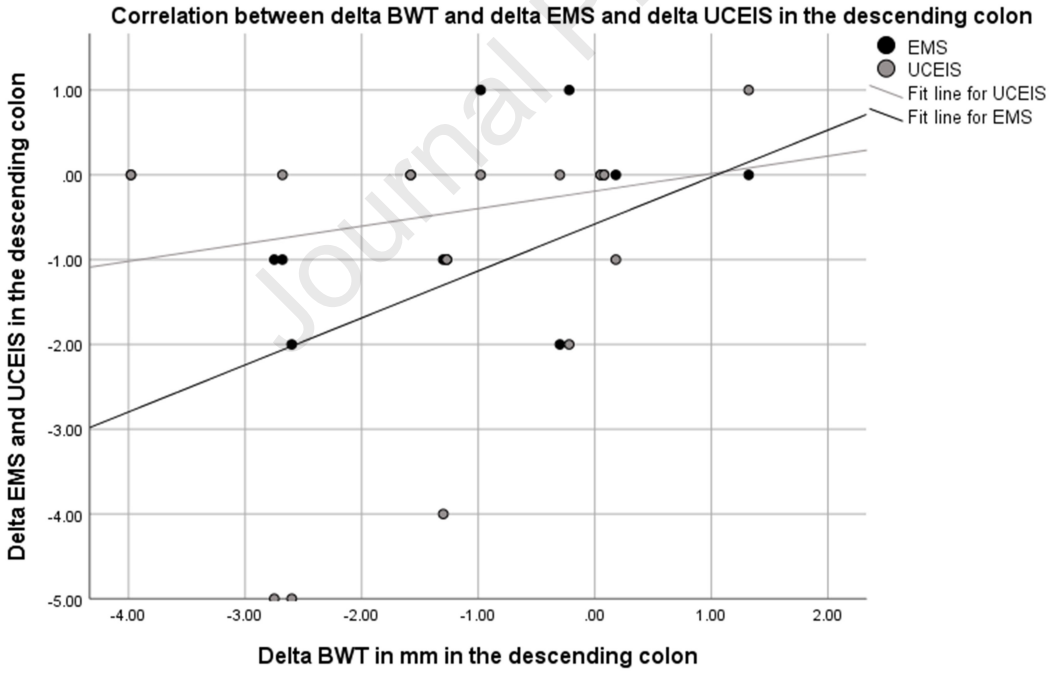
• Descending colon	
○ Not reached	2 (7%)
○ EMS=0	9 (30%)
○ EMS=1	3 (10%)
○ EMS=2	9 (30%)
○ EMS=3	7 (23%)
• Most affected segment	
○ Rectum	7 (24%)
○ Sigmoid colon	22 (73%)
○ Descending colon	1 (3%)
IUS parameters at baseline	
• Sigmoid colon	
○ Bowel wall thickness in mm [IQR]	5.1 [3.7-6.2]
○ Pathological bowel wall thickness (≥4.0 mm)	25 (89%)
○ Colour Doppler Signal (≥2)	19 (63%)
○ Loss of wall layer stratification (≥1)	10 (33%)
○ Loss of haustrations (=1)	20 (74%)
○ Presence of lymph nodes (=1)	15 (27%)
○ Fatty wrapping (=1)	9 (33%)
• Descending colon	
○ Bowel wall thickness [IQR]	4.0 [2.8-5.5]
○ Pathological bowel wall thickness (≥3.0 mm)	18 (64%)
○ Colour Doppler Signal (≥2)	13 (48%)
○ Loss of wall layer stratification (≥1)	7 (26%)
○ Loss of haustrations (=1)	14 (52%)
○ Presence of lymph nodes (=1)	8 (30%)
○ Fatty wrapping (=1)	5 (19%)
Robarts Histopathology Index at baseline [IQR]	16 [10-26]

Table 1: Baseline characteristics [IQR: interquartile range; SD: standard deviation; TNF: tumour necrosis factor; SCCAI: simple clinical colitis activity index; EMS: endoscopic Mayo score]

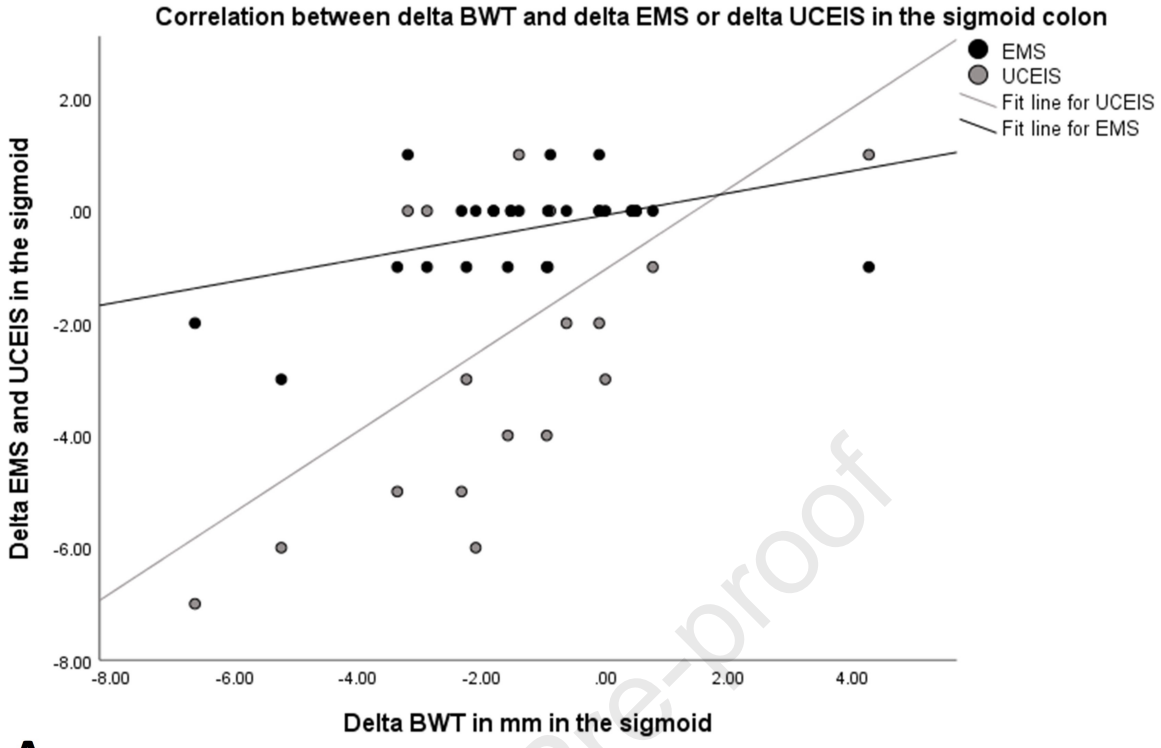




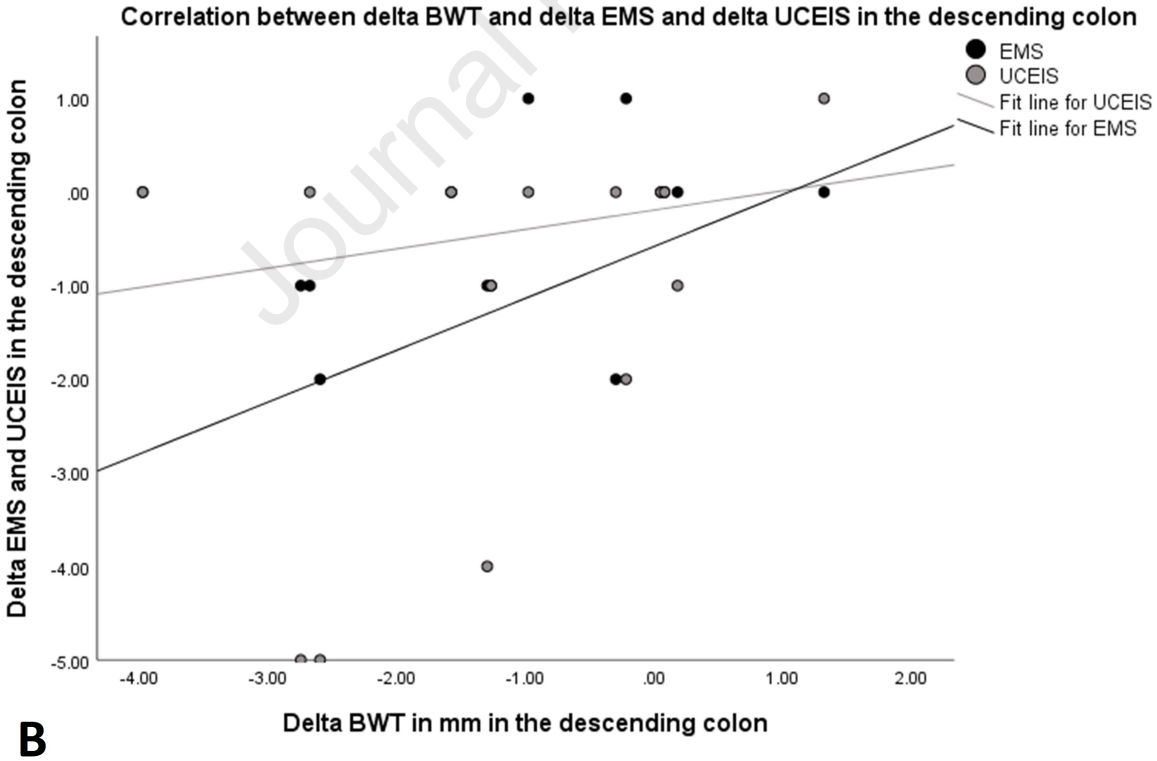
A



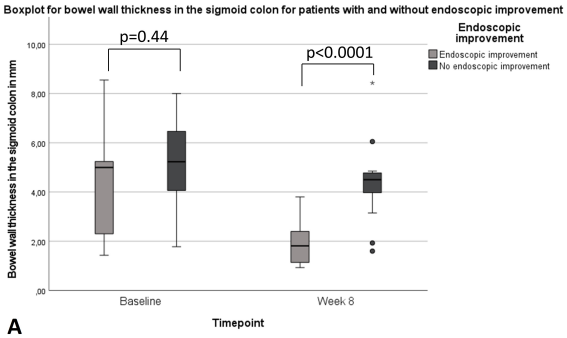
B



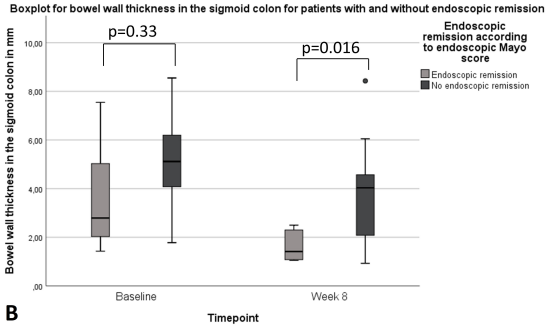
A



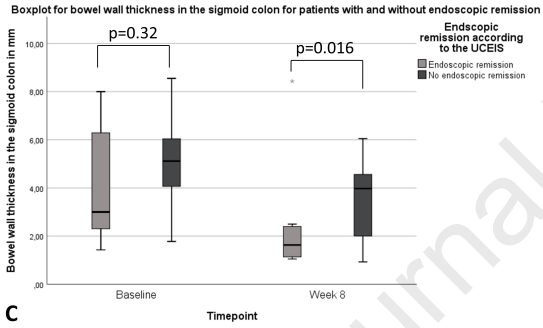
B



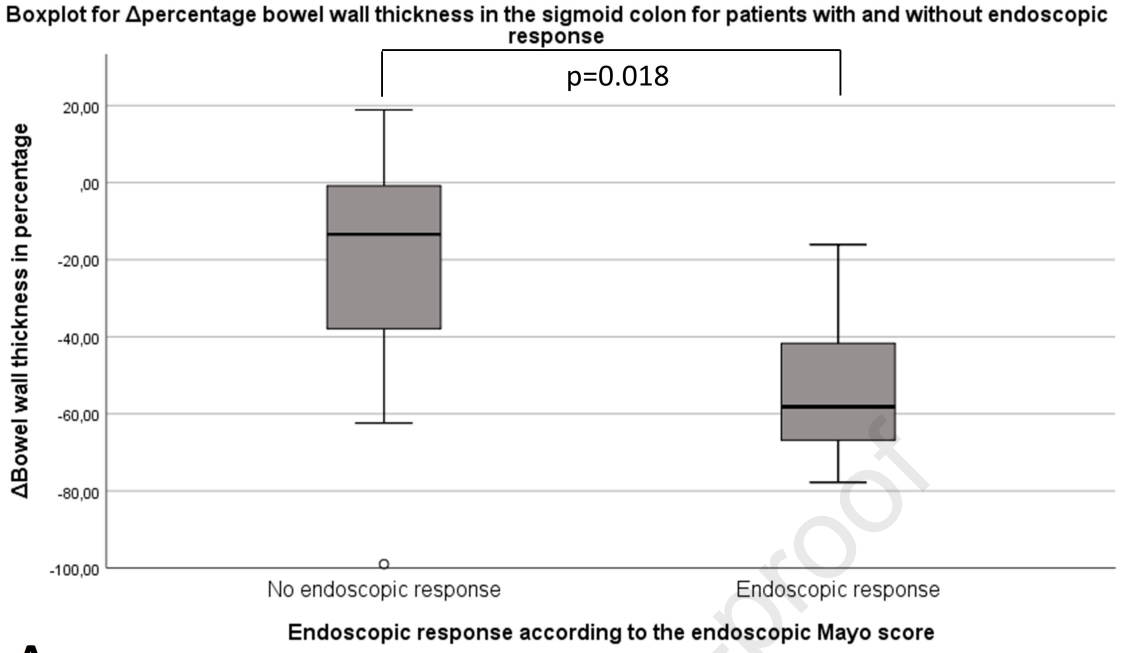
A



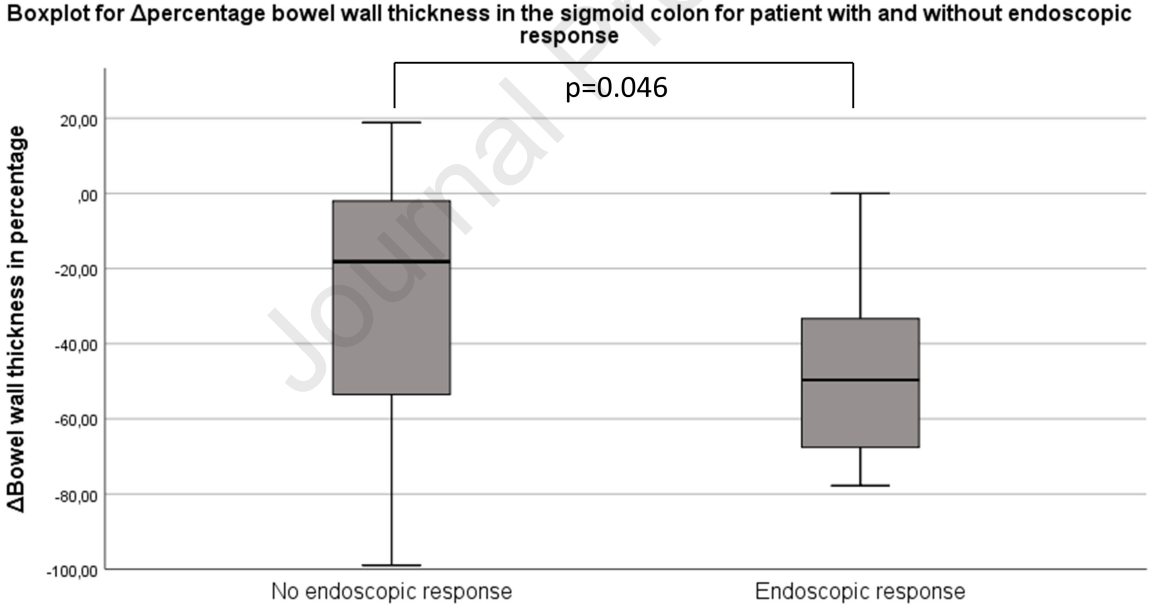
B



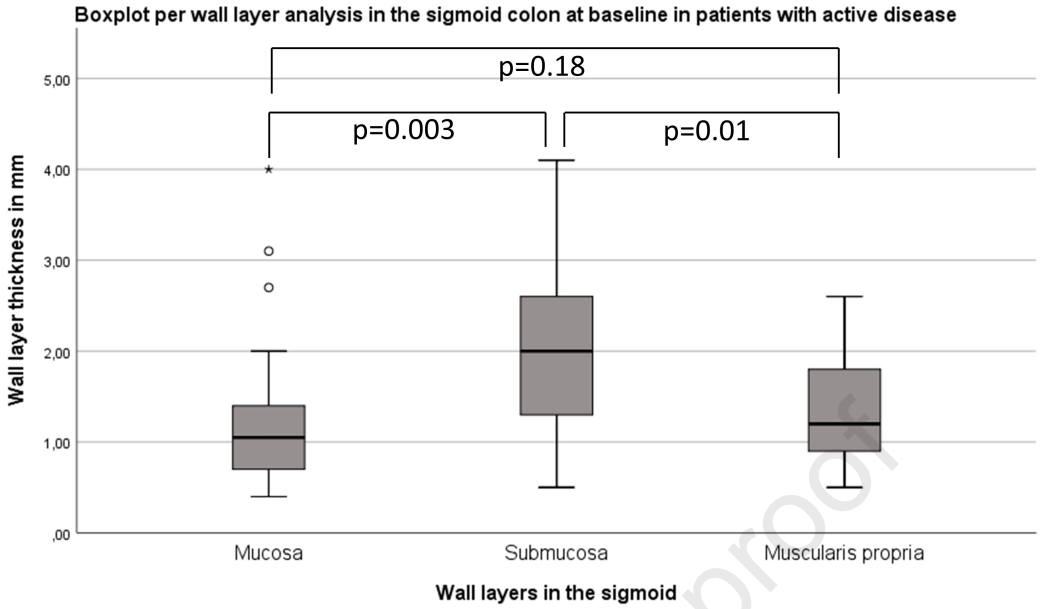
C

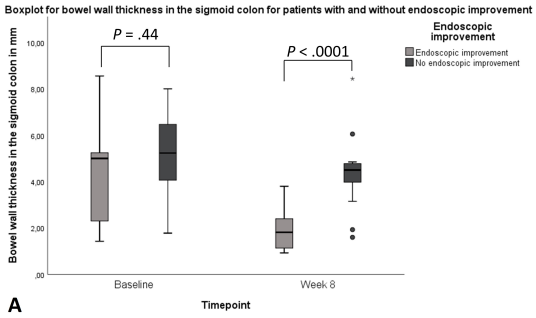


A

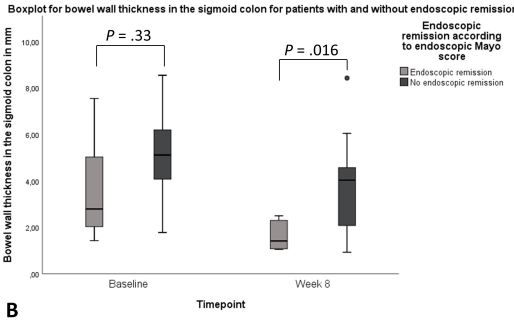


B

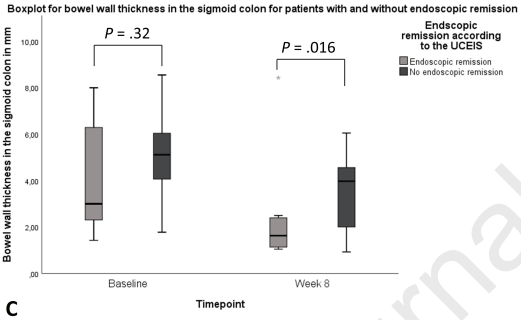




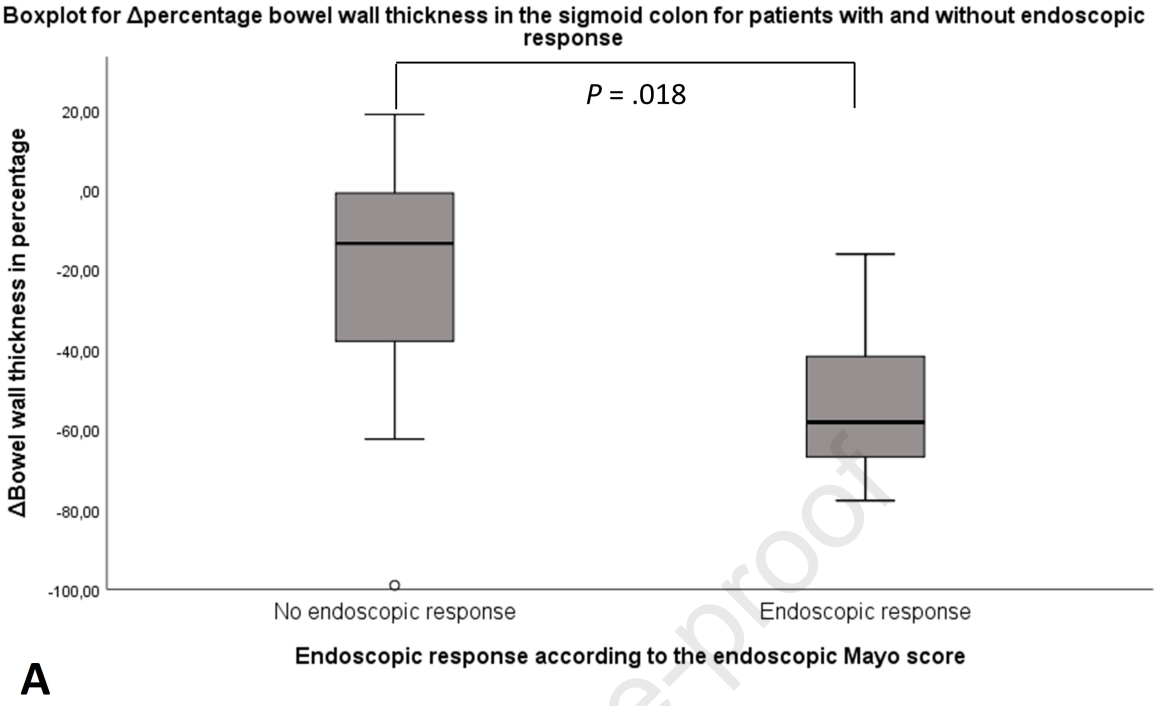
A



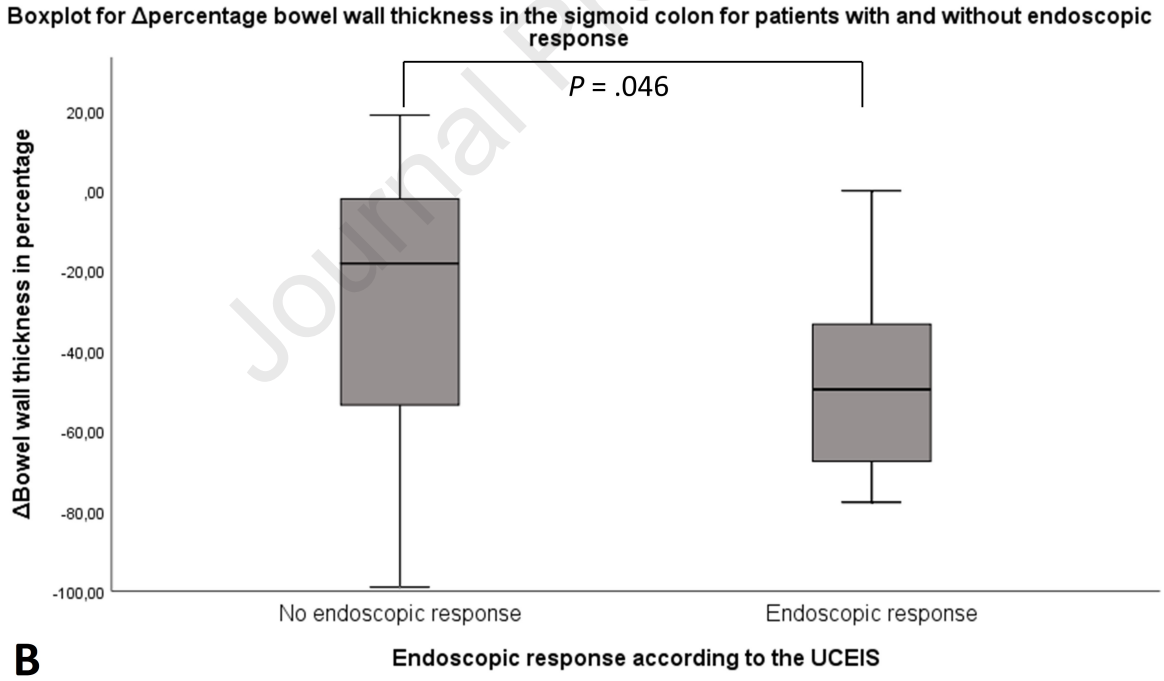
B



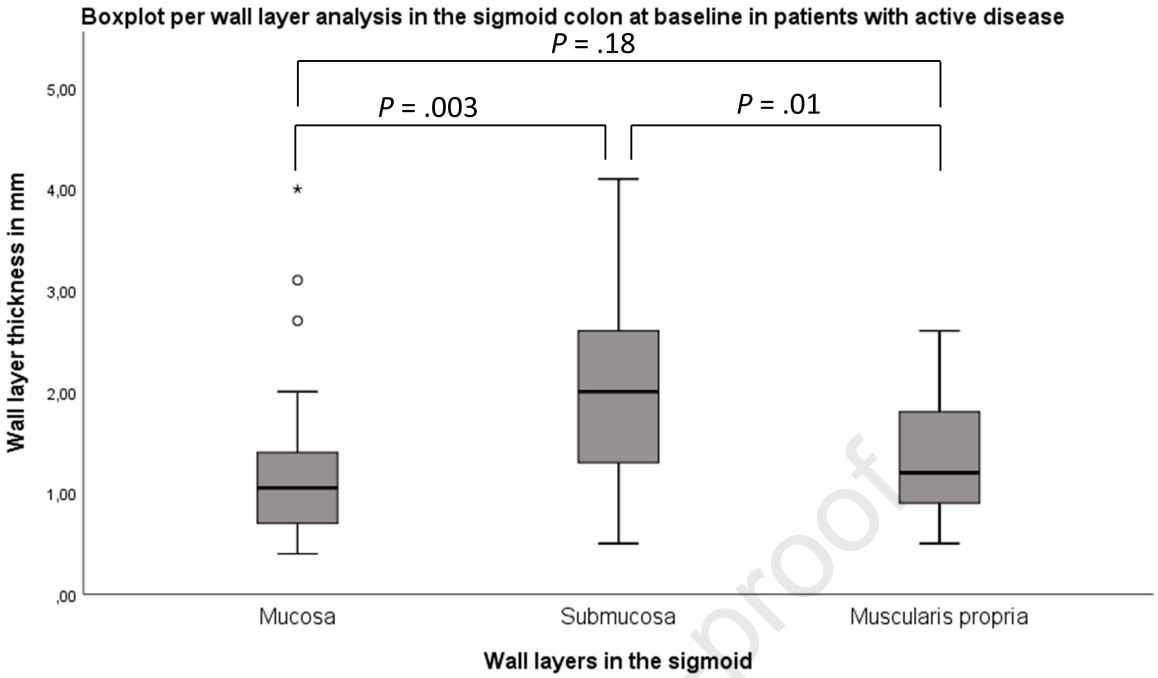
C

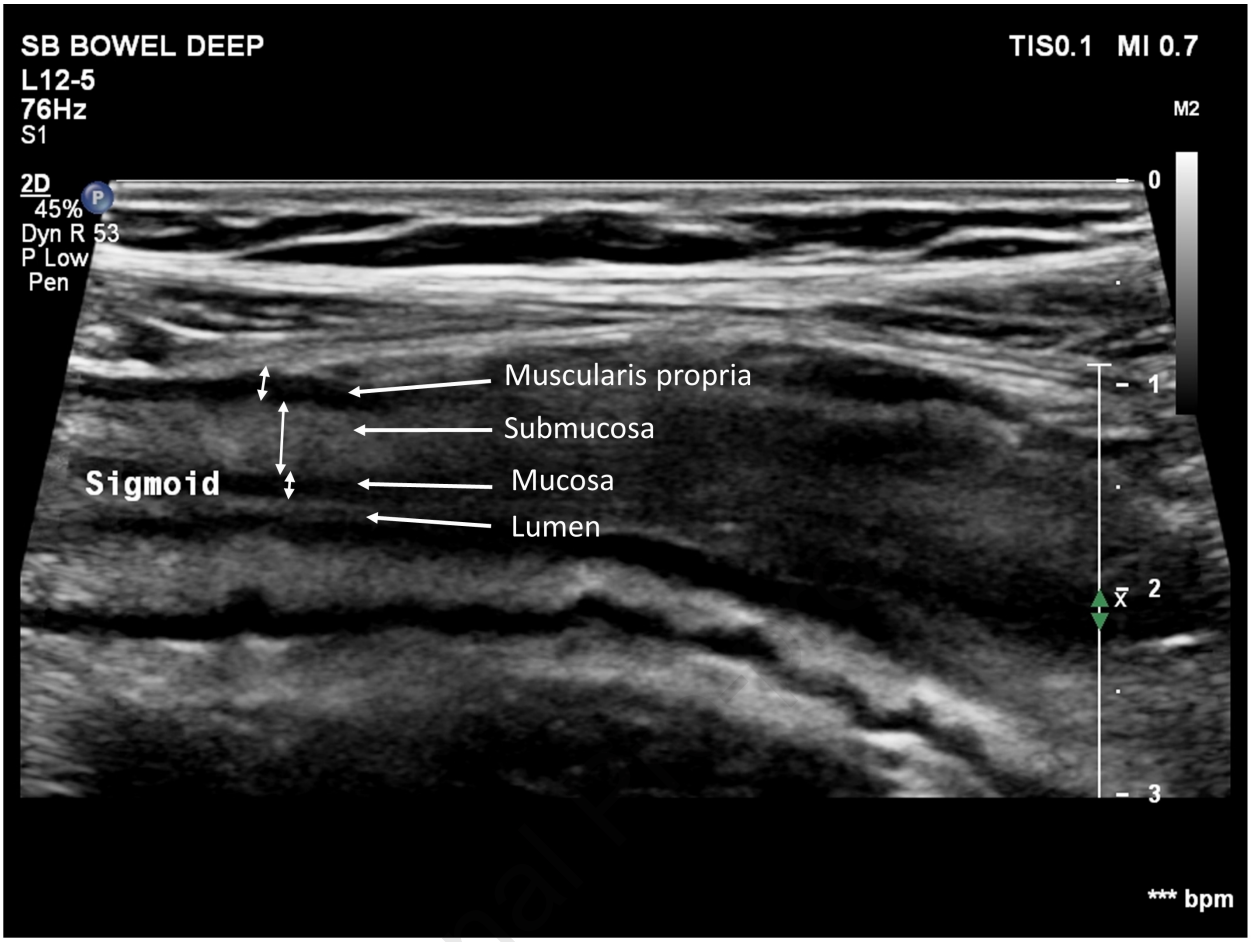


A



B




SB BOWEL DEEP
L12-5
76Hz
S1
TIS0.1 MI 0.7
M2
2D
45%
Dyn R 53
P Low
Pen
Muscularis propria
Submucosa
Mucosa
Lumen
Sigmoid
0
1
2
3
*** bpm

What you need to know

BACKGROUND AND CONTEXT

Intestinal ultrasound is non-invasive and accurate to determine disease activity in ulcerative colitis. Here we prospectively evaluated intestinal ultrasound for treatment response on tofacitinib against endoscopic and histologic reference standards.

NEW FINDINGS

Intestinal ultrasound, and predominantly bowel wall thickness, correlates with endoscopic and histologic scoring indices and accurately detects treatment response. The submucosa is the most responsive wall layer.

LIMITATIONS

The patient group was small and one mechanism of action was studied.

IMPACT

Intestinal ultrasound with bowel wall thickness as most important parameter has potential to evaluate treatment response in a non-invasive and accurate fashion.

Lay summary

Bowel ultrasound detects response to treatment in patients with ulcerative colitis and corresponds well with endoscopy. It could be used as non-invasive alternative to endoscopy in treatment response evaluation.

Intestinal ultrasound parameter	**Technique/categories**	**Pathological at baseline**
Bowel wall thickness in mm	(2x longitudinal plane + 2x cross-sectional plane)/4	> 4.0 mm in the SC > 3.0 mm in the DC
Colour Doppler Signal	0: absent; 1: small spots (single vessels) within the wall; 2: long stretches within the wall; 3: long stretches extending into the mesentery	≥2
Loss of wall layer stratification	0: preserved; 1: focal loss; 2: extensive loss	≥1
Loss of haustration	0: preserved; 1: loss	=1
Fatty wrapping	0: absent; 1: present	=1
Presence of lymph nodes	0: absent; 1: visible and larger than 4 mm in the shortest axis	=1

Supplementary Table 1: Assessed intestinal ultrasound parameters. [mm=millimeter, SC=sigmoid colon, DC=descending colon]

Median bowel wall thickness and IQR for endoscopic outcomes	Yes	No	p-value
Remission [EMS]	1.5 mm [1.1-2.6]	4.4 mm [3.0-5.5]	**0.009**
Remission [UCEIS]	1.2 mm [1.0-2.0]	4.3 mm [3.0-5.4]	**<0.0001**
Endoscopic improvement [EMS]	2.5 mm [1.3-4.4]	4.4 mm [4.1-6.0]	**0.027**

Median Δpercentage bowel wall thickness and IQR for endoscopic response	Yes	No	p-value
Response [EMS]	-34.1% [-62.9- -14.8]	-16.7% [-35.0-18.0]	0.08
Response [UCEIS]	-41.9% [-62.1- -18.3]	-16.7% [-20.4-18.0]	**0.008**

Supplementary Table 2: Median bowel wall thickness [IQR] in the descending colon for endoscopic remission, endoscopic improvement and endoscopic response according to the endoscopic Mayo score or UCEIS.

	SC
UCEIS remission (UCEIS=0)	Cut-off: 2.83 mm, 70% sensitivity, 100% specificity (AUROC: 0.88, 95% CI: 0.72-1.00, p=0.02)
UCEIS response (UCEIS≥2 points decrease)	Decrease of 32%, 63% sensitivity and 80% specificity (AUROC: 0.71, 95% CI: 0.51-0.92, p=0.07)

Supplementary Table 3: Cut-off values of BWT for endoscopic remission and response according to the UCEIS in the sigmoid [UCEIS: Ulcerative Colitis Endoscopic Index or Severity, SC=sigmoid colon, DC=descending colon, BWT=bowel wall thickness, AUROC=area under the receiver operating characteristic curve]

Endoscopic outcomes	Cut-off values BWT in the descending colon
EMS remission (EMS=0)	Cut-off: 2.92 mm, 93% sensitivity, 78% specificity (AUROC: 0.91, 95% CI: 0.79-1.00, p=0.001)
Endoscopic improvement (EMS≤1)	Cut-off: 3.6 mm, 86% sensitivity, 69% specificity (AUROC: 0.81, 95% CI: 0.64-0.99, p=0.019)
EMS response (EMS≥1 point decrease)	Decrease of 20%, 75% sensitivity and 78% specificity (AUROC: 0.78, 95% CI: 0.58-0.98, p=0.012)
UCEIS remission (UCEIS=0)	Cut-off: 2.92 mm, 80% sensitivity, 100% specificity (AUROC: 0.98, 95% CI: 0.94-1.00, p<0.0001)
UCEIS response (UCEIS≥2 points decrease)	Decrease of 23%, 89% sensitivity and 77% specificity (AUROC: 0.82, 95% CI: 0.64-1.00, p=0.012)

Supplementary Table 4: Cut-off values of BWT for endoscopic remission, endoscopic improvement and response according to the EMS and UCEIS in the descending colon [UCEIS: Ulcerative Colitis Endoscopic Index or Severity, DC=descending colon, BWT=bowel wall thickness, AUROC=area under the receiver operating characteristic curve]

Segment	Endoscopic score	SCCAI	Lichtiger score	Fecal calprotectin
Rectum	EMS	ρ=0.67, p<0.0001	ρ=0.54, p=0.003	ρ=0.34, p=0.08
	UCEIS	ρ=0.69, p<0.0001	ρ=0.57, p=0.002	ρ=0.42, p=0.03
Sigmoid colon	EMS	ρ=0.78, p<0.0001	ρ=0.68, p<0.0001	ρ=0.72, p<0.0001
	UCEIS	ρ=0.84, p<0.0001	ρ=0.74, p<0.0001	ρ=0.79, p<0.0001
Descending colon	EMS	ρ=0.54, p=0.004	ρ=0.57, p=0.003	ρ=0.47, p=0.016
	UCEIS	ρ=0.57, p=0.002	ρ=0.61, p=0.001	ρ=0.56, p=0.002

Supplementary Table 5: Correlation coefficients for endoscopic scores per segment with SCCAI, Lichtiger score and fecal calprotectin. [EMS: endoscopic Mayo score; UCEIS: Ulcerative Colitis Endoscopic Index of Severity, SCCAI: Simple Clinical Colitis Activity Index]

IUS parameter in corresponding segment	UCEIS in SC	EMS in SC	UCEIS in DC	EMS in DC
CDS	ρ=0.66, p<0.0001	ρ=0.59, p<0.0001	ρ=0.78, p<0.0001	ρ=0.80, p<0.0001
WLS	ρ=0.505, p<0.0001	ρ=0.395, p=0.003	ρ=0.504, p<0.0001	ρ=0.476, p<0.0001
Loss of haustrations	ρ=0.706, p<0.0001	ρ=0.708, p<0.0001	ρ=0.504, 0<0.0001	ρ=0.442, p=0.002
Lymph nodes	ρ=0.457, p=0.001	ρ=0.466, p<0.0001	ρ=0.416, p=0.004	ρ=0.355, p=0.011
Fatty wrapping	ρ=0.527, p<0.0001	ρ=0.456, p=0.001	ρ=0.340, p=0.017	ρ=0.447, p=0.002

Supplementary Table 6: Correlation intestinal ultrasound parameters with endoscopic scores in the sigmoid and descending colon.

Parameter	**Univariate analysis for endoscopic remission in SC**		**Univariate analysis for endoscopic remission in DC**	
	OR [95%CI]	p-value	OR [95%CI]	p-value
CDS ≤1	11.20 [1.23-101.89]	**0.017**	OR: 10.93 [1.23-97.08]	**0.014**
Loss of WLS	n.a.	1.00	n.a.	1.00
Loss of haustrations	0.57 [0.40-0.83]	**<0.0001**	n.a.	1.00
Presence of lymph nodes	n.a.	1.00	n.a.	1.00
Presence of fatty wrapping	n.a.	1.00	0.66 [0.03-12.27]	0.66
Faecal calprotectin	0.06 [0.02-1.44]	0.08	0.99 [0.94-1.04]	0.62

IUS parameter	**Univariate analysis for endoscopic improvement in SC**		**Univariate analysis for endoscopic improvement in DC**	
	OR [95%CI]	p-value	OR [95%CI]	p-value
CDS ≤1	11.56 [2.64-50.50]	**0.001**	16.13 [3.61-71.43]	**<0.0001**
Loss of WLS	0.60 [0.46-0.79]	**0.003**	0.18 [0.02-1.45]	0.11
Loss of haustrations	0.24 [0.11-0.51]	**<0.0001**	n.a.	1.00
Presence of lymph nodes	n.a.	1.00	0.16 [0.74-48.90]	0.09
Presence of fatty wrapping	n.a.	1.00	n.a.	1.00
Faecal calprotectin	0.79 [0.63-1.00]	0.052	0.95 [0.89-0.99]	**0.048**

IUS parameter	**Univariate analysis for endoscopic response in SC**		**Univariate analysis for endoscopic response in DC**	
	OR [95%CI]	p-value	OR [95%CI]	p-value
Per decrease in 1 CDS-category	2.50 [1.08-5.78]	**0.03**	2.60 [0.94-7.17]	0.06
Normalization of WLS	1.79 [0.35-9.13]	0.49	1.40 [0.11-18.62]	0.80
Normalization of haustrations	22.50 [2.55-198.38]	**0.005**	n.a.	1.00
Absence of lymph nodes	3.47 [0.56-21.35]	0.18	0.64 [0.07-5.61]	0.68
Absence of fatty wrapping	0.76 [0.28-2.03]	0.56	n.a.	1.00
Decrease in faecal calprotectin	0.96 [0.92-1.01]	0.08	0.95 [0.90-1.01]	0.056

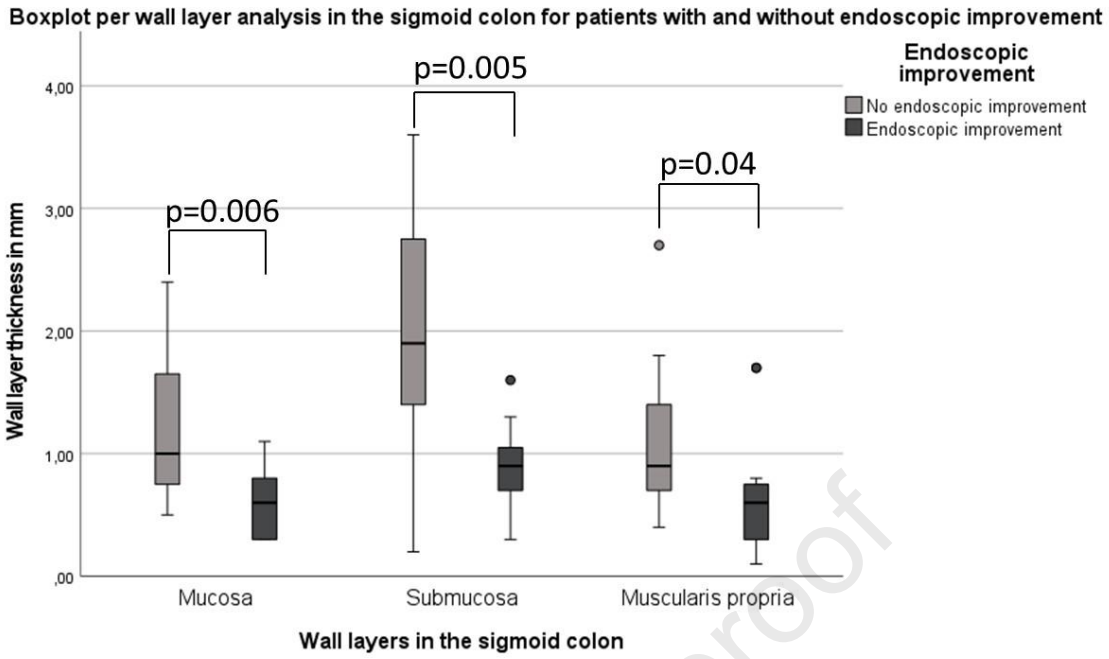
Supplementary Table 7: Odds ratios for individual intestinal ultrasound parameters in the sigmoid and descending colon and faecal calprotectin with endoscopic outcomes according to the endoscopic Mayo score after 8 weeks of treatment [CDS: Colour Doppler signal, WLS: wall layer stratification, OR: Odds ratio, CI: Confidence interval]

Endoscopic outcome	BWT cut-off value	Odds ratio	95% CI	p-value
Endoscopic remission	2.9 mm	12.83	1.70-97.19	0.013
Endoscopic improvement	3.6 mm	13.20	1.24-140.68	0.033
Endoscopic response	-20%	6.75	0.93-49.23	0.06

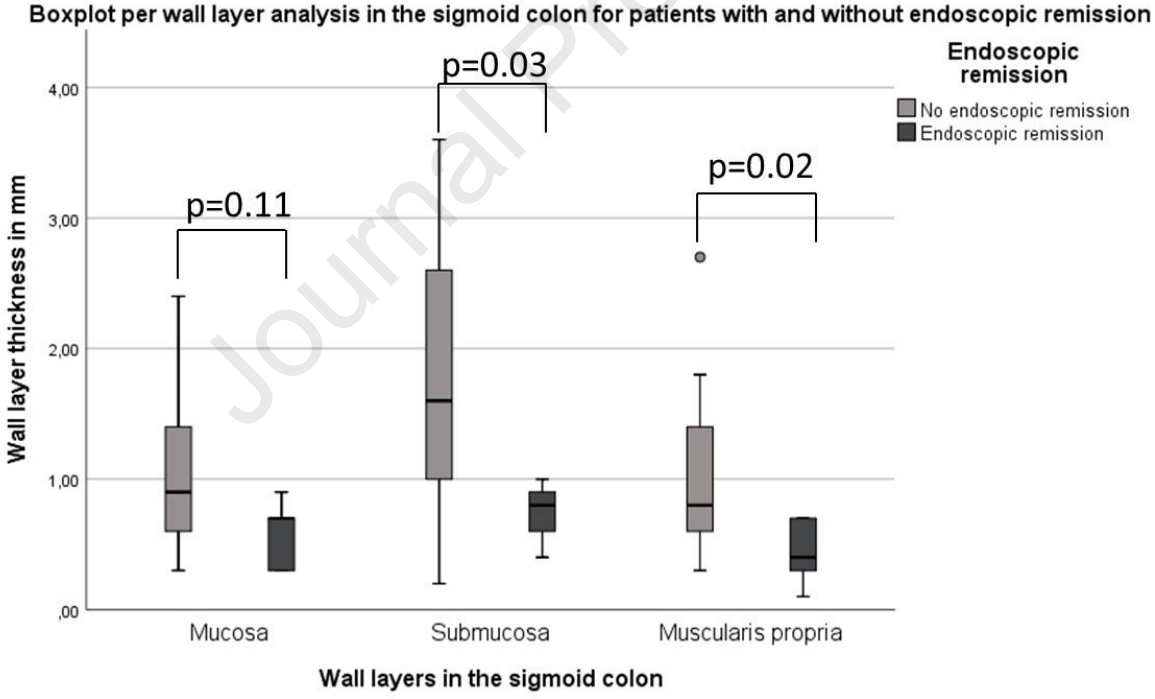
Supplementary Table 8: Association of different BWT cut-off values with endoscopic outcomes according to the EMS in the descending colon. [BWT: Bowel wall thickness, CI: Confidence value, EMS: endoscopic Mayo Score]

	CDS	Loss of WLS	Loss of haustrations	Presence of LN
Sigmoid colon	Kappa=0.57, p<0.0001	Kappa=0.31, p=0.035	Kappa=0.34, p=0.004	Kappa=0.41, p=0.012
Descending colon	Kappa=0.51, p=0.002	Kappa=0.46, p=0.005	Kappa=0.58, p=0.001	Kappa=0.17, p=0.197

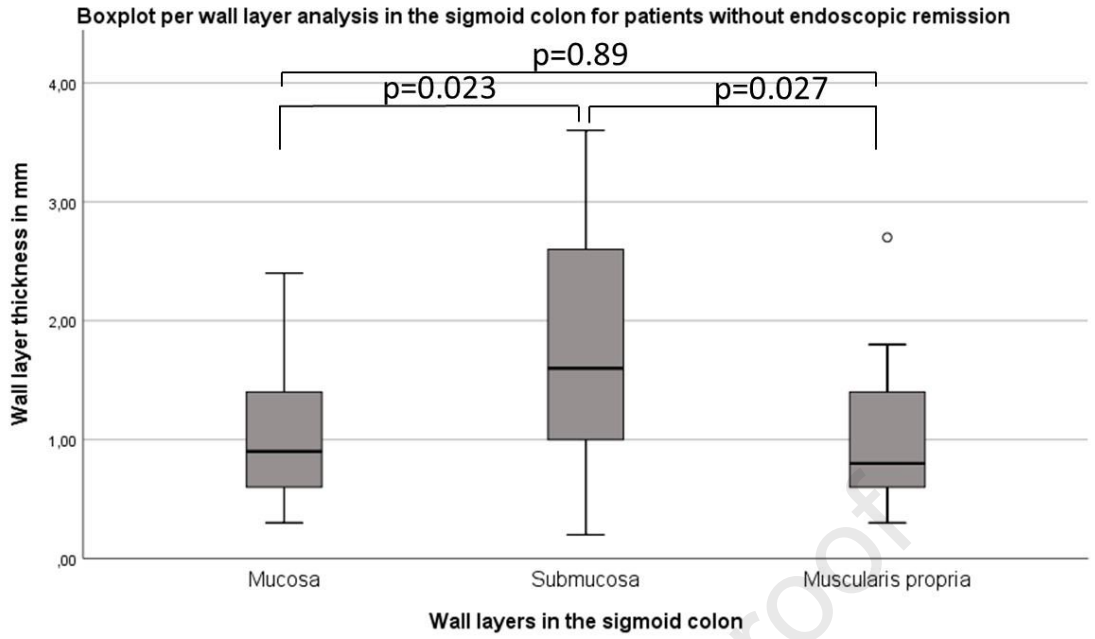
Supplementary Table 9: Inter-observer agreement per IUS parameter per segment [CDS=Colour Doppler Signal, WLS=wall layer stratification, LN=lymph nodes]



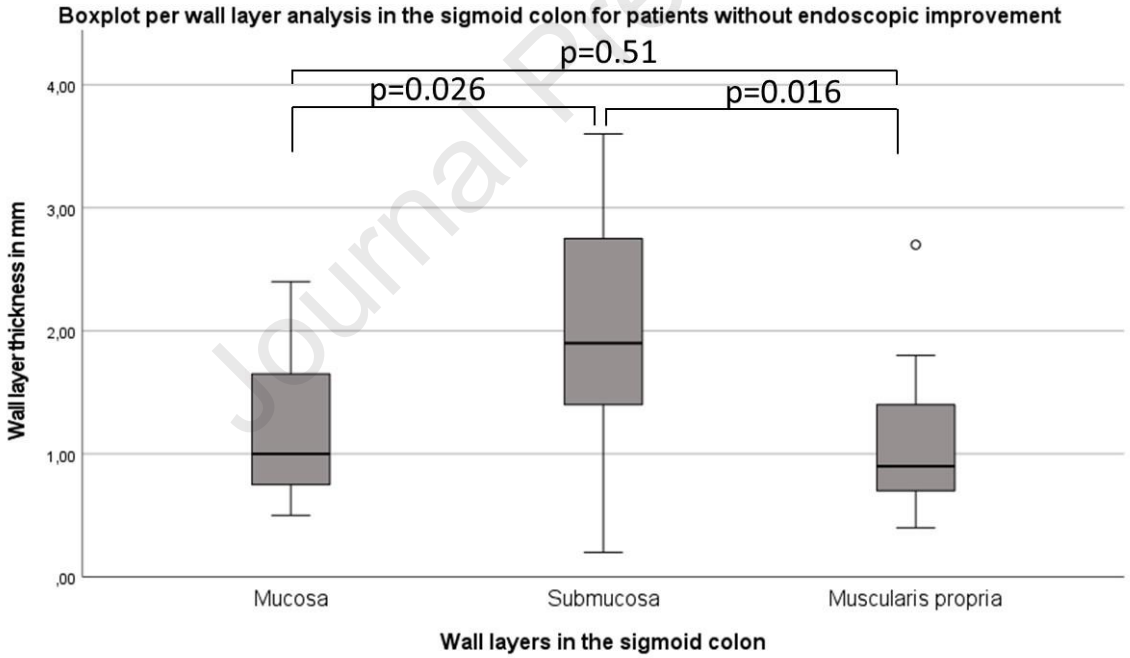
A



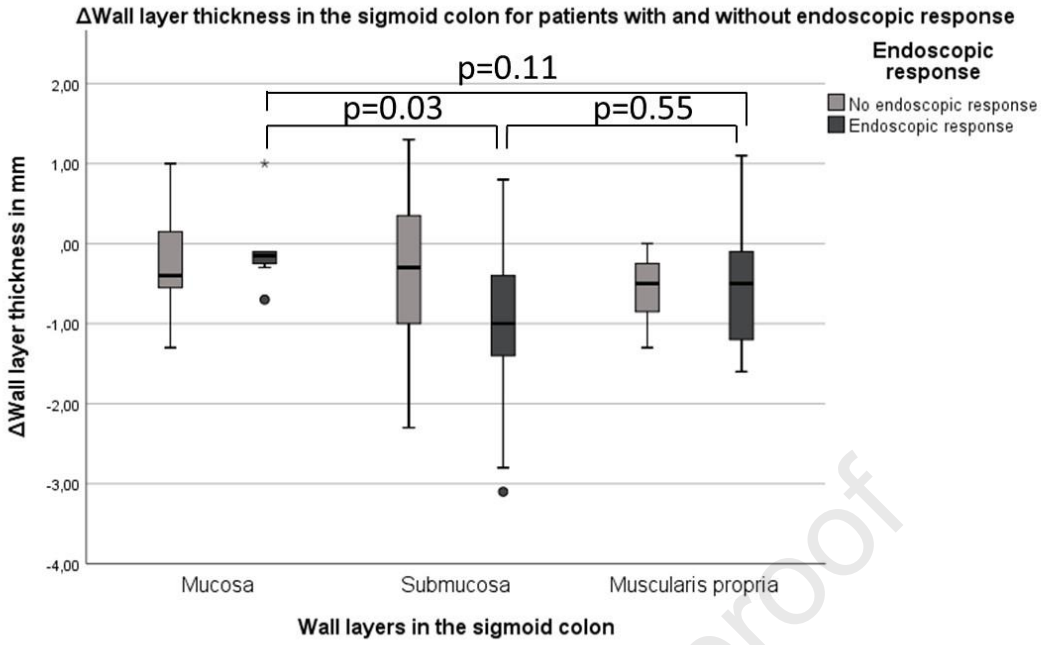
B



C

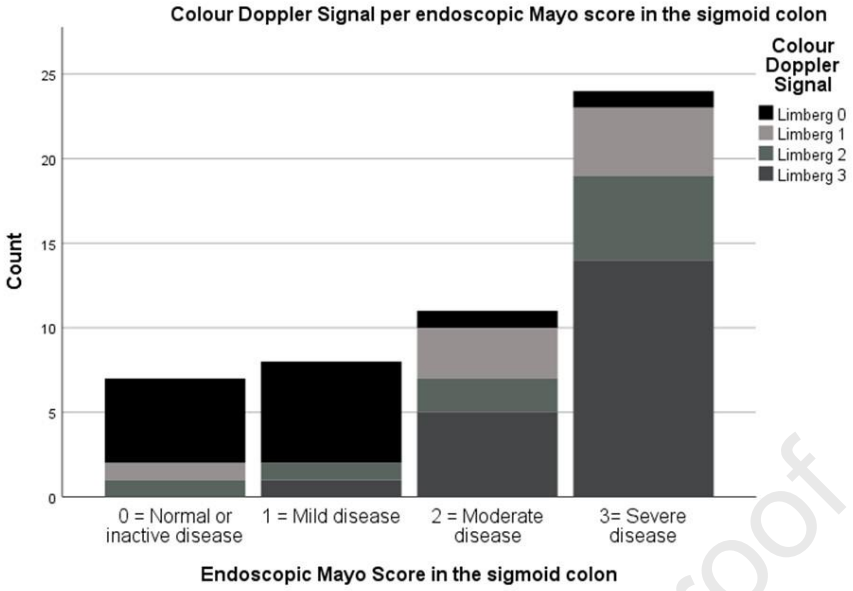


D

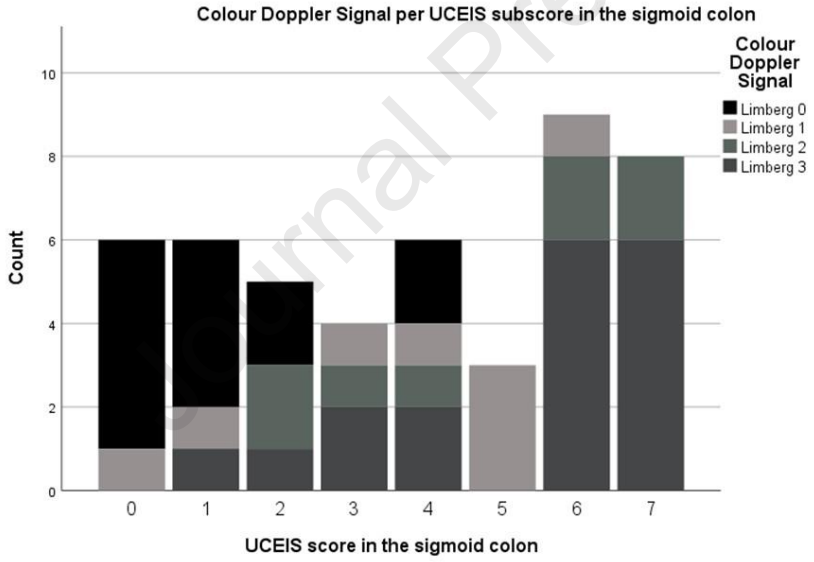


E

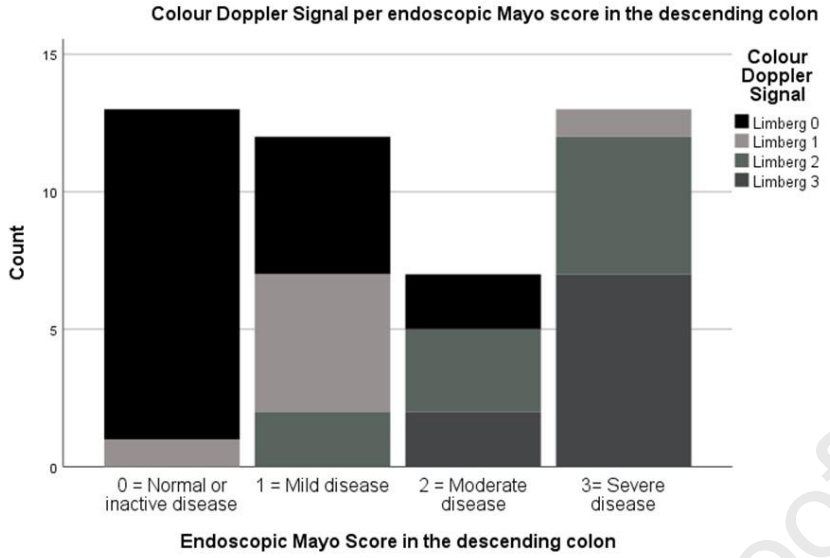
Supplementary Figure 1: Analysis for individual wall layers and endoscopic outcomes in the sigmoid colon. **A:** Median [IQR] wall layer thickness for the mucosa (0.6 mm [0.30-0.90] vs 1.0 mm [0.7-1.7], p=0.006), submucosa (0.9 mm [0.6-1.1] vs 1.9 mm [1.4-2.8], p=0.005) and muscularis propria (0.6 mm [0.3-0.8] vs 0.9 mm [0.7-1.4], p=0.04) for patients with and without endoscopic improvement after 8 weeks of treatment; **B:** Median [IQR] wall layer thickness for the mucosa (0.7 mm [0.3-0.9] vs 0.9 mm [0.6-1.5], p=0.11), submucosa (0.8 mm [0.5-1.0] vs 1.6 mm [0.9-2.7], p=0.03) and muscularis propria (0.4 mm [0.2-0.7] vs 0.8 mm [0.6-1.5], p=0.02) for patients with and without endoscopic remission after 8 weeks of treatment according to the EMS; **C:** Median [IQR] wall layer thickness of the mucosa (0.9 mm [0.6-1.5]), submucosa (1.6 mm [0.9-2.7]) and muscularis propria (0.8 mm [0.6-1.5]) in patients without endoscopic remission according to the EMS after 8 weeks of treatment; **D:** Median [IQR] wall layer thickness of the mucosa (1.0 mm [0.7-1.7]), submucosa (1.9 mm [1.4-2.8]) and muscularis propria (0.9 mm [0.7-1.4]) in patients without endoscopic improvement after 8 weeks of treatment; **E:** Change in median [IQR] wall layer thickness for the mucosa (-0.1 mm [-0.3-0.5] vs -0.4 mm [-0.6-0.3], p=0.24), submucosa (-1.0 mm [-2.1- -0.2] vs -0.3 [-1.1-0.4], p=0.17) and muscularis propria (-0.5 [-1.4-0.3] vs -0.5 [-0.9- -0.1], p=0.95) in the sigmoid colon in patients with and without endoscopic response according to the EMS for the sigmoid colon [EMS: endoscopic Mayo score]



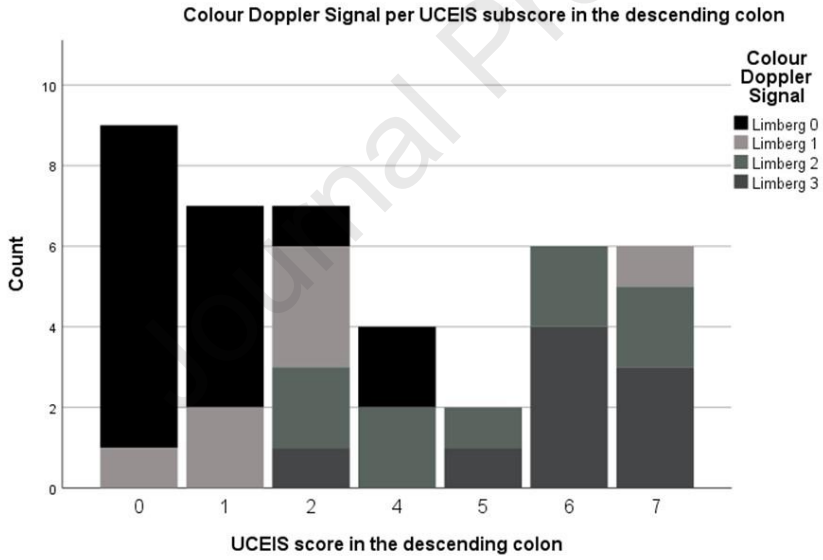
A.



B.

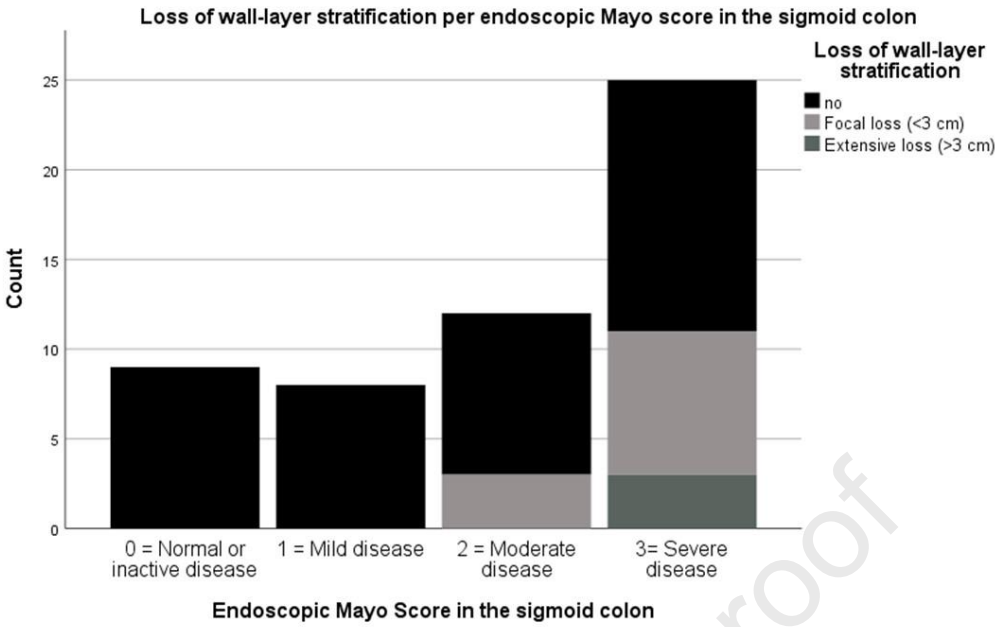


C.

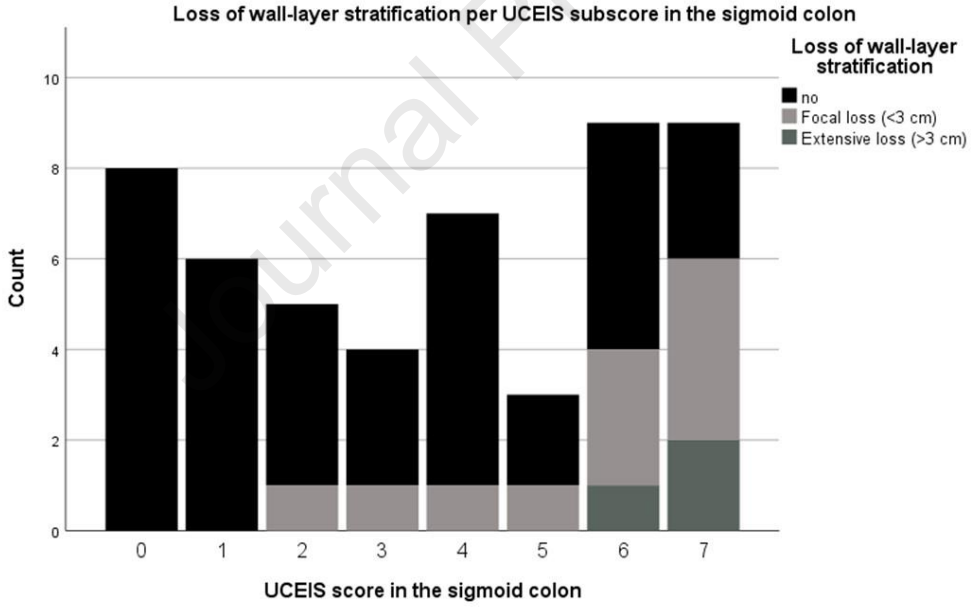


D.

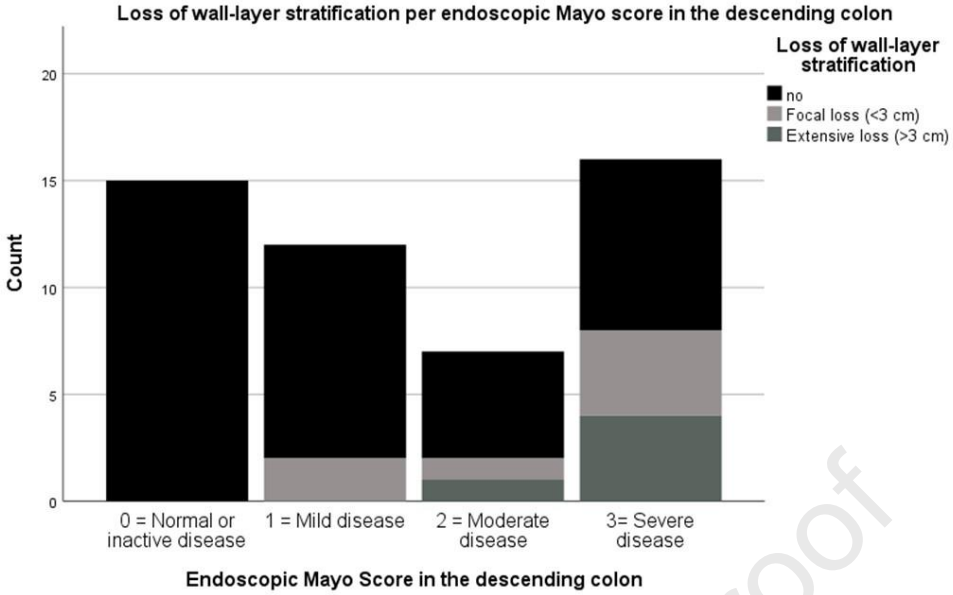
Supplementary Figure 2 (A-D): distribution of Colour Doppler Signal among EMS and UCEIS categories in the sigmoid (A and B) and descending colon (C and D).



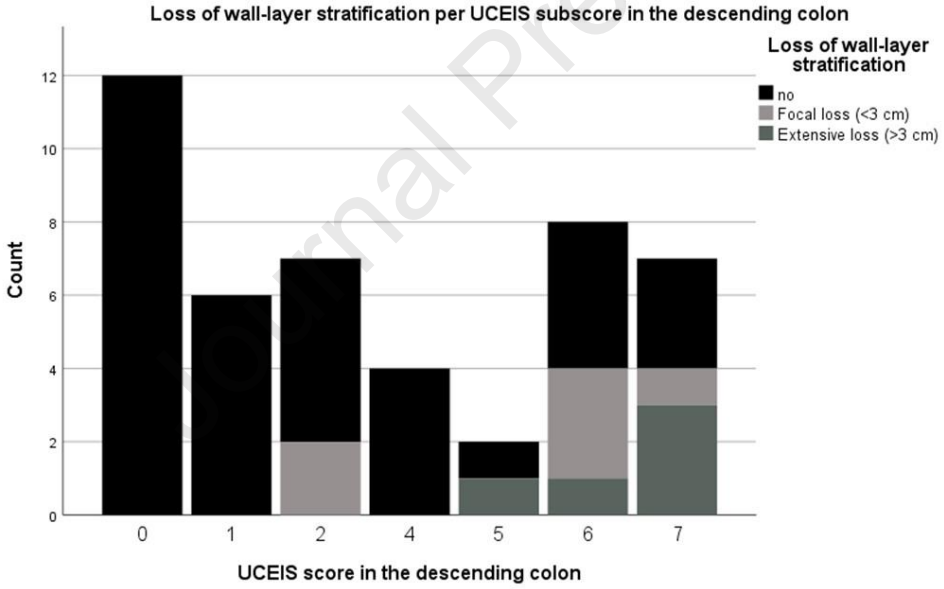
A.



B.

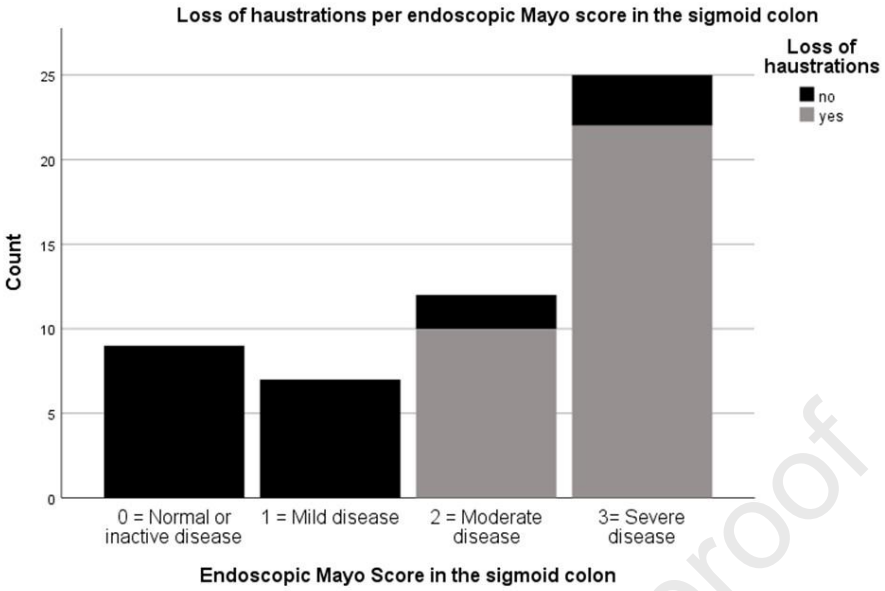


C.

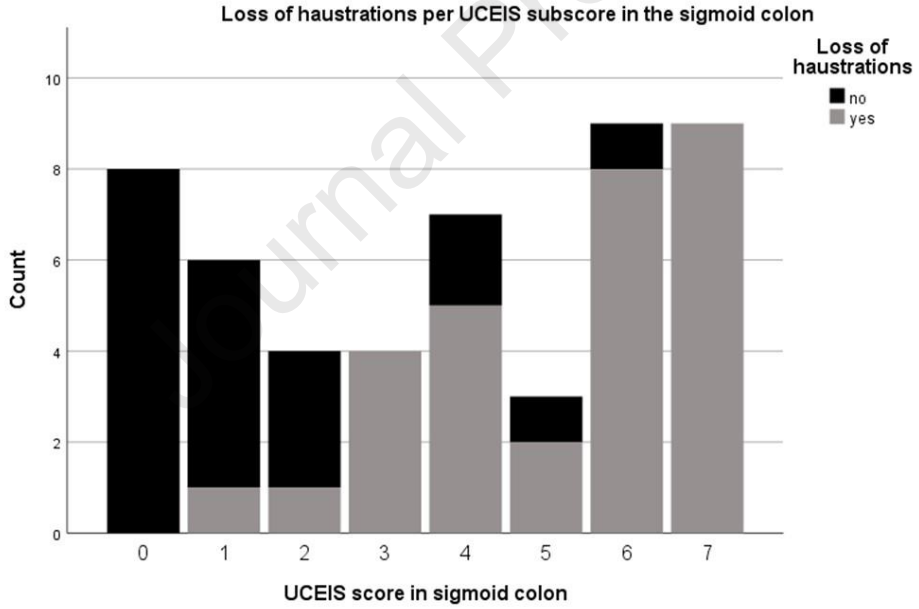


D.

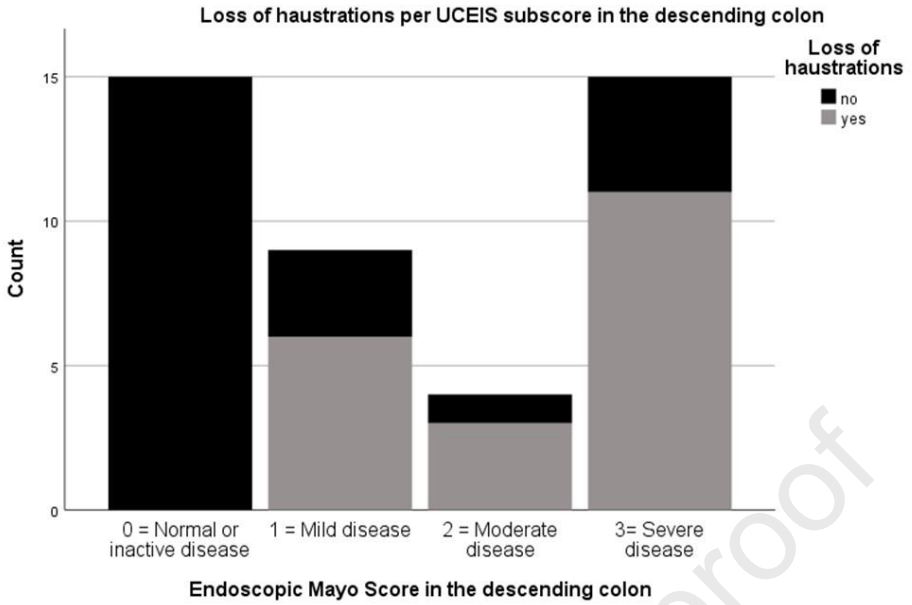
Supplementary Figure 3 (A-D): distribution of loss of wall layer stratification among EMS and UCEIS categories in the sigmoid (A and B) and descending colon (C and D).



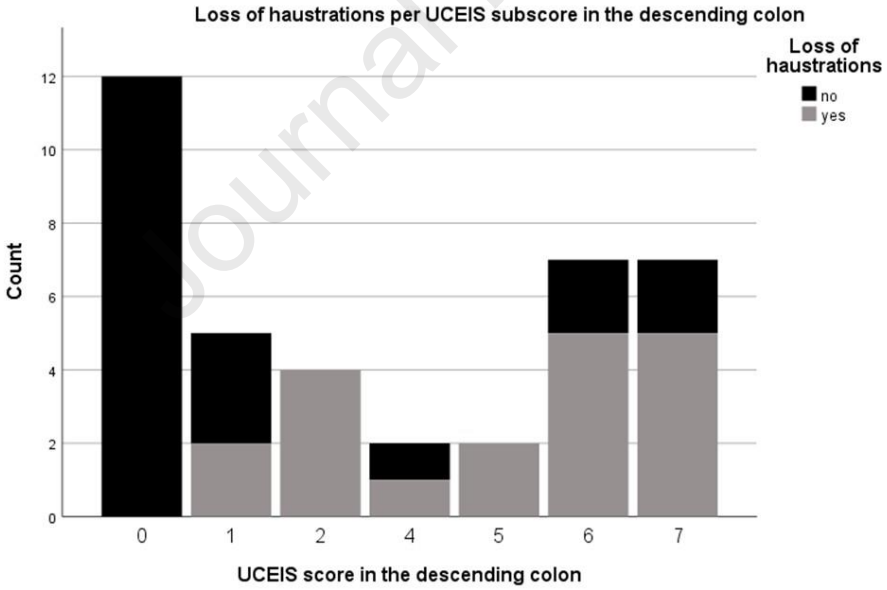
A.



B.

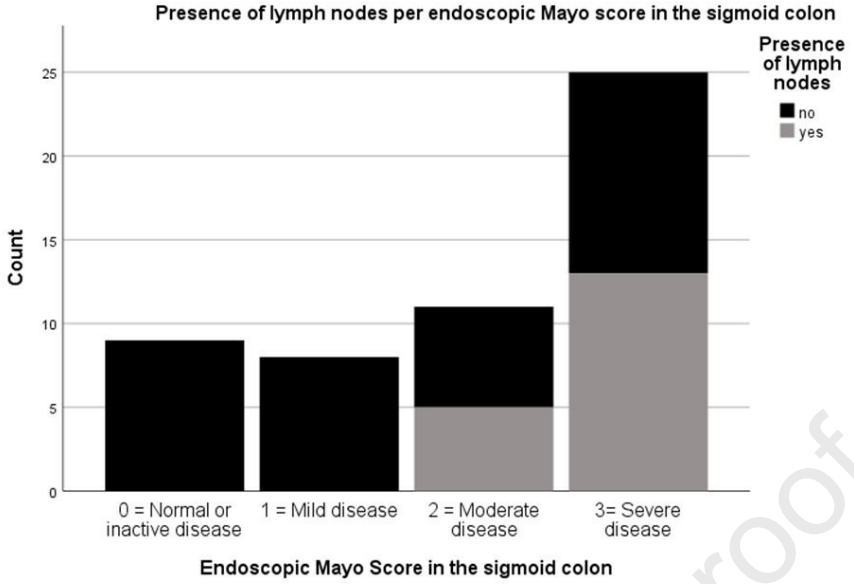


C.

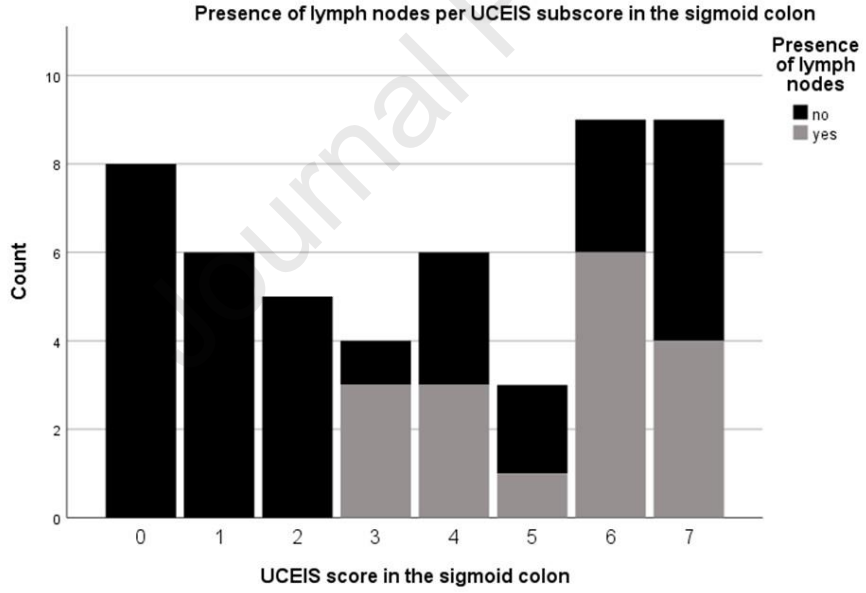


D.

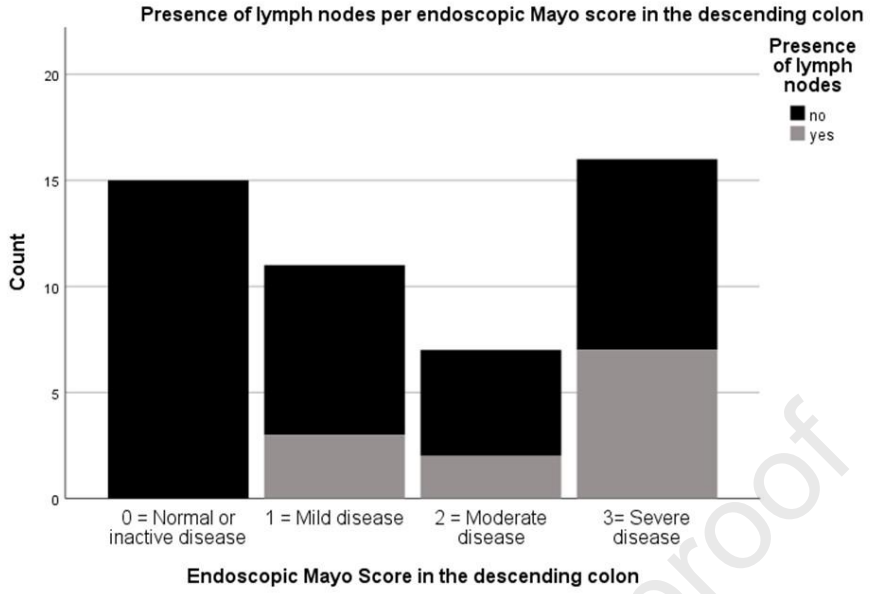
Supplementary Figure 4 (A-D): distribution of loss of haustrations among EMS and UCEIS categories in the sigmoid (A and B) and descending colon (C and D).



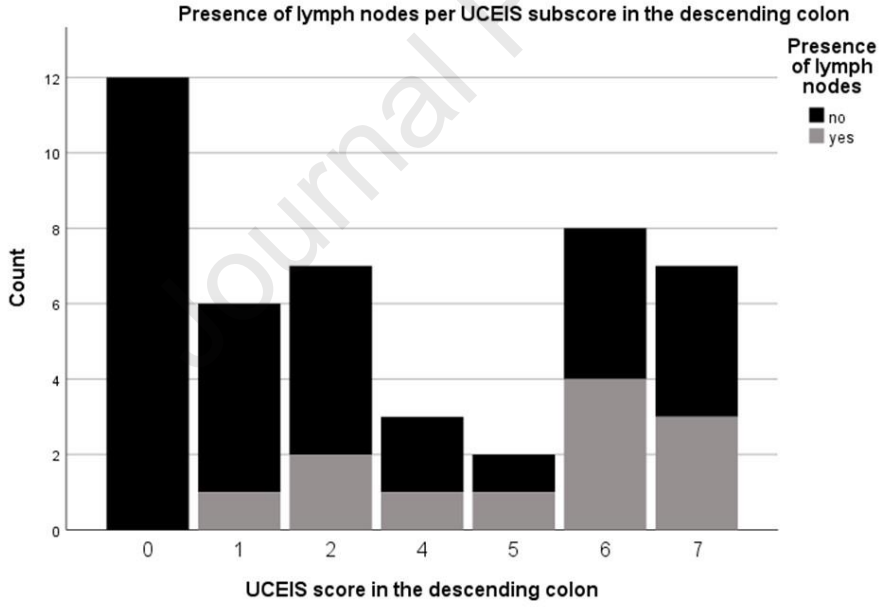
A.



B.

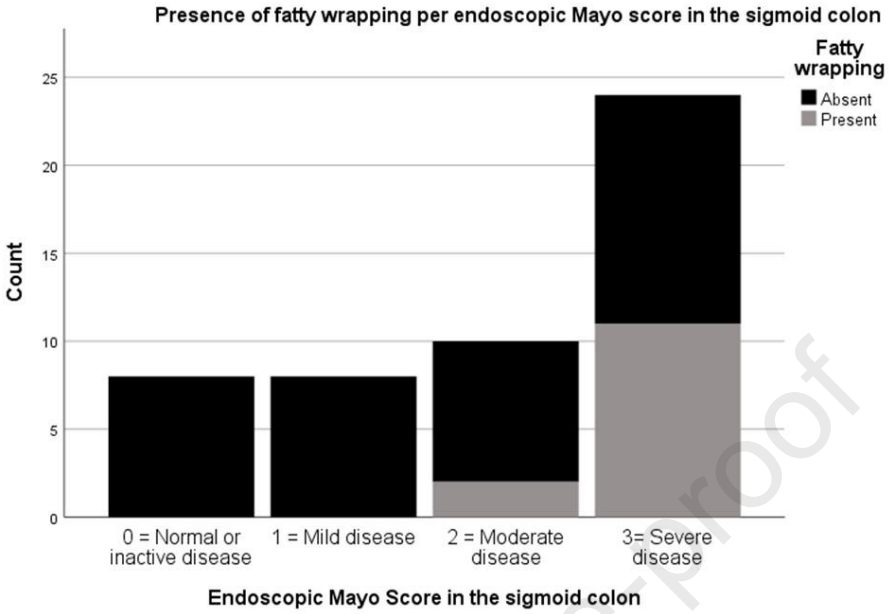


C.

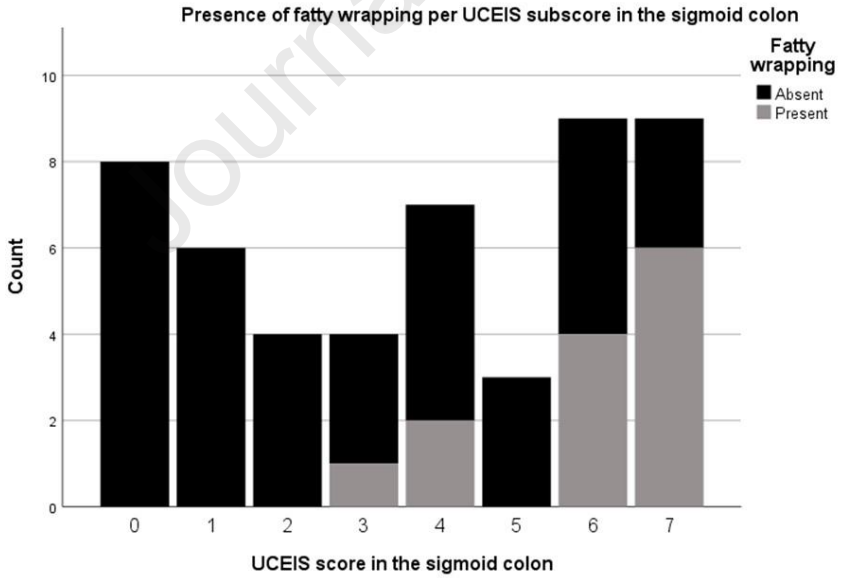


D.

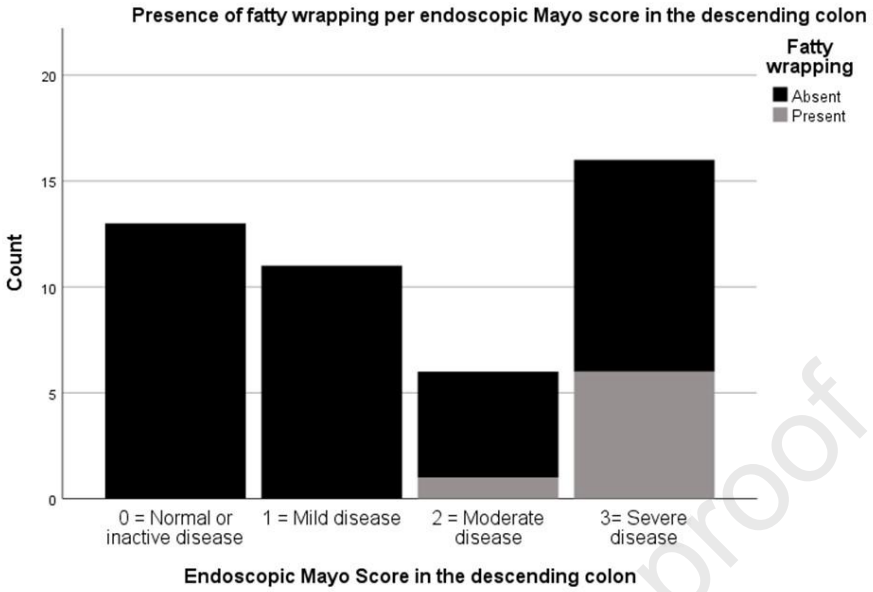
Supplementary Figure 5 (A-D): distribution for presence of lymph nodes among EMS and UCEIS categories in the sigmoid (A and B) and descending colon (C and D).



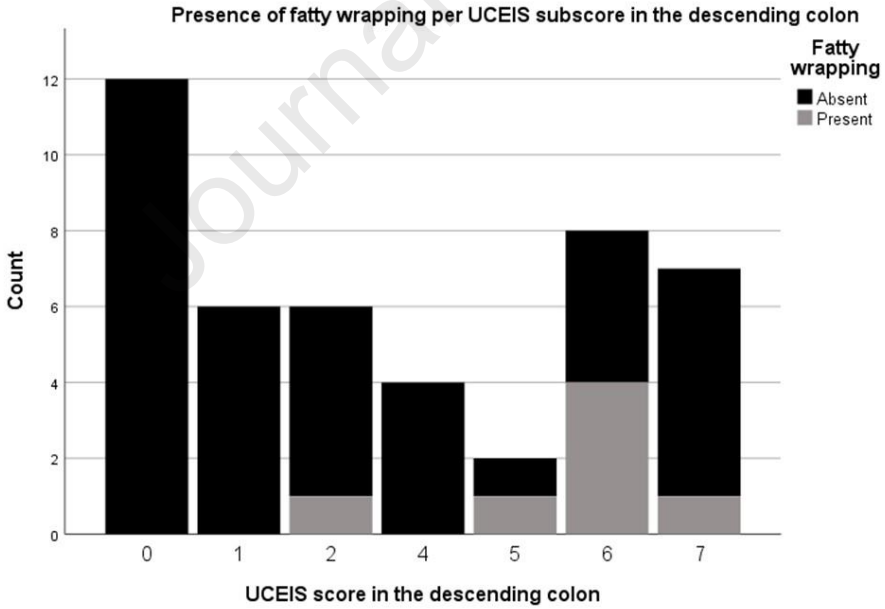
A.



B.

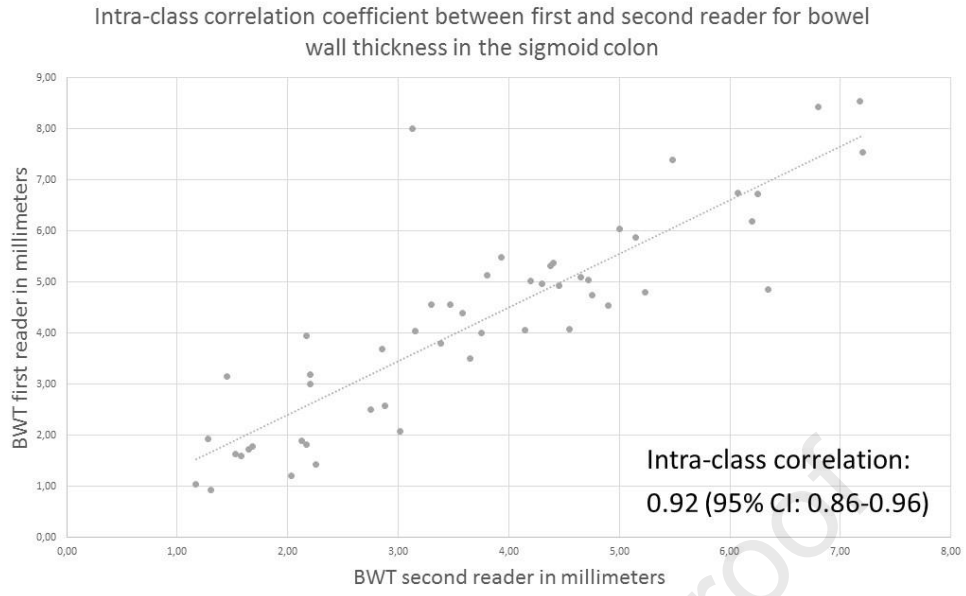


C.



D.

Supplementary Figure 6 (A-D): distribution of fatty wrapping among EMS and UCEIS categories in the sigmoid (A and B) and descending colon (C and D).



A.

B.

Supplementary Figure 7: Inter-observer agreement for BWT in the sigmoid (A) and descending colon (B)