



Title	VALIDATION OF US EVALUATION OF ULCERATIVE COLITIS ACTIVITY
Author(s)	Omotehara, Satomi; Nishida, Mutsumi; Kinoshita, Kenji; Onishi, Reizo; Onodera, Aki; Suya, Mitsuhisa; Hasegawa, Toru; Mitsumori, Daiki; Katsurada, Takehiko; Teshima, Takanori
Citation	Ultrasound in medicine and biology, 45(7), 1537-1544 <a href="https://doi.org/10.1016/j.ultrasmedbio.2019.02.018">https://doi.org/10.1016/j.ultrasmedbio.2019.02.018</a>
Issue Date	2019-07
Doc URL	<a href="http://hdl.handle.net/2115/78766">http://hdl.handle.net/2115/78766</a>
Rights	© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Rights(URL)	<a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Type	article (author version)
File Information	Ultrasound Med Biol_45_1537.pdf



[Instructions for use](#)

## VALIDATION OF US EVALUATION OF ULCERATIVE COLITIS ACTIVITY

SATOMI OMOTEHARA,<sup>\*,y</sup> MUTSUMI NISHIDA,<sup>\*,y</sup> KENJI KINOSHITA,<sup>z</sup> REIZO ONISHI,<sup>z</sup>  
AKI ONODERA,<sup>x</sup> MITSUHISA SUYA,<sup>†</sup> TORU HASEGAWA,<sup>‡</sup> DAIKI MITSUMORI,<sup>#</sup>  
TAKEHIKO KATSURADA,<sup>\*\*</sup> and TAKANORI TESHIMA<sup>\*</sup>

\* Division of Laboratory and Transfusion Medicine, Hokkaido University Hospital, Kita-ku, Sapporo, Japan; <sup>y</sup>Diagnostic Center for Sonography, Hokkaido University Hospital, Kita-ku, Sapporo, Japan; <sup>z</sup>Department of Gastroenterology and Hepatology, Hokkaido University Hospital, Kita-ku, Sapporo, Japan; <sup>x</sup> Department of Clinical Laboratory, Tomakomai City Hospital, Shimizutyou, Tomakomai, Japan; <sup>†</sup>Department of Clinical Laboratory, Sapporo Higashi Tokushukai Hospital, Higashi-ku, Sapporo, Japan;  
<sup>‡</sup> Department of Clinical Laboratory, Ohguro Gastroenterological Hospital, Kita-ku, Sapporo, Japan; <sup>#</sup>Department of Clinical Laboratory, NTT East Japan Sapporo Hospital, Chuou-ku, Sapporo, Japan; and <sup>\*\*</sup> Division of Endoscopy, Hokkaido University Hospital, Kita-ku, Sapporo, Japan

\*: Corresponding author

Corresponding author: Mutsumi Nishida

E-mail: [mutuni@gmail.com](mailto:mutuni@gmail.com)

Address: N14 W5, Kita-ku, Sapporo 060-8648, Japan

Tel: +81 011 7065697; Fax: +81 011 7067614

1 **Abstract-** This study was aimed at validating the inter-rater grading agreement  
2 for assessing disease activity in patients with established ulcerative colitis (UC)  
3 using transabdominal ultrasonography (US) versus colonoscopy (CS). Fifty-  
4 seven patients underwent US and CS at four facilities. UC disease activity was  
5 assessed using the original US grading system and CS Matts classification.  
6 Initially, the US and CS grades were assessed at each examining facility, and still  
7 images and movie clips were re-assessed at the central facility. Grading  
8 agreement between the examining and central facilities was evaluated. Grading  
9 agreement for US and CS were 0.75 and 0.72 in all segments and 0.82 and  
10 0.70 in the maximum grade of each patient, respectively (all  $p < 0.001$ ). US  
11 grading agreement was “almost perfect” for the maximum grade and “moderate”  
12 to “substantial” for other assessments. The inter-rater US grading agreement was  
13 good and not inferior to that of CS for evaluating UC disease activity.

14 **Keywords:** Ultrasound, Colonoscopy, Inter-rater agreement, Ulcerative colitis,  
15 Disease activity

16

## 1 INTRODUCTION

2 Ulcerative colitis (UC) is an inflammatory bowel disease of unknown etiology  
3 characterized by inflammation of the mucosa and submucosa (Danese and  
4 Fiocchi 2011; Ungaro et al. 2017). The inflammation usually involves the rectum  
5 and extends proximally in the colon for a variable distance. Additionally, it is  
6 characterized by alternating periods of remission and relapse. The incidence of  
7 UC has been increasing (Ng SC et al. 2018), and the accurate evaluation of  
8 disease extent and activity is necessary for disease monitoring and treatment  
9 response. The current UC therapeutic goal is to achieve and sustain mucosal  
10 healing evaluated using colonoscopy (CS) (Dignass et al. 2012). However, CS is  
11 invasive, and difficult to perform frequently; additionally, the risk of perforation is  
12 of concern (Civitelli et al. 2014).

13 Recently, transabdominal ultrasonography (US) has been identified as a useful  
14 diagnostic tool for gastrointestinal diseases, especially UC (Allocca et al. 2018;  
15 Parente et al. 2010). US can non-invasively observe the detailed wall layer  
16 structure in high resolution, as a result of recent advances in the ultrasound  
17 scanner (Kucharzik et al. 2015). We have reported moderate concordance  
18 between US and CS for evaluating UC in a multicenter prospective study

1 (Kinoshita et al. 2018). However, if only US experts can diagnose correctly, US  
2 cannot be used widely. Therefore, the diagnostic reliability of US requires  
3 validation. This study aimed to assess whether US could be used at different  
4 facilities to evaluate UC disease activity. We verified the inter-rater grading  
5 agreement of US compared with CS as the reference standard.

## 6 **METHODS**

### 7 ***Patients and study design***

8 From June 2013 to August 2016, established UC patients who had undergone  
9 US and CS within 2 d at Tomakomai City Hospital, Sapporo Higashi Tokushukai  
10 Hospital, Ohguro Gastroenterological Hospital, and NTT East Japan Sapporo  
11 Hospital were enrolled. Patients were excluded if they were younger than 19 y of  
12 age or had acute complications such as severe bleeding and toxic megacolon,  
13 and histologically proven cytomegalovirus colitis. Clinical disease activity was  
14 evaluated using the Rachmilewitz clinical activity index (CAI) and classified as  
15 inactive, mild, moderate or severe disease (Rachmilewitz 1989). The variables of  
16 this index include bowel frequency, blood in stool, general well-being, abdominal  
17 pain, body temperature  $>38.0^{\circ}\text{C}$ , erythrocyte sedimentation rate  $>50$  mm/h and  
18 hemoglobin level  $<10.0$  g/dL. Disease severity was originally scored as follows:

1 0-2, inactive disease; 3-4, mild disease; 5-7, moderate disease; 8, severe disease.

2 This study was approved by the institutional review board in each facility, and  
3 written informed consent was obtained from all the patients in accordance with  
4 the Declaration of Helsinki.

### 5 ***Transabdominal ultrasonography***

6 Ultrasonography was performed using the PVT-375 BT (center frequency: 3.75  
7 MHz), 674 BT (center frequency: 6 MHz) and 704 AT/BT (center frequency: 7.5  
8 MHz) equipped with Xario and Aplio XV/XG/500 (Canon Medical Systems Corp.,  
9 Otawara, Japan). Six segments of the colon and rectum were sequentially and  
10 individually assessed.

11 US was performed by 17 sonographers, of whom 5, 4 and 8 had <5, 5-10 and  
12 >10 y of experience, respectively. Patients generally remained in the supine  
13 position during examinations or were moved to the decubitus position as needed.  
14 First, convex transducers (frequency: 3.75/6.0 MHz) were used, followed by a  
15 high-frequency linear-array transducer (7.5 MHz) for detailed evaluation. The  
16 colon and rectum were sequentially assessed first to identify the hepatic flexure,  
17 and then the ascending colon toward the cecum was scanned to recognize the  
18 terminal ileum and Bauhin's valve. The examination proceeded to the sagittal

1 plane of the midline to identify the transverse colon located on the caudal side of  
2 the gastric antrum. The probe was then rotated 90° to trace the transverse colon  
3 from the hepatic to the splenic flexure. The descending colon was identified on  
4 the left flank. Finally, the colon was traced from the sigmoid colon to the rectum,  
5 which was visualized through the urinary bladder. Convex probes (center  
6 frequency: 3.75/ 6 MHz) were used to study the rectum. Patients were not asked  
7 to fill their bladders before US.

8 We divided the colon into six segments, including the cecum, ascending colon  
9 (A-colon), right-sided trans- verse colon (right T-colon), left-sided transverse  
10 colon (left T-colon), descending colon (D-colon), sigmoid colon (S-colon) and  
11 rectum, which were examined. The sonographers at the examination facilities  
12 saved still images and movie clips if needed.

13 The US severity was graded on a scale of 1-4 created according to a previously  
14 reported study (Hata et al. 1992). We verified this scale in our preliminary study  
15 (Wada et al. 2015) and used it in recently published prospective study (Kinoshita  
16 et al. 2018) as follows: grade 1 = normal thickening of the colonic wall, <4 mm  
17 (Hagiu and Badea 2007); grade 2 = thickened mucosa and submucosa  
18 without hypo-echoic change of the submucosa; grade 3 = bowel wall thickening

1 with loss of stratification; and grade 4 = bowel wall thickening with loss of  
2 stratification and irregular mucosa or hyper-echogenic areas in the mucosa,  
3 suggesting ulceration (Fig. 1). Furthermore, colonic segments that were difficult  
4 to observe were defined as grade “x”.

5 Initially, the US grade was assessed at each examining facility. All sonographers  
6 were aware of the UC diagnosis but were blinded to the patients’ clinical infor-  
7 mation and CS findings. Next, still images and movie clips were sent to and re-  
8 assessed at the central facility by two affiliated registered medical sonographers  
9 (M.N. and S. O.) specialized in gastroenterology and hepatology and having more  
10 than 5 y of experience. These sonographers interpreted the US grades via  
11 consensus. Central readers were also blinded to the interpretations of the US  
12 grades made by examining facilities. The grading agreement between the  
13 examining and central facilities was then evaluated. Before study initiation,  
14 sonographers at all facilities attended four training sessions to ensure a  
15 consistent interpretation of the US techniques and grading severity.

16

### 17 ***Colonoscopy***

18 Colonoscopy was performed by five expert endoscopists who had conducted at

1 least 2000 CSs using standard CS tools (PCF-Q260, CF-HQ290 I, CF-H260 AZI,  
2 CF-Q240 I; Olympus, Tokyo, Japan) before conducting the study procedures. A  
3 polyethylene glycol-based bowel preparation was provided to the patients. For  
4 patients with severe UC, CS was performed without bowel preparation to avoid  
5 the risk of perforation. Six colonic segments and the rectum were examined by  
6 CS.

7 In each segment, disease activity was assessed according to Matts' endoscopic  
8 classification (CS grades1-4) as follows: grade 1 = remission; grade 2 = mild  
9 activity; grade 3 = moderate activity; grade 4 = severe activity (Hozumi et al. 2013;  
10 Matts 1961; Walsh et al. 2014) (Fig. 2). The physicians performing the endo  
11 scopic examinations were aware of the UC diagnosis but were blinded to the  
12 results of the US examinations. First, CS grade was assessed at each examining  
13 facility. Next, still images were sent to and re-assessed at the central facility.  
14 Three endoscopists (K.K, R.O. and T.K) affiliated with the central facility who had  
15 more than 4 y of experience in performing CS interpreted CS grading by  
16 consensus. They were also blinded to the patients' clinical information.

17

18 ***Statistical analysis***

1 The US and CS visualization rates for each colonic segment were compared  
2 using McNemar tests. Grading agreement for US and CS between the examining  
3 and central facilities was assessed for each segment and all colonic segments,  
4 and the maximum US and CS grades, which represented the disease activity of  
5 each patient, were evaluated using weighted k statistics. The k values were  
6 interpreted as follows:  $k \leq 0.20$ , slight agreement; 0.21-0.40, fair agreement;  
7 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; 0.81-1.00,  
8 almost perfect agreement (Landis and Koch 1977). p Values < 0.05 were  
9 considered to indicate statistical significance. The statistical analysis was  
10 performed using standard statistical software (SPSS Statistics Version 23.0 [IBM,  
11 Armonk, NY, USA] and the Bell Curve for Excel Version 2.11 [Social Survey  
12 Research Information Corp., Tokyo, Japan]).

13

## 14 **RESULTS**

15 During the study period, 72 consecutive UC patients underwent US and CS; of  
16 these patients, 14 did not undergo US and CS within 2 d, and 1 was suspected  
17 of having Crohn's disease. The characteristics of the remaining 57 patients  
18 included in this study are summarized in Table 1. Forty-four patients underwent

1 US and CS on the same day, and all the patients underwent CS after US. Total  
2 CS was achieved in 48 (84.2%) patients; in the remaining 9 (15.8%) patients, CS  
3 was incomplete because of disease severity or insufficient preparation of the  
4 colon. In 57% and 20.8% of patients, the Rachmilewitz CAI assessment indicated  
5 inactive and moderately active disease, respectively (Table 1).

6 Table 2 outlines the details of the site-specific visualization rates, which were  
7 higher with US than with CS. Visualization rates for the A-colon, right T-colon and  
8 rectum differed significantly between US and CS. US grading agreement for  
9 evaluation of UC disease activity was substantial in the majority of the segments,  
10 but moderate in the cecum and rectum. CS grading agreement was substantial  
11 except for the rectum (0.43) (Table 3). For the maximum grade in each patient,  
12 the US and CS grading agreement values were 0.82 and 0.70, respectively (Table  
13 4). Notably, US achieved almost perfect agreement. Representative US findings  
14 are illustrated in Figures 3-5.

15

## 16 **DISCUSSION**

17 Reports have described the usefulness of US for assessing disease activity  
18 and therapeutic response in patients with UC (Allocca et al. 2018; Parente et al.

1       2010). However, previous studies were conducted at single centers and  
2       excluded patients in whom the disease was localized to the rectum. Other studies  
3       have reported that US can be used to observe the ileocecum, sigmoid and  
4       descending colon in the majority of patients (Bremner et al. 2006; Panes et al.  
5       2013). However, the present study achieved excellent US visualization rates in  
6       each segment (>90%), as well as in the rectum (84%). Recent advances in US  
7       equipment have enabled exploration of nearly the entire colon. However, the  
8       slightly lower visualization rate for the rectum was likely attributable to its deeper  
9       location in the pelvis. In contrast, the CS visualization rate was lower in the right-  
10      sided colon, which was attributed to insufficient preparation of the colon to avoid  
11      possible risks of perforation or aggravation of disease activity in cases of severe  
12      UC. Our results suggest that US might provide a feasible alternative means of  
13      evaluating the right-sided colon in these severe cases of UC.

14      Previous studies reported moderate to substantial inter-observer agreement  
15      between computed tomography (Horvat et al. 2016) and/or magnetic resonance  
16      imaging (MRI) enterography (AlSabban et al. 2017) intended to image  
17      inflammatory bowel diseases. Another study reported inter-observer agreement  
18      with k values of 0.22-0.85 for US evaluation of Crohn's disease and suggested

1 that the poor reproducibility might be attributable to difficulties in recognizing the  
2 layered structure (Fraquelli et al. 2008). Allocca et al. (2018) recently reported  
3 almost perfect agreement ( $k = 0.86$ ) between two physicians who conducted US  
4 evaluations of UC activity. In that study, however, US was performed by expert  
5 physicians at the same facility.

6 As previously published reports described US evaluations performed mostly by  
7 expert physicians at single centers, this is the first inter-rater validation study  
8 involving sonographers with a wide range of US experience at four examining  
9 facilities and a central facility. Notably, we achieved good US grading agreement  
10 and convergence in the majority of segments ( $k = 0.72-0.78$ ), as well as moderate  
11 grading agreement and convergence in the cecum and rectum (both with  $k =$   
12  $0.54$ ). Our results suggest that the use of standardized US procedures and unified  
13 interpretations might be necessary to generalize the US grading system for  
14 assessing UC activity.

15 In our study, 20% ( $n = 9/44$ ) of the cases exhibited inconsistencies in the rectum,  
16 which were attributed to an inability to use the high-frequency probes required for  
17 a detailed evaluation. Intermediate findings can be difficult to interpret, especially  
18 because US grade 2 encompasses a wide range of imaging findings. Therefore,

1 bowel wall thickness with or without irregular mucosa or a hyper-echogenic  
2 shallow concavity in the mucosa were difficult to grade using only still images. In  
3 our study, CS grading agreement was substantial ( $k = 0.68-0.79$ ), with the  
4 exception of the rectum. In a previous study, expert endoscopists determined a  
5 weighted  $k$  value of 0.76 for Mats' activity indices (Osada et al. 2010). Their result  
6 was concordant with our findings.

7 We note that the decreased US and CS concordance rates in our study might  
8 be due to the enrollment of mostly patients with inactive to mild disease activity.  
9 In a previous review of MRI, Horsthuis et al. (2009) stated that severe disease  
10 has more pronounced features and is therefore easier to diagnose than mild  
11 disease.

12 This study had several limitations. First, the study population included a small  
13 proportion of patients with moderate to severe disease activity (28.3%). Second,  
14 20 patients did not undergo US and CS on the same day, which may have  
15 affected our results. Finally, we compared recorded images rather than repeating  
16 the examination for the same patients. There was a difference between repeated  
17 examinations and re-assessed recorded images. However, in a multicenter study  
18 setting, it would be difficult to repeat both US and CS examinations at four

1 different facilities during a short interval.

2 Despite these limitations, this study represents a meaningful contribution with  
3 respect to US evaluation of disease activity associated with UC. Our findings  
4 suggest that it is possible to apply US not only at expert facilities, but also in  
5 trained facilities, if the inspection procedures and US techniques are  
6 standardized.

7

## 8 **CONCLUSIONS**

9 We achieved good to excellent agreement between the US results obtained at  
10 the examining facilities and central facility. These findings suggest that US is not  
11 inferior to CS for evaluating UC disease activity. Therefore, US can be used to  
12 evaluate UC disease activity at different facilities.

13

## 14 **ACKNOWLEDGMENTS**

15 The authors would like to thank Mami Sato, Department of Clinical Laboratory,  
16 Tomakomai City Hospital; Kazunori Eto, Department of Gastroenterology,  
17 Tomakomai City Hospital; Atsuo Maemoto, Inflammatory Bowel Disease Center,  
18 Sapporo Higashi Tokushukai Hospital; Junji Yamamoto, Ohguro

- 1 Gastroenterological Hospital; Shinji Yoshii, Department of Gastroenterology, NTT
- 2 East Japan Sapporo Hospital; and Kota Ono, Clinical Research and Medical
- 3 Innovation Center, Hokkaido University Hospital, for their support.
- 4

1   **REFERENCES**

- 2   Allocca M, Fiorino G, Bonovas S, Furfaro F, Gilardi D, Argollo M, Magnoni P,  
3       Peyrin-Biroulet L, Danese S. Accuracy of Humanitas Ultrasound Criteria in  
4       assessing disease activity and severity in ulcerative colitis: A prospective  
5       study. *J Crohns Colitis* 2018;12:1385-91.
- 6   AlSabban Z, Church P, Moineddin R, Navarro OM, Greer ML, Walters T,  
7       Chavhan GB. Accuracy and interobserver agreement of diffusion-weighted  
8       imaging in pediatric inflammatory bowel disease. *Clin Imaging* 2017;41:14-  
9       22.
- 10  Bremner AR, Griffiths M, Argent JD, Fairhurst JJ, Beattie RM. Sonographic  
11       evaluation of inflammatory bowel disease: a prospective, blinded,  
12       comparative study. *Pediatr Radiol* 2006;36:947-53.
- 13  Civitelli F, Di Nardo G, Oliva S, Nuti F, Ferrari F, Dilillo A, Viola F, Pallotta N,  
14       Cucchiara S, Aloï M. Ultrasonography of the colon in pediatric ulcerative  
15       colitis: a prospective, blind, comparative study with colonoscopy. *J Pediatr*  
16       2014;165:78-84.
- 17  Danese S, Fiocchi C. Ulcerative colitis. *N Engl J Med* 2011;365:1713-25.

1 Dignass A, Lindsay JO, Sturm A, Windsor A, Colombel JF, Allez M, D'Haens G,  
2 D'Hoore A, Mantzaris G, Novacek G, Oresland T, Reinisch W, Sans M,  
3 Stange E, Vermeire S, Travis S, Van Assche G. Second European evidence-  
4 based consensus on the diagnosis and management of ulcerative colitis part  
5 2: current management. *J Crohns Colitis* 2012;6:991-1030.

6 Fraquelli M, Sarno A, Girelli C, Laudi C, Buscarini E, Villa C, Robotti D, Porta P,  
7 Cammarota T, Ercole E, Rigazio C, Senore C, Pera A, Malacrida V, Gallo C,  
8 Maconi G. Reproducibility of bowel ultrasonography in the evaluation of  
9 Crohn's disease. *Dig Liver Dis* 2008;40:860-6.

10 Hagi C, Badea R. Applicability of abdominal ultrasonography in inflammatory  
11 bowel diseases. *J Gastrointestin Liver Dis* 2007;16:205-9.

12 Hata J, Haruma K, Suenaga K, Yoshihara M, Yamamoto G, Tanaka S, et al.  
13 Ultrasonographic assessment of inflammatory bowel disease. *Am J*  
14 *Gastroenterol* 1992;87:443-7.

15 Horsthuis K, Bipat S, Stokkers PC, Stoker J. Magnetic resonance imaging for  
16 evaluation of disease activity in Crohn's disease: a systematic review. *Eur*  
17 *Radiol* 2009;19:1450-60.

1 Horvat N, Tavares CC, Andrade AR, Cabral JC, Leao-Filho HM, Caiado AH,  
2 Ueda SK, Leite AZ, Sipahi AM, Rocha MS. Inter- and intraobserver  
3 agreement in computed tomography enterography in inflammatory bowel  
4 disease. *World J Gastroenterol* 2016;22:10002-8.

5 Hozumi H, Hokari R, Kurihara C, Narimatsu K, Sato H, Sato S, Ueda T,  
6 Higashiyama M, Okada Y, Watanabe C, Komoto S, Tomita K, Kawaguchi A,  
7 Nagao S, Miura S. Endoscopic finding of spontaneous hemorrhage  
8 correlates with tumor necrosis factor alpha expression in colonic mucosa of  
9 patients with ulcerative colitis. *Int J Colorectal Dis* 2013;28:1049-55.

10 Kinoshita K, Katsurada T, Nishida M, Omotehara S, Onishi R, Mabe K, Onodera  
11 A, Sato M, Eto K, Suya M, Maemoto A, Hasegawa T, Yamamoto J,  
12 Mitsumori D, Yoshii S, Ono K, Sakamoto N. Usefulness of transabdominal  
13 ultrasonography for assessing ulcerative colitis: a prospective, multicenter  
14 study. *J Gastroenterol* 2018; [Epub ahead of print].

15 Kucharzik T, Petersen F, Maaser C. Bowel ultrasonography in inflammatory  
16 bowel disease. *Dig Dis* 2015;14:17-25.

17 Landis JR, Koch GG. The measurement of observer agreement for categorical  
18 data. *Biometrics* 1977;33:159-74.

1 Matts SG. The value of rectal biopsy in the diagnosis of ulcerative colitis. *Q J*  
2 *Med* 1961;30:393-407.

3 Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione  
4 R, Ghosh S, Wu JCY, Chan FKL, Sung JJY, Kaplan GG. Worldwide  
5 incidence and prevalence of inflammatory bowel disease in the 21st century:  
6 a systematic review of population-based studies. *Lancet* 2018;390:2769-78.

7 Osada T, Ohkusa T, Yokoyama T, Shibuya T, Sakamoto N, Beppu K, Nagahara  
8 A, Otaka M, Ogihara T, Watanabe S. Comparison of several activity indices  
9 for the evaluation of endoscopic activity in UC: inter- and intraobserver  
10 consistency. *Inflamm Bowel Dis* 2010;16:192-7.

11 Panes J, Bouhnik Y, Reinisch W, Stoker J, Taylor SA, Baumgart DC, Danese S,  
12 Halligan S, Marincek B, Matos C, Peyrin-Biroulet L, Rimola J, Rogler G, van  
13 Assche G, Ardizzone S, Ba-Ssalamah A, Bali MA, Bellini D, Biancone L,  
14 Castiglione F, Ehehalt R, Grassi R, Kucharzik T, Maccioni F, Maconi G,  
15 Magro F, Martín-Comín J, Morana G, Pendsé D, Sebastian S, Signore A,  
16 Tolan D, Tielbeek JA, Weishaupt D, Wiarda B, Laghi A. Imaging techniques  
17 for assessment of inflammatory bowel disease: joint ECCO and ESGAR  
18 evidence-based consensus guidelines. *J Crohns Colitis* 2013;7:556-85.

1 Parente F, Molteni M, Marino B, Colli A, Ardizzone S, Greco S, Sampietro G,  
2 Foschi D, Gallus S. Are colonoscopy and bowel ultrasound useful for  
3 assessing response to short-term therapy and predicting disease outcome of  
4 moderate-to-severe forms of ulcerative colitis? A prospective study. Am J  
5 Gastroenterol 2010;105:1150-7.

6 Rachmilewitz D. Coated mesalazine (5-aminosalicylic acid) versus  
7 sulphasalazine in the treatment of active ulcerative colitis: a randomised  
8 trial. BMJ 1989;298:82-6.

9 Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative  
10 colitis. Lancet 2017;389:1756-70.

11 Wada T, Nishida M, Kudou S, Omotehara S. Comparison Between  
12 Transabdominal Ultrasound and Colonoscopy in Evaluating Disease Activity  
13 of Ulcerative Colitis (in Japanese). Japanese journal of medical ultrasound  
14 technology 2015;40:141-9.

15 Walsh A, Palmer R, Travis S. Mucosal healing as a target of therapy for colonic  
16 inflammatory bowel disease and methods to score disease activity.  
17 Gastrointest Endosc Clin N Am 2014;24:367-78.

18

1 **Figure Captions List**

2 Figure 1. Ultrasonography Grade

3 (A) Ultrasonography Grade 1: normal thickness of the colonic wall and all wall  
4 layers are maintained (arrows). (B) Ultrasonography Grade 2: thickened  
5 mucosa and submucosa without hypo-echoic change in the submucosa  
6 (arrows). (C) Ultrasonography Grade 3: thickened mucosa and submucosa with  
7 loss of stratification (arrows). (D) Ultrasonography Grade 4: bowel wall  
8 thickening with loss of stratification (yellow arrows), and irregular mucosa or  
9 hyper-echogenic areas in the mucosa (pink arrowhead) suggesting ulceration.

10

11 Figure 2. Colonoscopy Grade

12 (A) Matts' grade 1: normal. (B) Matts' grade 2: mild granularity of the mucosa  
13 with mild contact bleeding. (C) Matts' grade 3: marked granularity and edema of  
14 the mucosa, contact bleeding and spontaneous bleeding. (D) Matts' grade 4:  
15 severe ulceration of the mucosa with hemorrhaging.

16

17 Figure 3. Forty-year-old woman with extensive ulcerative colitis and a

18 Rachmilewitz clinical activity index of 11. Agreement was determined between

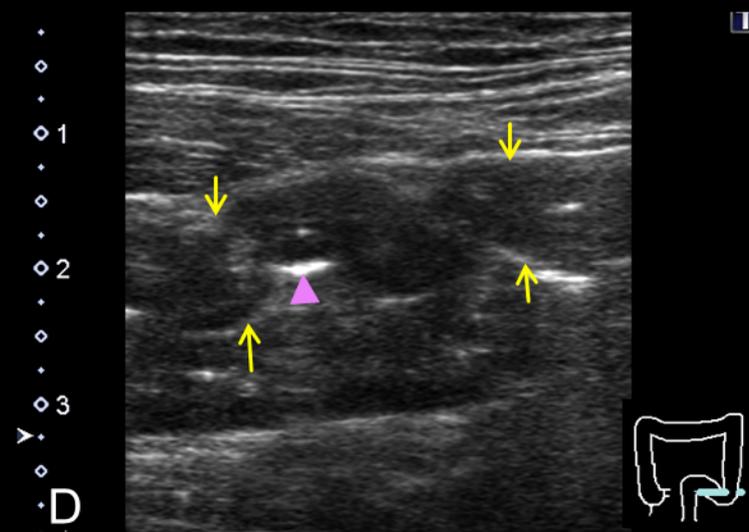
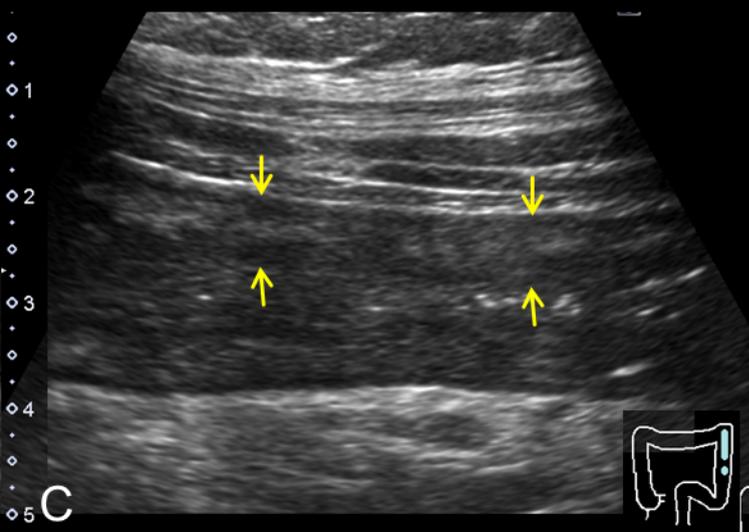
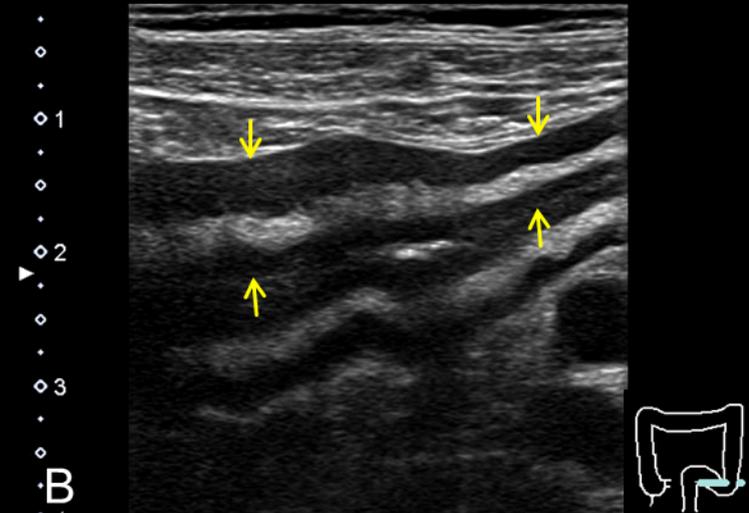
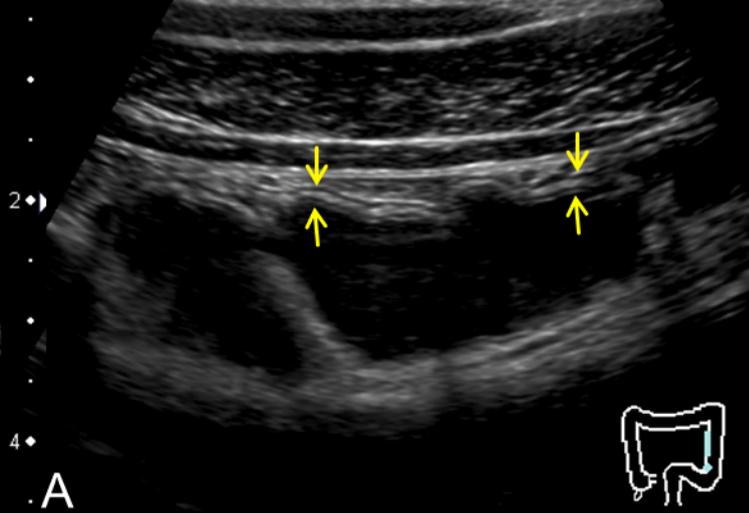
1 the examining and central facilities for the interpretations of US images of the  
2 sigmoid colon. (A) Both the central and examining facilities interpreted the  
3 ultrasonography image as grade 4; moderate to severe wall thickening, loss of  
4 wall layer stratification (hypo-echoic change of submucosa) and disruption of  
5 the mucosa or submucosal layer were identified. (B) Both the central and  
6 examining facilities interpreted the colonoscopy findings as Matts' grade 4.

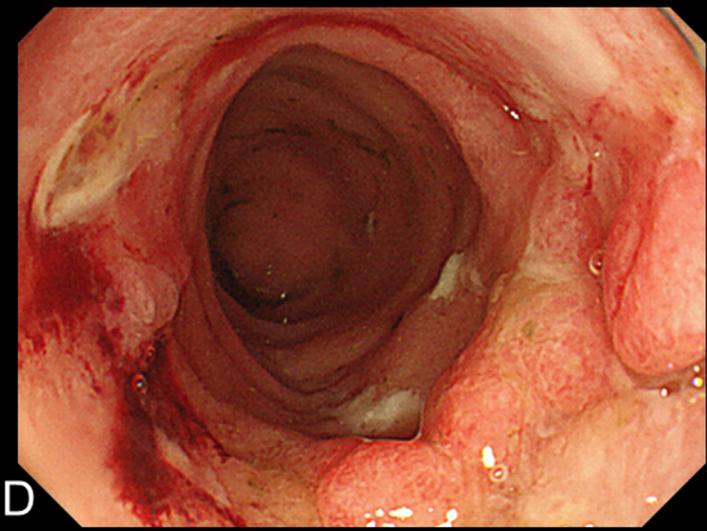
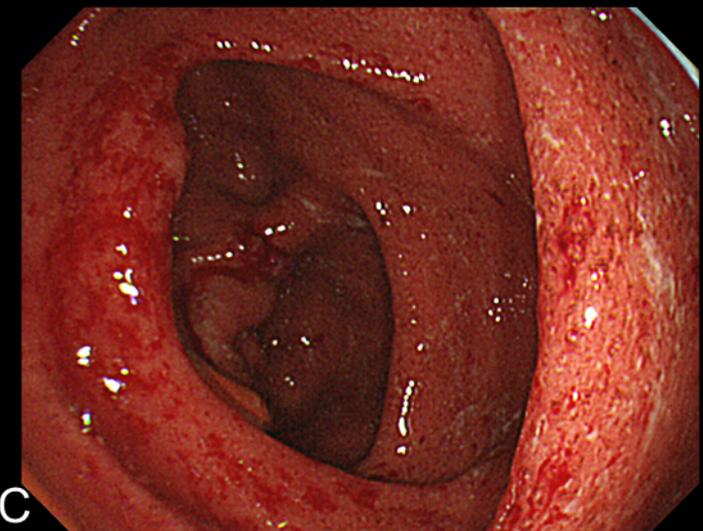
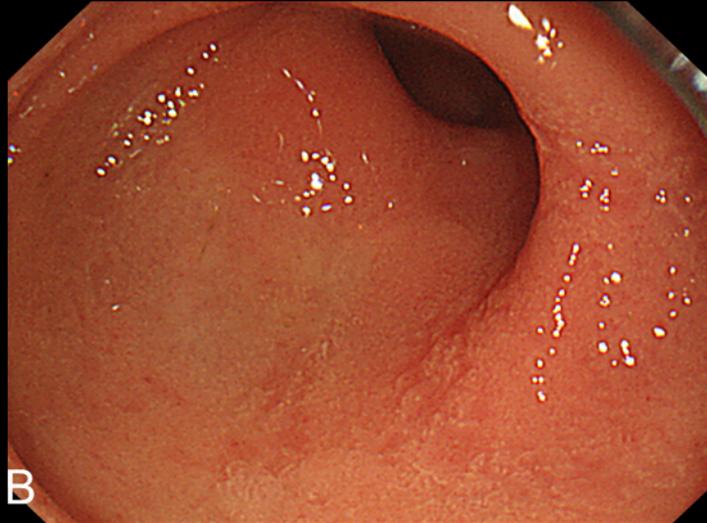
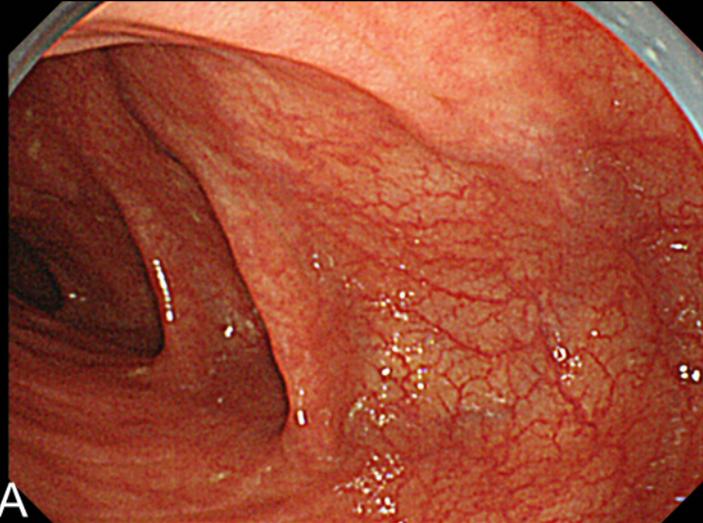
7

8 Figure 4. Thirty-four-year-old woman with extensive ulcerative colitis and a  
9 Rachmilewitz clinical activity index of 13. Disagreement occurred during  
10 interpretation of the US image of the cecum. (A) The examining facility interpreted  
11 the ultrasonography as grade 4, whereas the central facility interpreted it as grade  
12 2. Wall thickening without hypo-echoic changes in the submucosa was identified  
13 by ultrasonography. However, the image was misinterpreted as a loss of layer  
14 structure because of the ambiguous muscularis propria layer and serosal  
15 border of the posterior wall (yellow arrowhead). (B) The central and examining  
16 facilities both interpreted the colonoscopy as Matts' grade 2.

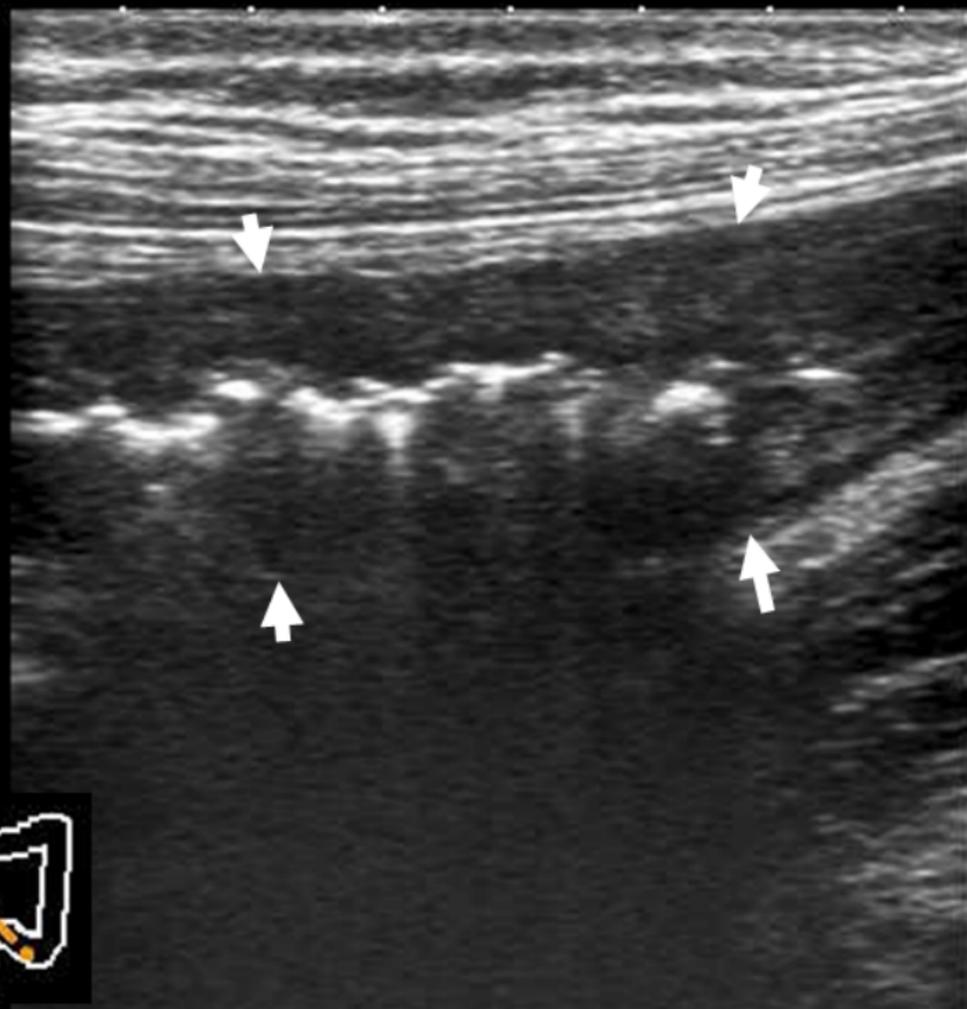
17

1 Figure 5. Fifty-seven-year-old man with extensive ulcerative colitis.  
2 Disagreement occurred during interpretation of the US image of the left  
3 transverse colon. (A) The examining facility interpreted the ultrasonography as  
4 grade 1, whereas the central facility interpreted the ultrasonography as grade 2.  
5 A very slightly thickened mucosa and submucosa with a layered structure were  
6 identified on the ultrasonography image. (B) Interpretation of the colonoscopy  
7 was also inconsistent between the facilities. The examining facility interpreted  
8 the colonoscopy as Mats' grade 1, whereas the central facility interpreted it as  
9 Mats' grade 2.

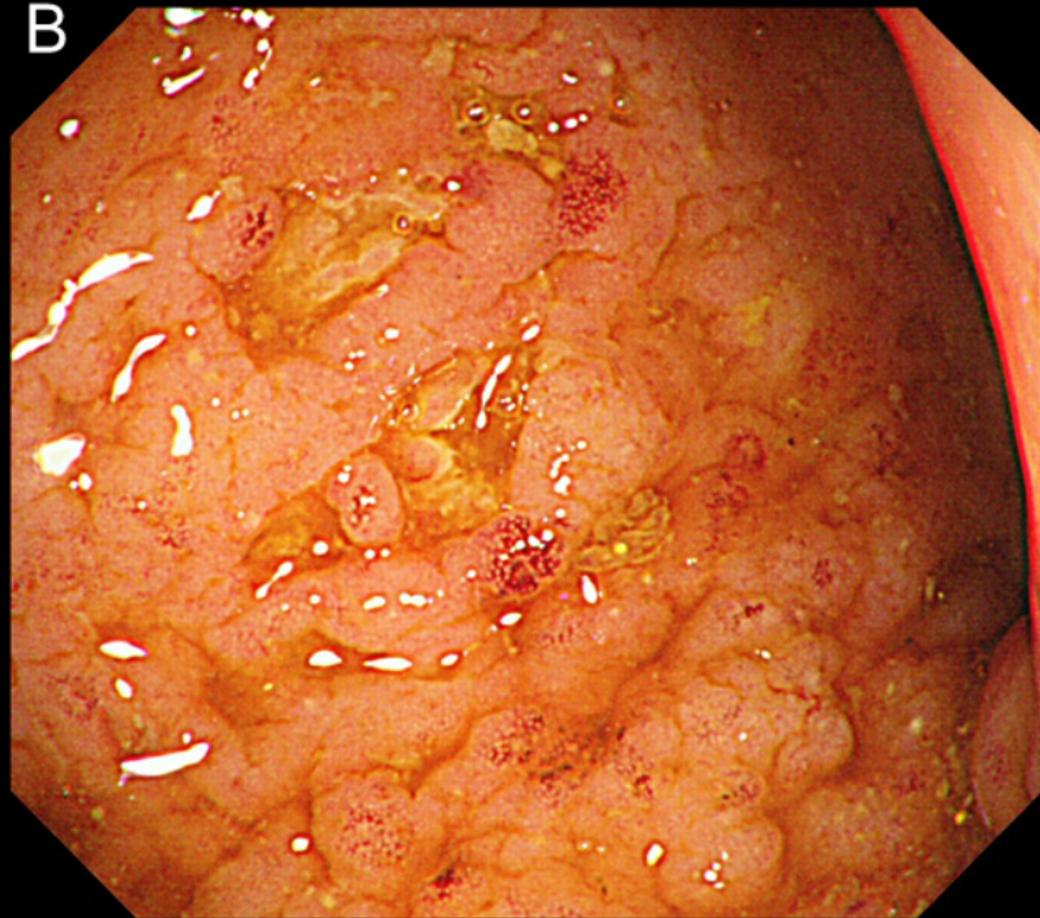


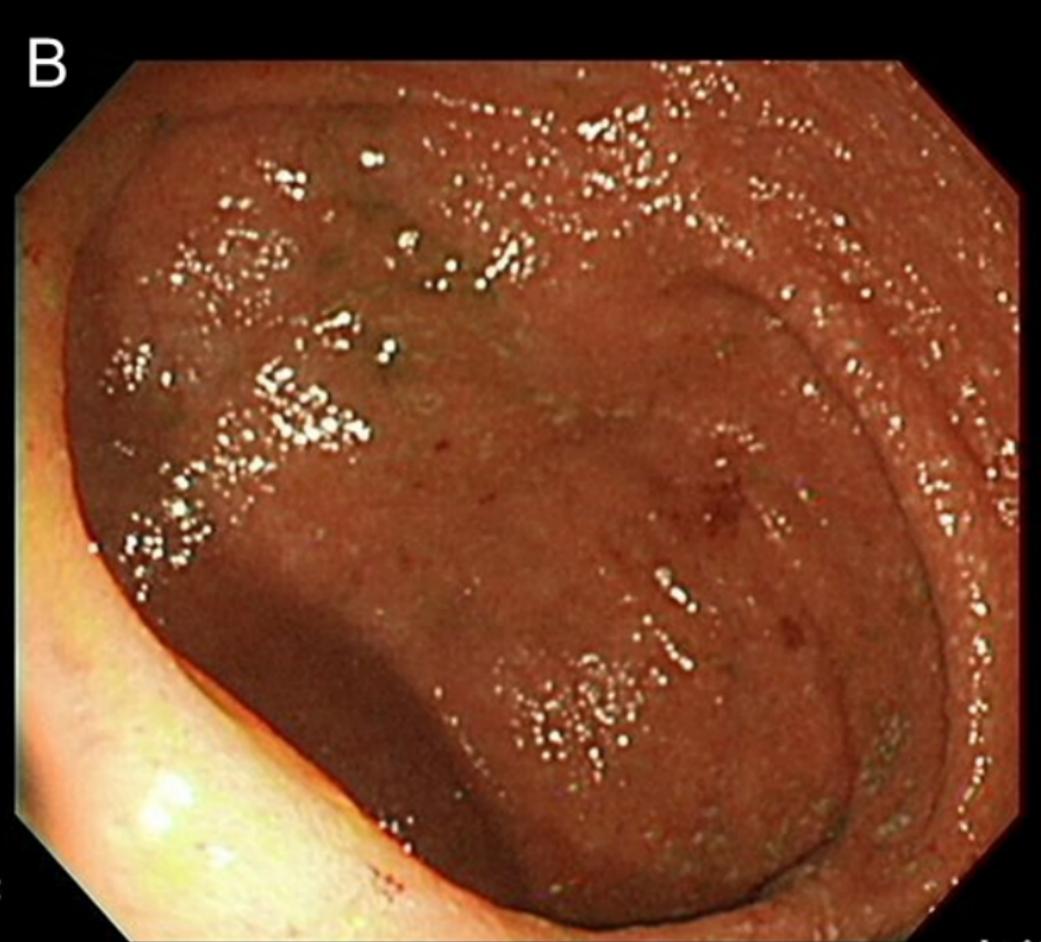
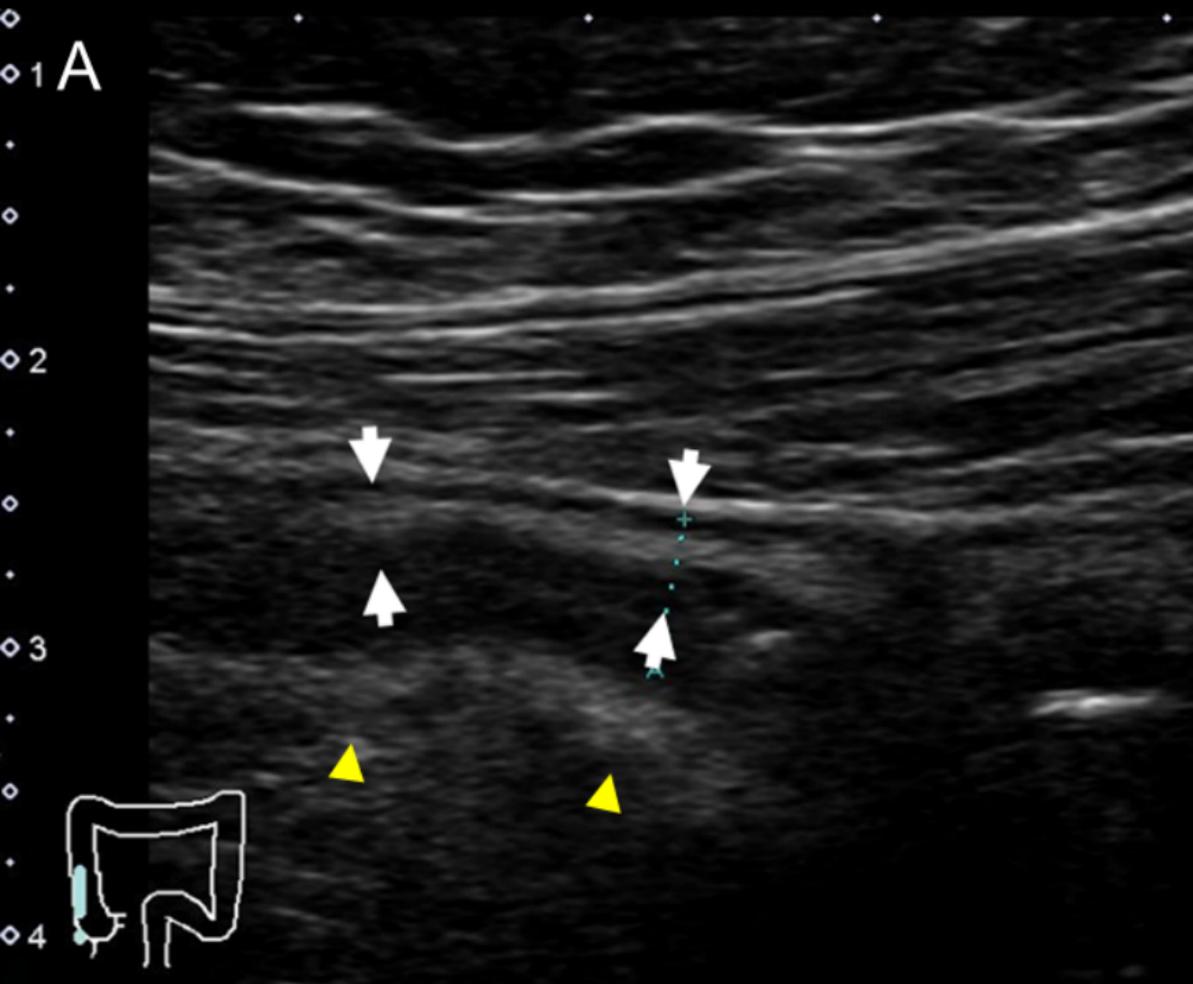


A

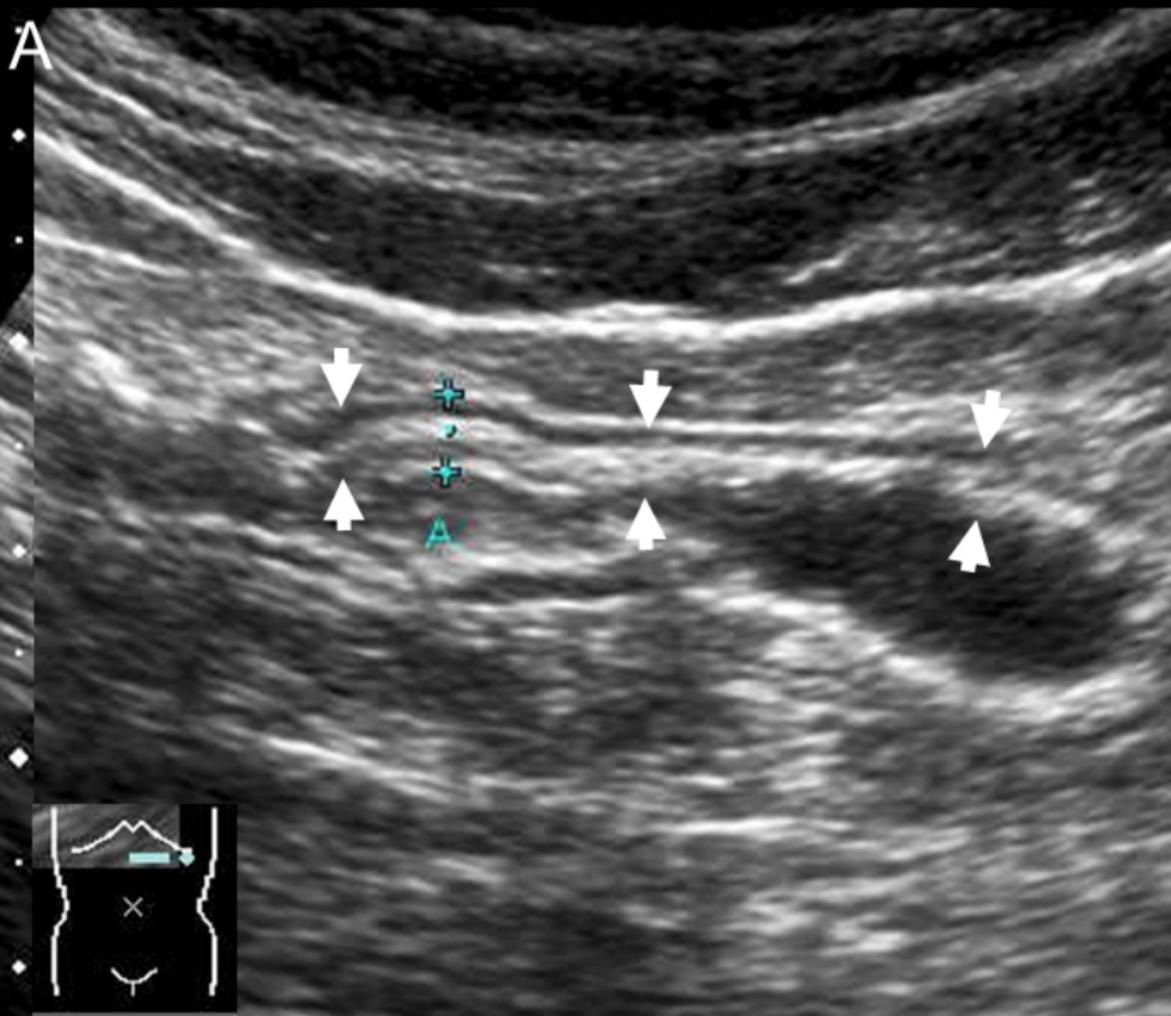


B





A



B

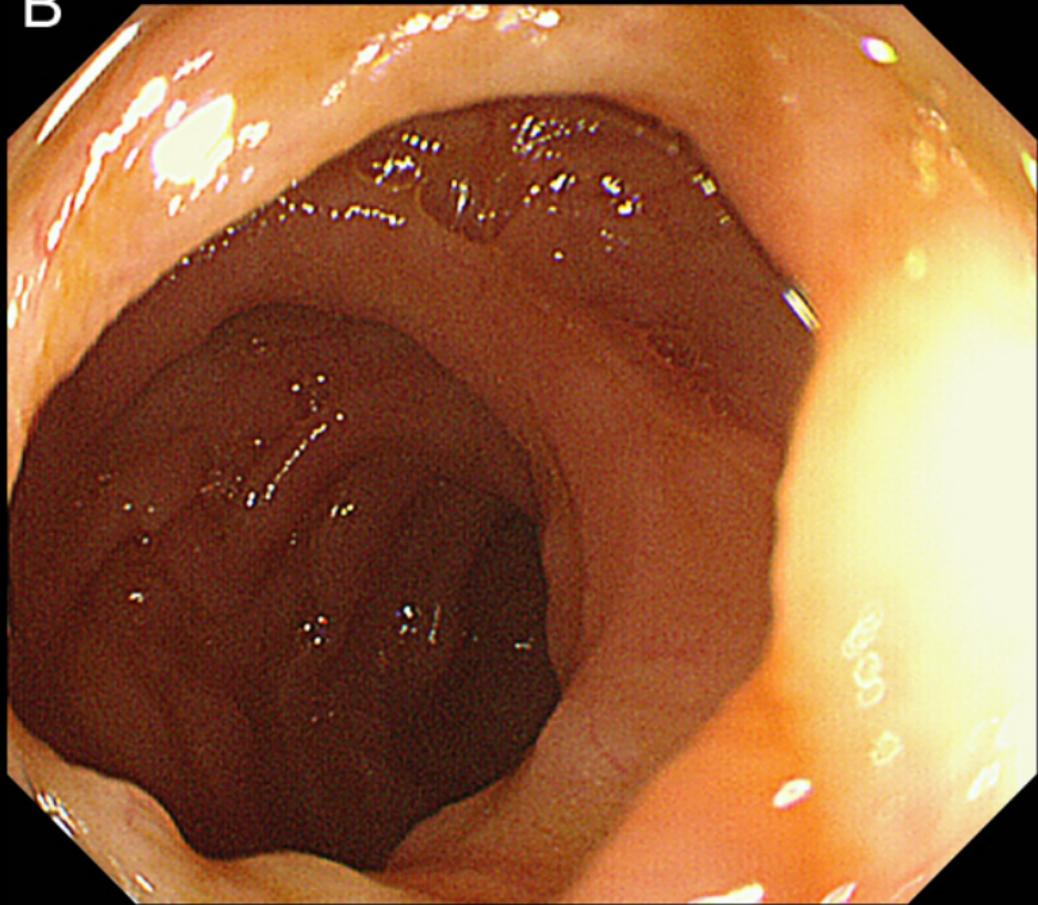


Table 1. Patient characteristic (n=57)

Median Age (years, range)	48 (19-80)
Male/Female	37/20
Disease duration (years, range) <sup>a</sup>	5.0 (0.1-32.3)
Disease extension	
Proctitis, % (n)	19.3 (11)
Left-sided colitis, % (n)	28.1 (16)
Extensive colitis, % (n)	52.6 (30)
Rachmilewits CAI, median (range) <sup>b</sup>	2.0 (0-14)
Inactive, % (n)	56.6 (30)
Mild, % (n)	15.1 (8)
Moderate, % (n)	20.8 (11)
Severe, % (n)	7.5 (4)
Concomitant medications, % (n)	
Mesalazine	86.0 (49)
Azathioprine/6-mercaptopurine	22.8 (13)
Oral or intravenous corticosteroids	15.8 (9)
Cytoapheresis	7.0 (4)
TNF- $\alpha$ antagonists	5.3 (3)
US maximum grade, median (range)	2 (1-4)
CS maximum grade, median (range)	2 (1-4)
CRP (mg/dL), median (range) <sup>c</sup>	0.04 (0.01-11.82)

Abbreviation: TNF- $\alpha$ , Tumor Necrosis Factor-alpha

<sup>a</sup>Data for the disease duration were missing for 16 patients.

<sup>b</sup>Data for the CAI score were missing for 4 patients.

<sup>c</sup>Data for the CRP were missing for 2 patients.

Table 2. Visualization rate in each colonic segment examined by US and CS

Colonic segment	US, n (%)	CS, n (%)	<i>P</i> -value
Cecum	54 (94.7)	48 (84.2)	0.070
Ascending colon	56 (98.2)	48 (84.2)	<u>0.008</u>
Right t-colon	55 (96.5)	47 (82.5)	<u>0.008</u>
Left t-colon	53 (93.0)	49 (86.0)	0.289
Descending colon	57 (100)	52 (91.2)	0.063
Sigmoid colon	57 (100)	56 (98.2)	1.000
Rectum	48 (84.2)	57 (100)	<u>0.004</u>
All segments	380 (95.2)	357 (89.5)	<u>0.001</u>

Abbreviations: Right t-colon, Right-sided transverse colon; Left t-colon, Left-sided transverse colon

Table 3. Grading agreement of evaluating UC disease activity in US and CS

Colonic segment	US			CS		
	n	kappa	95%CI	n	kappa	95%CI
Cecum	53	0.54 <sup>†</sup>	0.37-0.70	48	0.71 <sup>†</sup>	0.50-0.92
Ascending colon	54	0.74 <sup>†</sup>	0.57-0.92	48	0.73 <sup>†</sup>	0.60-0.86
Right T-colon	55	0.72 <sup>†</sup>	0.58-0.85	47	0.69 <sup>†</sup>	0.53-0.84
Left T-colon	52	0.75 <sup>†</sup>	0.61-0.89	48	0.79 <sup>†</sup>	0.65-0.93
Descending colon	57	0.78 <sup>†</sup>	0.68-0.89	50	0.68 <sup>†</sup>	0.53-0.82
Sigmoid colon	57	0.77 <sup>†</sup>	0.65-0.89	56	0.76 <sup>†</sup>	0.64-0.88
Rectum	44	0.54 <sup>†</sup>	0.34-0.74	57	0.43 <sup>†</sup>	0.15-0.72
All segments	372	0.75 <sup>†</sup>	0.70-0.80	354	0.72 <sup>†</sup>	0.66-0.79
Maximum grade in each patient	57	0.82 <sup>†</sup>	0.73-0.91	57	0.70 <sup>†</sup>	0.53-0.88

Abbreviations: Right T-colon, Right-sided Transverse colon; Left T-colon, Left-sided Transverse colon

†:  $p < 0.001$

Table 4. US and CS grading agreement of the maximum grade in each patient (n=57)

		Examining facilities			
US grade		1	2	3	4
Central facility	1	5	1	0	0
	2	6	24	4	0
	3	0	0	9	1
	4	0	0	3	4

		Examining facilities			
CS grade		1	2	3	4
Central facility	1	1	1	0	0
	2	12	17	3	0
	3	0	1	11	0
	4	1	0	3	7