Manuscript Doi: 10.1093/ecco-jcc/jjy048

Ultrasound for assessing disease activity in IBD patients: a systematic review of activity scores

S. Bots[1], K. Nylund[2], M. Löwenberg[1], K. Gecse[1], O.H. Gilja[2,3], G. D'Haens[1]

[1] Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands

[2] National Centre for Ultrasound in Gastroenterology, Haukeland University Hospital, Bergen, Norway

[3] Department of Clinical Medicine, University of Bergen, Bergen, Norway

Short title: Ultrasound for assessing disease activity in IBD.

Corresponding author:

Prof. dr. G. D'Haens
Meibergdreef 9 1105 AZ Amsterdam, The Netherlands
T: +31 20 5663534
F: +31 20 6917033
E: g.dhaens@amc.uva.nl

Authors contact details:

Steven Bots
T: +31 20 5661613
E: s.j.bots@amc.uva.nl

Kim Nylund
E: kim.Nylund@uib.no

Mark Löwenberg
E: m.lowenberg@amc.uva.nl

Krisztina Gecse
E: k.b.gecse@amc.uva.nl

Odd Helge Gilja
E: odd.helge.gilja@helse-bergen.no

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Conflicts of interest:

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

Kim Nylund has served as speaker for MEDA AS and Ferring Pharmaceuticals.

Mark Löwenberg has served as speaker and/or principal investigator for Abbvie, Covidien, Dr. Falk, Ferring Pharmaceuticals, Merck Sharp & Dohme, Receptos, Takeda, Tillotts and Tramedico. He has received research grants from AbbVie, Merck Sharp & Dohme, Achmea healthcare and ZonMW.

Krisztina Gecse has served as speaker and/or advisor for Amgen, AbbVie, Boehringer Ingelheim, Ferring, Hospira, MSD, Pfizer, Samsung Bioepis, Sandoz, Takeda, Tigenix and Tillotts.

Odd Helge Gilja has served as advisor for Abbvie, Bracco, Samsung and GE Healthcare and received speaker fees from Abbvie, Bracco, Almirall, GE Healthcare, Takeda AS, Meda AS, Ferring AS and Allergan.

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.

Authors contributions:

S.B.: Study design, study selection, data acquisition, data interpretation, writing first draft of the manuscript and final approval of the manuscript.

K.N.: Study design, study selection, data acquisition, data interpretation, revising the manuscript and final approval of the manuscript.

M.L.: Revising the manuscript and final approval of the manuscript.

K.G.: Revising the manuscript and final approval of the manuscript.

O.H.G.: Revising the manuscript and final approval of the manuscript.

G.D.: Revising the manuscript and final approval of the manuscript.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Abstract

Background and aims

Ultrasound (US) indices for assessing disease activity in IBD patients have never been critically reviewed. We aimed to systematically review the quality and reliability of available ultrasound (US) indices compared with reference standards for grading disease activity in IBD patients.

Methods

Pubmed, Embase and Medline were searched from 1990 until June 2017. Relevant publications were identified through full text review after initial screening by 2 investigators. Data on methodology and index characteristics were collected. Study quality was assessed with a modified version of the Quadas-2 tool for risk of bias assessment.

Results

Of 20 studies with an US index, 11 studies met the inclusion criteria. Out of these 11 studies, 7 and 4 studied CD and UC activity indices, respectively. Parameters that were used in these indices included bowel wall thickness (BWT), Doppler signal (DS), wall layer stratification (WLS), compressibility, peristalsis, haustrations, fatty wrapping, contrast enhancement (CE) and strain pattern. Study quality was graded high in 5 studies, moderate in 3 studies and low in 3 studies. Ileocolonoscopy was used as the reference standard in 9 studies. In 1 study a combined index of ileocolonoscopy and barium contrast radiography and in 1 study histology was used as the reference standard. Only 5 studies used an established endoscopic index for comparison with US.

Conclusions

Several US indices for assessing disease activity in IBD are available; however the methodology for development was suboptimal in most studies. For the development of future indices stringent methodological design is required.

Keywords: Imaging, gastrointestinal ultrasound, inflammatory bowel disease.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Introduction

Assessing disease activity in inflammatory bowel disease (IBD) patients is becoming increasingly important. Treatment targets in IBD patients are shifting from symptom control to intestinal repair, an endpoint that has been associated with improved long-term outcomes.[1, 2] Ileocolonoscopy is the gold standard for the assessment of disease activity in IBD patients. Therefore, it is increasingly being implemented to guide treatment decisions and to evaluate patients in clinical trials. Several endoscopic activity scores have been developed and validated and can be used to assess endoscopic disease activity.[3-8]

For optimal monitoring of disease activity in IBD patients, ileocolonoscopy should be performed on a regular basis. However, repeated colonoscopies represent a logistic and economic challenge, as well as significant burden for the patient. Moreover, there is a small risk of bowel perforation and transmural or extra-luminal disease activity and complications such as abscesses cannot be assessed. Finally, the ileum cannot be intubated in a significant proportion of patients due technical or anatomical difficulties.

Biomarkers such as serum C-reactive protein (CRP) and fecal calprotectin have limited reliability to asses and grade IBD disease activity.[9] Therefore, cross-sectional imaging modalities, such as trans-abdominal ultrasound (US), computed tomography (CT) and magnetic resonance imaging (MRI) are increasingly being used in the management of IBD.[10-12] These imaging techniques can be used to determine the extent and location of inflammation and to detect disease complication, such as stenosis, fistulas and abscesses in patients with Crohn's disease (CD).[2, 10, 11, 13-20] MRI and CT show good results for grading disease activity, but they are not ideal for repeated use due to logistical reasons (MRI) or (for CT) radiation exposure.[10, 11] Since US is rapid, non-invasive, relatively cheap and can even be performed in a point-of-care setting, it appears to be the most suitable modality for systematic monitoring in IBD patients.[21]

An accurate US index for grading disease activity would therefore be of great clinical value. Although various US activity indices for IBD patients exist, which have also been evaluated in previous reviews, the applicability of US in grading disease activity still remains uncertain.[11, 19, 22, 23] Still, a comprehensive evaluation of the characteristics and methods of all available studies focusing on US activity indices for assessing disease activity in IBD has never been conducted.

Here, we aim to critically review the quality and reliability of available US activity indices compared with reference standards for grading disease activity in IBD patients. This could serve as a basis to improve US activity indices and for the development of novel scoring systems.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Methods

This systematic review has been conducted in accordance with the Preferred Items for Systematic Reviews an Meta-analyses (PRISMA) guidelines[24] The protocol has not been published in advance.

Literature search

PUBMED, MEDLINE, CENTRAL and EMBASE were electronically searched from January 1990 until March 2017 for studies examining US for grading disease activity in CD and UC. Details of the search criteria are provided in the supplementary material (Appendix E1). All reference lists of the included studies were searched for potentially relevant records.

Inclusion and exclusion criteria

Study inclusion was based on the following criteria: (1) Study of an US index consisting of at least 3 categories for disease activity grading (i.e. quiescent, moderate or severe); (2) comparison with a reference test/standard such as ileocolonoscopy, MRI, barium contrast radiography or histology; (3) a sample size of at least 20 patients; (4) articles written in English; (5) full text available (i.e. no abstracts). Studies that used a clinical activity index as the reference standard were not included, since these instruments poorly correlate with inflammatory disease activity, especially in CD[25]

Study selection

All retrieved studies were assessed by one observer (SB). Irrelevant studies were excluded based on title, abstract and study type (i.e. review, case report, comment, letter). The remaining titles and abstracts were independently assessed by 2 observers (SB, KN) for eligibility for full text review. Subsequently, the selected full texts were assessed by both observers to identify studies with US indices. Finally, the remaining studies were assessed for inclusion by both observers. Disagreements were resolved through discussion after every phase in the selection process.

Data collection and analysis

The following data were collected on study characteristics: study design, diagnosis, number of included patients, number of US exams, segments analyzed, patient selection and inclusion methods, reference test and index used, blinding methods and time between reference and US exams. Additionally, the following data were collected on the US indices: index parameters, severity grades, cut-offs, index calculation methods, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and correlation coefficients with reference test. A meta-analysis was not performed due the heterogeneity in study methodology and index characteristics.

Study quality grading

All included studies were graded for methodological quality by two investigators (SB and KN) with a modified version of the QUADAS-2 tool.[26] The QUADAS-2 tool is designed to assess quality of diagnostic accuracy studies in with signaling questions in 4 domains (patient selection, index test, reference test and patient flow). The signaling questions of the modified tool are shown in table 1. Established reference indices were considered as good quality reference standards. If existing reference indices were modified for the purpose of the study they were considered as lower quality

reference standards. The questions of each domain could be answered with 'yes', 'no' or 'unclear'. Unclear answers were considered as 'no' for the final quality grading. Each subdomain was graded as high risk of bias if ≥50% of the signaling questions were answered with 'no'. A study was graded as high quality in case of a low risk of bias in at least 6 out of the 7 subdomains. A study was graded as low quality in case of a high risk of bias in 4 or more subdomains. All other studies were graded as moderate quality. Any disagreements were resolved through discussion.

Results

Study selection

A total of 2103 records were identified through electronic search and 1656 remained after removal of duplicates. One additional record was identified through other sources. This particular study was published after the search date, but we decided to include it due to its relevance.[27] After screening titles and abstracts, 140 potentially eligible studies were selected for full text review. After full text review, 20 records were identified that studied an US activity index (supplementary table 1). Out of these 20 studies, 11 met the inclusion criteria. A chart flow of the selection process is shown in figure 1.

Study characteristics

Study characteristics are shown in table 2. Eight studies used a prospective and two studies a retrospective design. One study consisted of a retrospective development phase and a prospective validation phase. The total number of studied subjects was 771 (mean 70.1; SD 56.2) and a total of 1088 (mean 98.9; SD 93.9) US exams were performed. In 4 studies, only the ileum was investigated. Ileocolonoscopy was used as the reference standard in 9 studies, in 1 study a combined index of ileocolonoscopy or barium contrast radiography was used as reference the standard and in 1 study histology was used as the reference standard.

Crohn's disease ultrasonographic activity indices

Seven CD indices were identified from 8 records. The parameters used in the CD indices included bowel wall thickness (BWT), Doppler signal (DS), wall layer stratification (WLS), compressibility, peristalsis, haustrations, fatty wrapping and contrast enhancement (CE). CD index details are provided in table 3.

Futagami et al. developed an US index with BWT and WLS as parameters.[28] The thresholds of the index were defined before the study. They compared the index with either endoscopy or barium contrast radiography in 55 patients. An endoscopic/radiological index was developed for comparison, thus not all patients received the same reference standard. The overall correlation with the reference index was average ($r^2 = 0.62$; $p<0.01$).

Neye et al. developed an US index with BWT and DS as parameters.[29] The thresholds of the index were defined before the study. The index was compared with a newly developed endoscopic activity index in 22 patients (i.e. for each bowel segment: 1 (no lesions), 2 (aphtes), 3 (aphtes and ulcers < 50%) to 4 (aphtes and ulcers >50%). The highest concordance was found in the descending colon (κ = 0.91; 96%CI 0.56-0.99) and the lowest in the ascending colon (κ = 0.75; 65% CI 0.56 – 0.94). Concordance for all bowel segments separately is shown in supplementary table 2.

Drews et al. conducted a retrospective study comparing the Limberg score with histologic inflammation in ileum biopsies obtained by ileocolonoscopy in 32 CD patients.[30] This index was first proposed by Limberg and semiquantitatively measures DS in thickened bowel segments (>4mm).[31] A histologic index for severity of inflammation was developed for the study. The association between the Limberg score and histologic grades of disease activity was poor (κ = 0.4375).

Sasaki et al. conducted a retrospective study comparing the Limberg score with the SES-CD score in 108 CD patients.[32] Only the ileum was investigated. The correlation between US and endoscopy was good (Þ = 0.709; p < 0.001).

Paredes et al. developed an US index with BWT and DS for grading of post-surgical recurrence in 33 patients.[33] The index was compared with the endoscopic Rutgeerts score for post-operative recurrence in 33 patients.[34] The Rutgeerts score is a prognostic score to predict post-operative disease course. The thresholds of the US index were determined before the study. The correlation of the US index with the Rutgeerts score was poor (κ = 0.29; p=unknown). For the diagnosis of moderate-severe recurrence, the correlation with endoscopy was average (κ = 0.57; p=0.009). A follow-up study with similar methods was conducted, combining the index with contrast enhanced ultrasound (CEUS).[35] Postoperative recurrence was assessed in 60 CD patients. A cut-off of 34.5% of maximum contrast enhancement predicted endoscopic recurrence most accurately. In combination with the other US parameters the accuracy was 94.4% and the correlation was good (κ =0.82; p<0.001). A cut-off >46% contrast enhancement was best for the prediction of moderate-severe endoscopic recurrence.

Pascu et al. developed an index with BWT, DS, compressibility, WLS and fatty wrapping as parameters.[36] The index was compared with ileocolonoscopy using a modified Baron score in 37 CD patients.[6] The thresholds of the index were defined before the study. The overall activity index was calculated by the sum of segmental indices. The overall correlation between US and ileocolonoscopy was good (r = 0.830; P < 0.001).

Novak et al developed an index with BWT and DS as parameters. The study consisted of a retrospective phase to develop the index and prospective phase to validate the index. The SES-CD and Rutgeerts score were used as reference standard. The index was developed using univariate and multivariate logistic regression model. Cut-offs to discriminate between inactive/mild endoscopic disease and moderate/severe endoscopic disease were determined with area under the receiver operating characteristic curve (AUROC). The SES-CD cut-off for active versus inactive disease was >5. Also, there were 7 UC patients in the development cohort. Additionally, there were 63 patients and 87 examinations in the validation cohort, thus for 24 patients 2 US examinations were used for the statistical calculations. In both phases, ultrasonographers and endoscopists were not blinded for the results of the other examinations. The final US score could be calculated with a formula (table 3). The AUROC was 0.836 for discerning disease activity in the validation cohort.

Ulcerative colitis ultrasonographic activity indices

Four US indices were identified. The parameters used in the indices included BWT, DS, WLS, compressibility, fatty wrapping and strain pattern. UC index details are provided in table 4.

Parente et al. developed an US index with BWT and DS for the assessment of mucosal healing.[2, 20] The index was compared with the endoscopic Baron score in 83 UC patients.[6] The thresholds of the US index were defined before the study. Patients were assessed at 0, 3 , 9 and 15 months. At baseline all patients had US scores and baron scores of 2-3. Concordance of the severity classes was average with a weighted κ coefficient of 0.59 (95 % CI: 0.40 – 0.78).

Ishikawa et al. 2011 proposed an US index with real-time elastography (RTE) based on normal, homogenous, random and hard patterns.[37] and compared it with ileocolonoscopy in 37 UC patients. Ileocolonoscopic findings were classified as (A) normal mucosa, (B) mucosal edema and erosion without ulcer, (C) punched-out ulcer, and (D) extensive ulcer. A significant correlation was reported between type A, B , C, D and normal, homogenous, random and hard respectively (chi-square $p<0.001$).

Civitelli et al. 2014 developed an US index for the assessment of disease activity in pediatric UC.[38] US parameters were compared with the endoscopic mayo score as dependent variable in 50 patients. Multiple regression analysis showed that BWT ($P = 0.0008$), increased vascularity ($P = 0.002$), loss of stratification ($P = 0.021$), and absence of colon haustrations ($P = 0.031$) were significantly associated with endoscopic disease severity. A US score >2 had a sensitivity of 100% and a specificity of 93% (AUC 0.98) for detecting severe endoscopic disease. The US index correlated strongly with endoscopic disease activity ($r = 0.94$; $P<0.0001$). Concordance between US and ileocolonoscopy for inactive, mild, moderate, and severe disease was very good ($\kappa = 0.94$; 95% CI 0.88-1).

Pascu et al. developed an US index with BWT, DS, compressibility, WLS and fatty wrapping as parameters.[36] The index was compared with a modified Baron score in 24 UC patients. The US activity index showed a strong correlation with ileocolonoscopy ($r = 0.974$, $P < 0.001$).

Grading of study quality

Study quality was graded high in 5 studies, moderate in 3 studies and low in 3 studies. Most concerns were raised in the sub-domains regarding the index test and the reference standard. Blinding was performed properly in most studies, but in 9 studies the thresholds of the index were defined before the study was performed. Civitelli et al. developed the US index using the reference standard as a dependent variable. Novak et al. developed the index in a retrospective study and validated it in a prospective study. Both studies were therefore used for quality grading. Five studies used an established endoscopic reference index (i.e. SES-CD, Mayo, Rutgeerts score). In the other studies, either a newly developed index or a modified Baron index was used. Methods for patient selection were suboptimal in 3 studies. Flow and timing was good in all studies. The results of the Quadas-2 assessment are shown in table 5. There were no studies that used central reading or inter and intra-observer variability assessment and only the study performed by Novak et al. used a development and validation phase.

Discussion

To our knowledge, this is the first comprehensive systematic review on US scoring indices that can be used to assess disease activity in IBD patients. The methods that were used for the development of these indices were suboptimal in most studies. Although 20 studies were identified that studied an US activity index, 9 were excluded due to small patient numbers or because clinical activity indices were used as the reference standard, indicating poor methodology. Out of 11 included studies, only 5 of them were graded as high quality using the modified Quadas-2 tool. Based on these findings we conclude that the methodology for the development of US indices for grading disease activity in IBD patients should be improved in future studies.

Important criteria for the development of a diagnostic index are appropriate patient selection, a proper sample size, implementation of blinding, use of an established reference index, inclusion of patients with different disease activity and proper study flow and timing (i.e. time between index and reference test and comparison of all patients with the same reference standard).[26] In addition, a diagnostic index should ideally be developed using the reference index as the dependent variable. Parameters of the imaging modality that can predict outcomes of the reference index should be determined and used for further development of the index. Subsequently, the most predictive cut-off values should be determined with appropriate statistical methods.[39] The methods that were used for the development of the so-called simple endoscopic indices for Crohn's disease (CDEIS and SES-CD) are good examples of such an approach.[3, 8]

The most commonly used parameters in both the CD and UC indices were BWT, DS and WLS (10, 9 and 3 indices in CD and 3, 3, and 2 indices in UC, respectively). BWT is the only quantifiable measurement and in theory probably the easiest to reproduce. However, it is important to standardize measurements methods in order to get reproducible results (i.e. measurement location and probe handling). DS is usually measured semi-quantitatively and thus more prone to interpretation. Additionally, the amount of DS is influenced by equipment and patient characteristics such as the amount of body fat and location of inflammation. To optimize reproducibility, clear definitions should be used and settings on the US scanner should be optimized and remain constant when assessing different patients (i.e. slow-flow settings). The assessment of WLS is also more subjective and thus clear definitions should be used. FW, haustrations, compressibility and peristalsis were rarely used as index parameters. However, FW is considered as an important finding and should be considered for score development in the future, especially in CD patients.

Ileocolonoscopy was used as the reference standard in most of the included studies (n=9), but only 5 studies compared US with an established endoscopic index (i.e. SES-CD, Mayo, Rutgeerts' score). In the other 4 studies a newly developed or a modified index was used as reference standard. Pascu et al. used for example the modified Baron score for assessing disease activity in both CD and UC. Since CD and UC are different entities, activity cannot be scored with the same scoring system. Futugami et al. used an activity score that was based on both endoscopic and barium contrast radiography findings in CD patients. It is likely that the comparison with these non-established reference indices has biased the results in these studies. This is also reflected by the wide range in statistical association between US and endoscopic indices in these studies.

Additionally, in all these studies, the thresholds for ultrasonographic parameters were determined before the study. Establishment of index thresholds prior to a study is likely to result in overestimation of the diagnostic value.[39] Civitelli et al. used an endoscopic index (Mayo endoscopic score) as dependent variable in order to determine thresholds of US parameters for the development of an US index for pediatric UC patients.[38] Additionally, Novak et al. conducted a retrospective study in which they determined parameters, cut-off values and the formula to calculate the activity score.[27] As a next step, they validated the index formula prospectively. However, a major limitation of this study was that ultrasonographers and endoscopists were not blinded for the results of the other examinations. Moreover, the SES-CD cut-off that was used for active disease was quite liberal (SES-CD >5) and there were 7 UC patients in the development cohort.

Drews et al. compared the Limberg score (see table 3 for index characteristics) with histologic inflammation in biopsies in CD patients. Correlation between this score and the histology index was poor to average, depending on the cut-off values that were used. This could be explained by the fact that the location or small amount of tissue obtained through biopsies may not accurately reflect disease activity. Additionally, a non-validated histology index was used. The Limberg score does seem to correlate better with endoscopic disease activity, as was shown by Sasaki et al.[32] However, the data for this study were collected retrospectively which may have introduced bias. Additionally, only ileal disease was compared in these studies since the Limberg score was initially developed to assess the ileum.

Interestingly, we found no studies that used an alternate cross-sectional imaging modality (e.g. MRI or CT) as the reference standard. This could be explained by the fact that disease activity indices for these modalities are also relatively rare and that no standard and widely used activity index exists (i.e. such as the SES-CD or Mayo score). A comprehensive systematic review by Puylaert et al. described 11 studies on MRI and 3 studies on CT for grading of disease activity, which all used endoscopy, biopsies or surgical specimens as the reference standard.[11] This confirms our finding that thus far, US has not been compared with activity indices from other cross-sectional modalities. Such comparisons could be of value and should be conducted in future studies.

Small intestine contrast ultrasonography has also been studied for the grading of disease activity in IBD. We identified 2 studies describing a SICUS activity index.[40, 41] However, both studies used clinical disease activity as the reference standard and therefore did not meet the inclusion criteria. Some studies have shown higher sensitivity and specificity of SICUS for the detection of inflammation than regular US.[42-44] The development of SICUS indices with use of a good reference standard could therefore be of important value. SICUS is however more time consuming than regular US and thus probably less useful in a point-of-care setting.

The value of contrast enhancement for the assessment of disease activity in IBD is increasingly being studied. It seems to have promising potential for the assessment of disease activity.[45-47] For instance, the pattern of bowel wall enhancement and perfusion quantification may have value for disease activity assessment. [35, 46, 48-51] The only index using CEUS that met our inclusion criteria was developed by Paredes et al.[35] They showed a high accuracy of CEUS for the assessment of postoperative recurrence in 33 patients. We identified 1 other index using CEUS.[52] However, this study was excluded because a clinical activity index was used as reference standard. It is to be expected that CEUS will be increasingly used for the development of new indices in the future. However, it is

important to note that CEUS parameters are more equipment dependent than classical US parameters. Additionally, results from perfusion quantification can currently not be compared between different ultrasound scanners.[53] It has also been postulated that CEUS could be useful for the differentiation between fibrosis and inflammation. However, results from different studies regarding this topic are conflicting.[52, 54-56] Therefore, it remains to be seen if CEUS truly will have additional value for differentiation of disease activity and fibrosis. Finally, CEUS is more expensive and time consuming than regular US.

We identified 1 index using real time elastography for the assessment of disease activity in UC patients.[37] Although the concept seems interesting, many factors in this study may have introduced bias. For instance, endoscopic findings from specific locations were compared with US, but in reality it is difficult to compare precise locations between 2 modalities. The elastographic patterns also seemed difficult to interpret. This complicates the applicability and reproducibility of the index. Finally, no established endoscopic index was used as a reference standard. Elastography may probably have more value for the detection of fibrotic intestinal tissue, as was shown in several studies.[57, 58]

US for the grading disease activity in IBD has been reviewed by other groups. Rimola et al. evaluated 4 US studies in a systematic review on different imaging modalities in CD patients.[23] They reported good accuracy between the different indices, but they did not assess the quality of these studies. Puylaert et al. reviewed several imaging modalities for grading of disease activity in CD, but they included only 2 US studies.[11] They concluded that US has low accuracy for disease activity grading in CD, but the number of patients (n=86) used in their analysis was relatively low. Panes et al. discussed 12 US studies for grading of disease severity 1231 patients and concluded that US findings correlate well with endoscopy and histology, but not with clinical activity indices and biomarkers.[19] However, study and index quality were not assessed. Moreover, most studies that were reviewed used clinical and/or biochemical activity as a reference standard. Calabrese et al. recently reviewed a variety of aspects of US in CD, but only briefly elaborated on the use of US for grading CD activity.[22] They stated that the role of US in the evaluation of inflammatory activity remains controversial. Hence, the contradictory conclusions of these reviews are exemplary for the uncertainties regarding US for disease activity grading in IBD and are probably caused by the heterogeneity of different US activity indices that have been developed so far.

Our study has some limitations. Firstly, we decided not to perform a meta-analysis. In our opinion, a meta-analysis could not be performed due to the considerable differences between the studies and would probably result in highly biased results. Secondly, some factors that are important for the development of diagnostic indices, such as implementation of central reading, inter-observer variability and the conduction of a development and validation study are not part of the Quadas-2 tool. However, there were no studies that used central reading or inter-observer variability assessment and only the study performed by Novak et al. used a development and validation phase.

In conclusion, gastrointestinal US seems a promising tool for the assessment of disease activity in IBD patients, but most available activity indices have been developed with suboptimal methodology. New indices should be developed with better methods in future studies. A reliable and standardized US activity index would be useful in order to facilitate the clinical decision making process and to assess and monitor treatment outcomes in daily practice and clinical trials.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Funding

No external funding was obtained.

Acknowledgements

We would like to thank Faridi van Etten for her help with the search strategy.

References

1. Schnitzler F, Fidder H, Ferrante M, Noman M, Arijs I, Van Assche G, et al. Mucosal healing predicts long-term outcome of maintenance therapy with infliximab in Crohn's disease. Inflammatory bowel diseases. 2009;15(9):1295-301.

2. Parente F, Molteni M, Marino B, Colli A, Ardizzone S, Greco S, et al. Bowel ultrasound and mucosal healing in ulcerative colitis. Digestive diseases (Basel, Switzerland). 2009;27(3):285-90.

3. Daperno M, D'Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointestinal endoscopy. 2004;60(4):505-12.

4. Travis SP, Schnell D, Krzeski P, Abreu MT, Altman DG, Colombel JF, et al. Developing an instrument to assess the endoscopic severity of ulcerative colitis: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS). Gut. 2012;61(4):535-42.

5. Khanna R, Nelson SA, Feagan BG, D'Haens G, Sandborn WJ, Zou GY, et al. Endoscopic scoring indices for evaluation of disease activity in Crohn's disease. The Cochrane database of systematic reviews. 2016(8):Cd010642.

6. Baron JH, Connell AM, Lennard-Jones JE. VARIATION BETWEEN OBSERVERS IN DESCRIBING MUCOSAL APPEARANCES IN PROCTOCOLITIS. Br Med J. 1964;1(5375):89-92.

7. Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. The New England journal of medicine. 1987;317(26):1625-9.

8. Mary JY, Modigliani R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). Gut. 1989;30(7):983-9.

9. De Vos M, Louis EJ, Jahnsen J, Vandervoort JG, Noman M, Dewit O, et al. Consecutive fecal calprotectin measurements to predict relapse in patients with ulcerative colitis receiving infliximab maintenance therapy. Inflammatory bowel diseases. 2013;19(10):2111-7.

10. Horsthuis K, Bipat S, Bennink RJ, Stoker J. Inflammatory bowel disease diagnosed with US, MR, scintigraphy, and CT: meta-analysis of prospective studies. Radiology. 2008;247(1):64-79.

11. Puylaert CA, Tielbeek JA, Bipat S, Stoker J. Grading of Crohn's disease activity using CT, MRI, US and scintigraphy: a meta-analysis. European radiology. 2015;25(11):3295-313.

12. Nylund K, Odegaard S, Hausken T, Folvik G, Lied GA, Viola I, et al. Sonography of the small intestine. World journal of gastroenterology : WJG. 2009;15(11):1319-30.

13. Maconi G, Parente F, Bollani S, Cesana B, Bianchi Porro G. Abdominal ultrasound in the assessment of extent and activity of Crohn's disease: clinical significance and implication of bowel wall thickening. The American journal of gastroenterology. 1996;91(8):1604-9.

14. Maconi G, Bollani S, Bianchi Porro G. Ultrasonographic detection of intestinal complications in Crohn's disease. Digestive diseases and sciences. 1996;41(8):1643-8.

15. Maconi G, Carsana L, Fociani P, Sampietro GM, Ardizzone S, Cristaldi M, et al. Small bowel stenosis in Crohn's disease: clinical, biochemical and ultrasonographic evaluation of histological features. Alimentary pharmacology & therapeutics. 2003;18(7):749-56.

16. Maconi G, Ardizzone S, Parente F, Bianchi Porro G. Ultrasonography in the evaluation of extension, activity, and follow-up of ulcerative colitis. Scandinavian journal of gastroenterology. 1999;34(11):1103-7.

17. Kucharzik T, Wittig BM, Helwig U, Borner N, Rossler A, Rath S, et al. Use of Intestinal Ultrasound to Monitor Crohn's Disease Activity. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2016.

18. Martinez MJ, Ripolles T, Paredes JM, Blanc E, Marti-Bonmati L. Assessment of the extension and the inflammatory activity in Crohn's disease: comparison of ultrasound and MRI. Abdominal imaging. 2009;34(2):141-8.

19. Panes J, Bouzas R, Chaparro M, Garcia-Sanchez V, Gisbert JP, Martinez de Guerenu B, et al. Systematic review: the use of ultrasonography, computed tomography and magnetic resonance imaging for the diagnosis, assessment of activity and abdominal complications of Crohn's disease. Alimentary pharmacology & therapeutics. 2011;34(2):125-45.

20. Parente F, Molteni M, Marino B, Colli A, Ardizzone S, Greco S, et al. Are colonoscopy and bowel ultrasound useful for assessing response to short-term therapy and predicting disease outcome of moderate-to-severe forms of ulcerative colitis?: a prospective study. The American journal of gastroenterology. 2010;105(5):1150-7.

21. Novak K, Tanyingoh D, Petersen F, Kucharzik T, Panaccione R, Ghosh S, et al. Clinic-based Point of Care Transabdominal Ultrasound for Monitoring Crohn's Disease: Impact on Clinical Decision Making. Journal of Crohn's & colitis. 2015;9(9):795-801.

22. Calabrese E, Maaser C, Zorzi F, Kannengiesser K, Hanauer SB, Bruining DH, et al. Bowel Ultrasonography in the Management of Crohn's Disease. A Review with Recommendations of an International Panel of Experts. Inflammatory bowel diseases. 2016;22(5):1168-83.

23. Rimola J, Ordas I, Rodriguez S, Ricart E, Panes J. Imaging indexes of activity and severity for Crohn's disease: current status and future trends. Abdominal imaging. 2012;37(6):958-66.

24. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine. 2009;6(7):e1000097.

25. Falvey JD, Hoskin T, Meijer B, Ashcroft A, Walmsley R, Day AS, et al. Disease activity assessment in IBD: clinical indices and biomarkers fail to predict endoscopic remission. Inflammatory bowel diseases. 2015;21(4):824-31.

26. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155(8):529-36.

27. Novak KL, Kaplan GG, Panaccione R, Afshar EE, Tanyingoh D, Swain M, et al. A Simple Ultrasound Score for the Accurate Detection of Inflammatory Activity in Crohn's Disease. Inflammatory bowel diseases. 2017.

28. Futagami Y, Haruma K, Hata J, Fujimura J, Tani H, Okamoto E, et al. Development and validation of an ultrasonographic activity index of Crohn's disease. European journal of gastroenterology & hepatology. 1999;11(9):1007-12.

29. Neye H, Voderholzer W, Rickes S, Weber J, Wermke W, Lochs H. Evaluation of criteria for the activity of Crohn's disease by power Doppler sonography. Digestive diseases (Basel, Switzerland). 2004;22(1):67-72.

30. Drews BH, Barth TF, Hanle MM, Akinli AS, Mason RA, Muche R, et al. Comparison of sonographically measured bowel wall vascularity, histology, and disease activity in Crohn's disease. European radiology. 2009;19(6):1379-86.

31. Limberg B. [Diagnosis of chronic inflammatory bowel disease by ultrasonography]. Zeitschrift fur Gastroenterologie. 1999;37(6):495-508.

32. Sasaki T, Kunisaki R, Kinoshita H, Yamamoto H, Kimura H, Hanzawa A, et al. Use of color Doppler ultrasonography for evaluating vascularity of small intestinal lesions in Crohn's disease: correlation with endoscopic and surgical macroscopic findings. Scandinavian journal of gastroenterology. 2014;49(3):295-301.

33. Paredes JM, Ripolles T, Cortes X, Reyes MD, Lopez A, Martinez MJ, et al. Non-invasive diagnosis and grading of postsurgical endoscopic recurrence in Crohn's disease: usefulness of abdominal ultrasonography and (99m)Tc-hexamethylpropylene amineoxime-labelled leucocyte scintigraphy. Journal of Crohn's & colitis. 2010;4(5):537-45.

34. Rutgeerts P, Geboes K, Vantrappen G, Beyls J, Kerremans R, Hiele M. Predictability of the postoperative course of Crohn's disease. Gastroenterology. 1990;99(4):956-63.

35. Paredes JM, Ripolles T, Cortes X, Moreno N, Martinez MJ, Bustamante-Balen M, et al. Contrast-enhanced ultrasonography: usefulness in the assessment of postoperative recurrence of Crohn's disease. Journal of Crohn's & colitis. 2013;7(3):192-201.

36. Pascu M, Roznowski AB, Muller HP, Adler A, Wiedenmann B, Dignass AU. Clinical relevance of transabdominal ultrasonography and magnetic resonance imaging in patients with inflammatory bowel disease of the terminal ileum and large bowel. Inflammatory bowel diseases. 2004;10(4):373-82.

37. Ishikawa D, Ando T, Watanabe O, Ishiguro K, Maeda O, Miyake N, et al. Images of colonic real-time tissue sonoelastography correlate with those of colonoscopy and may predict response to therapy in patients with ulcerative colitis. BMC gastroenterology. 2011;11:29.

38. Civitelli F, Di Nardo G, Oliva S, Nuti F, Ferrari F, Dilillo A, et al. Ultrasonography of the colon in pediatric ulcerative colitis: a prospective, blind, comparative study with colonoscopy. The Journal of pediatrics. 2014;165(1):78-84.e2.

39. Leeflang MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. Clinical chemistry. 2008;54(4):729-37.

40. Zorzi F, Stasi E, Bevivino G, Scarozza P, Biancone L, Zuzzi S, et al. A sonographic lesion index for Crohn's disease helps monitor changes in transmural bowel damage during therapy. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association. 2014;12(12):2071-7.

41. Calabrese E, Zorzi F, Zuzzi S, Ooka S, Onali S, Petruzziello C, et al. Development of a numerical index quantitating small bowel damage as detected by ultrasonography in Crohn's disease. Journal of Crohn's & colitis. 2012;6(8):852-60.

42. Calabrese E, La Seta F, Buccellato A, Virdone R, Pallotta N, Corazziari E, et al. Crohn's disease: a comparative prospective study of transabdominal ultrasonography, small intestine contrast ultrasonography, and small bowel enema. Inflammatory bowel diseases. 2005;11(2):139-45.

43. Pallotta N, Vincoli G, Montesani C, Chirletti P, Pronio A, Caronna R, et al. Small intestine contrast ultrasonography (SICUS) for the detection of small bowel complications in crohn's disease: a prospective comparative study versus intraoperative findings. Inflammatory bowel diseases. 2012;18(1):74-84.

44. Pallotta N, Tomei E, Viscido A, Calabrese E, Marcheggiano A, Caprilli R, et al. Small intestine contrast ultrasonography: an alternative to radiology in the assessment of small bowel disease. Inflammatory bowel diseases. 2005;11(2):146-53.

45. Saevik F, Nylund K, Hausken T, Odegaard S, Gilja OH. Bowel perfusion measured with dynamic contrast-enhanced ultrasound predicts treatment outcome in patients with Crohn's disease. Inflammatory bowel diseases. 2014;20(11):2029-37.

46. Migaleddu V, Scanu AM, Quaia E, Rocca PC, Dore MP, Scanu D, et al. Contrast-enhanced ultrasonographic evaluation of inflammatory activity in Crohn's disease. Gastroenterology. 2009;137(1):43-52.

47. Quaia E, Cabibbo B, De Paoli L, Toscano W, Poillucci G, Cova MA. The value of time-intensity curves obtained after microbubble contrast agent injection to discriminate responders from non-responders to anti-inflammatory medication among patients with Crohn's disease. European radiology. 2013;23(6):1650-9.

48. Serra C, Menozzi G, Labate AM, Giangregorio F, Gionchetti P, Beltrami M, et al. Ultrasound assessment of vascularization of the thickened terminal ileum wall in Crohn's disease patients using a low-mechanical index real-time scanning technique with a second generation ultrasound contrast agent. Eur J Radiol. 2007;62(1):114-21.

49. Ripolles T, Rausell N, Paredes JM, Grau E, Martinez MJ, Vizuete J. Effectiveness of contrast-enhanced ultrasound for characterisation of intestinal inflammation in Crohn's disease: a comparison with surgical histopathology analysis. Journal of Crohn's & colitis. 2013;7(2):120-8.

50. Ripolles T, Martinez MJ, Paredes JM, Blanc E, Flors L, Delgado F. Crohn disease: correlation of findings at contrast-enhanced US with severity at endoscopy. Radiology. 2009;253(1):241-8.

51. De Franco A, Di Veronica A, Armuzzi A, Roberto I, Marzo M, De Pascalis B, et al. Ileal Crohn disease: mural microvascularity quantified with contrast-enhanced US correlates with disease activity. Radiology. 2012;262(2):680-8.

52. Schirin-Sokhan R, Winograd R, Tischendorf S, Wasmuth HE, Streetz K, Tacke F, et al. Assessment of inflammatory and fibrotic stenoses in patients with Crohn's disease using contrast-enhanced ultrasound and computerized algorithm: a pilot study. Digestion. 2011;83(4):263-8.

53. Zink F, Kratzer W, Schmidt S, Oeztuerk S, Mason RA, Porzner M, et al. Comparison of Two High-End Ultrasound Systems for Contrast-Enhanced Ultrasound Quantification of Mural Microvascularity in Crohn's Disease. Ultraschall in der Medizin (Stuttgart, Germany : 1980). 2016;37(1):74-81.

54. Quaia E, Gennari AG, van Beek EJR. Differentiation of Inflammatory from Fibrotic Ileal Strictures among Patients with Crohn's Disease through Analysis of Time-Intensity Curves Obtained after Microbubble Contrast Agent Injection. Ultrasound in medicine & biology. 2017;43(6):1171-8.

55. Nylund K, Jirik R, Mezl M, Leh S, Hausken T, Pfeffer F, et al. Quantitative contrast-enhanced ultrasound comparison between inflammatory and fibrotic lesions in patients with Crohn's disease. Ultrasound in medicine & biology. 2013;39(7):1197-206.

56. Wilkens R, Hagemann-Madsen RH, Peters DA, Nielsen AH, Norager CB, Glerup H, et al. Validity of Contrast-enhanced Ultrasonography and Dynamic Contrast-enhanced MR Enterography in the Assessment of Transmural Activity and Fibrosis in Crohn's Disease. Journal of Crohn's & colitis. 2018;12(1):48-56.

57. Baumgart DC, Muller HP, Grittner U, Metzke D, Fischer A, Guckelberger O, et al. US-based Real-time Elastography for the Detection of Fibrotic Gut Tissue in Patients with Stricturing Crohn Disease. Radiology. 2015;275(3):889-99.

58. Giannetti A, Biscontri M, Matergi M, Stumpo M, Minacci C. Feasibility of CEUS and strain elastography in one case of ileum Crohn stricture and literature review. Journal of ultrasound. 2016;19(3):231-7.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Figure legends

Figure 1 Flow chart of study selection process.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Tables

Domain 1	**Patient selection**
1A	- Was a consecutive or random sample used? - Was a case-control or retrospective design avoided? - Were inappropriate exclusions avoided? - Was the samplesize appropriate (10 patients per index parameter)?[a]
1B	- Did the patients match the review question? (confirmed IBD)
Domain 2	**Index test**
2A	- Blinding for the results of the reference test? - Were the thresholds not pre-specified?[b]
2B	- Concerns regarding applicability of the index (reproducibility)?
Domain 3	**Reference standard**
3A	- Was the reference standard used to classify the condition? - Blinding for results of index test? - Use of an established reference index?[a]
3B	- Concerns regarding applicability of the reference test (reproducibility)?
Domain 4	**Flow and timing**
4A	- Appropriate interval between index and reference test (≥ 1 month) ? - Did all patients receive reference test? - Did all patients receive the same reference test? - Were all patients included in the analysis?

Table 1. Modified QUADAS-2 risk of bias assessment tool.

[a]This item was not part of the original Quadas-2 tool

[b]This question was adapted from the original tool

Manuscript Doi: 10.1093/ecco-jcc/jjy048

Study (index)	Diagnosis	Design	Subjects	US nr.	Index PM	Segments	Ref.	Ref. index	Days index/ref
Futagami 1999	CD	Prospective	55	126	BWT, WLS Haustrations Compressibility Peristalsis	Jejunum Ileum Ascending Transverse Descending Sigmoid Rectum	BCR/ ICC	Developed	3
Neye 2004	CD	Prospective	22	22	BWT, DS	Ileum Cecum Ascending Transverse Descending Sigmoid	ICC	Developed	3
Drews 2009 (Limberg)	CD	Retrospective	32	32	BWT, DS	Ileum	Biopsies	Developed	5
Sasaki 2014 (Limberg)	CD	Retrospective	108	108	BWT, DS	Ileum	ICC	SES-CD	30
Paredes 2010	CD	Prospective	40 (33)	40	BWT, DS Complications	Ileum	ICC	Rutgeerts	3
Paredes 2013	CD	Prospective	60	60	BWT, DS, CE, complications	Ileum	ICC	Rutgeerts	3
Novak 2017	CD	Phase 1: retrospective Phase 2: Prospective	223	247	BWT, DS	Ileum Cecum Ascending Transverse Descending	ICC	SES-CD & Rutgeerts	Phase 1: 60 Phase 2: 14

						Sigmoid Rectum			
Pascu 2004	UC and CD	Prospective	37 CD 24 UC	61	BWT, DS, WLS, FW Compressibility	Ileum Cecum Ascending Transverse Descending Sigmoid	ICC	Modified Baron	5
Civitelli 2014	UC	Prospective	60 (50)	50	BWT, DS, WLS, haustrations	Right colon Transverse Left colon	ICC	Mayo	1
Parente 2009/2010	UC	Prospective	83	305	BWT, DS	Ascending Transverse Descending Sigmoid	ICC	Modified Baron	3
Ishikawa 2011	UC	Prospective	37	37	Strain patterns	Ascending Transverse Descending Sigmoid	ICC	Developed	1

BWT = bowel wall thickness; DS = Doppler signal; WLS = wall layer stratification; FW = fatty wrapping; CE = contrast enhancement; ICC = ileocolonoscopy; BCR = barium contrast radiography; pm = parameters; Ref.= reference; Developed = Reference index was newly developed for study

Table 2. Characteristics of included studies.

Index					
Limberg - Drews - Sasaki	**Grade 0**	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**
	- BWT < 4mm -no vessels	- BWT > 4mm - no vessels	- BWT >4mm - Spots of vascularity	- BWT > 4 mm - longer stretches of vascularity	- BWT > 4mm - long stretches of vascularity into mesentery
Futagami	**Normal**	**Type A**	**Type B**	**Type C**	-
	- BWT < 4mm - normal compressibility and peristalsis - Haustrations present	- BWT <4mm - reduced compressibility and peristalsis - Loss of haustrations	- BWT > 4mm - stratification intact	- BWT > 4mm - loss of stratification	-
	The formula: 1 point for type A lesions [BWT -2] * 2 for type B lesions [BWT-2] * 4 for type C lesions				
Neye	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**	-
	BWT < 5mm, no vessels/cm^2	- BWT < 5mm, 1-2 vessels/cm^2 - BWT > 5 mm, no vessels/ cm^2	- BWT < 5mm, >2 vessels/ cm^2 - BWT > 5mm, 1-2 vessels/cm^2	- BWT > 5mm, 2 vessels/cm^2	
Paredes 2010	**Normal**	**Recurrence**	**Mod/Sev recurrence**		
	- BWT < 3mm - No DS	BWT >3mm and/or positive DS	BWT >5mm and DS grade 2 or 3.		
Paredes 2013	**Normal**	**Recurrence**	**Mod recurrence**	**Sev recurrence**	
	- BWT < 3mm - CE <34.5	BWT 3-5 mm CE <46%;	BWT >5mm or CE >46%	BWT >5 mm, or CE>70%, or presence of fistula.	
Pascu	**Grade 0**	**Grade 1**	**Grade 2**	**Grade 3**	
	BWT < 3mm No DS	BWT 3-5mm Increased DS Loss of compressibility Accentuated WLS	BWT 5-8mm Increased DS Loss of compressibility Loss of WLS	BWT > 8mm Increased DS Loss of compressibility Loss of WLS Fatty wrapping	

Novak	**Grade 0**	**Grade 1**	**Grade 2**	**Grade 3**	
	BWT < 3mm No DS	BWT 3.1-6mm DS mild	BWT 6.1-7.0mm DS moderate/severe	BWT >7.0mm DS moderate/severe	
	Score = (0.0563 * bwt1) + (2.0047 * bwt2) + (3.0881 * bwt3) + (1.0204 * doppler1) + (1.5460 * doppler2)				

BWT = bowel wall thickness; DS = Doppler signal; WLS = wall layer stratification; CE = contrast enhancement; Mod = moderate; Sev = severe.

Table 3. Characteristics of Crohn's disease indices.

Index					
Parente	**Grade 0**	**Grade 1**	**Grade 2**	**-**	
	- BWT < 4 mm - no or scarce intramural blood flow	BWT 4 – 6 mm and blood flow	BWT6 – 8 mm and blood flow		
	Grade 0	**Grade 1**	**Grade 2**	**Grade 3**	
Ishikawa	Normal color pattern	Homogenous color pattern	Random color pattern	Hard color pattern	
	Grade 0	**Grade 1**	**Grade 2**	**Grade 3**	**Grade 4**
Civitelli	no findings	1 finding	2 findings	3 findings	4 findings
	Findings: BWT > 3mm, increased DS, loss of WLS, absence of haustrations				
	Grade 0	**Grade 1**	**Grade 2**	**Grade 3**	
Pascu	BWT < 3mm No DS	BWT 3-4.5mm Increased DS Loss of compressibility Accentuated WLS	BWT 4.5-6mm Increased DS Loss of compressibility Loss of WLS	BWT >6mm Increased DS Loss of compressibility Loss of WLS FW	

BWT = bowel wall thickness; DS = Doppler signal; WLS = wall layer stratification; FW = fatty wrapping

Table 4. Characteristics of ulcerative colitis indices.

Study	Domain1: patient selection	Domain 2: Index test	Domain 3: reference standard	Domain 4: Flow and timing	Overall quality
Futagami 1999	A: Low B: Low	A: High B: Low	A: High B: High	A: Low	Moderate
Neye 2004	A: Low B: Low	A: High B: Low	A: High B: High	A: Low	Moderate
Drews 2009	A: High B: Low	A: High B: High	A: High B: High	A: Low	Low
Sasaki 2014	A: High B: Low	A: High B: Low	A: Low B: Low	A: Low	Moderate
Paredes 2010	A: Low B: Low	A: High B: Low	A: Low B: Low	A: Low	High
Paredes 2013	A: Low B: Low	A: High B: Low	A: Low B: Low	A: Low	High
Novak 2017	A: Low B: Low	A: High B: Low	A: Low B: Low	A: Low	High
Pascu 2004	A: High B: Low	A: High B: Low	A: High B: High	A: Low	Low
Civitelli 2014	A: Low B: Low	A: Low B: Low	A: Low B: Low	A: Low	High
Parente 2009/2010	A: Low B: Low	A: High B: Low	A: Low B: Low	A: Low	High
Ishikawa 2011	A: High B: Low	A: High B: High	A: High B: High	A: Low	Low

Table 5. Quadas-2 assessment results: risk of bias in all subdomains.

Manuscript Doi: 10.1093/ecco-jcc/jjy048

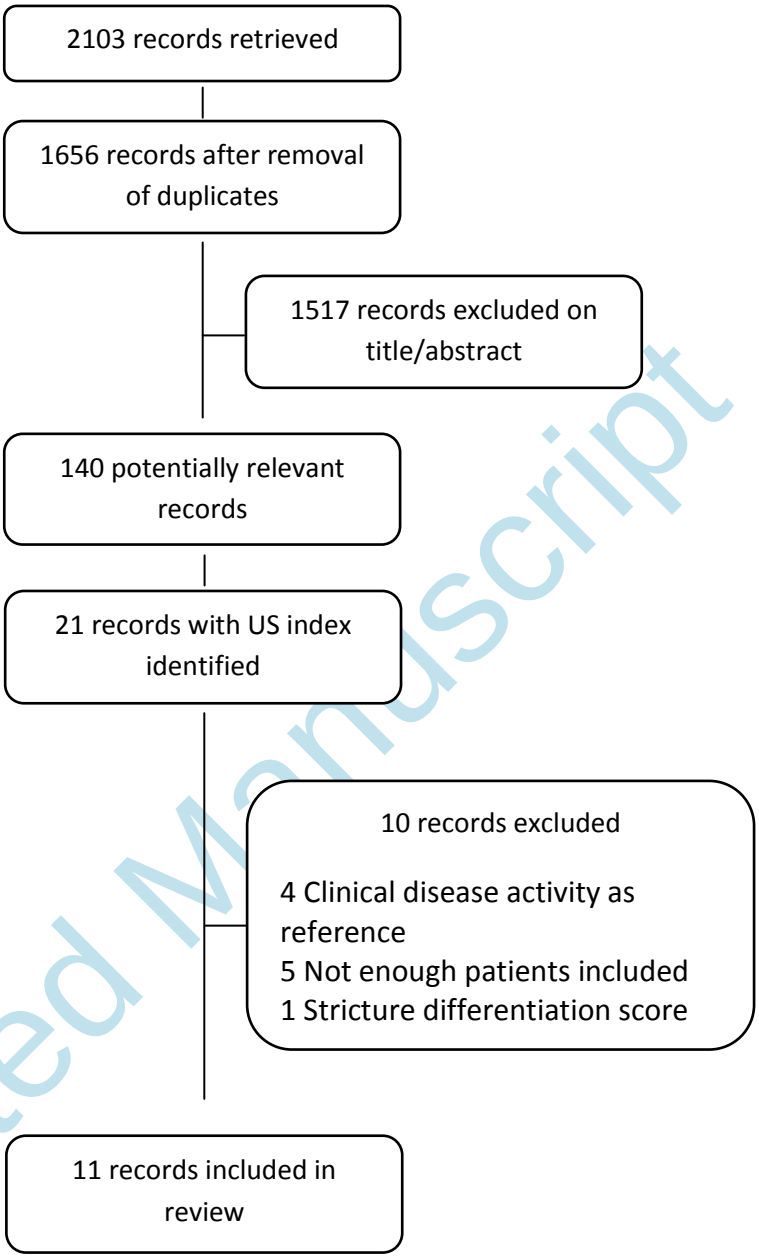

Figure 1 Flow chart of study selection process.