



HAL
open science

Understanding the learning curve of intestinal ultrasound in IBD: a comparative study between novice, regular and expert, the ULTRA-IBD study

Léo Bove

► To cite this version:

Léo Bove. Understanding the learning curve of intestinal ultrasound in IBD: a comparative study between novice, regular and expert, the ULTRA-IBD study. Human health and pathology. 2024. dumas-04759257

HAL Id: dumas-04759257

<https://dumas.ccsd.cnrs.fr/dumas-04759257v1>

Submitted on 29 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IMPORTANT : OBLIGATIONS DE LA PERSONNE CONSULTANT CE DOCUMENT

Conformément au *Code de la propriété intellectuelle*, nous rappelons que le document est destiné à un **usage strictement personnel**. Les "analyses et les courtes citations justifiées par le caractère critique, polémique, pédagogique, scientifique ou d'information" sont autorisées sous réserve de mentionner les noms de l'auteur et de la source (article L. 122-4 du *Code de la propriété intellectuelle*). Toute autre représentation ou reproduction intégrale ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit, est illicite.

De ce fait, nous vous rappelons notamment que, **sauf accord explicite** de l'auteur de la thèse ou du mémoire, **vous n'êtes pas autorisé** à rediffuser ce document sous quelque forme que ce soit (impression papier, transfert par voie électronique, ou autre). Tout contrevenant s'expose aux peines prévues par la loi.

NANTES UNIVERSITÉ

FACULTÉ DE MÉDECINE

Année : 2024

THÈSE

pour le

DIPLÔME D'ÉTAT DE DOCTEUR EN MÉDECINE

Hépatogastro-entérologie

par

Leo BOVE

Né le 23 août 1996 à Angers (49)

Présentée et soutenue publiquement le 3 juillet 2024

Understanding the Learning Curve of Intestinal Ultrasound in IBD: A Comparative Study Between Novice, Regular and Expert, the ULTRA-IBD study

Président : Monsieur le Professeur Arnaud BOURREILLE

Directeur de thèse : Madame le Docteur Catherine LE BERRE

Jury : Monsieur le Professeur Eric FRAMPAS
 Monsieur le Professeur Mathurin FUMERY
 Monsieur le Docteur Jérémie MEYER

Remerciements

Au Pr Arnaud BOURREILLE, je vous remercie de votre implication attentive durant la réalisation de ce travail, et de donner aux internes nantais la possibilité et les moyens d'évoluer et d'apprendre dans le domaine des MICI.

Au Dr Catherine LE BERRE, un grand merci de m'avoir proposé ce beau projet qui m'a permis de découvrir plein de choses, de ton encadrement sans faille durant sa réalisation, et de tout ce que tu as fait, et feras, pour le service et les internes.

Au Dr Jérémy MEYER, merci de t'être prêté à l'exercice et d'avoir apporté ton expertise et tes conseils.

Au Pr Éric FRAMPAS et au Pr Mathurin FUMERY, je vous remercie tous deux d'avoir accepté de faire partie de mon jury, et j'espère que vous trouverez un intérêt à la lecture de ce travail.

Au Dr Michael COLLINS, je te remercie de ta participation à cette thèse, pour tes conversations toujours intéressantes et pour ton investissement dans le service.

Aux patients inclus dans cette thèse et sans qui elle n'aurait bien sûr pas été possible, merci d'avoir accepté de participer et de prendre du temps en plus sur votre journée.

Table des matières

Abstract.....	3
Abbreviations.....	5
Introduction	7
Materials & Methods	13
Results	18
Discussion.....	36
Conclusion.....	43
References	44

Abstract

Background: Intestinal ultrasound (IUS) has emerged as a valuable tool for monitoring inflammatory bowel disease (IBD), but its reliability depends on the operator. This study aimed to assess the learning curve of IUS of a complete ultrasound novice and a gastroenterologist with formal general ultrasound training and regular IUS experience, using an expert radiologist as the gold standard.

Methods: A single-center, prospective study was conducted from April 2023 to January 2024, involving 50 consecutive IBD patients undergoing three successive IUS examinations by the three different operators, who assessed disease activity using quantitative and semi-quantitative scales, as well as IUS disease activity scores based on basic IUS parameters (bowel wall thickness and stratification, Doppler signal, mesenteric fat). Inter-rater agreement was calculated using the intraclass correlation coefficient.

Results: The experienced gastroenterologist demonstrated consistently good agreement with the radiologist throughout the study, while the novice showed poor agreement without a clear learning curve. A body mass index $>25 \text{ kg/m}^2$, the ileocolonic location of CD, and a history of IBD surgery significantly altered the novice's performance. In particular, excluding patients with a history of surgery highlighted a true learning curve for the novice, with moderate-to-good agreement at the end of the study.

Conclusion: A formal general ultrasound training and a few hundred IUS examinations of experience allow good agreement with an expert radiologist in IBD patients. However, novices require comprehensive training and more than 50 examinations of experience, although a strong learning curve effect is shown with less than 40 examinations in patients who had never been operated on. These findings have implications for the development of training guidelines and supervision practices in IUS for IBD assessment, highlighting the difficulty of the technical performance of IUS rather than the interpretation of its images.

Keywords: Crohn's disease; ulcerative colitis; IBUS-SAS; MUC score; monitoring.

Abbreviations

BMI: Body Mass Index

BWS: Bowel Wall Stratification

BWT: Bowel Wall Thickness

CD: Crohn's Disease

CDS: Color Doppler Signal

CI: Confidence Interval

CRP: C-reactive protein

EFSUMB: European Federation of Societies for Ultrasound in Medicine and Biology

EIM: Extra-Intestinal Manifestations

FC: Fecal Calprotectin

GI: Gastrointestinal

IBD: Inflammatory Bowel Disease

IBUS-SAS: International Bowel Ultrasound Segmental Activity Score

ICC: Intraclass Correlation Coefficient

i-fat: Inflammatory mesenteric fat

IUS: Intestinal Ultrasound

MUC: Milan Ultrasound Criteria

MRE: Magnetic Resonance Enterography

PRF: Pulse Repetition Frequency

PRO: Patient-Reported Outcomes

SD: Standard Deviation

UC: Ulcerative Colitis

VAS: Visual Analog Scale

Introduction

The burden of inflammatory bowel disease

Inflammatory bowel diseases (IBD), encompassing Crohn's disease (CD) and ulcerative colitis (UC), are lifelong digestive diseases affecting a growing number of patients, with today about 6.8 million people suffering from IBD around the world(1) and more than 270,000 in France according to the French public health insurance data in 2019. While traditionally considered as diseases affecting primarily high-income countries, the incidence is also sharply rising in newly industrialized countries(1–3), in which hospitalization rates for IBD are rapidly increasing(4), leading to an increasing demand on already struggling healthcare systems(5).

Both types of IBD may highly impact quality of life (more than half of patients) and lead to disability (rated as moderate to severe in one third of patients)(6). Even in the absence of active inflammation, many factors may affect daily life in patients with IBD, including extra-intestinal manifestations (EIM), fatigue, intestinal sequelae, the risk of surgery and adverse effects related to IBD treatments, but also the need for surveillance examinations(7). IBD has been associated with lower fulltime employment than the general population(8), and about one third of patients had lost or had to quit a job due to their disease(9).

Thus, IBD cause significant financial burden on society(10,11), especially because it mainly affects young working people(1). Recent studies estimate a mean cost per patient per year going from 2609€ in Europe(12) to 14,033\$ in the USA(13).

A more and more stringent surveillance

Therapeutic decisions in IBD are now ruled by the treat-to-target concept, based on regular and objective assessment of disease activity and subsequent therapeutic adjustment, of which the ultimate goal is to slow down the disease's natural history(14). Although its

efficacy in the long term is still unproven, this concept originates from years of step-wise therapeutic decisions made only based on clinical activity. Due to the poor correlation that may exist between symptoms and endoscopic disease activity(15,16), the STRIDE consensus switched in 2015 the ultimate therapeutic goal to deep remission including clinical and endoscopic remission(17).

In the wake of several studies showing that deep remission and tighter monitoring improve disease outcomes(18,19), the STRIDE-II consensus was then developed, introducing short-, mid- and long-term targets. The current therapeutic target for disease control in both CD and UC encompasses clinical response then remission, normalization of biomarkers (C-reactive protein [CRP] and fecal calprotectin [FC]) and endoscopic remission(20). Transmural healing for CD and histological healing for UC are suggested as additional targets(20), as it has been associated with better outcomes than endoscopic healing alone (21–24), paving the way for ever more stringent targets in the future.

Thus, close monitoring of disease activity is necessary to accurately treat to target and achieve the best outcomes, which means more and more examinations for patients to undergo.

The need for non-invasive monitoring tools

This tight monitoring with unappealing and/or time-consuming procedures such as flexible sigmoidoscopy, ileocolonoscopy or magnetic resonance enterography (MRE)(25) may reinforce the negative impact of the disease on the daily life of patients with IBD. Repeated expensive examinations also take part in the economic burden on health care systems through direct and indirect costs (e.g. sick leaves for colonoscopy)(12,26).

Thus, there is a need for the development of non-invasive and affordable monitoring tools. In the era of shared medical decision-making and patient-centered management,

intestinal ultrasound (IUS) appears to be the ideal tool to ensure proper surveillance adherence in IBD patients, including in newly industrialized countries.

Advantages of IUS

It has demonstrated excellent acceptability by the patient, at the same level of venipuncture(25,27). It is a low-cost, quick and easy monitoring procedure, that gastroenterologists can perform themselves in consultation to adapt their therapeutic strategy ("point-of-care")(28–30). IUS provides a real-time transmural evaluation of the intestinal wall and extra-digestive structures, allowing the assessment of the intestinal peristalsis in case of stricture. It is also of educational interest to patients, who can visualize the inflammation of their digestive tract by themselves, thus contributing to the empowerment process whereby patients strengthen their ability to take care of their own health. This is an infinitely repeatable examination, with no risks associated with irradiation or the injection of any contrast agent, and without any contraindication(31).

Indications of IUS in IBD

Several studies directly compared IUS to MRE for IBD diagnosis, either proving non-inferiority(32) or failing to show a significant difference in overall accuracy(33–35). Other studies were in line with these results, showing superior performance of MRE in small bowel CD, but suggesting superiority or similarity for IUS in terminal ileal CD or colonic CD(27,36). Regarding the detection of post-operative recurrence, a meta-analysis found that pooled sensitivity and specificity of IUS were 0.94 and 0.84, respectively(37).

Early treatment response can also be detected using IUS. In patients with UC, decrease of bowel wall thickness (BWT) in IUS after treatment initiation has demonstrated high correlation with clinical response(28) and is predictive of later endoscopic response(38). The

kinetics of this change is drug-specific, with a significant reduction in BWT already visible as early as week 2 for infliximab or tofacitinib therapy, and at week 6 for vedolizumab(38). Another study on 49 UC patients starting biologic therapy with evaluation by colonoscopy, IUS, and FC at baseline, week 12 and 1 year found that IUS at week 12 was the only predictor of 1-year endoscopic activity or relapse(39). Interestingly, this early IUS response also predicts intravenous corticosteroid response in hospitalized patients with severe UC(40). Similar results have been obtained in patients with CD. In a prospective study in 234 CD patients undergoing treatment intensification for a flare, all IUS parameters (BWT, bowel wall stratification [BWS], fibrofatty proliferation, and color Doppler signal [CDS]) were significantly improved at 3 and 12 months(30). In a sub study of the randomized controlled STARDUST trial designed to compare treat-to-target and standard-of-care treatment strategies in ustekinumab-treated CD patients, IUS non-response at week 4 was predictive of endoscopic non-response at week 48(41), and similar results have been found for anti-TNF therapy(29).

Conversely, clinical relapse could be predicted using IUS as demonstrated in a retrospective study performed in 58 patients with clinically inactive UC. Active IUS as defined by the Milan Ultrasound Criteria (MUC) score was associated with a higher risk and a shorter time to relapse(42).

The learning curve of IUS

Given the impact that IUS may have on therapeutic decision-making, it is essential to better define the level of training required before implementing it in routine practice. Indeed, the reliability of the findings greatly depends on the reproducibility of results between operators.

To date, the required experience to be able to reliably perform IUS has not been studied extensively. Most published studies focused on the inter-rater agreement of IUS being performed by well-experienced ultrasonographers. It has been shown to be substantial for most ultrasound parameters in two studies comparing 6 ultrasonographers with a median experience of 15 years (range 5-18 years) in one study (43) and 20.5 years (range 15-33 years) in another one(44).

The inter-rater agreement of IUS with ultrasonographers of varying experience has been studied more recently. In an Australian study performed in 49 IBD patients, a substantial to near perfect agreement between a novice and an experienced ultrasonographers was demonstrated, but the novice operator already had >200 examinations of experience, and the experienced operator >1000 examinations, not reflecting the reality of gastroenterologists at least in France, where the ultrasound training is not mandatory during residency(45). A recent Italian study involved 2 trainees with limited IUS experience (<50 examinations), 2 with extensive experience (>500 examinations) and one expert ultrasonographer (>50,000 examinations), and found a cut-off of 84 exams for trainees before demonstrating a great agreement on BWT with the expert ultrasonographer. However, this study was not limited to IBD patients (only half of the sample), and once again, this does not address the learning curve when starting from scratch, all the more so as in that study the trainees attended an 8-hour theoretical course and conducted 50 IUS together with the expert ultrasonographer(46).

Objectives

The primary objective of the ULtrasound inter-Reader Agreement in IBD (ULTRA-IBD) study was to assess, in a real-life setting among patients with IBD, the inter-rater agreement over time on IUS disease activity as measured by a visual analog scale (VAS) between three operators of varying experience: (i) a radiologist expert in ultrasound (gold standard); (ii) an IBD specialist with formal ultrasound training and regular IUS experience; and (iii) a gastroenterology resident with no prior ultrasound experience (IUS or any other type of ultrasound).

The secondary objectives were:

- to assess the inter-rater agreement on IUS disease activity over time as measured by (i) a semi-quantitative ordinal scale, (ii) validated scores of IUS disease activity in CD (International Bowel Ultrasound Segmental Activity Score [IBUS-SAS]) and UC (MUC score), and (iii) each item composing the scores;
- to analyze the impact of clinical characteristics that might impact IUS technical features on the inter-rater agreement between operators;
- to analyze the correlation between IUS disease activity and other markers of activity (PRO-2, biomarkers, physician global assessment).

Materials & Methods

Study design and setting

The ULTRA-IBD study was a single-center, prospective study led in the Nantes University Hospital from April 2023 to January 2024. Patients were recruited consecutively in the outpatient department.

Study population

Outpatients aged over 18, with an established diagnosis of IBD (CD or UC), coming in for their scheduled visit with their referent IBD specialist, could be included regardless of disease activity. Non-inclusion criteria included ongoing pregnancy, IBD-unclassified, pouchitis, or disease limited to the perineum or to the upper gastrointestinal (GI) tract.

The number of 50 patients to be included was chosen, due to the lack of data about IUS learning curve at the time of study design, based on the cut-off of 20 exams to be qualified being evidenced in studies studying the learning curve for residents in appendicitis or biliary ultrasound(47,48).

Ethics

Eligible patients were informed during the visit with their IBD specialist that they could participate in the study. Patient's oral non-opposition was collected and notified in the visit's report. This study was a non-interventional study approved by the local research ethics committee (*Groupe Nantais d'Éthique dans le Domaine de la Santé [GNEDS]*) on January 30th, 2023 (23-11-01-190).

Clinical data collection

Upon inclusion, the IBD specialist reported his own assessment of disease activity (physician global assessment) on a VAS between 0 and 10, and on a semi-quantitative 4-item ordinal scale (inactive, mildly active, moderately active, and severely active).

Clinical data were collected at baseline, including age, body mass index (BMI), type of IBD (CD or UC), disease duration, disease location and behavior according to the Montreal classification, history of intestinal resection, presence of stoma, patient-reported outcomes (PRO)-2, ongoing disease-related treatments, history of biologic therapy failure, endoscopic disease activity scores and level of biomarkers (CRP and/or FC) in the last 3 months if available.

The disease was considered as active based on the PRO-2 (abdominal pain >1 and/or stool frequency >3 /day in CD; rectal bleeding >0 and/or stool frequency >0 in UC) and/or CRP level (>5 mg/L) and/or FC level (>250 mg/kg) and/or endoscopically active disease and/or physician global assessment using the semi-quantitative ordinal scale rated as moderately or severely active.

IUS protocol

Each patient had 3 successive IUS examinations by 3 different operators always performing the IUS in the same order: first, a gastroenterology trainee with no formal training to general ultrasound and low IUS experience in IBD (<10 exams), hereafter named “GI Trainee”; second, a senior gastroenterologist specialized in IBD with formal training to general ultrasound (6 months including 1 month of supervised examinations) and good IUS experience in IBD (“on-the-job” training, >200 exams), hereafter named “IBD Specialist”; third, an expert GI senior radiologist with extensive general abdominal ultrasound experience (12 years) including IUS in IBD (>500 exams), hereafter named “Radiologist”.

Each IUS was performed in the same conditions with the same US machine (LOGIQ™ P9, GE Healthcare) equipped with two probes (abdominal convex, 1–6 MHz and linear, 5– 10 MHz), in the supine position without any preparation (no fasting required, no intravenous contrast agents administered).

All operators were blinded to the physician global assessment of disease activity, to patient's medical record and case history. Patients were asked not to give any information regarding their disease to the operators during the entire examination period. The only clinical information provided to the operators was the type of IBD and disease location in case of CD (ileal, colonic or ileocolonic). The second and third operators were also blinded to the previous operators' IUS findings for a given patient.

After having performed the 3 successive IUS to each patient, the 3 operators debriefed together in the patient's presence by reviewing ultrasound images notably in case of discordance between the operators.

IUS findings

IUS findings were reported by each operator on a case report form including the following criteria: disease activity as assessed by IUS on a VAS between 0 and 10 and on a semi-quantitative 4-item ordinal scale (inactive, mildly active, moderately active, severely active); maximal BWT (mm), CDS with pulse repetition frequency (PRF) around 5–7 m/s (none=0, short signals=1, long signals inside the bowel=2, long signals inside and outside the bowel=3 as in the modified Limberg score(49)), inflammatory mesenteric fat (i-fat) (absent, uncertain, present), BWS (normal, uncertain, focal < 3 cm, extensive > 3 cm), presence of an abscess (no/yes).

Based on these criteria, the validated IBUS-SAS (50) and MUC score (51) were calculated for patients with CD and UC, respectively, according to the following formulas:

- IBUS-SAS = 4 x BWT (quantitative measure in mm) + 15 x i-fat (absent=0, uncertain=1, present=2) + 7 x CDS (none=0, short signals=1, long signals inside the bowel=2, long signals inside and outside the bowel=3) + 4 x BWS (normal=0, uncertain=1, focal < 3 cm=2, extensive > 3 cm=3);
- MUC score = 1.4 x BWT (quantitative measure in mm) + 2 x CDS (present=1 or absent=0)

Based on the published literature, the cut-offs for defining endoscopic disease activity were >25.2 using the IBUS-SAS for patients with CD(52) and >6.2 using the MUC score for patients with UC(53).

Operators also reported the most inflamed segment (ileum, right colon, transverse colon, left colon/sigmoid, not applicable), the duration of IUS examination, and assessed the patient's echogenicity (good, moderate, bad).

Outcomes

The primary outcome was the inter-rater agreement over time as measured by the VAS of disease activity as scored by the ultrasonographers. Secondary outcomes were the inter-rater agreements over time as measured by (i) the ordinal scale of disease activity, (ii) the IBUS-SAS or MUC score, and (iii) each item composing the scores. The impact of clinical characteristics on the inter-rater agreement between ultrasonographers was also analyzed for the variables of interest that might impact IUS technical features (type of IBD, disease location and behavior in CD, history of intestinal resection, disease activity and BMI). The correlation between IUS disease activity and other markers of activity was also assessed (PRO-2, biomarkers, physician global assessment).

Statistical analysis

Continuous variables are described using means and standard deviations (SD) and were compared using the Student t-test when comparing 2 variables, or ANOVA one-way test when comparing 3 variables. Categorical variables are described as headcounts and percentages and were compared using Chi-squared test (or Fisher's exact test if statistically inappropriate).

The inter-rater agreement was assessed through the two-way random, single measure model of intra-class correlation coefficient (ICC(2,1)) for continuous variables, and through weighted kappa coefficient for categorical variables. All analyses were done with the expert Radiologist considered as the gold standard. ICC and kappa values are interpreted according to Cicchetti's values(54): <0.4 is poor agreement, 0.4 to 0.6 is moderate agreement, 0.6 to 0.75 is good agreement, >0.75 is excellent agreement.

The impact of IBD type, disease location and behavior in CD, history of intestinal surgery, disease activity and BMI, on the inter-rater agreement on IUS disease activity (as assessed by VAS) was analyzed through the comparison of ICCs in the presence and absence of these criteria. ICCs were compared by calculation of the Z-statistic and then of the right-tailed Student's distribution for that Z-statistic.

The correlation between IUS disease activity (as assessed by VAS) and other markers of disease activity (PRO-2, biomarkers [CRP and FC levels], physician global assessment) was assessed using the Spearman's correlation test.

Analyses were performed using Microsoft Excel 2023 (Microsoft Corporation, Redmond, USA) for descriptive statistics, and MedCalc® Statistical Software version 22.023 (MedCalc Software Ltd, Ostend, Belgium) for analytical statistics. For all statistical analyses, $\alpha = 0.05$ was considered as an acceptable threshold for type I error. Figures were done using GraphPad Prism version 10.0.0 for Windows (GraphPad Software, Boston, Massachusetts, USA).

Results

Study population

In total, 150 IUS exams were performed on 47 patients – three patients participated in the study twice. Clinical characteristics are described in **Table 1**. In total, 15 patients (31.9 %) had an active disease (3 UC, 12 CD), including one patient who participated twice and had an active disease at both time points. Only 7 patients (14.9%) had undergone an endoscopy in the 3 months preceding the ultrasound. All 11 patients with a history of intestinal resection were patients with CD, of whom 8 had ileocaecal resection, 2 had colectomy with permanent ileostomy, and 1 had colectomy with ileosigmoidal anastomosis and participated twice.

No significant difference was found between UC and CD patients in terms of BMI (mean [SD] 25.3 [4.3] and 24.1 [4.16], respectively, $p=0.39$), CRP level (3.6 [3.1] and 3.0 [2.5], respectively, $p=0.36$), FC level (1024.2 [1127.1] and 289.5 [1183.0], respectively, $p=0.41$), and physician global assessment of disease activity (using VAS, 0.9 [1.7] and 1.0 [1.6], respectively, $p=0.86$; using the semi-quantitative scale, 0.6 [0.7] and 0.5 [0.8], respectively, $p=0.47$).

Table 1. Clinical characteristics of the study population at baseline (n=47 patients).

Type of IBD, n (%)	
CD	36 (76.6)
UC	11 (23.4)
Gender, n (%)	
Female	24 (51.0)
Male	23 (49.0)
Disease duration [years], mean (SD)	15.3 (8.4)
Age at baseline [years], mean (SD)	44.5 (15.8)
BMI ^a [kg/m ²], mean (SD)	24.2 (4.2)
Disease location in patients with CD, n (%)	
L1 (ileal)	10 (27.8)
L2 (colonic)	10 (27.8)
L3 (ileocolonic)	16 (44.4)
L4 (upper digestive tract)	3 (8.3)
Perineal location	13 (36.1)

Table 1 (Continued). Clinical characteristics of the study population at baseline (n=47 patients).

Disease location in patients with UC, n (%)	
E1 (proctitis)	1 (9.1)
E2 (left-sided colitis)	4 (36.4)
E3 (extensive colitis)	6 (54.5)
History of biologic therapy, n (%)	
1 biologic	19 (40.4)
2 biologics	9 (19.1)
≥3 biologics	11 (23.4)
Disease behavior in patients with CD, n (%)	
B1 (inflammatory)	26 (72.2)
B2 (stricturing)	8 (22.2)
B3 (penetrating)	2 (5.6)
History of intestinal resection, n (%)	
11 (23.4)	
Presence of stoma, n (%)	
2 (4.3)	
PRO-2 in patients with CD ^b , mean (SD)	
Stool frequency	1.1 (0.7)
Abdominal pain	0.3 (0.6)
PRO-2 in patients with UC ^b , mean (SD)	
Stool frequency	0.3 (0.8)
Rectal bleeding	0.1 (0.3)
CRP ^{b, c} [mg/L], mean (SD)	
3.1 (2.6)	
Fecal calprotectin ^{b, d} , mean (SD)	
569.2 (1170.8)	
Physician global assessment ^b	
VAS [0-10], mean (SD)	1.0 (1.6)
Semi-quantitative scale, n (%)	
Inactive	29 (58.0)
Mildly active	15 (30.0)
Moderately active	6 (12.0)
Severely active	0 (0.0)
Treatment modification or optimization in the last 3 months ^b , n (%)	
7 (14.0)	
Ongoing IBD-related treatments ^b , n (%)	
No treatment or exclusive 5-ASA	9 (18.0)
Azathioprine or methotrexate	6 (12.0)
Anti-TNF agents	25 (50.0)
Other biologics	13 (26.0)
Jak inhibitor	3 (6.0)

^a: 3 missing data; ^b: calculations made on 50 patients as 3 patients participated twice in the study; ^c: 7 missing data; ^d: 20 missing data.

Abbreviations: ASA, Aminosalicylates; BMI, Body mass index; CD, Crohn's disease; CRP, C-reactive protein; IBD; Inflammatory bowel disease; SD, Standard deviation; TNF, Tumor necrosis factor; UC, Ulcerative colitis; VAS, Visual analog scale.

IUS findings

As designed, 150 US examinations (50 per operator) were performed in total. **Table 2** summarizes IUS findings by the expert Radiologist for the overall population. No abscess was found during the study.

Among the three patients who participated twice, one CD patient was judged as mildly active by the clinician at both timepoints (6 months apart); this was in agreement with the ultrasound assessment performed by Radiologist at the first timepoint but not at the second timepoint during which the patient was rated as inactive, with a decrease in BWT of 1 mm. Another CD patient went from mildly active to inactive according to the physician (3 months apart), and was considered as inactive by Radiologist at both timepoints. The last patient had UC assessed as moderately active at both timepoints by the physician (3 months apart), which was in agreement with the ultrasound assessment at the first timepoint but not at the second one during which the patient was rated as mildly active with a decrease in BWT of 2 mm.

Table 2. IUS findings by the expert Radiologist (gold standard) (n=50 examinations).

	Overall population (n=50)	CD (n=38)	UC (n=12)
IUS disease activity			
VAS [0-10], mean (SD)	1.1 (1.7)	1.1 (1.8)	1.0 (1.6)
Semi-quantitative scale, n (%)			
Inactive	16 (32.0)	12 (31.6)	4 (33.3)
Mildly active	6 (12.0)	5 (13.2)	1 (8.3)
Moderately active	0 (0.0)	0 (0.0)	0 (0.0)
Severely active			
Maximal BWT [mm], mean (SD)	4.1 (1.7)	4.2 (1.7)	3.9 (1.6)
Color Doppler signal, n (%)			
Limberg 0	37 (74.0)	29 (76.3)	8 (66.7)
Limberg 1	13 (26.0)	9 (23.7)	4 (33.3)
Limberg 2	0 (0.0)	0 (0.0)	0 (0.0)
Limberg 3	0 (0.0)	0 (0.0)	0 (0.0)
Inflammatory mesenteric fat, n (%)			
Absent	38 (76.0)	30 (78.9)	8 (66.7)
Uncertain	9 (18.0)	5 (13.2)	4 (33.3)
Present	3 (6.0)	3 (7.9)	0 (0.0)
BWS, n (%)			
Normal	34 (68.0)	25 (65.8)	9 (75.0)
Uncertain	1 (2.0)	1 (2.6)	0 (0.0)
Focal < 3 cm	4 (8.0)	4 (10.5)	0 (0.0)
Extensive > 3 cm	11 (22.0)	8 (21.1)	3 (25.0)
IBUS-SAS, mean (SD)	NA	26.3 (20.4)	NA
MUC score, mean (SD)	NA	NA	6.2 (3.1)

Abbreviations: BWS, Bowel wall stratification; BWT, Bowel wall thickness; CD, Crohn's disease; IBUS-SAS, International Bowel Ultrasound Segmental Activity Score; IUS, Intestinal ultrasound; MUC, Milan ultrasound criteria; NA, Not applicable; SD, Standard deviation; UC, Ulcerative colitis; VAS, Visual analog scale.

IUS technical features

Table 3 describes the differences between the three operators in terms of examination times, patient's echogenicity, and assessment of the most inflamed segment. Overall, IBD Specialist was the fastest operator especially for patients with UC, while GI Trainee was the slowest. The less experienced the operator, the poorer patient's echogenicity was judged to be. The assessment of the most inflamed segment was quite concordant between the three operators.

When considering only the gold standard Radiologist, there was a significant positive correlation between longer examination time and higher disease activity in patients with UC, with a Spearman's ρ at 0.62 (95% CI 0.07-0.88, $p=0.03$); such a correlation was not found in patients with CD ($\rho=0.21$ [95% CI -0.11-0.50, $p=0.21$]). The examination time was significantly lower in patients with UC than in patients with CD ($p=0.03$), but there was no significant difference in Radiologist's examination time according to CD location (L2 disease versus L1+L3 diseases), with mean examination times of 04:13 (SD 02:29) and 04:29 (SD 01:21), respectively ($p=0.76$). Radiologist took the same time for performing IUS in patients with a BMI >25 and in those <25 kg/m^2 ($p=0.64$). A history of intestinal resection did not affect the examination time for Radiologist, with mean examination times of 04:18 (SD 01:15) for patients with history of surgery versus 04:10 (SD 01:42) for patients who never had surgery ($p=0.78$).

Table 3. Differences in IUS technical features between the three operators.

	Radiologist (n=50 exams)	IBD Specialist (n=50 exams)	GI Trainee (n=50 exams)	p-value
Examination time [min:sec], mean (SD)				
Overall (n=50)	04:12 (01:35)	03:45 (01:46)	05:17 (01:16)	<0.001
Patients with CD (n=38)	04:24 (01:41)	04:12 (01:40)	05:32 (01:10)	0.001
Patients with UC (n=12)	03:31 (00:58)	02:17 (01:14)	04:27 (01:14)	0.05
BMI				
>25 kg/m ² (n=19)	04:21 (01:34)	03:51 (02:13)	05:34 (01:12)	0.005
<25 kg/m ² (n=28)	04:07 (01:40)	03:43 (01:33)	05:00 (01:04)	0.001
History of surgery (n=12)	04:18 (01:15)	03:39 (00:54)	05:29 (01:29)	0.004
Patient's echogenicity, n (%)				
Good	48 (96.0)	44 (88.0)	34 (68.0)	-
Average	2 (4.0)	5 (10.0)	14 (28.0)	
Bad	0 (0.0)	1 (2.0)	2 (4.0)	
Most inflamed segment, n (%)				
Ileum	14 (28.0)	12 (24.0)	15 (30.0)	-
Right colon	0 (0.0)	0 (0.0)	0 (0.0)	
Transverse colon	0 (0.0)	0 (0.0)	1 (2.0)	
Left colon/sigmoid	11 (22.0)	15 (30.0)	12 (24.0)	
Not applicable (no activity)	25 (50.0)	23 (26.0)	22 (44.0)	

Abbreviations: BMI, Body mass index; CD, Crohn's disease; GI, Gastrointestinal; IBD; Inflammatory bowel disease; IUS, Intestinal ultrasound; SD, Standard deviation; UC, Ulcerative colitis.

Inter-rater agreement on IUS disease activity assessed by VAS

When considering the VAS assessing IUS disease activity, the inter-rater agreement over time between IBD Specialist and the gold standard (Radiologist) was good to excellent throughout the study without any “learning curve” effect, with an ICC ranging from 0.61 to 0.79 during the whole study (**Figure 1A**). The overall agreement after 50 patients was good with an ICC at 0.71 (95% CI 0.54-0.83).

By contrast, the inter-rater agreement between GI Trainee and Radiologist was poor with an overall ICC after 50 patients at 0.33 (95% CI 0.05-0.55). There was no clear “learning curve” effect, with a steady improvement until the 30th patient, with an ICC going from 0.21 after 10 patients to 0.49 after 30 patients, but then agreement gradually decreased until 0.33 after 50 patients (**Figure 1B**).

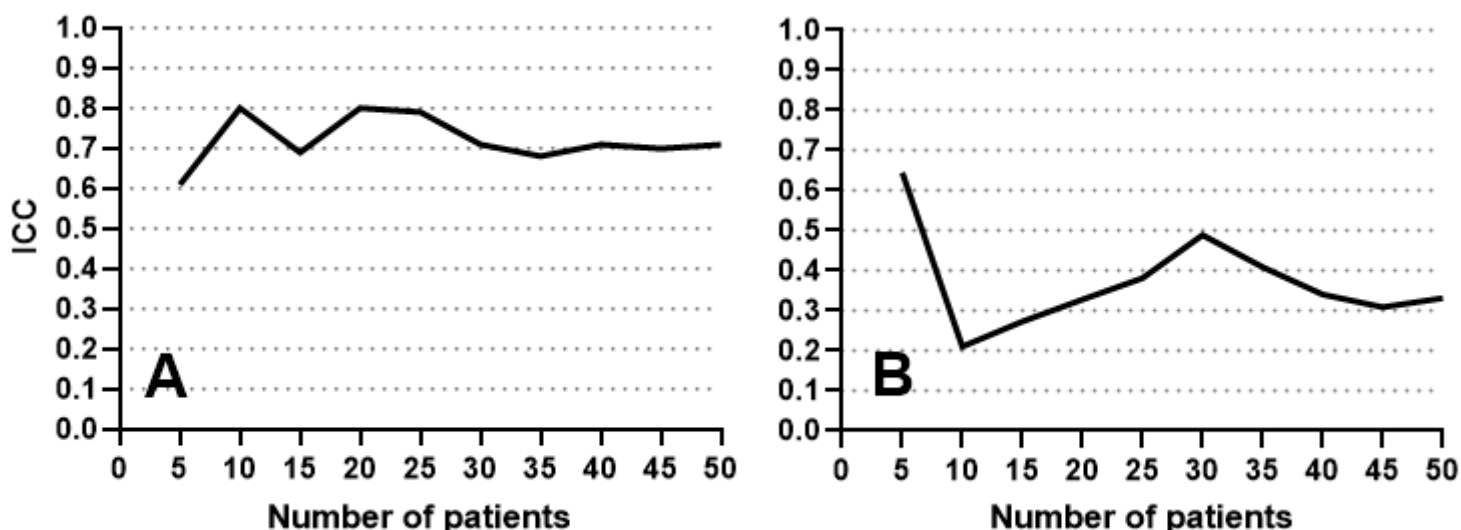


Figure 1. Evolution of inter-rater agreement over time on IUS disease activity assessed by VAS: (A) between IBD Specialist and Radiologist; and (B) between GI Trainee and Radiologist.

Abbreviations: GI, Gastrointestinal; IBD, Inflammatory bowel disease; ICC, Intraclass correlation coefficient; IUS, Intestinal ultrasound; VAS, Visual analog scale.

Inter-rater agreement on IUS disease activity assessed by a semi-quantitative ordinal scale

Similar results were achieved when using a semi-quantitative 4-item ordinal scale for assessing IUS disease activity. The overall weighted Cohen’s kappa for agreement between IBD Specialist and Radiologist after 50 patients was 0.68 (95% CI 0.55-0.80) (good agreement) (**Figure 2A**), while the overall weighted Cohen’s kappa for agreement between GI Trainee and Radiologist after 50 patients was 0.23 (95% CI -0.10-0.40) (poor agreement) (**Figure 2B**).

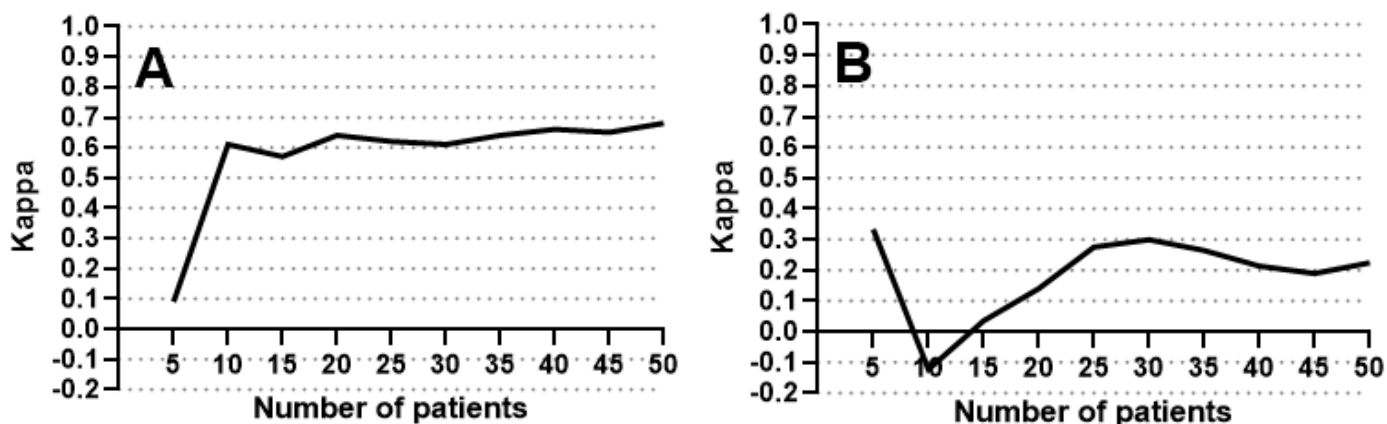


Figure 2. Evolution of inter-rater agreement over time on IUS disease activity assessed by a semi-quantitative ordinal scale: (A) between IBD Specialist and Radiologist; and (B) between GI Trainee and Radiologist.

Abbreviations: GI, Gastrointestinal; IBD, Inflammatory bowel disease; IUS, Intestinal ultrasound.

Inter-rater agreement on IUS disease activity scores

In patients with CD (n=38), the IBUS-SAS showed almost-excellent agreement between IBD Specialist and Radiologist, whether using the score as a continuous variable (ICC = 0.74 [95% CI 0.56-0.85]) or considering the cut-off >25.2 to indicate an active disease (kappa = 0.74 [95% CI 0.53-0.94]). By contrast, the agreement between GI Trainee and Radiologist was poor in both cases (ICC = 0.23 [95% CI -0.10-0.51] and kappa = 0.17 [95% CI -0.14-0.48]) (Figure 3).

Similarly, in patients with UC (n=12), the MUC score showed good-to-excellent agreement between IBD Specialist and Radiologist, whether using the score as a continuous variable (ICC = 0.80 [95% CI 0.20-0.94]) or considering the cut-off >6.2 (kappa = 0.62 [95% CI 0.21-1.00]). The agreement between GI Trainee and Radiologist was poor in both cases (ICC = 0.36 [95% CI -0.24-0.76] and kappa = 0.27 [95% CI -0.26-0.80]) (Figure 3).

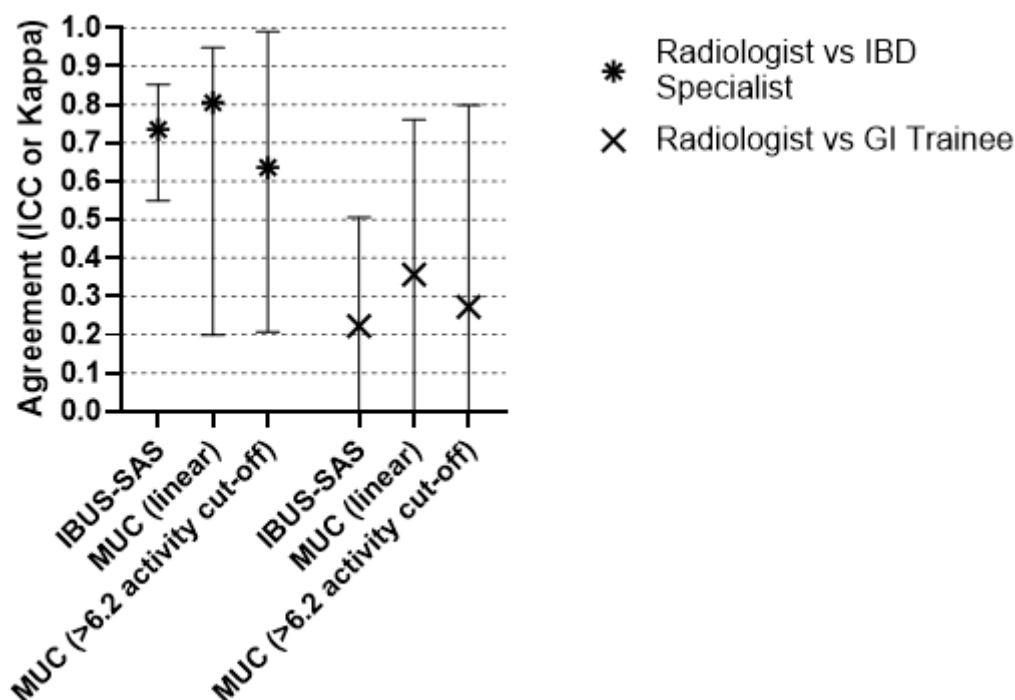


Figure 3. Inter-rater agreement between the three operators on IBUS-SAS in CD patients and MUC score in UC patients, as linear or categorical values.

Abbreviations: CD, Crohn's disease; GI, Gastrointestinal; IBD, Inflammatory bowel disease; IBUS-SAS, International Bowel Ultrasound Segmental Activity Score; ICC, Intraclass correlation coefficient; IUS, Intestinal ultrasound; MUC, Milan ultrasound criteria; UC, Ulcerative colitis.

Inter-rater agreement on IUS parameters

Out of all four IUS parameters scored, BWT and BWS had the best agreement between IBD Specialist and Radiologist. The inter-rater agreement on BWT stayed in a narrow range of good agreement throughout the study (range of ICC 0.66-0.74) and the overall ICC after 50 patients was at 0.74 (95% CI 0.59-0.85). The inter-rater agreement on BWS trended towards a “learning curve” effect, with a kappa coefficient steadily rising from 0.48 after 5 patients to 0.83 after 40 patients. The overall kappa coefficient after 50 patients was at 0.74 (95% CI 0.57-0.91) (good agreement). The other parameters showed poor-to-moderate agreement, with overall kappa values for CDS and i-fat at 0.39 (95% CI 0.21-0.58) and 0.53 (95% CI 0.23-0.82), respectively, without any “learning curve” effect (**Figure 4A**).

The inter-rater agreement between GI Trainee and Radiologist was poor on all four IUS parameters (ICC for BWT = 0.20 [95% CI -0.06-0.46]; kappa for CDS = 0.28 [95% CI 0.04-0.52]; kappa for i-fat = 0.06 [95% CI -0.21-0.32]; kappa for BWS = 0.11 [95% CI -0.13-0.34]), without any “learning curve” effect (**Figure 4B**).

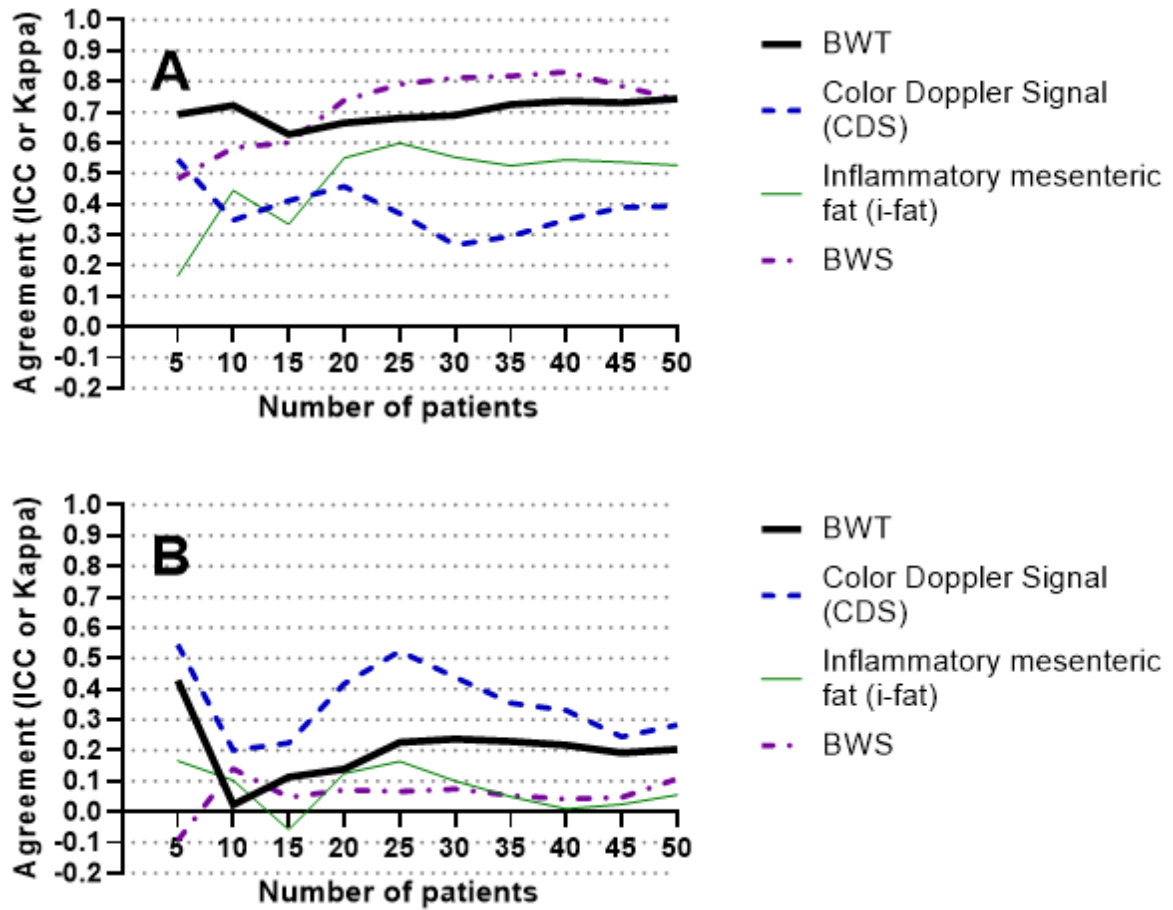


Figure 4. Inter-rater agreement on each IUS parameter: (A) between IBD Specialist and Radiologist; and (B) between GI Trainee and Radiologist.
 Abbreviations: GI, Gastrointestinal; ICC, Intraclass correlation coefficient; IUS, Intestinal ultrasound; BWT, Bowel Wall Thickness; BWS, Bowel Wall Stratification.

Impact of clinical characteristics on the inter-rater agreement

Type of IBD

The inter-rater agreement was significantly higher for UC patients than for CD patients when comparing IBD Specialist and Radiologist, with an ICC at 0.91 (95% CI 0.72-0.97) (excellent) and 0.67 (95% CI 0.44-0.82) (good), respectively ($p=0.003$). When comparing GI Trainee and Radiologist, the inter-rater agreement was quite similar for both types of IBD, with an ICC at 0.22 (95% CI -0.37-0.69) for UC patients and 0.36 (95% CI 0.03-0.60) for CD patients ($p=0.30$) (**Figure 5**).

CD location

For CD patients specifically, the inter-rater agreement differed depending on disease location. IBD Specialist and Radiologist had a good-to-excellent agreement for ileal disease L1 (ICC = 0.75 [95% CI 0.28-0.93]) and ileocolonic disease L3 (ICC = 0.68 [95% CI 0.27-0.88]), but only poor agreement for colonic disease L2 (ICC = 0.19 [95% CI -0.57-0.70]) ($p=0.02$ for L2 versus L1 and $p=0.04$ for L2 versus L3). The inter-rater agreement between GI Trainee and Radiologist had a different pattern, with moderate-to-good agreement for ileal L1 and colonic L2 diseases (ICC at 0.61 [95% CI -0.02-0.89] and 0.62 [95% CI 0.09-0.89], respectively), but poor agreement for ileocolonic disease L3 with an ICC at 0.08 (95% CI -0.41-0.52) ($p=0.03$ for L3 versus L1 and $p=0.02$ for L3 versus L2) (**Figure 5**).

History of IBD-related surgery

History of IBD-related surgery did not significantly affect the inter-rater agreement between IBD Specialist and Radiologist, with respective ICCs of 0.65 (95% CI 0.19-0.88) for patients with history of surgery and 0.74 (95% CI 0.55-0.86) for patients without history of

surgery ($p=0.75$). However, it highly impacted the inter-rater agreement between GI Trainee and Radiologist, with a negative agreement (ICC = -0.44 [95% CI -0.89-0.15]) for patients with history of surgery, but an almost-good agreement (ICC = 0.59 [95% CI 0.34-0.76]) for patients who never underwent surgery ($p<0.001$) (**Figure 5**).

Given this high impact of surgery on IUS accuracy, the assessment of inter-rater agreement in CD patients according to disease location was repeated after excluding patients with a history of surgery. This significantly impacted the results between GI Trainee and Radiologist, going up to excellent agreement for L1 ($n=7$) and L2 ($n=9$) diseases (ICC = 0.89 [95% CI 0.49-0.98] and 0.93 [95% CI 0.75-0.98], respectively) with a statistically significant increase as compared to the analysis on the total population of CD patients ($p=0.03$ and $p=0.02$, respectively), while the inter-rater agreement for L3 disease ($n=10$) went from poor to moderate agreement (ICC = 0.47 [95% CI -0.22-0.83]) although the difference was not significant ($p=0.09$). The inter-rater agreement according to CD location between IBD Specialist and Radiologist was not significantly impacted by the exclusion of patients with history of surgery: $ICC_{L1} = 0.79$ (95% CI 0.17-0.96) ($p=0.62$), $ICC_{L2} = 0.12$ (95% CI -0.65-0.72) ($p=0.43$), and $ICC_{L3} = 0.74$ (95% CI 0.04-0.94) ($p=0.66$) (**Figure 5**).

Following the same logic, the assessment of inter-rater agreement over time was repeated in the overall population after exclusion of patients with a history of surgery ($n=38$). Agreement between GI Trainee and Radiologist yielded a clear “learning curve” effect, with an ICC going from 0.21 (poor agreement) to 0.59 (almost-good agreement) with a steady increase throughout the study (**Figure 6**). The same learning curve was achieved for ordinal activity scale agreement, although both the nadir kappa ($k=-0.07$) and the end kappa ($k=0.46$) were lower.

CD behavior

For IBD Specialist and Radiologist, agreement was high regardless of CD behavior, with an ICC at 0.69 (95% CI 0.38-0.86) for B1 disease and an ICC at 0.70 (95% CI 0.12-0.92) for B2 disease (p=0.53). For GI Trainee and Radiologist, agreement was poor-to-moderate for B1 disease with an ICC at 0.44 (95% CI 0.08-0.70), and poor for B2 disease with an ICC at 0.12 (95% CI -0.44-0.68), although the difference was not significant (p=0.16). B3 disease could not be statistically analyzed because of the small sample size (**Figure 5**).

Disease activity

Disease activity did not impact the inter-rater agreement between operators. IBD Specialist and Radiologist had similar agreement for active patients (ICC = 0.68 [95% CI 0.30-0.87] and inactive patients (ICC = 0.70 [95% CI 0.47-0.84]) (p=0.57). GI Trainee and Radiologist had a non-significant trend of better agreement for active patients, with an ICC at 0.38 (95% CI -0.11-0.72) versus 0.21 (95% CI -0.14-0.51) for inactive patients (p=0.30) (**Figure 5**).

Body mass index

Quite surprisingly, the BMI affected the inter-rater agreement in opposite ways according to the operator, but without any statistical difference. The agreement between IBD Specialist and Radiologist was a bit higher in patients with overweight (BMI >25 kg/m²) than in non-overweight patients (n=28) (ICC = 0.73 [95% CI 0.42-0.89] and 0.60 [95% CI 0.26-0.81], respectively, p=0.17). Conversely, the agreement between GI Trainee and Radiologist was better in non-overweight patients (ICC = 0.45 [95% CI 0.09-0.70]) than in overweight patients (ICC = 0.25 [95% CI -0.24-0.62]) (p=0.19) (**Figure 5**).

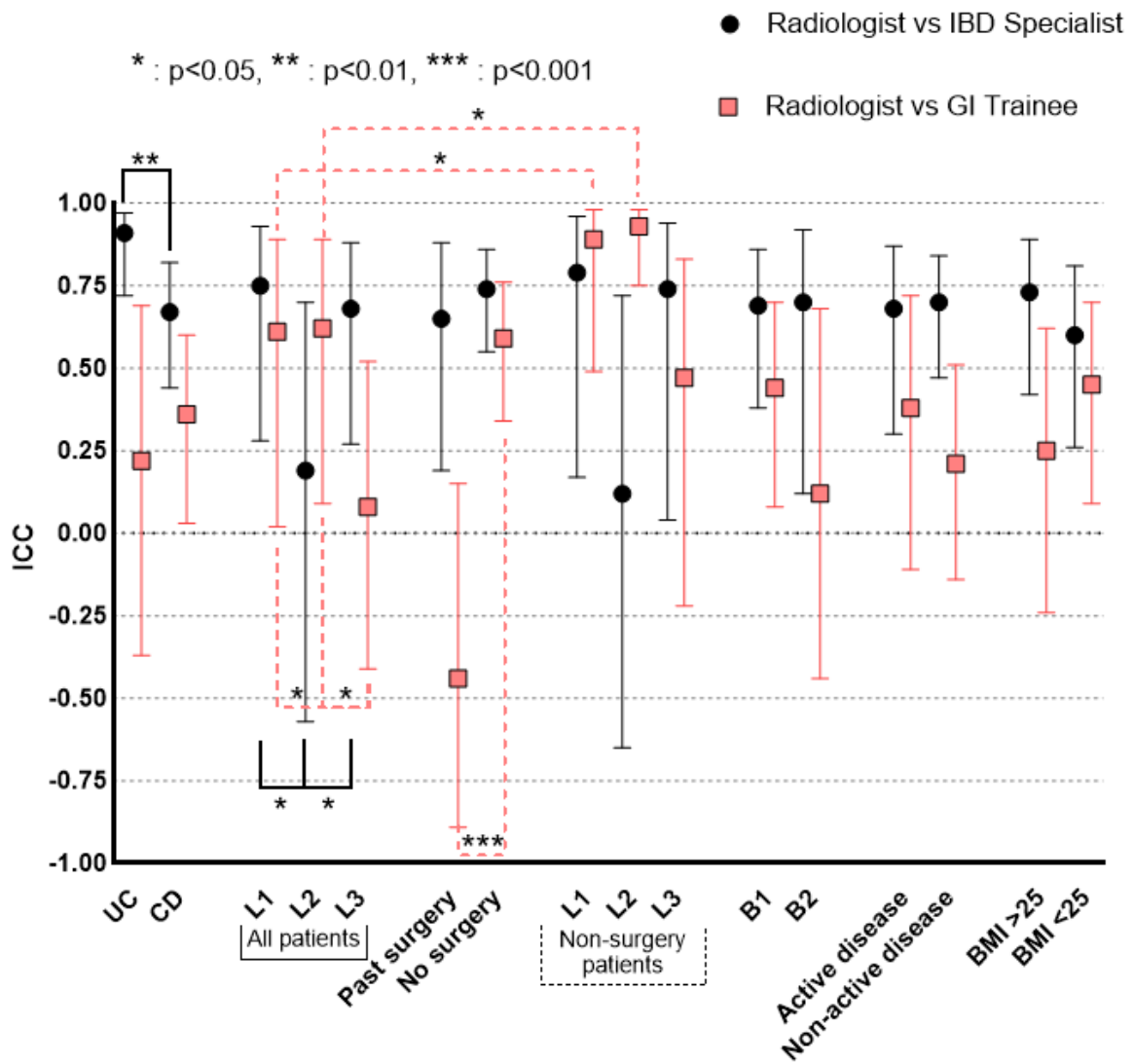


Figure 5. Variations of inter-rater agreement on IUS disease activity assessed by VAS according to IBD type, CD location, history of surgery, CD behavior, disease activity and BMI.

Abbreviations: BMI, Body mass index; CD, Crohn’s disease; IBD, Inflammatory bowel disease; UC, Ulcerative Colitis; ICC, Intraclass correlation coefficient.

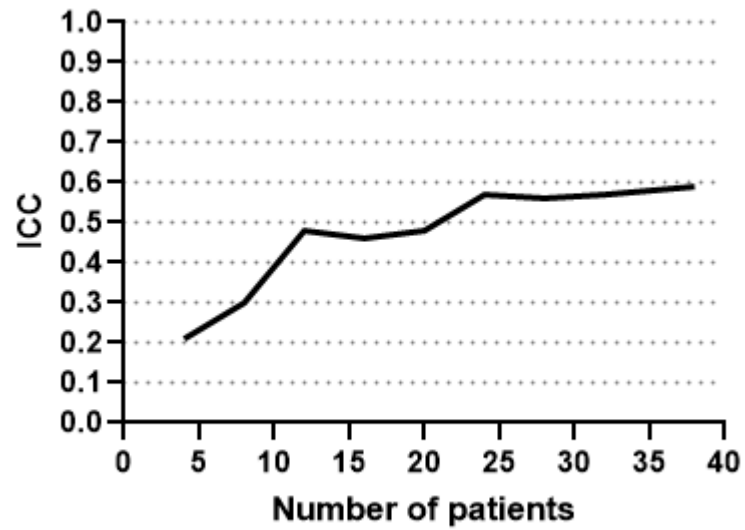


Figure 6. Evolution of inter-rater agreement over time on IUS disease activity assessed by VAS between GI Trainee and Radiologist after exclusion of patients with history of IBD-related surgery.

Abbreviations: GI, Gastrointestinal; IBD, Inflammatory bowel disease; ICC, Intraclass correlation coefficient; IUS, Intestinal ultrasound; VAS, Visual analog scale.

Correlation between IUS disease activity and other markers of disease activity

PRO-2

In patients with CD, there was a slight correlation between PRO-2 and Radiologist's assessment of disease activity through VAS, with Spearman's $\rho = 0.36$ (95% CI 0.05-0.61, $p=0.02$). Such correlation was not found for IBD Specialist and GI Trainee, with respective ρ values of 0.15 (95% CI -0.17-0.45, $p=0.35$) and 0.29 (95% CI -0.03-0.56, $p=0.07$).

In patients with UC, there was a good correlation between PRO-2 and Radiologist's and IBD Specialist's assessments of disease activity through VAS, with Spearman's $\rho = 0.77$ (95% CI 0.34-0.93, $p=0.004$) and 0.60 (95% CI 0.39-0.87, $p=0.04$), respectively. Such a correlation was not found for GI Trainee, with a ρ value of 0.34 (95% CI -0.29-0.76, $p=0.27$).

Biomarkers

For all three operators, Spearman's ρ showed a poor correlation between IUS disease activity assessment and biomarkers. For CRP (43 available data), ρ values were at 0.20 (95% CI -0.10-0.48, $p=0.19$), 0.14 (95% CI -0.16-0.43, $p=0.35$) and 0.09 (95% CI -0.22-0.38, $p=0.58$) for Radiologist, IBD Specialist and GI Trainee, respectively. For FC (30 available data), ρ values were at 0.12 (95% CI -0.25-0.46, $p=0.52$), 0.20 (95% CI -0.17-0.52, $p=0.29$), and -0.27 (95% CI -0.57-0.10, $p=0.15$) for Radiologist, IBD Specialist and GI Trainee, respectively.

Physician global assessment

There was a significant positive correlation between all three operators and the physician global assessment using the VAS, with the correlation being highest for Radiologist with ρ at 0.60 (95% CI 0.40-0.76, $p<0.0001$), moderate for IBD Specialist with ρ at 0.40

(95% CI 0.14-0.61, $p=0.004$), and weakest for GI Trainee with ρ at 0.32 (95% CI 0.05-0.55, $p=0.02$).

Similar results were found when using the semi-quantitative ordinal scale for the physician global assessment, with the correlation being highest for Radiologist with ρ at 0.45 (95% CI 0.20-0.65, $p<0.001$), moderate for IBD Specialist with ρ at 0.32 (95% CI 0.04-0.55, $p=0.02$), and weakest for GI Trainee with ρ at 0.29 (95% CI 0.01-0.53, $p=0.04$).

Discussion

The ULTRA-IBD study analyzed the inter-rater agreement on IUS findings in patients with IBD between three operators of varying experience, including an expert gastrointestinal radiologist, an IBD specialist and a gastroenterology resident starting from scratch. After 50 patients recruited consecutively, no “learning curve” effect was observed for both gastroenterologists but for totally opposite reasons. The IBD specialist with formal general ultrasound training and already experienced in IUS (“on-the-job” training) showed continuously good agreement with the expert radiologist throughout the study, while the unexperienced gastroenterology resident failed to go beyond poor agreement.

This primarily highlights that being an expert radiologist is not needed to have satisfying performance in point-of-care IUS for assessing disease activity in IBD, and gastroenterologists with sufficient experience can be confident in their results. This is consistent with other studies in which an experience of 200 examinations allowed a gastroenterologist to get accurate IUS results for the assessment of IBD activity as compared to MRE and ileocolonoscopy(35) or to experts with more than 1000 IUS examinations of experience(45). The other key point is the importance of a formal training in general ultrasound with supervised examinations before beginning IUS training. Indeed, for a complete beginner without proper ultrasound training, a total of 50 IUS examinations is not enough to achieve acceptable performance on IBD patients.

This study is the first one to analyze the “learning curve” from scratch in IUS performance in a specific population of IBD patients. It has been performed in a real-life setting and three operators of varying experience in IUS and IBD each performed an ultrasound examination on 50 patients.

Our study has limitations, mainly the fact that only one novice trainee and one experienced IUS operator were evaluated against the gold standard in a single center. Thus, we cannot rule out the individual operators' innate capabilities limiting the generalization of results to a wider population of gastroenterologists. Another weakness is the low number of active patients included in the study, although it does reflect the reality of IBD outpatients in a tertiary expert center, and makes our results maybe more significant as more active patients would have likely be easier to assess. Finally, few of the included patients had recent endoscopy, which remains the gold standard in IBD activity assessment, to compare our IUS results with.

The discrepancy in patterns of performance according to disease type, as well as disease location for CD, raise interesting questions about the learning of IUS. In our study, IBD Specialist and Radiologist achieved excellent agreement for UC patients with a similar ICC to the one in a study evaluating agreement between experts of IUS(55); yet, in CD it is for colonic-only disease that the agreement was poor, contrasting with good agreement for both CD disease locations involving the ileum.

One explanation is that since UC is a continuous disease with retrograde extension from the rectum, its examination is easy to standardize by finding the easily accessible sigmoid colon, which will always be affected in case of colonic disease activity. In contrast, colonic CD being potentially discontinuous, it necessitates careful examination of the whole colon from beginning to end, which is harder to achieve.

In contrast, GI Trainee and Radiologist achieving a slightly higher (but still poor) agreement for CD than for UC goes against the hypothesis of UC activity being easier to assess, at least at a complete beginner level. Even more surprisingly, looking at performance by CD disease location showed moderately good agreement for isolated ileal and colonic disease, but poor agreement for ileocolonic disease.

This suggests that at a very beginner level, evaluating the activity in a restrained and relatively pre-defined location is easier than finding the location of activity in the first place, meaning that successfully exploring and understanding dynamic probe handling is a key IUS skill that is more discriminating than successfully seeing if there is activity. In other words, performing the IUS examination is harder than interpreting it.

The important impact of surgery on novice performance is also interesting. The way surgery did not significantly alter agreement between IBD Specialist and Radiologist, while it was source of a stark difference in agreement for GI Trainee and Radiologist, can be interpreted as an example of the same issue. Overall performance of GI Trainee after excluding the patients with history of surgery increased from poor to moderate-to-good with a clear “learning curve” effect, hinting once more that performing the examination and working around non-modal anatomy is a more advanced IUS skill than seeing and interpreting standard “cuts” found in the same location in modal anatomy.

Thus, while less than 40 examinations on patients who never had surgery seems to be enough to achieve moderate-to-good agreement and a consistent trend of improvement, a total of 50 examinations on a general population of IBD patients including about one-quarter of patients operated on is not enough for a complete beginner to achieve consistent improvement up to acceptable agreement.

This echoes the recent study by Bezzio et al. (46) which showed that around 84 examinations were necessary for IUS trainees, and 101 examinations for complete novices like GI Trainee in our study, to achieve consistent agreement with an expert ultrasonographer on BWT measurement; and that is despite these trainees having 50 IUS examinations supervised by an expert, unlike in our study. Since patient characteristics are not detailed in

this study, it is also unclear in which proportion patients with factors that led to trouble for GI Trainee's performance in our study (BMI >25 kg/m², history of surgery, ileocolonic CD) were recruited.

However, the early gap in performance in Bezzio et al.'s study between trainees who were already well-experienced in general ultrasound and complete ultrasound novices reinforces the hypothesis that the key early-level ultrasound skill is *technically performing* the US examination itself and understanding the relationship between probe positioning and angulation and internal anatomy, rather than the interpretation of the examination and recognizing normal or pathological characteristics.

A roughly similar cut-off value for achieving good performances for trainees in IUS was found in a study by Brodersen et al. (34). Although the study concerned CD diagnosis and thus the gold standard was ileocolonoscopy, there was a trend of progress in performance from first 55 examinations to next 55 examinations for trainees who had a previous experience of less than 50 examinations. However, it should be noted that one of the trainees was following formal IUS training in parallel. In line with our own results, that improvement only happened for colonic-only and ileo-colonic CD. It is also noted that all CD diagnosis missed by trainees were short-extent disease of 6 cm or less, which reinforces the hypothesis that technical skills is more essential than interpretation skills.

The study by Kim et al.(47) about appendicitis diagnosis with ultrasound by emergency medicine residents tends to the same conclusion : a very low cut-off number of 20 examinations before reaching good-to-excellent agreement with a surgical or CT-scan diagnosis was reached. Since all included patients had right lower quadrant pain, we can postulate that most of the included appendicitis were of typical and easy to find retrocaecal or paracaecal location, making the learning curve mostly about interpretation of the abnormalities rather than about finding the abnormalities. Similarly, agreement between emergency medicine residents and expert radiologists was good after 20 examinations for

detecting gallstone, sonographic Murphy sign, or gallbladder wall thickening(48), all concerning an anatomically fixed structure.

Even after the novice level, this effect was evidenced by Bhatnagar et al.(56) in a sub-study of the METRIC trial involving ultrasonographers who are not described as experts but as well-trained and regularly practicing IUS: they found consistently higher agreement for disease activity when limiting the analysis to segments identified correctly by both ultrasonographers. Much like in our study for the results of IBD Specialist, agreement for colonic-only CD was poor-to-moderate while agreement for small-bowel-only CD was good.

Not all the IUS parameters analyzed in our study showed good agreement between IBD Specialist and Radiologist, and our results differ from the literature. While we confirm that BWT is a consistently reliable indicator, BWS was surprisingly the parameter with the best agreement, while CDS scored according to the modified Limberg scale had poor agreement and the aspect of mesenteric fat (i-fat) had moderate agreement.

Apart from BWT which is always evaluated as reproducible and reliable in studies, our other results are in contrast with most of the literature. In the development study of the IBUS-SAS(50), it was the BWS that had poor agreement with $k=0.39$ while CDS had moderate-to-good agreement with $k=0.60$. In Fraquelli et al.'s study (43), poor agreement was found for BWS while CDS had good agreement in CD patients. Calabrese et al.'s (44) also found higher reliability for CDS than for BWS, although the difference was less pronounced. Smith et al.'s(45) found higher agreement for i-fat than for other parameters, while BWS and CDS were at similar agreement. One explanation may be the low number of patients with active disease in our study, because CDS and i-fat have been reported as good signs of current inflammation and good predictors of future flare-ups (29,38,41,42,51,57), while BWT and BWS can also be altered in case of inactive disease (sequelae).

While all patients are not good fit for ultrasound examinations in general, we showed contrasting patterns of difficulty factors depending on experience level. The fact that BMI $>25 \text{ kg/m}^2$ impaired performance for GI Trainee is consistent with common experience of ultrasound being harder to perform on overweight patients; however, the opposite trend was shown for IBD Specialist, with a higher agreement for overweight patients. This difference is not down to examination time, since both GI Trainee and IBD Specialist spent the same time or longer with overweight patients.

Few studies published the analysis of such factors' impact on agreement, although it should be noted that in Kim et al.'s appendicitis study, the resident with the lower early performance had more overweight patients in his cohort than his two peers, even though the difference was not statistically significant(47).

In our study, no strong correlation was found between IUS disease activity assessment and common markers of disease activity (PRO-2, CRP, FC) each taken in isolation. However, we showed a good correlation between IUS activity and the physician global assessment of disease activity which is obviously based in part on PRO-2, CRP and FC. Keeping in mind that the IUS were performed in full blindness of current or past biomarkers, symptoms, and clinician assessment, this hints at blinded IUS examination being a good overall assessment tool and not just a surrogate for other individual markers of activity.

With the growing number of IBD patients and the need for accurate evaluation of disease activity, IUS is of the utmost interest both for its reliability, ease of access, time-efficiency, patient tolerability, and low cost. However, our study along with others outlines the need for proper training and experience before getting reliable results in IUS. But what is proper training, and its modalities, is today not well defined.

No formal consensus by a scientific society currently exists on IBD IUS training. The European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines for IUS(58) only mention that “dedicated training is necessary”, with a low quality of evidence. An international consensus using Delphi method was recently published about the objectives to attain in training for achieving competency(59), in which several gastroenterologists and only one radiologist participated. It is interesting to note that, contrary to what our results tend to show as being harder for novices, this consensus has more statements about *interpretation* skills than about *technical* skills, and no mention is made of being able to successfully identify post-surgical anatomical modifications. The number of examinations necessary to acquire proper experience is also voluntarily not defined, on the argument that this figure will differ from trainee to trainee.

Thus, a formal consensus by a scientific society, addressing these points, should be developed and published to streamline IUS training and ensure that all IBD patients get accurate assessment. A possible model would be the European Society of Gastrointestinal Endoscopy (ESGE) statement for endoscopic ultrasound training(60), which not only defines skills to acquire but actual training modalities, as well as measurable accuracy success and minimum number of procedures to complete. Such a consensus would also need to address the points raised earlier about the factors making the examination more difficult for novices, e.g. post-surgery or overweight patients.

Conclusion

The ULTRA-IBD study is the first study to analyze the inter-rater agreement on IUS findings in IBD patients between three operators of varying experience, with an expert radiologist as the gold standard. This study demonstrates the reliable results that can be achieved in IBD IUS by a gastroenterologist with a few hundred examinations of experience (“on-the-job” training) but with a formal general ultrasound training, while an IUS novice without any ultrasound experience only reached poor agreement on 50 patients in a real-life setting. Thus, while point-of-care IUS can be a valuable tool in any IBD Specialist’s toolbox and physicians willing to expand their practice should not censor themselves out of fear of not being expert radiologists, they should not either underestimate the learning curve and disregard proper ultrasound training. We also highlighted a few factors that pose significant challenge to an IUS novice, namely overweight patients, ileocolonic CD, and most importantly history of surgery. This, along with the literature, hints at the main difficulty of IUS learning being the technical performance of the IUS rather than the interpretation of its images, a consideration that should be kept in mind when developing training guidelines or supervising new trainees.

References

1. Alatab S, Sepanlou SG, Ikuta K, Vahedi H, Bisignano C, Safiri S, et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Gastroenterology & Hepatology*. 2020 Jan 1;5(1):17–30.
2. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet*. 2017 Dec 23;390(10114):2769–78.
3. Windsor JW, Kaplan GG. Evolving Epidemiology of IBD. *Curr Gastroenterol Rep*. 2019 Jul 23;21(8):40.
4. Buie MJ, Quan J, Windsor JW, Coward S, Hansen TM, King JA, et al. Global Hospitalization Trends for Crohn’s Disease and Ulcerative Colitis in the 21st Century: A Systematic Review With Temporal Analyses. *Clin Gastroenterol Hepatol*. 2023 Aug;21(9):2211–21.
5. Watermeyer G, Awuku Y, Fredericks E, Epstein D, Setshedi M, Devani S, et al. Challenges in the management of inflammatory bowel disease in sub-Saharan Africa. *Lancet Gastroenterol Hepatol*. 2022 Oct;7(10):962–72.
6. Williet N, Sarter H, Gower-Rousseau C, Adrianjafy C, Olympie A, Buisson A, et al. Patient-reported Outcomes in a French Nationwide Survey of Inflammatory Bowel Disease Patients. *J Crohns Colitis*. 2017 Feb;11(2):165–74.
7. Le Berre C, Ananthakrishnan AN, Danese S, Singh S, Peyrin-Biroulet L. Ulcerative Colitis and Crohn’s Disease Have Similar Burden and Goals for Treatment. *Clin Gastroenterol Hepatol*. 2020 Jan;18(1):14–23.

8. Netjes JE, Rijken M. Labor participation among patients with inflammatory bowel disease. *Inflamm Bowel Dis*. 2013 Jan;19(1):81–91.
9. Lönnfors S, Vermeire S, Greco M, Hommes D, Bell C, Avedano L. IBD and health-related quality of life -- discovering the true impact. *J Crohns Colitis*. 2014 Oct;8(10):1281–6.
10. Zhao M, Gönczi L, Lakatos PL, Burisch J. The Burden of Inflammatory Bowel Disease in Europe in 2020. *J Crohns Colitis*. 2021 Sep 25;15(9):1573–87.
11. Beard JA, Click BH. The burden of cost in inflammatory bowel disease: a medical economic perspective. *Curr Opin Gastroenterol*. 2020 Jul;36(4):310–6.
12. Burisch J, Vardi H, Schwartz D, Friger M, Kiudelis G, Kupčinskas J, et al. Health-care costs of inflammatory bowel disease in a pan-European, community-based, inception cohort during 5 years of follow-up: a population-based study. *Lancet Gastroenterol Hepatol*. 2020 May;5(5):454–64.
13. Singh S, Qian AS, Nguyen NH, Ho SKM, Luo J, Jairath V, et al. Trends in U.S. Health Care Spending on Inflammatory Bowel Diseases, 1996-2016. *Inflamm Bowel Dis*. 2022 Mar 2;28(3):364–72.
14. Bouguen G, Levesque BG, Feagan BG, Kavanaugh A, Peyrin-Biroulet L, Colombel JF, et al. Treat to target: a proposed new paradigm for the management of Crohn's disease. *Clin Gastroenterol Hepatol*. 2015 Jun;13(6):1042-1050.e2.
15. 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. *Inflamm Bowel Dis*. 2015 Apr;21(4):824–31.

16. Laterza L, Piscaglia AC, Minordi LM, Scoleri I, Larosa L, Poscia A, et al. Multiparametric Evaluation Predicts Different Mid-Term Outcomes in Crohn's Disease. *Dig Dis*. 2018;36(3):184–93.
17. Peyrin-Biroulet L, Sandborn W, Sands BE, Reinisch W, Bemelman W, Bryant RV, et al. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): Determining Therapeutic Goals for Treat-to-Target. *Am J Gastroenterol*. 2015 Sep;110(9):1324–38.
18. Colombel JF, Panaccione R, Bossuyt P, Lukas M, Baert F, Vaňásek T, et al. Effect of tight control management on Crohn's disease (CALM): a multicentre, randomised, controlled phase 3 trial. *Lancet*. 2017 Dec 23;390(10114):2779–89.
19. Ungaro RC, Yzet C, Bossuyt P, Baert FJ, Vanasek T, D'Haens GR, et al. Deep Remission at 1 Year Prevents Progression of Early Crohn's Disease. *Gastroenterology*. 2020 Jul;159(1):139–47.
20. Turner D, Ricciuto A, Lewis A, D'Amico F, Dhaliwal J, Griffiths AM, 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 Apr;160(5):1570–83.
21. Lafeuille P, Hordonneau C, Vignette J, Blayac L, Dapoigny M, Reymond M, et al. Transmural healing and MRI healing are associated with lower risk of bowel damage progression than endoscopic mucosal healing in Crohn's disease. *Aliment Pharmacol Ther*. 2021 Mar;53(5):577–86.
22. Geyl S, Guillo L, Laurent V, D'Amico F, Danese S, Peyrin-Biroulet L. Transmural healing as a therapeutic goal in Crohn's disease: a systematic review. *Lancet Gastroenterol Hepatol*. 2021 Aug;6(8):659–67.

23. Park J, Kang SJ, Yoon H, Park J, Oh HJ, Na HY, et al. Histologic Evaluation Using the Robarts Histopathology Index in Patients With Ulcerative Colitis in Deep Remission and the Association of Histologic Remission With Risk of Relapse. *Inflamm Bowel Dis*. 2022 Nov 2;28(11):1709–16.
24. Zenlea T, Yee EU, Rosenberg L, Boyle M, Nanda KS, Wolf JL, et al. Histology Grade Is Independently Associated With Relapse Risk in Patients With Ulcerative Colitis in Clinical Remission: A Prospective Study. *Am J Gastroenterol*. 2016 May;111(5):685–90.
25. Buisson A, Gonzalez F, Poullenot F, Nancey S, Sollellis E, Fumery M, et al. Comparative Acceptability and Perceived Clinical Utility of Monitoring Tools: A Nationwide Survey of Patients with Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*. 2017 Aug 1;23(8):1425–33.
26. Burisch J, Zhao M, Odes S, De Cruz P, Vermeire S, Bernstein CN, et al. The cost of inflammatory bowel disease in high-income settings: a *Lancet Gastroenterology & Hepatology* Commission. *The Lancet Gastroenterology & Hepatology*. 2023 May 1;8(5):458–92.
27. Taylor SA, Mallett S, Bhatnagar G, Morris S, Quinn L, Tomini F, et al. Magnetic resonance enterography compared with ultrasonography in newly diagnosed and relapsing Crohn's disease patients: the METRIC diagnostic accuracy study. *Health Technol Assess*. 2019 Aug;23(42):1–162.
28. Maaser C, Petersen F, Helwig U, Fischer I, Roessler A, Rath S, et al. Intestinal ultrasound for monitoring therapeutic response in patients with ulcerative colitis: results from the TRUST&UC study. *Gut*. 2020 Sep;69(9):1629–36.

29. de Voogd F, Bots S, Gecse K, Gilja OH, D'Haens G, Nylund K. Intestinal Ultrasound Early on in Treatment Follow-up Predicts Endoscopic Response to Anti-TNF α Treatment in Crohn's Disease. *Journal of Crohn's and Colitis*. 2022 Oct 1;16(10):1598–608.
30. Kucharzik T, Wittig BM, Helwig U, Börner N, Rössler A, Rath S, et al. Use of Intestinal Ultrasound to Monitor Crohn's Disease Activity. *Clin Gastroenterol Hepatol*. 2017 Apr;15(4):535-542.e2.
31. Allocca M, Furfaro F, Fiorino G, Peyrin-Biroulet L, Danese S. Point-of-Care Ultrasound in Inflammatory Bowel Disease. *J Crohns Colitis*. 2021 Jan 13;15(1):143–51.
32. Castiglione F, Mainenti PP, De Palma GD, Testa A, Bucci L, Pesce G, et al. Noninvasive diagnosis of small bowel Crohn's disease: direct comparison of bowel sonography and magnetic resonance enterography. *Inflamm Bowel Dis*. 2013 Apr;19(5):991–8.
33. Rispo A, de Sire R, Mainenti PP, Imperatore N, Testa A, Maurea S, et al. David Against Goliath: Direct Comparison of Handheld Bowel Sonography and Magnetic Resonance Enterography for Diagnosis of Crohn's Disease. *Inflamm Bowel Dis*. 2023 Apr 3;29(4):563–9.
34. Brodersen JB, Jensen MD, Juel MA, Kjeldsen J, Knudsen T, Rafaelsen SR. Intestinal ultrasound in patients with suspected Crohn's disease - results of a prospective evaluation by trainees. *Scand J Gastroenterol*. 2023;58(12):1405–11.
35. Wright EK, Wang I, Wong D, Bell SJ, Connell WR, Thompson AJ, et al. Accuracy of point-of-care intestinal ultrasound for Crohn's disease. *Australas J Ultrasound Med*. 2020 Jul 24;23(3):176–82.
36. Xu C, Li L, Zhang Y, Wang R, Zhang H. Diagnostic accuracy of different cross-sectional imaging techniques for disease location and activity in Crohn's disease and external

- validation and comparison of MARIAs and IBUS-SAS. *Abdom Radiol.* 2023 Mar 1;48(3):821–32.
37. Rispo A, Imperatore N, Testa A, Nardone OM, Luglio G, Caporaso N, et al. Diagnostic Accuracy of Ultrasonography in the Detection of Postsurgical Recurrence in Crohn's Disease: A Systematic Review with Meta-analysis. *Inflamm Bowel Dis.* 2018 Apr 23;24(5):977–88.
38. de Voogd FA, Bots SJ, van Wassenae EA, de Jong M, Pruijt MJ, D'Haens GR, et al. Early Intestinal Ultrasound Predicts Clinical and Endoscopic Treatment Response and Demonstrates Drug-Specific Kinetics in Moderate-to-Severe Ulcerative Colitis. *Inflamm Bowel Dis.* 2023 Nov 27;izad274.
39. Allocca M, Dell'Avalle C, Furfaro F, Zilli A, D'Amico F, Peyrin-Biroulet L, et al. Early Intestinal Ultrasound Predicts Long-Term Endoscopic Response to Biologics in Ulcerative Colitis. *Journal of Crohn's and Colitis.* 2023 Oct 1;17(10):1579–86.
40. Ilvemark JFKF, Wilkens R, Thielsen P, Dige A, Boysen T, Brynskov J, et al. Early Intestinal Ultrasound Predicts Intravenous Corticosteroid Response in Hospitalised Patients With Severe Ulcerative Colitis. *Journal of Crohn's and Colitis.* 2022 Nov 1;16(11):1725–34.
41. Kucharzik T, Wilkens R, D'Agostino MA, Maconi G, Le Bars M, Lahaye M, et al. Early Ultrasound Response and Progressive Transmural Remission After Treatment With Ustekinumab in Crohn's Disease. *Clin Gastroenterol Hepatol.* 2023 Jan;21(1):153-163.e12.
42. Maeda M, Sagami S, Tashima M, Yamana Y, Karashima R, Miyatani Y, et al. Milan Ultrasound Criteria Predict Relapse of Ulcerative Colitis in Remission. *Inflamm Intest Dis.* 2023 Jul 26;8(3):95–104.

43. Fraquelli M, Sarno A, Girelli C, Laudi C, Buscarini E, Villa C, et al. Reproducibility of bowel ultrasonography in the evaluation of Crohn's disease. *Dig Liver Dis.* 2008 Nov;40(11):860–6.
44. Calabrese E, Kucharzik T, Maaser C, Maconi G, Strobel D, Wilson SR, et al. Real-time Interobserver Agreement in Bowel Ultrasonography for Diagnostic Assessment in Patients With Crohn's Disease: An International Multicenter Study. *Inflamm Bowel Dis.* 2018 Aug 16;24(9):2001–6.
45. Smith RL, Taylor KM, Friedman AB, Su HY, Con D, Gibson PR. Interrater reliability of the assessment of disease activity by gastrointestinal ultrasound in a prospective cohort of patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol.* 2021 Oct 1;33(10):1280–7.
46. Bezzio C, Saibeni S, Venero M, Furfaro F, Monteleone M, Ribaldone D, et al. The learning curve for using intestinal ultrasonography. *Digestive and Liver Disease* [Internet]. 2024 Feb 5 [cited 2024 May 19];0(0). Available from: [https://www.dldjournalonline.com/article/S1590-8658\(24\)00222-6/abstract](https://www.dldjournalonline.com/article/S1590-8658(24)00222-6/abstract)
47. Kim J, Kim K, Kim J, Yoo J, Jeong W, Cho S, et al. The learning curve in diagnosing acute appendicitis with emergency sonography among novice emergency medicine residents. *J Clin Ultrasound.* 2018 Jun;46(5):305–10.
48. Seyedhosseini J, Nasrelari A, Mohammadrezaei N, Karimialavijeh E. Inter-rater agreement between trained emergency medicine residents and radiologists in the examination of gallbladder and common bile duct by ultrasonography. *Emerg Radiol.* 2017 Apr;24(2):171–6.
49. Limberg B. [Diagnosis of chronic inflammatory bowel disease by ultrasonography]. *Z Gastroenterol.* 1999 Jun;37(6):495–508.

50. Novak KL, Nylund K, Maaser C, Petersen F, Kucharzik T, Lu 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. *J Crohns Colitis*. 2021 Apr 6;15(4):609–16.
51. Allocca M, Fiorino G, Bonovas S, Furfaro F, Gilardi D, Argollo M, et al. Accuracy of Humanitas Ultrasound Criteria in Assessing Disease Activity and Severity in Ulcerative Colitis: A Prospective Study. *J Crohns Colitis*. 2018 Nov 28;12(12):1385–91.
52. Dragoni G, Gottin M, Innocenti T, Lynch EN, Bagnoli S, Macrì G, et al. Correlation of Ultrasound Scores with Endoscopic Activity in Crohn's Disease: A Prospective Exploratory Study. *J Crohns Colitis*. 2023 Oct 20;17(9):1387–94.
53. Allocca M, Filippi E, Costantino A, Bonovas S, Fiorino G, Furfaro F, et al. Milan ultrasound criteria are accurate in assessing disease activity in ulcerative colitis: external validation. *United European Gastroenterol J*. 2021 May;9(4):438–42.
54. Cicchetti DV. The precision of reliability and validity estimates re-visited: distinguishing between clinical and statistical significance of sample size requirements. *J Clin Exp Neuropsychol*. 2001 Oct;23(5):695–700.
55. De Voogd F, Wilkens R, Geese K, Allocca M, Novak K, Lu C, et al. A Reliability Study: Strong Inter-Observer Agreement of an Expert Panel for Intestinal Ultrasound in Ulcerative Colitis. *J Crohns Colitis*. 2021 Aug 2;15(8):1284–90.
56. Bhatnagar G, Quinn L, Higginson A, Plumb A, Halligan S, Tolan D, et al. Observer agreement for small bowel ultrasound in Crohn's disease: results from the METRIC trial. *Abdom Radiol*. 2020 Oct 1;45(10):3036–45.

57. Allocca M, Dal Buono A, D'Alessio S, Spaggiari P, Garlatti V, Spinelli A, et al. Relationships Between Intestinal Ultrasound Parameters and Histopathologic Findings in a Prospective Cohort of Patients With Crohn's Disease Undergoing Surgery. *J Ultrasound Med.* 2023 Feb 6;
58. Nylund K, Maconi G, Hollerweger A, Ripolles T, Pallotta N, Higginson A, et al. EFSUMB Recommendations and Guidelines for Gastrointestinal Ultrasound. *Ultraschall Med.* 2017 Jun;38(3):e1–15.
59. Madsen GR, Wilkens R, Boysen T, Burisch J, Bryant R, Carter D, et al. The knowledge and skills needed to perform intestinal ultrasound for inflammatory bowel diseases-an international Delphi consensus survey. *Aliment Pharmacol Ther.* 2022 Jul;56(2):263–70.
60. Badaoui A, Teles De Campos S, Fusaroli P, Gincul R, Kahaleh M, Poley JW, et al. Curriculum for diagnostic endoscopic ultrasound training in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy.* 2024 Mar;56(03):222–40.

Titre de Thèse : Comprendre la courbe d'apprentissage de l'échographie intestinale dans les MICI : Une étude comparative entre novice, gastro-entérologue entraînée et radiologue expert, l'étude ULTRA-IBD

RESUME

Contexte : L'échographie intestinale (EI) semble un outil précieux pour le suivi des maladies inflammatoires chroniques de l'intestin (MICI), mais est opérateur-dépendant. Cette étude visait à évaluer la courbe d'apprentissage de l'EI chez un novice en matière d'échographie et une gastro-entérologue ayant reçu une formation générale en échographie et ayant une expérience régulière de l'EI, en utilisant un radiologue expert comme *gold-standard*.

Méthodes : Une étude prospective monocentrique a été menée d'avril 2023 à janvier 2024, portant sur 50 patients consécutifs atteints de MICI soumis à trois EI successives, chaque opérateur évaluant l'activité de la maladie à l'aide d'échelles quantitatives et de scores d'activité échographiques basés sur les éléments de base de l'EI (épaisseur et stratification de la paroi intestinale, signal Doppler, graisse mésentérique). La concordance entre les opérateurs a été calculée à l'aide du coefficient de corrélation intraclasse.

Résultats : La gastroentérologue expérimentée avait une bonne concordance avec le radiologue tout au long de l'étude, tandis que le novice avait une mauvaise concordance, sans courbe d'apprentissage. Le surpoids, la localisation iléo-colique de la MC et des antécédents de chirurgie de MICI ont significativement altéré les performances du novice. L'exclusion des patients ayant des antécédents chirurgicaux a permis l'obtention d'une véritable courbe d'apprentissage pour le novice, avec une concordance modérée à bonne à la fin de l'étude.

Conclusion : Une formation en échographie générale et quelques centaines EI d'expérience permettent un bon accord avec un radiologue expert chez les patients atteints de MICI. Les novices ont besoin d'une formation complète et de plus de 50 examens d'expérience, bien qu'une forte courbe d'apprentissage soit observée avec moins de 40 examens chez des patients qui n'ont jamais été opérés. Ces résultats ont un intérêt pour la formation, en soulignant la difficulté des performances techniques des MICI plutôt que de l'interprétation de leurs images.

MOTS-CLES

Maladie de Crohn, rectocolite hémorragique, IBUS-SAS, score MUC, surveillance

Titre de Thèse : Understanding the Learning Curve of Intestinal Ultrasound in IBD: A Comparative Study Between Novice, Regular and Expert, the ULTRA-IBD study

RESUME

- *Background:* Intestinal ultrasound (IUS) has emerged as a valuable tool for monitoring inflammatory bowel disease (IBD), but its reliability depends on the operator. This study aimed to assess the learning curve of IUS of a complete ultrasound novice and a gastroenterologist with formal general ultrasound training and regular IUS experience, using an expert radiologist as the gold standard.

Methods: A single-center, prospective study was conducted from April 2023 to January 2024, involving 50 consecutive IBD patients undergoing three successive IUS examinations by the three different operators, who assessed disease activity using quantitative and semi-quantitative scales, as well as IUS disease activity scores based on basic IUS parameters (bowel wall thickness and stratification, Doppler signal, mesenteric fat). Inter-rater agreement was calculated using the intraclass correlation coefficient.

Results: The experienced gastroenterologist demonstrated consistently good agreement with the radiologist throughout the study, while the novice showed poor agreement without a clear learning curve. A body mass index >25 kg/m², the ileocolonic location of CD, and a history of IBD surgery significantly altered the novice's performance. In particular, excluding patients with a history of surgery highlighted a true learning curve for the novice, with moderate-to-good agreement at the end of the study.

MOTS-CLES

Crohn's disease, ulcerative colitis, IBUS-SAS, MUC score, monitoring.