Diagnostic Accuracy of Transabdominal Ultrasound in Detecting Intestinal Inflammation in Paediatric IBD Patients-a Systematic Review

van Wassenaer, Elsa A.; de Voogd, Floris A. E.; van Rijn, Rick R.; van der Lee, Johanna H.; Tabbers, Merit M.; van Etten-Jamaludin, Faridi S.; Gecse, Krisztina B.; Kindermann, Angelika; de Meij, Tim G. J.; D'haens, Geert R.; Benninga, Marc A.; Koot, Bart G. P.

Published in:
Journal of Crohn s & colitis

DOI:
10.1093/ecco-jcc/jjz085

Published: 01/01/2019

Citation for pulished version (APA):
van Wassenaer, E. A., de Voogd, F. A. E., van Rijn, R. R., van der Lee, J. H., Tabbers, M. M., van Etten-Jamaludin, F. S., Gecse, K. B., Kindermann, A., de Meij, T. G. J., D'haens, G. R., Benninga, M. A., & Koot, B. G. P. (2019). Diagnostic Accuracy of Transabdominal Ultrasound in Detecting Intestinal Inflammation in Paediatric IBD Patients-a Systematic Review. *Journal of Crohn s & colitis*, *13*(12), 1501-1509. https://doi.org/10.1093/ecco-jcc/jjz085

Journal of Crohn's and Colitis, 2019, 1501–1509
doi:10.1093/ecco-jcc/jjz085
Advance Access publication April 23, 2019
Original Article

Original Article

Diagnostic Accuracy of Transabdominal Ultrasound in Detecting Intestinal Inflammation in Paediatric IBD Patients—a Systematic Review

Elsa A. van Wassenaer,[a] Floris A. E. de Voogd,[b] Rick R. van Rijn,[c] Johanna H. van der Lee,[d] Merit M. Tabbers,[a] Faridi S. van Etten-Jamaludin,[e] Krisztina B. Gecse,[b] Angelika Kindermann,[a] Tim G. J. de Meij,[f] Geert R. D'Haens,[b] Marc A. Benninga,[a] Bart G. P. Koot[a]

[a]Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Pediatric Gastroenterology, Amsterdam, The Netherlands [b]Amsterdam UMC, University of Amsterdam, Gastroenterology and Hepatology, Amsterdam, The Netherlands [c]Amsterdam UMC, University of Amsterdam, Radiology, Amsterdam, The Netherlands [d]Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Pediatric Clinical Research Office, Amsterdam, The Netherlands [e]Amsterdam UMC, University of Amsterdam, Medical Library, Amsterdam, The Netherlands [f]Emma Children's Hospital, Amsterdam UMC, Vrije Universiteit Amsterdam, Pediatric Gastroenterology, Amsterdam, The Netherlands

Corresponding author: Elsa A. van Wassenaer, MD, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Pediatric Gastroenterology, Amsterdam, The Netherlands. Email: e.a.vanwassenaer@amc.nl

Abstract

Background and Aims: Currently used non-invasive tools for monitoring children with inflammatory bowel disease [IBD], such as faecal calprotectin, do not accurately reflect the degree of intestinal inflammation and do not provide information on disease location. Ultrasound [US] might be of added value. This systematic review aimed to assess the diagnostic test accuracy of transabdominal US in detecting intestinal inflammation in children with IBD in both diagnostic and follow-up settings.

Methods: We systematically searched PubMed, Embase [Ovid], Cochrane Library, and CINAHL [EBSCO] databases for studies assessing diagnostic accuracy of transabdominal US for detection of intestinal inflammation in patients diagnosed or suspected of IBD, aged 0–18 years, with ileo-colonoscopy and/or magnetic resonance enterography [MRE] as reference standards. Studies using US contrast were excluded. Risk of bias was assessed with QUADAS-2.

Results: The search yielded 276 records of which 14 were included. No meta-analysis was performed, because of heterogeneity in study design and methodological quality. Only four studies gave a clear description of their definition for an abnormal US result. The sensitivity and specificity of US ranged from 39-93% and 90–100% for diagnosing *de novo* IBD, and 48–93% and 83–93% for detecting active disease during follow-up, respectively.

Conclusions: The diagnostic accuracy of US in detecting intestinal inflammation as seen on MRE and/or ileo-colonoscopy in paediatric IBD patients remains inconclusive, and there is currently no consensus on defining an US result as abnormal. Prospective studies with adequate sample size and methodology are needed before US can be used in the diagnostics and monitoring of paediatric IBD.

Key Words: Ultrasound; monitoring; diagnostic accuracy; paediatric inflammatory bowel disease

1. Introduction

Children with inflammatory bowel disease [IBD] need to be monitored regularly in order to detect disease activity timely and to prevent complications, such as abscesses or stenoses. Gastrointestinal endoscopy and magnetic resonance enterography [MRE] are the gold standard diagnostic tests to detect disease activity and location, for large and small bowel resepectively.[1] However, both tests are invasive, expensive, and time-consuming, and it is therefore not feasible to perform them frequently in children. Hence, paediatricians frequently rely on non-invasive tools, such as clinical disease activity scales and biomarkers, such as C-reactive protein [CRP] and faecal calprotectin. However, these tools have limited accuracy[2,3] and do not provide information on disease location, whereas disease location is relevant—both in the diagnostic setting and during monitoring. As a consequence, paediatricians often still rely on endoscopy and MRE to be informed on disease activity and disease location. Therefore, IBD patients would benefit from the implementation of additional accurate non-invasive monitoring tools.

Ultrasound [US] is a promising example of such a tool. It is safe, fast, and cheap, and does not require any anaesthesia, and it can be used both for monitoring children with IBD and for children suspected of IBD.[4] US examinations provide information on, for instance, bowel wall thickness [BWT] and vascularity of the bowel wall, two important measures of inflammation.[5] In adult IBD patients, US has already shown its usefulness to assess disease activity. Meta-analyses showed pooled sensitivities and specificities ranging 84–90% and 92–97%, respectively.[6–9] However, due to the difference in habitus and disease behaviour between adults and children, these data cannot be directly extrapolated to the paediatric IBD population. It is known that US correlates well with clinical disease activity scales and biomarkers in paediatric IBD,[10–12] but as mentioned above, these monitoring tools have their pitfalls. Knowledge about the diagnostic accuracy compared with an adequate reference standard is therefore crucial, but literature addressing this topic is limited. If US could improve non-invasive monitoring by accurately detecting disease location and severity, health care and quality of life of IBD patients could be improved; targeted treatment could be started quicker, potentially preventing complications; and the need for colonoscopies and MREs could be reduced.[13]

We performed a systematic review of currently published literature to answer the following question: 'What is the diagnostic accuracy of US in detecting IBD in suspected children and in detecting intestinal inflammation in paediatric IBD patients using MRE and/or ileo-colonoscopy as reference standard?' Secondary aims were to assess which items of the US examination were used for disease assessment, and to assess mean bowel wall thickness in paediatric IBD patients.

2. Methods

We conducted a search, with help of a clinical librarian [FE], in the PubMed, Embase [Ovid], Cochrane Library, and CINAHL [EBSCO] databases, looking for articles published from 1990 to July 2018. The search terms are available as Supplementary data at *ECCO-JCC* online. Inclusion criteria were: studies investigating the diagnostic accuracy of US in children [aged 0–18 years] with IBD or suspicion of IBD, using ileo-colonoscopy and/or MRE as the reference standard. Exclusion criteria were articles that did not report sensitivity/specificity or data to produce a 2 × 2 contingency table, and articles not published in English, Dutch, French, German, Spanish, or Italian. Furthermore, studies only using contrast enhanced ultrasonography [CEUS] or small intestine ultrasonography [SICUS] were excluded.

The titles and/or abstracts of the studies retrieved using the search strategy were screened independently by two reviewers [EW and FV] to identify studies that potentially met the inclusion criteria outlined above. The full texts of these potentially eligible studies were then retrieved and independently assessed for eligibility by the same two reviewers. Any disagreements were resolved through discussion with a third reviewer [BK].

A standardised, pre-piloted form was used to extract data and assess risk of bias of the included studies. The extracted information included: study setting; study population; participants' baseline characteristics; details of the ultrasound, MRE, and ileo-colonoscopy examinations; study design; recruitment and study completion rates; measures of accuracy; and patient satisfaction. We also noted the definition used for defining a US as positive, and references this choice was based on. In case a study did not report a specific definition, we noted the US items used [i.e., BWT]. Studies were subdivided into studies analysing accuracy at patient level [i.e., whether there is any disease activity or not] and studies analysing accuracy at segment level [i.e., whether there is disease activity in a specific segment]. Where needed, authors were contacted to retrieve additional information.

2.1. Risk of bias assessment

The risk of bias was assessed using the QUADAS-2 tool.[14] This is a generic tool developed specifically for use in diagnostic test accuracy reviews. It contains four domains [patient selection, index test, reference test, and flow and timing] with different signalling questions to identify study characteristics that could result in bias. We modified the QUADAS-2 tool, adding two signalling questions recommended in the Cochrane Handbook for Systematic Reviews on Diagnostic Test Accuracy [see Table 3][15]: 'Did test operators have appropriate training?' and 'Was treatment withheld between performance of the index- and reference test?'. We defined an acceptable interval between US and reference standard as a maximum of 7 days, and we defined an appropriate training as having finished the radiology specialisation or another ultrasonography education.

2.2. Meta-analysis

Criteria for inclusion in the meta-analysis were: 1] interval between US and reference standard of 7 days or less; 2] blinding of reference test [i.e., endoscopy/MRE] results during interpretation of US and blinding of US results during interpretation of reference standard; 3] appropriate training of operators; and 4] withholding of treatment in between US and reference standard examinations. For the meta-analysis, we planned to pool eligible studies in four separate groups: 1] diagnostic studies comparing US with ileo-colonoscopy; 2] diagnostic studies comparing US with MRE for small bowel disease; 3] follow-up studies comparing US with ileo-colonoscopy; and 4] follow-up studies comparing US with MRE for small bowel disease. In case a group contained two or more studies, we planned to calculate pooled sensitivity and specificity using Review Manager® version 5.0 and SAS [SAS Institute, Cary, NC] macro METADAS version 1.3, applying the bivariate method as recommended in the Cochrane Handbook for Systematic Reviews on Diagnostic Test Accuracy.[15]

We also planned to perform a meta-analysis of all the studies describing bowel wall thickness [BWT] measurements assessed with

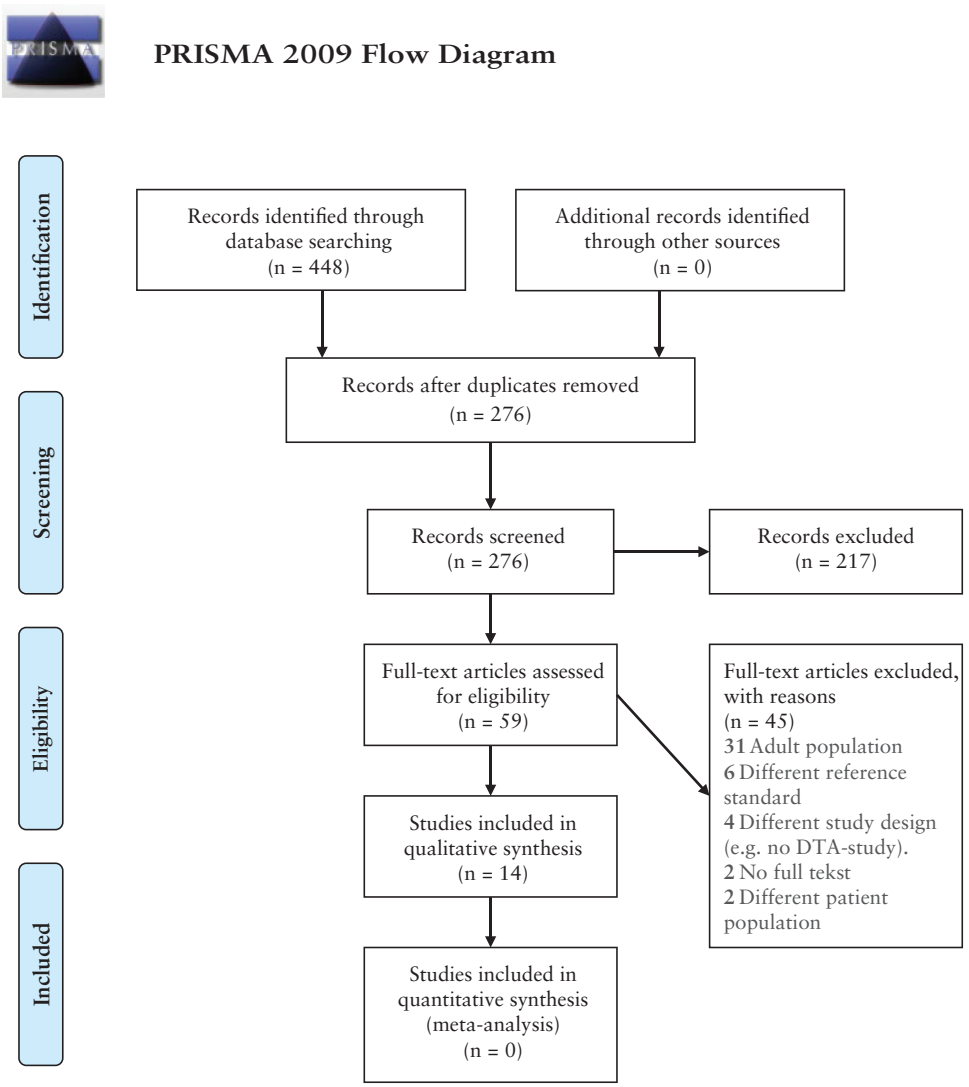


Figure 1. PRISMA flow diagram. DTA, diagnostic test accuracy.

US in different endoscopic disease activity categories [i.e., remission, mild, moderate, and severe disease activity]. The sample-size weighted pooled mean and pooled standard deviation [SD] scores per endoscopic disease activity category were calculated with Microsoft Excel® version 2016, and the differences between the categories were first assessed with analysis of variance [ANOVA] and subsequently with Student's t test for independent samples—corrected for multiple testing with Bonferroni—using GraphPad Prism® version 7; *p*-values of less than 0.017 were considered statistically significant [= 0.05/3].

3. Results

A total of 448 studies were found and, after exclusion of duplicates, 276 studies remained [see Figure 1]. After title and abstract screening, 59 full texts were screened for eligibility. Reasons for exclusion were: different study population [e.g., adult] [*n* = 33]; different reference standard [*n* = 4]; different study design [e.g., no accuracy study] [*n* = 4]; and unavailability of full text [*n* = 2].

Fourteen studies including a total of 424 patients were finally included in the review: 12 prospective and two retrospective studies. The number of included patients per study that underwent both US and reference standard examinations ranged from nine to 50. Eight studies compared US with endoscopy[16–23] and six studies US with MRE.[24–29] Characteristics of the included studies are depicted in Table 1, and the specifications of the US and MRE technique are available as Supplementary data at *ECCO-JCC* online. Two out of eight studies comparing US with endoscopy were performed in children with a clinical suspicion of IBD, four included both children with suspicion and those with known IBD, and two included only children with known IBD. From the studies comparing US with MRE, three included both children with suspicion and diagnosis of IBD, and the other three only included patients with known IBD. As depicted in Table 1, nine studies included both CD and UC patients, three studies included only CD patients, and one study included only UC patients.

3.1. Definition used for assessing US

Although 13 out of 14 studies did describe the US items they used for assessing the US examination, only four studies gave a clear description of their definition for an abnormal US result [see Table 2]. In all

Table 1. Characteristics of included studies.

First author *Country*	n* [m/f]	Mean age (range)	Aim (patient population)	Design	Cut-off value BWT	Items used for Disease Assessment	Definition of positive US + reference	Time-interval US-reference	Segments evaluated	Analysis at segment or patient level ****
Ileo-colonoscopy as reference test										
Alberini 2001 *France* 16	23 [17/11]	10 (2-15)	Diagnosis *(CD&UC)***	R	2 mm (colon), 2,5 mm (ileum)	BWT/ free fluid / lymph nodes	Increased BWT AND/ OR (free fluid AND lymph nodes) 23,37	Unspecified	Colon, ileum	Patient
Bothe 2006 *Norway* 17	15 [23/20]	12 (5-16)	Diagnosis *(CD&UC)*	P	Unspecified	BWT/ DS/ lymph nodes/ complications	Unspecified	<3 weeks	Terminal ileum	Segment
de Ridder 2012*** *Netherlands* 20	19 [14/6]	15 (11-18)	Diagnosis *((suspicion) CD)*	P	3 mm	BWT / SMA DS/ abscess	BWT AND/OR SMA DS (peak systole > 100 cm/s in fasting children)11	2-25 days (median 8)	Small bowel (duodenum, jejunum, ileum)	Patient
Ziech 2014 *Netherlands* 22	24 [15/13]	14 (10-17)	Diagnosis *(Suspicion CD & UC)*	P	3 mm	BWT/ WLS / lymph nodes / DS	Unspecified	1-40 days (median 7.5)	Terminal ileum, colon	Patient
Bremner 2006 *UK* 18	33	12 (4-17)	Follow up *((suspicion) CD & UC)*	P	3 mm	BWT + SMA DS	Unspecified	<10 days	Colon	Both
Civitelli 2014 *Italy* 19	50 [24/20]	13 (3-18)	Follow up *(known UC)*	P	3 mm	BWT, WLS, haustra visibility, DS	2 of the following: BWT, WLS, haustra visibility, DS 18,38–40	1 day	Large bowel	Both
Faure 1997 *France* 23	38	11 (4-18)	Follow up *((suspicion) CD & UC)*	P	2 mm (colon), 2,5 mm (ileum)	BWT & WLS	Unspecified	1 day	Terminal ileum, large bowel	Both
Haber 2002 *Germany* 21	41 [43/35]	11 (0-17)	Follow up *(Known CD & UC)***	P	1,5 mm (ileum), 2 mm (colon)	BWT & WLS	BWT 41	Unspecified	Terminal ileum, colon	Both
MRE as reference test										
Ahmad 2016 *Canada* 24	33	15 (7-18)	Follow up ((*suspicion) CD & UC)*	P	4 mm	BWT/ DS/ stricture/ fistula/ creeping fat / lymph nodes.	Unspecified	2 hours	Jejunum, ileum, terminal ileum, cecum, colon, sigmoid, rectum.	Both
Barber 2017 *UK* 27	49 [33/16]	4 (7-17)	Follow up *((suspicion) CD & UC)*	R	3 mm	BWT/ abnormal echogenicity/ DS/ creeping fat	Unspecified	IQR: 1–8 days	Jejunum, ileum, terminal ileum, colon.	Patient
Tsai 2017 *USA* 28	41 [22/19]	14 (5- 19)	Follow up *((suspicion) CD & UC)*	P	1.9 (rater1) / 1.8 (rater2) mm	BWT/ abnormal echogenicity/ DS/ creeping fat	Unspecified	<1 day	Terminal ileum	Segment
Dagia 2008 *Australia* 29	9	Unspecified	Follow up (*Known CD)*	P	Unspecified	BWT/ strictures/ mural abnormality/ creeping fat/ lymph nodes	Unspecified	0–54 weeks (median 7 weeks)	Unspecified	Patient
Dillman 2016 *USA* 26	29 [18/11]	15 (9-18)	Follow up *(Known CD)*	P	Unspecified	Unspecified	Unspecified	1 hour	Distal small bowel	Patient
Magnano 2003 *Italy* 25	20 [10/12]	15(8-18)	Follow up (*Known CD)*	P	Unspecified	BWT + DS	BWT + DS 42	<7 days	Unspecified	Patient

BWT, bowel wall thickness; CD, Crohn's disease; DS, Doppler signal; P, prospective; R, retrospective; SMA, superior mesenteric artery; UC, ulcerative colitis; US, ultrasound; WLS, wall layer stratification; MRE, magnetic resonance enterography; IBD, inflammatory bowel disease; IQR, interquartile range.

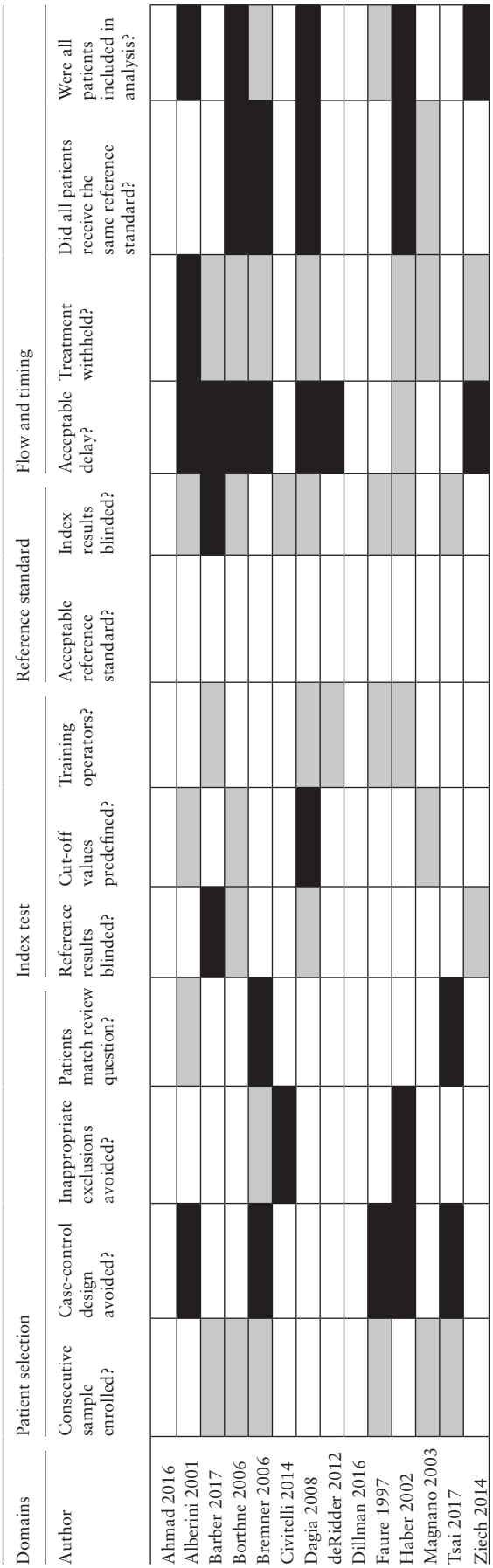
*Number of patients that underwent both US and reference standard, number of males/females in whole study.

**Also included some non-IBD patients [e.g., Behcet's disease, granulomatosis, and gastroenteritis].

*** Used single-balloon enteroscopy as reference standard.

****.See method section for explanation.

Table 2. Risk of bias assessment using the QUADAS-2 tool.



White, low risk of bias; gray, unclear risk of bias; black, high risk of bias.
Criteria for inclusion in meta-analysis are underlined.

studies except for one, increased BWT was used as one of the criteria for an abnormal US result. Cut-off values for BWT ranged from 1.5 mm to 4 mm for all ages. Two out of 14 studies used different cut-off values for ileum and colon.[16,23] Other reported US items were: increased vascularity measured with Doppler [*n* = 9]; presence of enlarged lymph nodes [*n* = 6]; absence of wall layer stratification or presence of creeping fat [*n* = 4]; and abscesses or other complications [*n* = 2]. One study in UC patients also used absence of visible colonic haustrations as a criterion.

3.2. Risk of bias assessment

The details of the risk of bias assessment according to the QUADAS-2 tool are shown in Table 3. There was a wide spread in time interval between US and the reference test, from 1 h up to 54 weeks. In eight studies, the interval between US and reference standard was longer than 7 days or undefined. Most studies [8/14] did not describe whether US results were blinded during the interpretation of the reference standard, and in eight out of 14 studies it was not specified whether treatment was withheld in between US and reference test. Only two studies fulfilled the predefined criteria for inclusion in the meta-analyses.[24,26] However, the two studies included different types of patients [suspicion of CD and UC versus known CD], and different bowel segments were included in the analysis, and thus no meta-analysis could be performed.

3.4. Accuracy of US in diagnosing IBD

3.4.1. Endoscopy

The results of the diagnostic accuracy at patient level are summarised in Table 3. Three studies assessed the accuracy of US in diagnosing IBD with ileo-colonoscopy as reference standard,[16,17,22] and one study used single-balloon enteroscopy as reference standard assessing the small bowel.[20] Point estimates of sensitivity of US in diagnosing endoscopically active IBD [UC and CD combined] ranged from 39% to 55% and specificity from 90% to 100%. One study assessed the accuracy of US for disease localisation in the terminal ileum at diagnosis of IBD; sensitivity was 93%. Sensitivity of US in diagnosing active small bowel CD was 54% and specificity 100% in the one study that used single-balloon enteroscopy.

3.4.2. MRE

No study used MRE as reference standard to assess accuracy of US in diagnosing IBD.

3.5. Accuracy of US in follow-up of IBD

3.5.1. Endoscopy

Ten studies assessed the accuracy of US in detecting disease activity or assessing disease extent in IBD for purpose of follow-up. Four of those used ileo-colonoscopy as reference standard and performed analysis both at patient level and at segment level. In these four studies, sensitivity and specificity at patient level, including patients with both CU and CD, ranged from 48% to 93% and from 83% to 93%, respectively. The results for analysing each segment separately are depicted in Table 4. In nearly all segments, the specificity was found to be higher than the sensitivity, with a 4% to 46% difference.

One study used a self-designed US activity index to detect colonic inflammation.[19] This study included only patients with UC and a suspected disease flare-up. The overall sensitivity and specificity for detection of severe disease were excellent [100% and 93%, respectively], the sensitivity in all different segments for detecting colonic

Table 3. Diagnostic accuracy of US compared with ileo-colonoscopy or MRE at patient level.

Author	n	Aim (patient population)	Sensitivity (95% CI)**	Specificity (95% CI)**
Ileo-colonoscopy as reference test				
Alberini 2001	23	Diagnosis *(CD&UC)*	39% (20-61)*	90% (68-99)*
de Ridder 2012	19	Diagnosis *((suspicion) CD)*	54% (25-81)	100% (54-100)
Ziech 2014	24	Diagnosis *(suspicion CD & UC)*	55%	100%
Bremner 2006	33	Follow up *((suspicion) CD & UC)*	48%	93%
Civitelli 2014	50	Follow up *(known UC)*	100%	93%
Faure 1997	38	Follow up *((suspicion) CD & UC)*	88%	93%
Haber 2002	41	Follow up *(known CD & UC)*	77%	83%
MRE as reference test				
Ahmad 2016	33	Follow up ((*suspicion) CD & UC)*	64%* (45-80)*	Not available
Barber 2017	49	Follow up *((suspicion) CD & UC)*	81% (70-89)	95% (92-97)
Dagia 2008	9	Follow up (*known CD)*	71%* (29-96)*	100%* (16-100)*
Magnano 2003	20	Follow up (*known CD)*	93%* (68-100)*	100%* (48-100)*

CD, Crohn's disease; UC, ulcerative colitis; MRE, magnetic resonance enterography; US, ultrasound; CI, confidence interval.
*Calculated from raw data presented in article.
**95% CI only calculated where possible

Table 4. Diagnostic accuracy of US compared with ileo-colonoscopy or MRE at segment level sorted on reference standard and study aim.

Author	n	Aim *(patient population)*	Segment	Sensitivity (95% CI)**	Specificity (95% CI)**
Ileo-colonoscopy as reference test					
Borthne 2006	15	Diagnosis *(CD&UC)*	Term. ileum	93% (70–99)	-
Bremner 2006	33	Follow up *((suspicion) CD & UC)*	Cecum	33%	100%
			Asc. colon	46%	88%
			Tran. colon	67%	90%
			Desc. colon	54%	100%
			Sigmoid	50%	88%
Civitelli 2014	50	Follow up *(known UC)*	Asc. colon	75% (42-93)	100% (74-100)
			Tran. colon	86% (60-97)	100% (70-100)
			Desc. colon	96% (80-100)	100% (62-97)
Faure 1997	38	Follow up *((suspicion) CD & UC)*	Term. ileum	100%	92%
			Asc. colon	88%	92%
			Tran. colon	80%	90%
			Desc. colon	93%	100%
			Rectum	89%	-
Haber 2002	41	Follow up *(Known CD & UC)*	Ter. ileum	100%	72%
			Asc. colon	72%	81%
			Tran. colon	74%	94%
			Desc. colon	74%	89%
MRE as reference test					
Dillman 2016	29	Follow up *(Known CD)*	Distal small Bowel	83%*	71%*
Tsai 2017	41	Follow up *((suspicion) CD & UC)*	Term. ileum	67% (rater1)	78% (rater1)
				83% (rater2)	78% (rater2)

Asc., ascending; CD, Crohn's disease; desc., descending; MRE, magnetic resonance enterography; ran., transverse; term., terminal; UC, ulcerative colitis; US, ultrasound.
*Calculated from raw data presented in article.
**95% CI only calculated where possible.

lesions ranged from 75–96%, and the specificity was 100% in each part of the colon.

3.5.2. MRE

A total of six studies used MRE as reference standard for assessing the accuracy of US in detecting disease activity or assessing disease extent in IBD for monitoring patients [Tables 3 and 4]. Four studies evaluated small bowel only, and in two studies the colon was also included in the analysis. As these two studies analysed their results at patient level, we could not report small bowel results only. Two out of six studies included patients with CD and patients with UC[24,27,28]; the sensitivity at patient level ranged from 64% to 81%, and specificity was 95% in one study[27] but could not be calculated in the other study.[24] Two studies only included CD patients[25,29]; sensitivity ranged from 71% to 93% and specificity was 100% in both studies. In two studies, the analysis was segment-based[26,28] [see Table 4].

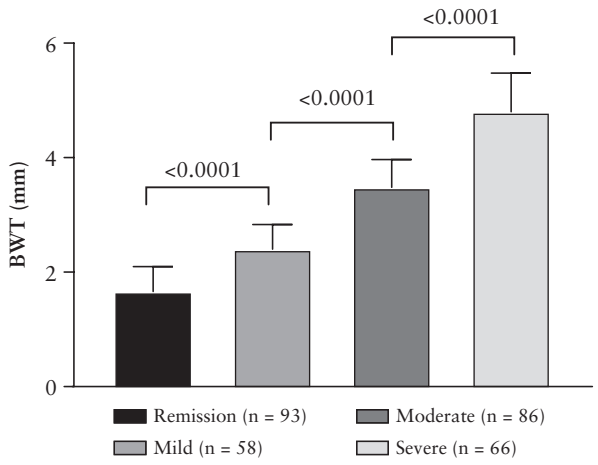


Figure 2. Mean (+standard deviation [SD]) colonic bowel wall thickness [BWT] as assessed with US in different categories of endoscopic disease activity. Differences are tested with analysis of variance [ANOVA] and subsequently Student's t test, corrected for multiple testing with Bonferroni.

3.6. Bowel wall thickness

Three studies reported the differences in BWT between different categories of disease severity.[18,19,21] All studies described a statistically significant increase in BWT in moderately to severely inflamed segments. The sample-size weighted pooled means and pooled SD scores of the two studies that described BWT as measured by US in different categories of endoscopic disease activity in colon are displayed in Figure 2 [18,19]. Mean BWT [and SD] in patients in remission or with mild, moderate, or severe endoscopic disease activity, respectively, was 1.7 mm [0.4], 2.4 [0.4] mm, 3.5 [0.5] mm, and 4.8 [0.7] mm. The other study used clinical disease activity as reference standard and was thus not included in this figure.

4. Discussion

This systematic review shows that the diagnostic accuracy of US in detecting intestinal inflammation as seen on MRE and/or ileocolonoscopy in paediatric IBD patients remains inconclusive. Most studies had important methodological limitations, such as an inefficient time flow and unclear blinding procedures, and the reported accuracy varied widely between studies. In addition, we establish that there is no generally accepted or applied definition of an abnormal US result.

The role of US in detection of inflammation in the adult population has been studied extensively.[5] In a meta-analysis of Fraquelli *et al.* studying the role of US in the detection of CD and including seven studies, a pooled sensitivity and specificity of 88% and 93% [95% CI not reported] were found.[7] Similar results were found in a subsequent systematic review of Dong *et al.* (sensitivity 88% [95% CI: 85–91%], specificity 97% [95% CI: 96–98%]) which included 15 studies and described the diagnostic accuracy of US in the diagnosis of CD and the evaluation of disease activity.[9] As in the adult studies, our results show higher specificity rates compared with sensitivity. This may be explained by the fact that superficial endoscopic lesions do not always cause intramural changes such as an increased bowel wall thickness, as was also pointed out by Magnano *et al.*[25]

However, we are unable to reliably interpret the diagnostic accuracy results reported in this systematic review because of the major methodological shortcomings of the included studies. Most of the studies included in this review did not meet the predefined criteria for inclusion in the meta-analysis, which raises questions about the validity of the presented results. The limited quality can be partly explained by the fact that we included two retrospective studies and two studies that were not designed to study diagnostic accuracy of US, but from which raw data could be derived. The primary aim of these studies was to assess the accuracy of MRE or scintigraphy.[16,25] The most common reasons for exclusion from the meta-analysis were an unacceptable time interval between the US and the reference standard, and lack of data on whether treatment was withheld during this interval. However, these two are very important, since BWT can already change significantly after 2 weeks of treatment.[30]

In addition, an important finding in this systematic review is that only few studies defined what was considered an abnormal US result, although most did describe the US features they took into account, such as BWT or increased vascularity. Only abnormal BWT was often defined; however, there was a wide spread in cut-off values for BWT, ranging from 1.5 to 4 mm. According to several studies assessing the bowel wall in healthy children using US, the mean colonic wall thickness is 1.2 mm and does not reach 2 mm.[31–33] As we show [Figure 2], the mean colonic wall thickness in paediatric IBD patients in remission is 1.6 mm [SD 0.4] and is 2.4 mm [SD 0.4] in patients with mild inflammation. This suggests that for this US feature, the optimal cut-off may be 2–3 mm. However, this presumption remains to be confirmed in future prospective studies. It is necessary to formulate standard criteria for an abnormal US result, and preferably also to develop a validated activity index, combining several US items in a score.[34] To date, only Civitelli *et al.* aimed to design such an index for children with UC, showing promising results.[19] However, this index has not been externally validated.

A limitation in this systematic review was that we could not distinguish between CD and UC in most studies, although both illnesses are characterised by different features.[35] Additionally we aimed to assess MRE solely as reference standard for small bowel, as endoscopy is still the gold standard for large bowel assessment.[1] However, this was not possible in every study, as some studies using MRE only reported analysis at patient level and not at segment level.[25,27,29] Another discussion point is our definition used for 'appropriate training of test operators', which we defined as having finished the radiology specialisation or another ultrasonography education. Training is relevant, because an important and frequently reported downside of US is its operator dependency.[36] The results of Barber *et al.* suggest the presence of a learning curve when assessing an intestinal US, as the accuracy of the first 10 examinations was lower in comparison with the last 10 examinations.[27] Although we cannot determine whether our definition used is strict enough, we do not think it biased the results presented in this systematic review, as we did not pool our data. It would be valuable to define quality criteria for performing intestinal US, both in the research setting and in clinical practice.

To our knowledge, this is the first systematic review summarising the literature on the diagnostic accuracy of US to assess intestinal inflammation in paediatric IBD. The strengths of this review are: that we solely included studies using a proper reference standard; the use of a validated tool for risk of bias assessment[14]; and the strict criteria for inclusion in the meta-analyses, to avoid pooling biased data. Although US seems a promising non-invasive diagnostic tool for children with IBD, this review shows that the accuracy of US in detecting intestinal inflammation in this group of patients has not been accurately established yet. More evidence is needed to accurately determine the diagnostic value of US in IBD. Future studies: should be specific on the definition of an abnormal US result; should

explore the use of indices combining several US features; should guarantee an efficient time flow between US and reference test; and should be strict in blinding procedures in order to prevent bias. Additionally, studies should include sufficient participants to reach a proper power.

4.1. Conclusion

The diagnostic accuracy of US in children with IBD has not been established. High-quality prospective studies of adequate sample size are needed before US can be reliably used for the purposes of paediatric IBD diagnosis and monitoring disease activity.

Funding

No specific funding has been received for this work.

Conflict of Interest

MAB reports fees as consultant or PI for Norgine, Takeda, Shire, Coloplast, Danone, Friesland Campina, Novalac, Sensus, Allergan, and Johnson and Johnson, outside the submitted work; KBG reports personal fees from Amgen, AbbVie, Biogen, Boehringer Ingelheim, Ferring, Hospira, Immunic Therapeutics, Janssen, MSD, Pfizer, Sandoz, Samsung Bioepis, Takeda, Tigenix, and Tillotts, outside the submitted work; other authors have no conflicts of interest to declare.

Author Contributions

Concept and design of the study: EW, RR, MT, BK. Acquisition of data: EW, FV, FE. Analysis and interpretation of data: EW, FV, JL, KG, AK, TM, GD, MA, BK. Drafting the article or revising it critically for important intellectual content: all authors. Final approval of the version to be submitted: all authors.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Panes J, Bouhnik Y, Reinisch W, *et al*. Imaging techniques for assessment of inflammatory bowel disease: joint ECCO and ESGAR evidence-based consensus guidelines. *J Crohns Colitis* 2013;**7**:556–85.
2. Lahiff C, Safaie P, Awais A, *et al*. The Crohn's disease activity index [CDAI] is similarly elevated in patients with Crohn's disease and in patients with irritable bowel syndrome. *Aliment Pharmacol Ther.* 2013;**37**:786–94.
3. Henderson P, Anderson NH, Wilson DC. The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease: a systematic review and meta-analysis. *Am J Gastroenterol* 2014;**109**:637–45.
4. Schreiber-Dietrich D, Chiorean L, Cui XW, *et al*. Particularities of Crohn's disease in pediatric patients: current status and perspectives regarding imaging modalities. *Expert Rev Gastroenterol Hepatol* 2015;**9**:1313–25.
5. Calabrese E, Maaser C, Zorzi F, *et al*. Bowel ultrasonography in the management of Crohn's disease. A review with recommendations of an international panel of experts. *Inflamm Bowel Dis* 2016;**22**:1168–83.
6. Panés J, Bouzas R, Chaparro M, *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. *Aliment Pharmacol Ther* 2011;**34**:125–45.
7. Fraquelli M, Colli A, Casazza G, *et al*. Role of US in detection of Crohn disease: meta-analysis. *Radiology* 2005;**236**:95–101.
8. 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**:64–79.
9. Dong J, Wang H, Zhao J, *et al*. Ultrasound as a diagnostic tool in detecting active Crohn's disease: a meta-analysis of prospective studies. *Eur Radiol* 2014;**24**:26–33.
10. Dillman JR, Smith EA, Sanchez RJ, *et al*. Pediatric small bowel Crohn disease: correlation of US and MR enterography. *Radiographics* 2015;**35**:835–48.
11. Scholbach T, Herrero I, Scholbach J. Dynamic color Doppler sonography of intestinal wall in patients with Crohn disease compared with healthy subjects. *J Pediatr Gastroenterol Nutr* 2004;**39**:524–8.
12. Thomson M, Rao P, Berger L, Rawat D. Graded compression and power Doppler ultrasonography versus endoscopy to assess paediatric Crohn disease activity pre- and posttreatment. *J Pediatr Gastroenterol Nutr* 2012;**54**:404–8.
13. Bremner AR, Pridgeon J, Fairhurst J, Beattie RM. Ultrasound scanning may reduce the need for barium radiology in the assessment of small-bowel Crohn's disease. *Acta Paediatr* 2004;**93**:479–81.
14. Whiting PF, Rutjes AW, Westwood ME, *et al*.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36.
15. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMGDJ. In: Deeks J, editor. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. London:* The Cochrane Collaboration; 2009.
16. Alberini JL, Badran A, Freneaux E, *et al*. Technetium-99m HMPAO-labeled leukocyte imaging compared with endoscopy, ultrasonography, and contrast radiology in children with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 2001;**32**:278–86.
17. Borthne AS, Abdelnoor M, Rugtveit J, Perminow G, Reiseter T, Kløw NE. Bowel magnetic resonance imaging of pediatric patients with oral mannitol MRI compared with endoscopy and intestinal ultrasound. *Eur Radiol* 2006;**16**:207–14.
18. Bremner AR, Griffiths M, Argent JD, Fairhurst JJ, Beattie RM. Sonographic evaluation of inflammatory bowel disease : a prospective, blinded, comparative study. *Pediatr Radiol* 2006;**36**:947–53. doi:10.1007/s00247-006-0245-8
19. Civitelli F, Di Nardo G, Oliva S, *et al*. Ultrasonography of the colon in pediatric ulcerative colitis: a prospective, blind, comparative study with colonoscopy. *J Pediatr* 2014;**165**:78–84.e2.
20. de Ridder L, Mensink PB, Lequin MH, *et al*. Single-balloon enteroscopy, magnetic resonance enterography, and abdominal US useful for evaluation of small-bowel disease in children with [suspected] Crohn's disease. *Gastrointest Endosc* 2012;**75**:87–94.
21. Haber HP, Busch A, Ziebach R, Dette S, Ruck P, Stern M. Ultrasonographic findings correspond to clinical, endoscopic, and histologic findings in inflammatory bowel disease and other enterocolitides. *J Ultrasound Med* 2002;**21**:375–82.
22. Ziech ML, Hummel TZ, Smets AM, *et al*. Accuracy of abdominal ultrasound and MRI for detection of Crohn disease and ulcerative colitis in children. *Pediatr Radiol* 2014;**44**:1370–8.
23. Faure C, Belarbi N, Mougenot JFF, *et al*. Ultrasonographic assessment of inflammatory bowel disease in children: comparison with ileocolonoscopy. *J Pediatr* 1997;**130**:147–51.
24. Ahmad TM, Greer ML, Walters TD, Navarro OM. Bowel Sonography and MR Enterography in Children. *AJR Am J Roentgenol* 2016;**206**:173–81.
25. Magnano G, Granata C, Barabino A, *et al*. Polyethylene glycol and contrast-enhanced MRI of Crohn's disease in children: preliminary experience. *Pediatr Radiol* 2003;**33**:385–91.
26. Dillman JR, Smith EA, Sanchez R, *et al*. Prospective cohort study of ultrasound-ultrasound and ultrasound-MR enterography agreement in the evaluation of pediatric small bowel Crohn disease. *Pediatr Radiol* 2016;**46**:490–7.
27. Barber JL, Maclachlan J, Planche K, *et al*. There is good agreement between MR enterography and bowel ultrasound with regards to disease location and activity in paediatric inflammatory bowel disease. *Clin Radiol* 2017;**72**:590–7.

28. Tsai TL, Marine MB, Wanner MR, *et al*. Can ultrasound be used as the primary imaging in children with suspected Crohn disease? *Pediatr Radiol* 2017;**47**:917–23.
29. Dagia C, Ditchfield M, Kean M, Catto-Smith T. Imaging for Crohn disease: use of 3-T MRI in a paediatric setting. *J Med Imaging Radiat Oncol* 2008;**52**:480–8.
30. Dillman JR, Dehkordy SF, Smith EA, *et al*. Defining the ultrasound longitudinal natural history of newly diagnosed pediatric small bowel Crohn disease treated with infliximab and infliximab-azathioprine combination therapy. *Pediatr Radiol* 2017;**47**:924–34.
31. Haber HP, Stern M. Intestinal ultrasonography in children and young adults : bowel wall thickness is age dependent. *J Ultrasound Med* 2000;**19**:315–21.
32. Ramsden WH, Moya EF, Littlewood JM. Colonic wall thickness, pancreatic enzyme dose and type of preparation in cystic fibrosis. *Arch Dis Child* 1998;**79**:339–43.
33. Pohl M, Krackhardt B, Posselt HG, Lembcke B. Ultrasound studies of the intestinal wall in patients with cystic fibrosis. *J Pediatr Gastroenterol Nutr* 1997;**25**:317–20.
34. Bots S, Nylund K, Löwenberg M, Gecse K, Gilja OH, D'Haens G. Ultrasound for assessing disease activity in IBD patients: a systematic review of activity scores. *J Crohns Colitis* 2018;12:920–9.
35. Domènech E, Mañosa M, Cabré E. An overview of the natural history of inflammatory bowel diseases. *Dig Dis* 2014;**32**:320–7.
36. Chiorean L, Schreiber-Dietrich D, Braden B, *et al*. Ultrasonographic imaging of inflammatory bowel disease in pediatric patients. *World J Gastroenterol* 2015;**21**:5231–41.
37. Picoraro J, Winberry G, Siegel CA, *et al*. Premedication use before infliximab administration: a cross-sectional analysis. *Inflamm Bowel Dis* 2017;**23**:174–80.
38. Epifanio M, Baldisserotto M, Spolidoro JV, Gaiger A. Grey-scale and colour Doppler sonography in the evaluation of children with suspected bowel inflammation: correlation with colonoscopy and histological findings. *Clin Radiol* 2008;**63**:968–78.
39. Darge K, Anupindi S, Keener H, Rompel O. Ultrasound of the bowel in children: How we do it. *Pediatr Radiol* 2010;**40**:528–36.
40. Alison M, Kheniche A, Azoulay R, Roche S, Sebag G, Belarbi N. Ultrasonography of Crohn disease in children. *Pediatr Radiol* 2007;**37**:1071–82.
41. Haber HP, Benda N, Fitzke G, *et al*. Colonic wall thickness measured by ultrasound: striking differences in patients with cystic fibrosis versus healthy controls. *Gut* 1997;**40**:406–11.
42. Spalinger J, Patriquin H, Miron MC, *et al*. Doppler US in patients with Crohn disease: vessel density in the diseased bowel reflects disease activity. *Radiology* 2000;**217**:787–91.