Magnetic resonance imaging in Crohn's disease
Horsthuis, K.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Magnetic Resonance Enterography for suspected IBD in a pediatric population

Karin Horsthuis
Lissy de Ridder
Anne M. J. B. Smets
Maarten S. van Leeuwen
Marc A. Benninga
Roderick H. J. Houwen
Annemieke S. Littooij
Rutger A.J. Nievelstein
Jaap Stoker

Submitted
ABSTRACT

**Purpose:** To determine the accuracy of Magnetic Resonance Enterography (MRE) in diagnosing and differentiating pediatric inflammatory bowel disease (IBD). Secondary aims were to determine the accuracy of MRE in grading disease activity and to determine the interobserver agreement for individual MRE parameters.

**Methods:** Pediatric patients scheduled to undergo esophagogastroduodenoscopy, ileocolonoscopy (CS) with biopsies and barium enteroclysis (BE) for suspected IBD were included and underwent MRE. MRE images were evaluated by 3 observers. The accuracy of MRE was calculated using the clinical diagnosis based on endoscopic, histopathological and BE examinations, as reference standard.

**Results:** 33 patients were available for analysis. IBD was correctly diagnosed in respectively 61%, 61% and 91% of the patients by the 3 observers, with a specificity of 80%, 90% and 60%. Differentiation between CD and UC was accurately done in respectively 67%, 53% and 80% of CD patients and 0%, 14% and 43% of UC patients. Disease activity was understaged on MRE in the majority of patients. Intraclass correlation coefficients for measurements of bowel thickness were 0.52 (observer 1-2;observer 1-3) and 0.34 (observer 2-3). Interobserver agreement on bowel wall enhancement and stenosis was moderate to good (kappa 0.59; 0.56 and 0.56 and kappa 0.62, 0.32, 0.30 respectively).

**Conclusion:** Sensitivity and specificity values of MRE for diagnosing pediatric IBD were moderate to good. CD, but not UC, was accurately diagnosed by MRE in a large proportion of patients. Activity was understaged in a large proportion of patients. Interobserver agreement for individual MRE parameters was fair to good.
INTRODUCTION

Crohn’s disease (CD) and ulcerative colitis (UC), the two main subtypes of inflammatory bowel disease (IBD), are diagnosed frequently in the pediatric age group, accounting for 25% of all cases of IBD (1). Distinction between CD and UC is important as prognosis, clinical course and treatment options vary. In addition, assessment of the degree of disease severity is important in optimizing treatment.

The diagnostic tests that are currently used for initial evaluation of suspected IBD are esophagogastroduodenoscopy (EGD), ileocolonoscopy (CS) with biopsies and imaging of the small bowel with enteral contrast medium (2, 3). However, several drawbacks are associated with these examinations (e.g. extensive bowel preparation, invasiveness, ionizing radiation, need for anesthesia). As Magnetic Resonance Imaging (MRI) is non-invasive and does not use ionizing radiation, the diagnostic potential of this technique has been investigated in many studies in adult IBD patients (4-9). In adults MRI has been used to diagnose active IBD (4, 5), to distinguish between CD and UC (6, 7) and to determine the degree of disease activity in CD (8, 9).

In children only a small number of studies have been conducted, mainly in CD patients (10-13). Conflicting results have been reported on the ability of MRI to differentiate between CD and UC in children (14, 15), which is also true for staging disease activity in CD (15, 16). In addition, data on reliability of abdominal MRI in pediatric IBD are scarce; only one study including 14 patients has reported interobserver agreement for diagnosing IBD and differentiation between CD and UC (15).

Therefore, we conducted a prospective study to determine the accuracy of MR enterography (MRE) in diagnosis and differentiation of IBD in pediatric patients. Secondary aims were to determine the accuracy of MRE in grading disease activity and to determine the interobserver agreement for individual MRE parameters.

MATERIALS AND METHODS

Study population

This study was conducted in two tertiary care hospitals. Eligible patients were all patients aged 8-18 years scheduled to undergo EGD, CS and barium enteroclysis (BE) to confirm or exclude IBD. Exclusion criteria were general contraindications to MRI, pregnancy, or inability to tolerate a 25-second breath hold.

Between November 2004 and November 2006, 37 consecutive patients participated in this study after oral and written informed consent was obtained from all patients and their parents. The study was approved by the national Central Committee on Research Involving Human Subjects (CCMO).
MRI examination

All MR examinations were planned within two weeks of the endoscopic examinations (mean 8±6 days, median 7 days, range 0-34 days) and the small bowel radiology examinations (mean 6±6 days, median 6 days, range 0-21 days).

Starting four hours prior to the MR examination patients were requested to take 3.4 grams of Isphagula Husk Anhydricum (Metamucil, Procter & Gamble Cincinnati, OH) each hour, dissolved in 250 ml of water. During this period no ingestion of food or fluids was allowed, with the exception of additional water if wished.

MR imaging was performed on 3.0 Tesla-MRI scanners (Achieva [hospital 1] and Intera [hospital 2], Philips Medical Systems, Best, the Netherlands) using a torso phased-array surface coil. T2-weighted 2D Turbo Spin Echo (TSE) coronal (TE 60; TR 1699.9; TSE factor 81; FOV 485; RFOV 100%; scan matrix 304 x 274; SENSE factor 2; slice thickness 5 mm; gap 0; 27-35 slices) and axial (TE 60; TR 1042.9; TSE factor 49; FOV 455; RFOV 60%; scan matrix 304 x 243; SENSE factor 2; slice thickness 5 mm; gap 0; 72 slices) respiratory triggered sequences were performed.

Prior to acquisition of T1-weighted sequences 0.1 ml/ kg of bodyweight contrast medium (Gadodiamide, Omniscan, General Electric Healthcare, Chalfont St. Giles, United Kingdom) and a spasmolytic agent (butylscopolamine bromide, Buscopan; Boehringer-Ingelheim, Ingelheim, Germany) were administered intravenously. Fat saturated 3D T1-weighted breath hold sequences were performed in the coronal and axial plane (TE 1.5; TR 3.0; flip angle 10º; FOV 395; RFOV 100% (coronal) or 85% (axial); matrix scan 192 x 154; SENSE factor 2; 80 slices (1 stack; coronal) or 160 slices (2 stacks; axial) slice thickness 2 mm (interpolated); SPAIR inversion delay 100ms).

Image reading and interpretation

Observers

MR images were evaluated on two different workstations (Philips Easy Vision, release 10.2P5, Philips Medical Systems, Best, The Netherlands [hospital 1, University Medical Center Utrecht] and IMPAX, IMPAX SP4 SU4 DS3000, AGFA, Mortsel, Belgium [hospital 2, Academic Medical Center Amsterdam]) by three observers, who were blinded to clinical and endoscopic findings as well as to findings at BE. Observer 1 was a pediatric radiologist [RJN] with 11 years experience in reading abdominal MRI and three years experience in reading MRI of the small bowel; observer 2 was an abdominal radiologist [MvL] with no previous experience in pediatric abdominal imaging, but with twenty years experience in abdominal imaging in adults and three years experience in reading MRI of the small bowel; observer 3 was a pediatric radiologist [AS] with twenty years experience in pediatric abdominal imaging and no previous experience in reading MRI of the small bowel. Observer 1 and 2 were from hospital 1, observer 3 was from hospital 2. Directed training of the three observers in reading MRE using our standardized evaluation sheet was done via a database of 25 MRE examinations validated
by endoscopy and/or surgery. After the first ten examinations feedback was provided, as was done after all examinations had been read.

**Image evaluation**

Images were evaluated both qualitatively and morphologically. To this end the gastrointestinal tract was divided in 8 segments: duodenum, jejunum, proximal ileum, terminal ileum, cecum/right colon, transverse colon, left colon/sigmoid, and rectum. All bowel loops left of an imaginary line from the junction of the liver dome and sinus pleurae to the middle of the acetabulum were considered jejunal loops. Bowel loops located right of this imaginary line were considered ileal bowel loops. The distal 20 centimeters of the ileum were considered the terminal ileum.

1) **Image quality**

To determine image quality the degree of bowel distension and the conspicuity of the bowel wall were scored per bowel segment. Bowel distension was scored as good, moderate, or poor (only slight distension or collapsed bowel). Bowel wall conspicuity was scored as good (wall clearly seen), moderate (wall seen without clear delineation) or poor (wall not seen or only portions seen).

Overall image quality was scored on a per-patient basis as poor, moderate or good.

2) **Pathological findings**

Morphological evaluation was performed per bowel segment for the following items: 1) maximum wall thickness was measured in mm by using electronic calipers; 2) on T2-weighted images was noted if stratification (i.e. a layered appearance) of the bowel wall was seen; 3) on post-contrast T1-weighted fat saturated images bowel wall enhancement was scored subjectively as either normal, or as mild (just visible thin layer), moderate (relatively thick and/or enhancing layer) or strong (very intense enhancement) pathological enhancement. If pathological enhancement was observed, the enhancement pattern was noted (mucosal, mucosal and serosal, or transmural); 4) the outer contour of the bowel wall was described as either sharp, fuzzy, spiked or continuous with surrounding soft tissue; 5) the length of pathological bowel was measured using hand-held calipers; 6) stenosis was recorded as either absent, low-grade stenosis (lumen reduction > 50% without prestenotic dilatation), or high-grade stenosis (lumen reduction >50% with prestenotic dilatation).

On a per-patient basis the observers recorded presence or absence of fibrofatty proliferation, increased mesenterial vascularization (‘comb sign’), mesenteric lymphadenopathy (short axis >10 mm), lymphadenopathy at other locations, infiltrate, abscess and fistula.

A diagnosis of CD, UC, Indeterminate Colitis (IC) or no IBD was made in accordance with the definition previously used by Darbari et al. (14): CD was defined as a transmural pathological enhancement of the colon with either the involvement of the terminal ileum or a pathological enhancement of proximal small bowel, with wall thickening. UC was diagnosed if mucosal enhancement with submucosal sparing was present, extending...
from the rectum with contiguous involvement of the proximal colon. Diagnosis of CD was suspected, but not considered definitive if full thickness enhancement of the large bowel was noted with relative sparing of the rectum and sigmoid colon. Enhancement of small bowel without bowel wall thickening was also considered suggestive of CD, but not definitive.

The severity of disease was subjectively graded (no disease, mild disease, moderate disease, severe disease) per bowel segment and per patient, taking the abovementioned items into account.

Reference standard

Endoscopy

In all patients EGD and CS were performed under general anesthesia by either a pediatric gastroenterologist or a pediatric gastroenterology fellow. The colon was adequately cleansed prior to endoscopy according to the protocol prevailing in each of the two participating hospitals.

EGD was performed using a standard pediatric gastroscope (Olympus Medical Systems Europe, Hamburg, Germany). Both random biopsies were taken as well as biopsies of suspect lesions. CS was performed using a standard pediatric colonoscope (Olympus Medical Systems Europe, Hamburg, Germany). At CS tissue sampling was performed at predefined locations (rectum, descending colon, splenic flexure, transverse colon, hepatic flexure, ascending colon, cecum, terminal ileum). If suspect lesions were present at these locations, these were biopsied; otherwise random tissue sampling was performed per bowel segment.

The endoscopist was asked to record endoscopic findings on a standardized form; findings were reported as no pathology, CD, UC, or indeterminate colitis. If pathological findings indicative of IBD were present, but a definitive diagnosis could not be made based on the endoscopic appearance, the endoscopist scored disease as indeterminate colitis. If UC was present disease severity was graded using predefined endoscopic criteria (17). If CD or IC was diagnosed, the endoscopist was asked to rate disease severity as no disease or as mild, moderate or severe disease, based on his/her subjective judgment.

Histopathology

The diagnosis of CD or UC was made based on well-defined histological criteria (18).

Barium enteroclysis (BE)

For small bowel barium examinations a standard pediatric patient preparation was used, incorporating both dietary restrictions as well as purgation of the bowel by laxatives. After nasojejunal intubation barium sulphate suspension was injected into the small bowel, in hospital 2 followed by infusion of a 0.5% methylcellulose solution. Barium examinations were performed by experienced pediatric radiologists or by a senior resident radiology under close supervision of the pediatric radiologist.
Fluoroscopy and compression radiography were used to visualize all segments of the small bowel. In two patients in hospital 1 a small bowel follow-through (SBFT) was performed instead of a barium enteroclysis. Results of BE or SBFT were reported in a standardized manner in accordance with clinical routine.

Statistical analysis
The accuracy of MRI in the diagnosis of IBD and in differentiating between CD and UC was calculated using the final clinical diagnosis as reference standard. The final clinical diagnosis was established by the treating pediatric gastroenterologist, based on a combination of endoscopic, histopathological and radiological results, as recommended by the Working Group of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (2) and the IBD Working Group of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) (3). With regard to assessment of disease activity, agreement between grading at MRI and endoscopic grading was separately determined for the patients with CD and the patients with UC. In addition, the accuracies of MRI versus BE for assessment of disease in the terminal ileum were compared. For this purpose, the endoscopic results were used as reference standard.

Linear regression analysis was performed for rating of segmental bowel distension. Interobserver agreement was calculated for the following MRI parameters: measured bowel wall thickness (in mm), the degree of bowel wall enhancement and diagnosis of stenosis. Bowel wall measurements of the three observers were compared using the intraclass correlation coefficient. Interobserver agreement for the categorical data (enhancement, stenosis) was quantified using weighted kappa (κ) statistics. Interpretation of κ values was done according to Altman: 0 – 0.2 was considered ‘poor’, 0.21 – 0.40 was considered ‘fair’, 0.41 - 0.60 was considered ‘moderate’, 0.61 - 0.80 was considered ‘good,’ and > 0.8 as ‘very good’ agreement (19). For statistical analysis we used SPSS 14.0 (SPSS Inc., Chicago, IL) and Cytel StatXact (Cytel Inc., Cambridge, MA).

RESULTS
A total of thirty-seven patients were included in the study. Of these patients four had to be excluded (one patient did not want to undergo the MRI examination after an MRI simulation due to anxiety, in another patient MRI acquisition was halted after performance of the first sequence also due to patient anxiety, two patients had fixed dental braces that were affected by the strong magnetic field). Therefore, a total of 33 patients were available for analysis (15 male: 18 female, mean age 13.5±2.4 years, median 14, range 8-17).
Table 1: Agreement between clinical diagnosis and Magnetic Resonance Enterography (MRE)

<table>
<thead>
<tr>
<th></th>
<th>No disease</th>
<th>CD</th>
<th>UC</th>
<th>IC</th>
<th>Infectious colitis</th>
<th>Other disease</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No disease</td>
<td>4 / 5 / 4</td>
<td>5 / 5 / 0</td>
<td>3 / 3 / 2</td>
<td>1 / 1 / 0</td>
<td>4 / 4 / 2</td>
<td>0 / 0 / 0</td>
<td>17 / 18 / 8</td>
</tr>
<tr>
<td>CD</td>
<td>0 / 0 / 0</td>
<td>10 / 8 / 12</td>
<td>2 / 1 / 1</td>
<td>0 / 0 / 1</td>
<td>0 / 0 / 2</td>
<td>1 / 1 / 1</td>
<td>13 / 10 / 17</td>
</tr>
<tr>
<td>UC</td>
<td>0 / 0 / 1</td>
<td>0 / 0 / 3</td>
<td>0 / 1 / 3</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 1 / 7</td>
<td></td>
</tr>
<tr>
<td>IC</td>
<td>1 / 0 / 0</td>
<td>0 / 2 / 0</td>
<td>2 / 2 / 1</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>3 / 4 / 1</td>
<td></td>
</tr>
<tr>
<td>Infectious colitis</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td></td>
</tr>
<tr>
<td>Other disease</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>15</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>33</td>
</tr>
</tbody>
</table>

Results for observer 1 are shown on the left, for observer 2 in the middle, and for observer 3 on the right. Numbers in bold indicate agreement between MRE and the clinical diagnosis.

† 1 patient suffered from congenital chloride diarrhoea (CCD). At the time of endoscopy severe abdominal complaints (abdominal pain, nausea, vomiting) were present, causing the suspicion of IBD. At CS, deep ulcerations of the terminal ileum were seen. At surgery, this patient proved to have multiple adhesions after previous exploratory surgery. Complaints were caused by an exacerbation of CCD combined with adhesions.

Image quality

Observer 1 rated image quality as poor in 2, moderate in 18 and good in 13 of the examinations. Observer 2 rated image quality as poor in 1, moderate in 25 and good in only 7 patients. Observer 3 rated none of the examinations as being of poor quality, 6 of moderate quality and 27 of good quality.

Bowel distension was scored as poor or moderate in respectively 60.3, 57.9% and 13.2% of segments by observer 1, 2 and 3. Best distension was observed in the proximal ileum, the terminal ileum and the ascending colon (Fig 1).
Clinical diagnosis
In 23 patients (69.7%) IBD was diagnosed (CD in 15, UC in 7, IC in 1 patient). In 5 patients other diseases were diagnosed (infectious colitis in 4 patients, exacerbation of congenital chloride diarrhea with ulcerations in 1 patient) and in 5 patients no pathological findings were seen at CS, EGD, histopathological examination and BE (Table 1).

MRE results

Diagnosis of IBD
Sensitivity values for diagnosing IBD were 60.9%, 60.9% and 91.3% for the three observers, respectively. Specificity values were 80.0%, 90.0% and 60.0% (Table 1).

Differentiation between CD and UC
Differentiation between CD and UC based on MRE was accurately done in respectively 67% (10/15), 53.3% (8/15) and 80% (12/15) of CD patients and 0% (0/7), 14.3% (1/7) and 42.9% (3/7) of the UC patients by the three observers.

Determination of disease activity
With regard to staging disease activity in only a minority of patients all three observers agreed with CS on the degree of disease severity; for CD disease activity was understaged on MRE in 60% (observer 1), 80% (observer 2) and 33% (observer 3) (Table 2a). In most UC patients disease activity was also understaged (Table 2b).

Table 2a: Grading of disease activity in CD patients

<table>
<thead>
<tr>
<th>MRI grading</th>
<th>No</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0 / 0 / 0</td>
<td>3 / 4 / 0</td>
<td>2 / 1 / 0</td>
<td>0 / 0 / 0</td>
<td>5 / 5 / 0</td>
</tr>
<tr>
<td>Mild</td>
<td>0 / 0 / 0</td>
<td>1 / 0 / 3</td>
<td>1 / 4 / 3</td>
<td>1 / 1 / 0</td>
<td>3 / 5 / 6</td>
</tr>
<tr>
<td>Moderate</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 1</td>
<td>3 / 1 / 1</td>
<td>2 / 2 / 2</td>
<td>5 / 3 / 4</td>
</tr>
<tr>
<td>Severe</td>
<td>0 / 0 / 0</td>
<td>1 / 1 / 1</td>
<td>0 / 0 / 2</td>
<td>1 / 1 / 2</td>
<td>2 / 2 / 5</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 2b: Grading of disease activity in patients with UC

<table>
<thead>
<tr>
<th>MRI grading</th>
<th>No</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0 / 0 / 0</td>
<td>1 / 1 / 1</td>
<td>1 / 1 / 0</td>
<td>2 / 1 / 0</td>
<td>4 / 3 / 2</td>
</tr>
<tr>
<td>Mild</td>
<td>0 / 0 / 0</td>
<td>1 / 1 / 0</td>
<td>1 / 1 / 0</td>
<td>0 / 1 / 1</td>
<td>2 / 3 / 1</td>
</tr>
<tr>
<td>Moderate</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 0</td>
<td>1 / 0 / 0</td>
<td>1 / 0 / 0</td>
</tr>
<tr>
<td>Severe</td>
<td>0 / 0 / 0</td>
<td>0 / 0 / 1</td>
<td>0 / 0 / 1</td>
<td>0 / 1 / 2</td>
<td>0 / 1 / 4</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

Results for observer 1 are shown on the left, for observer 2 in the middle, and for observer 3 results are displayed on the right. Numbers in bold indicate agreement between MRI grading and endoscopic grading.
Comparison of BE and MRE

BE was performed in 32 of the 33 included patients. In four of these 32 patients the examination was not diagnostic (due to nausea (n=2) and/or stasis of contrast medium (n=3). Pathological findings of the small bowel were observed in 12 patients; in nine patients inflammation was restricted to the terminal ileum, while in three patients besides the terminal ileum more proximal ileal loops were affected as well.

The sensitivity of BE for the detection of inflammation of the terminal ileum was 50%, the specificity was 53%. MRE was more accurate in diagnosing terminal ileitis than BE; sensitivity values were 60% (observer 1), 50% (observer 2) and 80% (observer 3). Specificity values were 94% (observer 1), 94% (observer 2) and 63% (observer 3) (Table 3).

With regard to findings of the small bowel proximal of the terminal ileum no endoscopic verification was available. However, in all three patients with pathological findings of the proximal ileum on BE, the ileum was rated as inflamed on MRE.

Interobserver agreement

Intraclass correlation coefficients for measurements of bowel wall thickness were 0.52 (observer 1-2, 0.52 (observer 1-3) and 0.34 (observer 2-3)(fig 2). Moderate agreement was found between the observers for evaluation of the degree of bowel wall enhancement:

### Table 3a: Agreement between ileocolonoscopy (CS) and barium examinations (BE) regarding pathological changes of the terminal ileum (TI)

<table>
<thead>
<tr>
<th></th>
<th>CS</th>
<th>BE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ileitis terminalis</td>
<td>No pathology TI</td>
</tr>
<tr>
<td>Ileitis terminalis</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>No pathology TI</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>TI not visualized</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>

§ In one patient no small bowel barium examination was performed.
Numbers in bold indicate agreement between CS and BE regarding assessment of the terminal ileum.

### Table 3b: Agreement between ileocolonoscopy (CS) and Magnetic Resonance Enterography (MRE) regarding pathological changes of the terminal ileum (TI)

<table>
<thead>
<tr>
<th></th>
<th>CS</th>
<th>MRE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ileitis terminalis</td>
<td>No pathology TI</td>
</tr>
<tr>
<td>Ileitis terminalis</td>
<td>6 / 5 / 8</td>
<td>1 / 0 / 6</td>
</tr>
<tr>
<td>No pathology TI</td>
<td>4 / 5 / 2</td>
<td>15 / 15 / 10</td>
</tr>
<tr>
<td>TI not visualized</td>
<td>0 / 0 / 0</td>
<td>0 / 1 / 0</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>16</td>
</tr>
</tbody>
</table>

Results for observer 1 are shown on the left, for observer 2 in the middle, and for observer 3 results are displayed on the right. Numbers in bold indicate agreement between CS and MRE regarding assessment of the terminal ileum.
kappa 0.59 (85.5% agreement; observer 1-2), kappa 0.56 (79.2% agreement; observer 1-3) and kappa 0.56 (79.7% agreement; observer 2-3). For bowel stenosis agreement was fair to moderate: respectively kappa 0.46 (96.4% agreement; observer 1-2), kappa 0.35 (94.2% agreement; observer 1-3,) and kappa 0.24 (95.3% agreement; observer 2-3).

For the terminal ileum interobserver agreement was separately calculated: High intraclass correlation coefficients were observed for bowel wall thickness measurements of the terminal ileum between the three observers: (icc 0.82, observer 1-2; icc 0.90, observer 1-3; icc 0.82, observer 2-3).

Interobserver agreement was good to very good for the evaluation of the degree of enhancement: kappa-values were 0.78 (74.2% agreement; observer 1-2), 0.76 (69.7% agreement; observer 1-3) and 0.87 (74.2% agreement; observer 2-3). Interobserver agreement for grading of stenosis was fair to good: kappa 0.62 (87.1%; observer 1-2), kappa 0.32 (75.8%; observer 1-3), kappa 0.30 (77.4%; observer 2-3).

DISCUSSION

In our study we have shown that the accuracy of MRE for the diagnosis of IBD was moderate to good, with sensitivity values ranging from 61 to 91% and specificity values ranging from 60 to 100%. MRE proved to be able of diagnosing CD in a large proportion of patients, but UC was not diagnosed accurately in most of the patients. MRE was more accurate than BE in diagnosing inflammation of the terminal ileum.

Accuracy values of MRE reported in this study are in accordance with earlier findings: reported sensitivity values for the diagnosis of IBD were in the same range as the value we reported for observer 3 (11, 12, 16) and specificity values were in the same range as the highest specificity value of 90% reported in our study (11, 12, 16). However, in our
study none of the three observers obtained both a high sensitivity and a high specificity. While the two observers from hospital 1 both correctly scored absence of disease at the expense of missing disease in several patients, observer 3 from hospital 2 correctly scored many patients with disease, but at the expense of many false positives. This might reflect an institution dependency in scoring. This also suggests that interpretation and scoring of MRE findings is difficult and subjective to some degree, even for observers with previous experience in interpreting MRE. It is not yet known how many MRE examinations are needed to attain competence in reading these for suspected IBD.

It is known that poor reproducibility of a test method between scorers will negatively affect diagnostic accuracy (20). Although we tried to diminish the influence of the subjective component of the MRI evaluation by providing precise criteria for evaluation of disease and pre-study training sessions, we found the interobserver agreement taking into consideration all bowel segments to be rather low. In a recently published study by Negaard et al., concerning adult patients with CD, interobserver agreement was evaluated for individual MRI items. In this study, agreement was very good for most items including bowel wall thickening and pathological enhancement (21). However, Negaard et al scored these items dichotomously as either pathological or normal, while we have tried to subcategorize these items to reflect clinical practice. In clinical practice it is important to know whether a bowel loop is just slightly thickened or very severely affected as bowel wall thickness is significantly correlated with clinical disease activity (22, 23). Also, in literature correlations have been described between the degree of bowel wall enhancement and disease severity (5, 8, 23-25). A possible explanation for this lower interobserver agreement could be that bowel distension was poor or moderate in a substantial number of bowel segments. This hypothesis is substantiated by the observation that interobserver agreement was high for the terminal ileum, a bowel segment that was adequately distended in almost all patients. For adequate evaluation of disease, bowel distension is mandatory, as in collapsed bowel loops bowel wall measurements cannot be adequately performed, while it also is difficult to diagnose and grade obstruction. We have used a solution of psyllium fibers as oral contrast medium in our study, as had been previously reported by Patak et al (26). In a study by Lauenstein et al different oral contrast media were compared, with better results for solutions with mannitol or sorbitol than for a solution with psyllium fibers (27). Also, bowel distension has been reported to be greater when contrast medium is administered by enteroclysis (after nasojejunal intubation) than after oral administration of contrast medium. However, in the two available studies comparing MR enterography with MR enteroclysis in Crohn’s disease no significant differences were observed regarding diagnostic accuracy of the two methods (21, 28). We have decided to administer the enteral contrast medium orally, as a major disadvantage of MR enteroclysis is the nasojejunal intubation needed; which causes discomfort and requires the use of ionizing radiation for catheter placement.

We have performed our study using high field strength of 3.0 Tesla. While theoretically disease detection might have been better compared with field strength of 1.5 Tesla due to
the higher Signal to Noise Ratio that can be used for better image quality or shortening of acquisition times, in practice our accuracy values were not higher than in previous studies performed at 1.5 Tesla. However, a drawback of field strength of 3.0 Tesla is an increased sensitivity to inhomogeneities of the magnetic field. Furthermore, image acquisition is more susceptible to various intrinsic artifacts. Image quality was scored as moderate in a relatively large number of patients; these findings are in accordance with an earlier study by Schmidt et al who reported that 3.0 T whole-body MRI showed significantly more artifacts with a mild to moderate impact on image assessment (29).

We used predefined criteria to differentiate between CD and UC, but although CD was diagnosed correctly in a large percentage of patients, UC was either diagnosed as no disease or as CD or IC. A possible explanation is that rectal sparing can be seen in a rather large percentage of children with newly diagnosed, untreated UC, making the diagnosis of UC less straightforward (30, 31). More research is needed to determine which MR imaging features are compatible with UC and with CD in children in order to be able to provide a correct definition for radiologists to base their diagnosis on.

In our study MRE understaged disease severity in CD. Only in a few studies in children an attempt to grade disease severity has been described. In the study by Laghi et al MRE understaged disease activity as seen on CS in almost 20% of patients (16), while in the study by Durno et al. no significant correlation was found between the degree of bowel wall enhancement and disease severity (15). We have tried to determine disease severity by taking into account all the MRE parameters that could possibly be affected in active IBD as in CS appraisal of disease activity is also done taking into account all the manifestations of disease (e.g. stenosis, ulceration, erythema). A likely explanation for the fact that MRE understages disease is that superficial pathology (e.g. erythema, aphthoid ulcers) which are very well depicted at endoscopy are not well visualized at MRE due to the limitations in spatial resolution. Also, endoscopy and MRE are essentially two different modalities; with endoscopy only the mucosal surface of the bowel can be evaluated, while with MRE transmural and extramural assessment is possible for which no endoscopic verification is available.

In conclusion, MRE can be used to diagnose IBD in children. For assessment of the small bowel, MRE is more accurate than BE.

REFERENCES


Magnetic Resonance Enterography for suspected IBD in a pediatric population


