Interobserver Agreement in MR Enterography for Diagnostic Assessment in Patients with Crohn's Disease

Interobserver-Übereinstimmung bei der MR-Enterografie in der Diagnostik bei Patienten mit Morbus Crohn

Abstract

**Purpose:** To evaluate the extent to which MRE can be used as an observer-independent, objective imaging method for the diagnosis and evaluation of CD with respect to the detection of inflammatory changes of the small bowel and lymphadenopathy as diagnostic criterion and bowel distension as a quality criterion.

**Materials and Methods:** MRE scans of 84 patients (42 female, median age 37 years) were evaluated independently by 4 experienced radiologists. Analysis of inflammatory changes of the bowel wall, lymphadenopathy and adequate bowel distension was conducted separately for the jejunum, ileum and terminal ileum. The Kendall’s W-test was used for the statistical comparison of concordance.

**Results:** In a total of 55 patients, inflammatory activity of the bowel wall was detected and MRE was found to have a high interobserver reproducibility concerning inflammatory changes of the intestinal wall (Kendall’s W 0.527 – 0.823). Concerning lymphadenopathy (31 cases, 36.9%), a low to moderate consensus could be shown with a Kendall’s W value of 0.402 – 0.505. For the assessment of adequate bowel distension, a moderate concordance between the operators could be found (Kendall’s W 0.497 – 0.581).

**Conclusion:** MRE has proven high interobserver agreement with respect to the diagnosis of inflammatory bowel wall activity of the bowel as a diagnostic criterion in patients with CD. Concerning adequate bowel distension as a quality criterion of the examination itself and lymphadenopathy as a diagnostic criterion, moderate interobserver agreement could be found. This is thought to have a rather small effect on the diagnostic significance and conclusiveness of the method in the daily routine.

**Key Points:**
- MR enterography as observer independent diagnostic procedure in patients with Chron’s Disease.
- Highest interobserver concordance for the criterion of inflammatory bowel wall affection.
- Moderate interobserver concordance for lymphadenopathy and bowel distension.

**Zusammenfassung**

**Ziel:** Zu evaluieren, inwieweit die MR-Enterografie (MRE) bezüglich der Beurteilung entzündlicher Dünndarmwandveränderungen und der Lymphadenopathie als Diagnosekriterium und der Darmdistension als Qualitätskriterium eine untersucherunabhängige objektive Methode in der Diagnostik des Morbus Crohn (MC) darstellt.

**Material und Methoden:** Es wurden die MRE von 84 konsekutiven Patienten (42 weiblich, Median 37 Jahre) von 4 erfahrenen Radiologen unabhängig voneinander analysiert. Dabei wurde von jedem Auswerter für das Jejunum, Ileum und terminale Ileum beurteilt, ob eine entzündliche Darmwandveränderung, eine Lymphadenopathie und eine diagnostisch adäquate Darmdistension vorlagen. Um die Übereinstimmung der Diagnose der einzelnen Radiologen untereinander zu überprüfen, wurde der Kendall-W-Test verwendet.

**Ergebnisse:** Bei insgesamt 55 Patienten (65,5%) wurde entzündliche Aktivität festgestellt, die MRE weist bezüglich entzündlicher Darmwandveränderungen eine hohe Untersucherunabhängigkeit
In addition to conventional radiological double-contrast examination such as enteroclysis [1, 2], abdominal sonography (US) [3, 4], computed tomography (CT) [5, 6], and magnetic resonance imaging (MRI) [7–11] can be used for the primary or relapse diagnosis of patients with Crohn’s disease (CD) via radiological imaging. Each of the indicated examination methods has various advantages and disadvantages. With respect to the primary diagnosis of CD, the current German S3 guideline regarding the CD diagnosis states that no single examination method can represent the gold standard [12, 13].

The current European guidelines also describe the importance of all of the methods with preference being given to MRI [14], which with the current technical developments has the advantage of being able to image the entire small and large bowel with high spatial and temporal resolution without the use of radiation. This method is therefore specified in the guidelines of the German Society for Digestive and Metabolic Diseases (DGVS) in addition to US as the preferred radiological basic diagnostic method for CD.

MRI of the small bowel with oral contrast, referred to in the literature as MR enterography (MRE), is a method that has the advantage of being able to image the entire small and large bowel with high spatial and temporal resolution without the use of radiation. This method is therefore specified in the guidelines of the German Society for Digestive and Metabolic Diseases (DGVS) in addition to US as the preferred radiological basic diagnostic method for CD.

In the present study, 84 consecutive patients (42 women, 42 men, mean age: 37 years) were included prospectively and were examined in the case of histologically proven CD with MRE for initial diagnosis or for follow-up. There was no contraindication for MRE in any patient. A study entitled “Sequence-Evaluation für die MR-Enterografie für Patienten mit Morbus Crohn” [Sequence Evaluation for MR Enterography for Patients with Crohn’s Disease] [25] was already performed using the same collective.

Patient collective

All patients had been fasting for at least 12 hours at the time of the examination. To distend the bowel and to generate a hypointense intraluminal signal in the T1-weighted sequence (“dark lumen technique”), the patients drank two liters of water with an additional 25 g mannitol and 5 g carob bean gum (Nestargel; Nestlé Deutschland AG, Frankfurt am Main, Germany) per liter within an hour. Just prior to the examination, the patients received 200–300 ml NaCl 0.9 % as a rectal contrast for additional distension of the rectum [26, 27]. Moreover, all patients received 40 mg butylscopolamine (Buscopan, Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim am Rhein, Germany) intravenously just prior to the examination. In our experience the duration of effect of the butylscopolamine over all acquired sequences was adequate to reduce movement and to distend the bowel so that we were able to use a single intravenous dose. There were no contraindications for any of the evaluated patients. In the contrast-enhanced T1-weighted sequences, a dose of 0.1 mmol/kg body weight of Gd-DTPA (Magnevist, Schering AG, Berlin, Germany) was administered intravenously.

The examination was performed on a 1.5 T MRI unit with a gradient field strength of 20 mT/m (Symphony, Siemens Medical Systems, Erlangen, Germany). Coronal T2/T1-weighted sequences (TrueFISP, True Fast Imaging With Steady Precession, TR/TE 4.76/2.38 msec, flip angle 60°, matrix 256, FOV 450 mm, measurement time 19 s) with a slice thickness of 5 mm were first acquired followed by axial T2-weighted turbo spin echo sequences (HASTE, Half Fourier-Acquired Single Shot Turbo Spin Echo, TR/TE 1070/77 msec, flip angle 150°, matrix 256, FOV 400 mm, breath-triggered, measurement time 109 s) with a slice thickness of 8 mm. In addition, axial, fat-saturated, T1-weighted gradient echo sequences (2D-FLASH, Fast Low Angle Shot, TR/TE 124/4.7 msec, flip angle 70°, matrix 256, FOV 380 mm) with a slice thickness of 6 mm were acquired before (measurement time 72 s) and after Gd-DTPA administration (measurement time 119 s). The sequence protocol ended with coronal, fat-saturated, T1-weighted gradient echo sequences (2D-FLASH, Fast Low Angle Shot, TR/TE 86/4.7 msec, flip angle 150°, matrix 256, FOV 400 mm, breath-triggered, measurement time 109 s) with a slice thickness of 8 mm.
flip angle 70°, matrix 256, FOV 380 mm, measurement time 90 s) with a slice thickness of 5 mm.

Both the selected sequences and the indicated sequence parameters correspond to the routine MRI protocol at our institute and have been optimized and established over years with more than 200 patients with chronic inflammatory bowel diseases being examined by MRE each year.

**Diagnosis and evaluation**

All MR acquisitions were analyzed independently by four experienced radiologists. The four evaluators were radiologists with at least three years of abdominal diagnostics experience.

**Inflammatory small bowel change**

Evaluation was performed under consideration of all available MRI sequences. Every examiner evaluated whether there was an inflammatory change of the jejunum, ileum, and terminal ileum. For this purpose bowel wall thickening with a diameter of more than 3 mm and contrast uptake of the bowel wall were defined as direct signs of inflammation. Increased vascular injections and the presence of pathological lymph nodes were indirect signs of inflammation (Table 1). In the case of a lack of direct and indirect signs of inflammation, a “0: no inflammation” was assigned, in the case of a corresponding single finding and a positive lymph node finding, a “1: uncertain inflammation” was assigned, in the case of two or more of the radiological signs of inflammation, a “2: definite inflammation” was assigned as the MRE evaluation.

**Lymphadenopathy**

The presence of a perifocal mesenterial lymphadenopathy was detected independently of the above criteria in the case of the presence of lymph nodes with a short axis diameter greater than 1 cm. In addition, clustered lymph nodes were evaluated as lymphadenopathy independently of their diameter. An accumulation of 3 or more lymph nodes was evaluated as a cluster. The presence of a lymphadenopathy was assigned to the jejunum, ileum, and the terminal ileum according to the above described small bowel topology.

**Bowel distension**

The distension of the small bowel as a quality criterion of the examination was evaluated with “0: none” in the case of adequate distension of the bowel segment in less than 50% of the segment (jejunum, ileum, terminal ileum). In the case of adequate distension of the segment in approximately 50–80% of the segment length, the segment was evaluated with “1: minor” and with “2: good” in the case of 80% distension in the segment [28, 29].

**Statistical methods**

The program SPSS version 16.0 (SPSS Inc., Chicago, IL, USA) for Windows (Microsoft, Redmond, WA, USA) was used for the statistical tests. To check the agreement of the diagnoses of the individual radiologists among one another, the Kendall’s W test was used. The calculated concordance coefficient W is in a value range of 0 – 1. A value W = 0 indicates no agreement and a value W = 1 indicates complete agreement. W = between 0 and 0.482 indicates minor agreement, between 0.482 – 0.7 indicates moderate agreement and a value W = 0.7 or higher indicates significant agreement.

**Results**

**Inflammatory small bowel change**

According to the division of the small bowel into the jejunum, ileum, and terminal ileum, all four radiologists excluded inflammation in 29 of 84 (34.5%) patients; in 2 (2.4%) cases uncertain signs of inflammation were diagnosed by all four radiologists; and definite inflammatory activity was determined in 18 (21.5%) more of the evaluated cases by all four radiologists (Fig. 1). In the remaining 35 (41.7%) cases, no agreement regarding the presence of inflammation was achieved.

For the criterion “inflammatory small bowel change”, the Kendall’s W values were as follows: 0.527 for the jejunum, 0.724 for the ileum, and 0.823 for the terminal ileum. The

**Table 1** Criteria for diagnosing inflammation of the small bowel.

<table>
<thead>
<tr>
<th>criteria for diagnosing inflammation</th>
</tr>
</thead>
<tbody>
<tr>
<td>bowel wall thickening of more than 3 mm</td>
</tr>
<tr>
<td>increased contrast uptake of the bowel wall compared to the surrounding bowel</td>
</tr>
<tr>
<td>vascular injections</td>
</tr>
<tr>
<td>lymphadenopathy (lymph nodes ≥ 1 cm or cluster ≥ 3 lymph nodes)</td>
</tr>
</tbody>
</table>

**Fig. 1** Coronal contrast-enhanced fat-saturated T1-weighted gradient echo sequence: Inflammatory stenosed ileal loop with prestenotic and poststenotic dilation marked by a white arrow.
examination method and the evaluation therefore exhibit high observer independence in this regard.

Lymphadenopathy
In 53 (63.1 %) cases, the presence of lymph nodes was unanimously dismissed and was equally unanimously confirmed in 3 (3.6 %) cases. In the remaining 28 (33.3 %) examinations, the four radiologists produced differing evaluations (Fig. 2). The Kendall’s W value for the criterion “lymphadenopathy” is on the border to low or average agreement and is 0.402 for the jejunum, 0.505 for the ileum, and 0.505 for the terminal ileum.

Bowel distension
At least 50 – 80 % bowel distension (Fig. 3) was achieved in all segments. A distension of over 80 % was even achieved in the jejunum and terminal ileum according to the evaluations of the four radiologists (Table 2). The median value for the relevant segment was “2: good”. The evaluations of the four radiologists yielded Kendall’s W values of 0.497 for the jejunum, 0.581 for the ileum and 0.566 for the terminal ileum.

Discussion
The results of this study show a high level of objectivity and observer independence for MRE for the criterion of inflammatory small bowel changes. This is particularly true for the ileum (Kendall’s W 0.724) and the terminal ileum (Kendall’s W 0.823) and also for the jejunum (Kendall’s W 0.527) with moderate observer independence. However, the inferiority of MRE in the jejunum compared to other diagnostic methods, such as capsule endoscopy, is already known [30]. With respect to the examination quality, there is only average agreement of the evaluations for the criterion of bowel distension with Kendall’s W values between 0.497 for the jejunum and 0.581 for the ileum. The lowest level of agreement was achieved regarding a perifocal lymphadenopathy. The Kendall’s W values are only between 0.402 and 0.505 for this case.

In contrast to the previously published literature, not only the presence or absence of inflammatory bowel wall changes was analyzed in the present study but also the inflammatory affection of bowel segments according to anatomical small bowel segments was evaluated as definite, uncertain, or not present by four different experienced evaluators. As in previous studies, there was no gold standard for evaluating the correctness of the results of MRI diagnoses. Rather the goal of the study was to evaluate the objectivity of the radiological examination modality already established in the guidelines. The merely moderate agreement of the evaluation of the examination quality in relation to bowel distension is difficult to interpret. The intentionally subjectively selected evaluation criterion of “adequate” distension in the Materials and Methods section was probably the reason for the partially divergent interpretation of the term “adequate”. Therefore, this subjective and divergent evaluation may be due to the differ-

---

Table 2 Evaluation of the “bowel distension” criterion by four radiologists.

<table>
<thead>
<tr>
<th>in bowel segment</th>
<th>evaluation of distension</th>
<th>Kendall’s W 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25 % quartile</td>
<td>Median</td>
</tr>
<tr>
<td>jejunum</td>
<td>50 – 80 %</td>
<td>50 – 80 %</td>
</tr>
<tr>
<td>ileum</td>
<td>50 – 80 %</td>
<td>&gt; 80 %</td>
</tr>
<tr>
<td>terminal ileum</td>
<td>50 – 80 %</td>
<td>&gt; 80 %</td>
</tr>
</tbody>
</table>

1 A Kendall’s W value = 0 indicates no agreement and a value W = 1 means complete agreement. W = between 0 and 0.482 indicates minor agreement, between 0.482 and 0.7 indicates moderate agreement and a value W = 0.7 or higher indicates significant agreement.
ent intraindividual demands and experience with respect to examination quality. However, the selected categorization (adequate distension of the relevant bowel segment to approximately 50 – 80% of the segment length was evaluated with “1: minor” and distension of the relevant bowel segment to over 80% was evaluated with “2: good” [28, 29]) allows significant room for subjectivity.

With respect to the decisive question regarding inflammatory bowel affection, the four experienced evaluators in this study were able to achieve a high level of agreement for the terminal ileum and ileum. In the case of CD, small bowel involvement is present in these bowel segments in most cases. The merely moderate agreement in the jejunum should only be of secondary clinical relevance in clinical practice due to the extremely rare isolated occurrence of the disease in this segment. The poor ability to evaluate small bowel affections in the proximal small bowel was already able to be shown by Gölder et al. in a comparison study of MRE with capsule endoscopy [30].

Previous studies of the interobserver variability of MRE by Jensen et al. involving 50 patients with confirmed or suspected CD focusing on the comparison of MRE with CT enterography already showed moderate observer independence (κ = 0.48) for the diagnosis of small bowel involvement in CD [23]. These results are confirmed by the present study and are even exceeded with an overall high observer independence. A more precise analysis of the study data shows that the slightly lower results of κ = 0.48 was probably due to a single outlier of the four evaluators. The different small bowel segments were also not sufficiently differentiated in this study. Schmidt et al. also calculated a moderate observer independence for the thickening of the bowel wall (κ = 0.48) for 50 patients examined with MR enteroclysis due to unclear intestinal complaints. The study was also primarily designed to compare MR enteroclysis with the CT-based technique. However, the study did not exclusively include patients with CD [31]. Only 34% of the examined patients had a chronic inflammatory bowel disease. The remaining patients were examined due to small bowel tumors, bleeding, postoperative complications, or for other reasons.

With respect to the criterion of lymphadenopathy, there was no adequate observer independence (κ = 0.15) in the study by Schmidt et al. [31]. However, the criterion of lymphadenopathy was selected to be very small with a lymph node diameter starting at 0.5 cm. In our study a significantly higher observer independence with an average Kendall’s W value of 0.482 was able to be achieved with respect to the diagnosis of lymphadenopathy using other criteria (lymph nodes ≥ 1 cm, cluster) but this still does not yield a reliable degree of agreement. Overall, the limited ability to assess lymphadenopathy in CD compared to CT examination has already been shown in other studies [31 – 34].

Conclusion

In patients with Crohn’s disease, MR enterography has proven to be an objective and stable radiological examination method with excellent interobserver agreement regarding the diagnosis of inflammatory small bowel segments. The merely moderate agreement of different evaluators regarding bowel distension as an indicator of examination quality and a lymphadenopathy as an indirect inflammation criterion should not impact the high and stable objectivity of the method in clinical practice.

References

6 Fletcher JC. CT enterography technique: theme and variations. Abdom Imaging 2009; 34: 283 – 288
8 Schreyer AG, Seitz J, Feuerbach S et al. Modern imaging using computer tomography and magnetic resonance imaging for inflammatory bowel disease (IBD). AJR. Inflamm Bowel Dis 2004; 10: 45 – 54
12 Schreyer AG, Ludwig D, Koletzko S et al. Updated German S3-guideline regarding the diagnosis of Crohn’s disease – implementation of radiological modalities. Fortschr Röntgenstr 2010; 182: 116 – 121
20 Egger L, Fikentscher T, Poschenrieder F et al. Klinischer Stellenwert der Durchleuchtung der Magen-Darm-Passage (MDP) an einem Universitätsklinikum. Fortschr Röntgenstr 2012; 184: 893 – 898
23 Jensen MD, Ormstrup T, Vagn-Hansen C et al. Interobserver and intermodality agreement for detection of small bowel Crohn's disease with MR enterography and CT enterography. Inflamm Bowel Dis 2011; 17: 1081 – 1088
30 Golder SK, Schreyer AG, Endlicher E et al. Comparison of capsule endoscopy and magnetic resonance (MR) enteroclysis in suspected small bowel disease. Int J Colorectal Dis 2006; 21: 97 – 104