Crohn’s disease, advances in MRI
Ziech, M.L.W.

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.
Chapter 4

Retrospective Comparison of Magnetic Resonance Imaging Features and Histopathology in Crohn’s Disease Patients.

M.L.W. Ziech
S. Bipat
J.J.T.H. Roelofs
C.Y. Nio
B. Mearadji
S. van Doorn
A.M. Spijkerboer
J. Stoker

Published in European Journal of Radiology
Abstract
Purpose:
To retrospectively compare histopathological findings of surgically resected bowel segments with MRI findings on Crohn’s disease activity.

Materials and methods:
Patients who underwent a MR enterography or enteroclysis before surgery were included after informed consent. MRI features (T1-enhancement, T1 and T2 stratification, T2 signal intensity, bowel wall thickness, presence of ulcerations, comb sign, creeping fat, and disease activity) were assessed by three experienced abdominal radiologists. An acute inflammatory score based on histopathology (parameters: mucosal ulceration, oedema, depth and degree of neutrophils) was calculated. Interobserver variability for subjective MRI features was also assessed.

Results:
Thirty-nine segments in 25 patients (mean age 38 years) were included. Of the MRI features, disease activity per segment and bowel wall thickness had a positive association with the acute inflammatory score (p<0.05). T1-enhancement had a positive correlation with disease chronicity. All other MRI features did not have an association with the acute inflammatory score. Interobserver agreement between the three observers was weak to moderate.

Conclusion:
Bowel wall thickness and disease activity per-segment reflect disease severity in Crohn’s disease patients.

Introduction
Crohn’s disease is a chronic inflammatory bowel disease that often relapses and remits. It is known that the patient’s complaints do not always accurately represent disease activity. In clinical practice, MRI is used for assessment of Crohn’s disease in combination with endoscopy, biochemical markers and patient’s clinical symptoms. Specific signs at MRI are used as disease activity markers although there is no consensus which markers are the best indicators of disease activity. A meta-analysis reported that bowel wall enhancement after intravenous contrast and bowel wall thickening are most often used as MRI features in grading disease activity. Until now most studies compared overall assessment of disease activity obtained by MRI with clinical scoring systems such as CDAI and endoscopic scoring systems such as the CDEIS as a reference standard. A drawback of the CDAI is that this scoring system evaluates disease activity on a per patient basis, instead of per segment. The CDEIS on the other hand does assess individual bowel segments but inherent to the endoscopic technique it does not include mural and extraluminal findings.

Ideally a histopathology-based reference standard should be used as this overcomes the limitations of CDAI and CDEIS. This analysis can then be performed on a segmental basis and includes assessment of mural and extramural disease. Borley et al developed a histological scoring system that quantifies acute inflammation (acute inflammation score; AIS) specifically in Crohn’s disease patients. Recently, a study has been published that correlates MRI features with histopathology. This study demonstrated that the AIS was positively correlated with bowel wall signal intensity on T2-weighted fat-saturated images, mural wall thickness and a layered pattern of enhancement but not with mural enhancement. As that study was limited to 18 patients, it would be interesting to know if the results would be similar in other datasets. Current data on interobserver agreement of MRI features in Crohn’s disease is scarce.

The purpose of the current retrospective study was therefore to compare histopathological findings of resected bowel segments with the MRI findings on Crohn’s disease activity and to study the interobserver agreement of MRI.

Materials and methods
We searched the hospital’s pathological database for patients with Crohn’s disease, who underwent small bowel surgery for Crohn’s disease related complications between January 2001 and August 2008. Included were patients who had small bowel surgery and a MRI
examination up to 4 months prior to surgery. Permission by the medical ethics committee was waived as this was a retrospective study. Written informed consent was obtained from all patients.

**Histopathology**

All resected bowel specimens had been formalin-fixed and photographed. All tissue sections had been taken according to the hospital’s standard protocol; sections were taken of all macroscopically abnormal areas and in addition at 10 cm intervals, from the proximal towards the distal end of the specimen. Sections were stained with haematoxylin and eosin. A pathologist with ample experience in gastrointestinal pathology evaluated all available sections and graded the AIS in each section using the modified method of Borley et al\(^4\) in the most affected sections (see table 1 for items in scoring list). This system scores acute inflammation up to a maximum score of 13.

<table>
<thead>
<tr>
<th>Mucosal ulceration score</th>
<th>Oedema score</th>
<th>Depth of neutrophil penetration</th>
<th>Degree of neutrophil infiltration</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
<td>None</td>
<td>No increase</td>
<td>0</td>
</tr>
<tr>
<td>Aphthous ulceration only, &lt;7 mm diameter</td>
<td>Mild</td>
<td>Mucosa</td>
<td>Mild (0-25% of inflammatory population composed of neutrophils)</td>
<td>1</td>
</tr>
<tr>
<td>Linear or rake ulceration</td>
<td>Moderate</td>
<td>Submucosa</td>
<td>Moderate (25-50%)</td>
<td>2</td>
</tr>
<tr>
<td>Confluent or large ulceration</td>
<td>Severe</td>
<td>Muscularis propria</td>
<td>Marked (&gt;50%)</td>
<td>3</td>
</tr>
<tr>
<td>Serosa or extramural fat</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

**MRI**

During the study period MR enterography was performed in patients with known Crohn’s disease. MR enteroclysis was used in patients without a history of Crohn’s disease at the time of imaging. The MRI preparation was as follows: all patients fasted four hours prior to the MRI. MR enterography patients ingested 1600 ml of a 2.9% mannitol solution (Baxter, Utrecht, the Netherlands) in 60 minutes before the MRI. MR enteroclysis patients underwent nasoduodenal intubation under fluoroscopic guidance. Mannitol was administered through the nasoduodenal tube at a speed 60-120 ml/min until adequate luminal distension was achieved. All patients received a spasmyolytic agent (butylscopolamine bromide, Buscopan, Boehringer-Ingelheim, Ingelheim, Germany) intravenously to suppress small bowel peristalsis and either Omniscan (0.1 ml/kg, Amersham Health BV, Eindhoven, The Netherlands) or Magnevist (0.2 ml/kg, Schering, Berlin, Germany) at 60-70 seconds before the T1-weighted sequence. MRI examinations were performed at 1.5T (Signa Horizon Echospeed, LX 9.0, General Electric Medical Systems, Milwaukee, WI and Avanto, Siemens, Erlangen, Germany) and at 3T (Intera, Philips Medical Systems, Best, the Netherlands).

The following sequences were used for acquisition of images: Interpolated 3D T1-weighted gradient echo, Single-shot half-Fourier RARE (rapid acquisition with relaxation enhancement) and Balanced Steady State Free Precession (see table 2 and 3).
Matching

The surgical report, histopathology report, and photographs of the resected bowel were scrutinized to match the location of the resected bowel segments to the pre-operative MRI. Important features used during matching included anatomical landmarks (e.g., ileoacceal valve), strictures and length of the resected small bowel. This matching was performed by a research fellow and an abdominal radiologist in consensus. This radiologist was not involved with the image scoring. The former performs multiple studies on MRI in Crohn’s disease, including a prospective matching study. The latter has 15 years of experience in abdominal imaging with emphasis on MRI and CT of the gastrointestinal tract. The matching was performed on a segmental basis; sections were matched with the same segment on MRI. The bowel segment was marked on one balanced gradient echo sequence or if not present on one
...Features and Histopathology in Crohn’s Disease Patients. Retrospective Comparison of Magnetic Resonance Imaging...

Image evaluation
All images were independently evaluated on an IMPAX workstation by an experienced abdominal radiologist (observer 1; experience 14 years, 20,000 MRI, 1,000 MR small bowel). Two other abdominal radiologists (observer 2 and observer 3 with respectively 14 and 2 years of experience in abdominal imaging also scored the same subjective MR imaging features. (observer 2: 20,000 MRI, 200 MR small bowel, observer 3: 600 MRI, 100 MR small bowel) All observers knew that patients were operated for complications of Crohn’s disease, but had no knowledge of other clinical information. Image evaluation was performed of the bowel segments that were marked on the MRI on the IMPAX workstation, to make sure that each radiologist scored the same bowel segment. All measurements were conducted as one per segment; placement was on the site with the visually most deviate lesion.

The following features were assessed: Maximum bowel wall thickness was measured in millimetres with the use of electronic callipers. The observers were allowed to perform multiple measurements but had to choose one final measurement to be used for this study. Maximal bowel wall enhancement was measured on post-contrast T1-weighted images as compared to the enhancement of vascular structures (aorta or iliac arteries): enhancement ratio = signal intensity bowel wall / signal intensity vessel. Stratification pattern on post-contrast T1-weighted images was assessed as previously described: type A, multilayered mural stratification; type B, composed of two layers with strong mucosal enhancement and prominent low-intensity submucosa; type C, composed of two layers, one prominent serosal enhancement and a layer of low-intensity mucosa; type D, homogeneous enhancement of the whole bowel wall. Mural T2 signal intensity was measured and compared to the signal intensity of cerebral spinal fluid: mural signal intensity ratio = mural signal intensity / signal intensity of cerebral spinal fluid. Stratification on T2-weighted sequences was defined as: type A, multilayered mural stratification; type B, mucosal hyperintensity; type C, serosal hyperintensity; type D, homogeneous enhancement of the whole bowel wall.

Presence or absence of mucosal ulceration was determined. Overall segmental disease activity was judged by the radiologist as no, mild, moderate or severe, based on all MR features present though the most important factors were considered enhancement (on post-contrast T1) and wall thickness (on a T2 or balanced gradient echo series), based on the literature regarding which features were most often used for determining disease activity.

On a per patient basis, presence of the comb sign, defined as tortuosity and dilatation with a prominent comb-like arrangement of the vasa recta in the small-bowel mesentery, and presence of creeping fat (or fibrofatty proliferation) were determined. Overall disease activity per patient was also subjectively scored as no, mild, moderate or severe, based on the segmental and per patient MR features that were present, including extraluminal complications.

Statistical analysis
Interobserver variability.
For both per segment and per patient agreements were calculated for each feature (in percentages). For the wall thickness a difference of 1 mm was considered as agreement. Kappa statistics were calculated, a linear weight was applied for ordinal values, for wall thickness spearman’s correlations was calculated. Values were interpreted as follows: 0.0 no agreement, 0.2 weak agreement, 0.5 moderate agreement, 0.8 strong agreement, and 1.0 perfect agreement.

MRI features versus AIS score.
On per-segment basis:
Association between AIS and bowel wall thickness, enhancement ratio and mural intensity ratio were analyzed by means of Spearman’s correlation. Association between AIS and all other MR features were assessed by the Chi-2 intraclass correlation.

In some patients, more than 1 segment was evaluated. We selected the segment with the highest AIS score for comparison between MRI features and AIS. Association was tested with the Chi-2 intraclass correlation. Statistical analysis was performed by using software SPSS (15.0 Statistics UK). A p-value < 0.05 was considered statistical significant.

Results
Population
Twenty-five patients were included in our study. Median interval between MRI and surgery was 47 days (range 2-125). MR enterography was performed in 23 patients and two patients underwent MR enteroclysis. Our study population constituted 14 women (56%) and 11 men (44%); mean age was 38 years (range 13-61 years). Disease chronicity ranged between 0 and 30 years (mean 9 years).
Twenty-two patients (88%) used Crohn’s disease-related medication when the MRI and surgery were performed. Anti-TNFα antibodies were used by eight patients (32%), corticosteroids by 17 patients (68%), thiopurines in 11 patients (44%), aminosalicylates in three patients (12%) and methotrexate in two patients (8%).

Indications for surgery were as follows: unresponsiveness to medical therapy in seven patients (28%), suspected small bowel stenosis in 15 patients (60%) and suspected enteric fistulas in three patients (12%). Surgery concerned ileocecal resection in 14 patients (56%), neo-ileocoecal resection in eight patients (32%), jejunal resection in one patient (4%), right sided colectomy in one patient and subtotal colectomy in one patient. Mean length of resected bowel was 30 cm (range 12 – 79 cm).

In total, 39 bowel segments were scored at histopathology: ileum 23 (59%), coecum and ascending colon 14 (36%), transverse colon one (3%) and jejunum one (3%).

**MRI sequences**
Eight patients were scanned on a 3T scanner (32%), all other 17 patients on a 1.5 Tesla scanner (68%). Sequences that were performed were Single-shot half-Fourier RARE in 17 patients (68%), balanced Steady State Free Precession (SSFP) in 21 patients (84%) and Interpolated 3D T1-weighted sequences with intravenous contrast in all patients (100%). Quality of all MRI examinations was sufficient for grading.

In eight of the 25 cases one of the scoring radiologists already had seen the MR previously in routine patient care. However, there was a very large time interval (median 579 days before the scoring, range 133-2220 days).

**Acute inflammatory score**
AIS ranged from 0 to 11 with a median of 5 (see figure 1). The most found item was mild oedema (in 20 segments) and the least found item neutrophil infiltration into the serosa (only in 2 segments).

**Interobserver variability of MRI-based score for severity Crohn’s disease**
All agreement and kappa scores are presented in table 4. Overall interobserver agreement for disease activity was highest between observer 1 and 3 (agreement 203/301; 67%). Agreement between observer 1 and 2 was similar to the agreement between observer 2 and 3 (191/301 (63%) versus 186/301 (62%) respectively). Interobserver agreement was weak to moderate for most subjective features, although for some there was no agreement (creeping fat between observer 2 and 3).

**MR imaging features**

**Per segment analysis.**
MRI estimates of bowel wall thickness ranged from 2 to 19 mm (median 8 mm). Comparison with the AIS score: MRI measured mural thickness was significantly positively correlated with AIS for all observers (observer 1: correlation 0.452, p=0.004; observer 2: r=0.409 p=0.010; observer 3: r=0.469 p=0.003) (figure 2 and 3).

Bowel wall enhancement after intravenous contrast was present in all segments. Enhancement was not significantly associated with AIS (p>0.05) but had a significant correlation with disease chronicity (r=0.413, p=0.009) (figure 2, 3 and 4).

Stratification pattern at T1-weighted sequences after enhancement was not associated with AIS (p=0.301). Specifically pattern type A (multilayered mural stratification) was not associated with the AIS score (p=0.590). In segments with an AIS <3 a stratified pattern of enhancement was not seen. T2-weighted images were available in 28 of 39 segments (in 18 patients). Measured T2 mural signal intensity was not correlated with AIS. Stratification type on T2-weighted was also not associated with AIS (p=0.919). The presence of ulcerations were not significantly associated with AIS for all observers (p=0.054, p=0.074, p=0.287). Disease activity per segment was significantly associated with AIS for all observers (p=0.017, p=0.002, p=0.001).

**Per patient analysis.**
The presence of the comb sign and presence of creeping fat were not associated with AIS for all observers. Disease activity per patient had no correlation with AIS (p=0.778).

**Discussion**
This study demonstrates that post-contrast bowel wall thickness and subjective disease activity assessment positively correlate with histopathological disease activity. The mural wall thickness on MRI was significantly positively correlated with the AIS for all observers, which is likely due to oedema and inflammation. This was also found by Punwani et al who compared MRI to histopathology but also in other studies comparing with non-histological reference standards.\(^7,10,12\).
Earlier studies have shown that enhancement is associated with clinical grade, CDAI or clinical assessment with endoscopy, barium examinations and surgery. The study by Punwani et al. was the first to correlate MRI features with histology; however did not find a correlation with the AIS as we did in our study. Another study by Zappa et al. did find a correlation with an histological reference standard but used subjective assessment of enhancement. We did find a weak correlation of enhancement ratio with disease chronicity. This was also found by Taylor et al. The hypothesis behind this is that chronic disease activity causes an increased permeability of the affected microvascular surface area and therefore increased enhancement.

We did not find an association of AIS with a layered type of enhancement. Three-layered enhancement was seen in the study of Punwani et al. in patients with high AIS, indicating that a stratified appearance of bowel wall reflects disease activity. Although we did not find that association, in bowel segments with an AIS <3 no stratified appearance was found. This could indicate that a layered enhancement pattern is associated with active Crohn’s disease in general but not with the severity of disease, as was hypothesised in another study.

In tissues where oedema or inflammation is present, a high T2 mural signal intensity is present. In our study a high T2 signal intensity was not significantly associated with increased AIS. It can be rather difficult to distinguish high bowel wall signal intensity when no fat saturation and a positive luminal contrast agent are used, which was the case in our dataset. The study of Punwani did find a correlation with AIS when using relative measurements on fat saturated images which could explain this discrepancy.

Depiction of ulcerations can be difficult on MRI because small mucosal lesions can be missed due to low special resolution (sensitivity reported to be 56%). There was a trend that the presence of ulcerations seen on MR enterography was associated with the AIS, but this was not significant at the preassigned level (p = 0.054) and only for observer 1.

Disease activity per segment was associated with AIS for all observers (p < 0.05). The assessment of disease activity is subjective. It is based on multiple features (most heavily on T1 enhancement and wall thickness) and the radiologist has to combine all MR features to judge what the disease activity is for the whole segment.

The presence of the comb sign was not associated with the AIS. The comb sign indicates increased blood flow in the vasa recta, which leads to more perfusion in active disease. This could mean that the comb sign merely indicates active disease and does not give an indication of the severity of the disease. The presence of creeping fat was also not associated with higher AIS. This corresponds with the hypothesis that creeping fat is not a marker of how severe the disease activity is but of a past episode of active disease.

Most of the subjective features had weak to moderate interobserver agreement. This is in accordance with other studies comparing interobserver variability of subjective MRI features. Highest correlation was for wall thickness and highest kappa’s for disease activity per segment. As these were also the features correlating with AIS, this gives extra weight to the importance of these features. Lowest kappa values were there for presence of creeping fat and the comb sign, which are very subjective features. The value of MRI features with a relative low interobserver variability is questionable. We recommend the use of objective measurements when grading to overcome this problem.

Our study does have several limitations. The matching of the MR segments with histopathology slices was performed retrospectively using pathology reports and photographs. However, the matching was performed with meticulous precision and in consensus by two observers. We compared the MRI measurements with histological sections in the same segment, which can be used as a representative of the whole diseased area.

Most patients underwent an (neo)ileocoecal resection were only a two segments were scored (ileum and coecum/ascending colon). The difference between ileal and colonic sections could be seen when the pathologist evaluated the sections and care was taken to evaluate the appropriate segment. No identical scan protocol was used in all patients and therefore no comparison of all sequences could be made in all patients. This specifically concerned the signal intensity on T2-weighted images, because these were only available in 25 patients. All other imaging features could be assessed in all patients. The patients in this study were all surgical patients of who can be expected to have more severe disease than patients not considered for surgery. To overcome this bias towards more severely diseased patients we included all bowel wall sections that were available, to ensure that the whole disease severity spectrum (normal - severe disease) was present in our study spectrum.

The time between the MR examination and the MRI ranged from 2 to 125 days. In this timeframe disease activity could have been altered because of natural course or medication. However, since all patients were operated, it is unlikely that the disease activity was minimalised in the aforementioned period.

One of the scoring radiologists already had seen the MR previously in routine patient care in eight of the 25 cases. However, the radiologists had seen many other examinations...
during this time interval and the time interval between the scoring and the routine patient care was large (mean 837 days before the scoring, range 133-2220 days). All scoring radiologists knew the patients were surgically treated. As in previous studies with histopathology verification this created bias, although the radiologists did not know what kind of operation was performed.

Conclusions
In this retrospective study a positive association existed between wall thickness measured and subjective disease activity per-segment at MRI and AIS, indicating that these features correlate with disease severity. Interobserver variability was weak to moderate.

Reference List


