Title

Monitoring Crohn's disease during anti-TNF-α therapy: validation of the magnetic resonance enterography global score (MEGS) against a combined clinical reference standard.

Abstract

Objectives: To assess the ability of the magnetic resonance enterography global score (MEGS) to characterise Crohn's disease (CD) response to anti-TNF-α therapy.

Methods: Thirty-six CD patients (median age 26 years, male 20) commencing anti-TNF-α therapy with concomitant baseline MRI enterography (MRE) were identified retrospectively. Clinical course was followed and correlated with subsequent MRE's. Scan order was randomised and MEGS (a global activity score) applied by two blinded radiologists. A physician global assessment of disease activity (remission, mild, moderate and severe) at the time of MRE was assigned. The cohort was divided into clinical responders and non-responders and MEGS compared according to activity status and treatment response. Interobserver agreement was assessed.

Results: Median MEGS decreased significantly between baseline and first follow-up in responders (28 versus 6, P<0.001) but was unchanged in non-responders (26 versus 18, P=0.28). Median MEGS was significantly lower in clinical remission (9) than in moderate (14) or severe (29) activity, P<0.001. MEGS correlated significantly with clinical activity (r=0.53; P<0.001). Interobserver Bland-Altman limits of agreement (BA LoA) were -19.7 to 18.5.

Conclusions: MEGS reduces significantly in clinical responders to anti-TNF-α therapy but not in non-responders, demonstrates good inter observer agreement and moderate correlation with clinical disease activity.

Keywords

Magnetic resonance imaging; Crohn's disease; therapeutic monitoring; biological therapy; enterography.

Key Points

- MRI scores of Crohn's activity are used increasingly in clinical practice and therapeutic trials.
- Such scores have been advocated as biomarkers of therapeutic response.
- MEGS reflects clinical response to anti-TNF-α therapy and the clinical classification of disease activity.
- MEGS demonstrates good interobserver agreement.

Abbreviations and acronyms

ANOVA = analysis of variance

BA LoA = Bland-Altman limits of agreement

CD = Crohn's disease

CRP = C-reactive protein

fC = faecal calprotectin

ICC = intra-class correlation coefficient

IQR = interquartile range

IBD = inflammatory bowel disease

MEGS = magnetic resonance enterography global score

MRE = magnetic resonance enterography

MRI = magnetic resonance imaging

TNF-α = tumor necrosis factor α

Introduction

Anti-tumour necrosis factor α (anti-TNFα) agents infliximab and adalimumab are used widely to treat severe active Crohn's disease (CD) (1). The ultimate therapeutic goal in CD is to control disease and stop progression. The currently favoured marker of disease control is mucosal healing assessed endoscopically (2), which is associated with better clinical outcomes (3). Data suggests treatment with anti-TNFα agents achieves mucosal healing in a significant proportion of patients (4). Conversely, approximately 10-30% of patients commencing anti-TNFα agents fail at induction and a further 20-50% lose response by one year (5).

Accurate and timely assessment of treatment response is therefore paramount to guide clinical management.

Magnetic resonance enterography (MRE) is now recommended by consensus guidelines as a first line test for the diagnosis and monitoring of CD (6). It is noninvasive and provides an overview of the total disease burden, not only in terms of distribution and complications, but also in terms of full transmural extent. Morphological magnetic resonance imaging (MRI) observations such as wall thickness, contrast enhancement and T2 mural signal are validated as markers of activity and MRI activity scores are therefore entering clinical practice (7-9).

To date there is relatively limited data regarding the utility of MRI activity scores for assessment of response to anti-TNFα agents. Van Assche et al., for example, reported significant reductions in a morphological MRI activity score 26 weeks after infliximab induction; persistence of transmural abnormality on MRE was common, despite good clinical response (10). Tielbeek et al. found that changes in an MRI activity score mirrored long term clinical response (11) while more recently Ordás et al. have shown that the MaRIA score mirrors endoscopic disease response after 12 weeks of therapy, achieving 80% sensitivity for mucosal healing (9).

The MEGS (magnetic resonance enterography global score) (12) was developed to better capture the full disease burden and includes longitudinal disease extent and extra-enteric findings. However, its ability to monitor disease response to anti-TNF-α agents has not been investigated.

The primary purpose of this study was to assess the ability of MEGS to characterise disease response to anti-TNF-α therapy compared to a global physician assessment based on all available clinical data. Our secondary purpose was to evaluate inter-observer agreement for MEGS.

Materials and methods

The local ethics committee granted a waiver for retrospective review of imaging and clinical data acquired as part of normal clinical care.

A single-centre departmental *BLINDED* Audit database, compiled as part of the *BLINDED* Audit (13) was searched, at a single tertiary center for IBD, for patients undergoing MRE as part of usual clinical care at the time of commencing anti-TNFα therapy and with at least one follow up MRE, between October 2006 and November 2013.

Inclusion and exclusion criteria

Patients were potentially eligible for inclusion if aged ≥14 years and commencing either infliximab (Remicade, Schering-Plough) or adalimumab (Humira, AbbVie) for active small bowel or colonic CD. Cross reference was then made with the hospital Radiology Information System and the cohort restricted to those patients who had completed: (a) a baseline MRE examination within 3 months of starting anti-TNFα therapy (up to two months before or one month after); (b) at least one follow-up MRE examination no earlier than 3 months after baseline.

Patients were excluded if they: did not receive IV contrast as part of their MRE (n = 1); they had no assessable disease on baseline MRE (based on review by the study readers - see below), had

undergone surgery within the last 6 months, or had perianal disease only (n = 5). Patients were also excluded if, in the opinion of the gastroenterologist undertaking the global physician assessment of disease activity (described in detail below), there was insufficient clinical data, either at baseline or follow-up, to determine clinical status (n = 5). Disease distribution and behaviour based on the Montreal classification (14), surgical history, disease duration and patient age at diagnosis were recorded for each patient. Concomitant treatment with other medications was recorded. If patients had undergone more than one follow up MRE, all follow up scans were included.

Imaging Protocol

Magnetic resonance enterography (MRE) was performed using a standardised clinical protocol on one of two static magnets: 1.5 Tesla (Avanto; Siemens Medical Systems, Erlangen, Germany) or 3 Tesla (Achieva; Philips, Best, The Netherlands). Patients fasted for 4 hours and then ingested 1 to 1.5 L of 0.2 % locust bean gum/2.5 % mannitol solution over 45 min immediately before imaging (15). Twenty milligrams of intravenous hyoscine butylbromide (Buscopan; Boehringer Ingelheim, Ingelheim, Germany) were administered together with 0.1 mmol/kg gadolinium (3 mL/sec injection using a power injector). Full sequence parameters are shown in **Appendix 1 and 2**.

MR enterography global score (MEGS)

MEGS represents the evolution of a score proposed initially by Steward et al. (7), a segmental MRI activity scoring system derived using surgical resection specimens as reference standards and validated against endoscopic biopsies. The original score included the sum of qualitative grades for segmental mural thickness, mural T2 signal (mural oedema), mural contrast enhancement and perimural T2 signal. In order to better reflect the global burden of disease (i.e. beyond point estimates of segmental activity) the score was expanded to include segmental disease length, evaluation of colonic haustral loss (16) and evaluation of extra-enteric complications, such as enlarged mesenteric lymph nodes (16), abscesses and fistulae. MEGS has been previously validated

prospectively against stool calprotectin, blood CRP and Harvey-Bradshaw index (12) and is detailed in **Table 1**. As part of the validation of MEGS, a simplified model based on the average score across all segments was also proposed (1.8.wall thickness + 0.08.mural T2 signal + 0.19. length -0.192) (7). Sample images from the current dataset are presented in **Figure 1** and **Figure 2**.

Image analysis

Two gastrointestinal radiologists, with 3 and 4 years of experience respectively in reading MRE, evaluated the scans independently using all available sequences. Anonymised scans were presented in randomised order on an Impax 5.0 (AGFA Healthcare, Agfa-Gevaert) PACS workstation and readers were blinded to all clinical information. Readers derived the total MEGS score (**Table 1**) for each scan presented, from which the simplified model was also calculated.

Combined clinical reference standard

A physician global assessment incorporating all available clinical information was used to define disease activity. Specifically, a consultant gastroenterologist with subspecialty interest in IBD and 15 years of experience reviewed all clinical data available at the time of the MRE, including all inpatient episodes, clinic letters, endoscopy, blood and stool results such as C-reactive protein (CRP) and calprotectin, but blinded to the MRI scan report. Based on these clinical data patients were classified into four ranks of disease activity, graded in keeping with the Second European Based Consensus of diagnosis and management of Crohn's disease (17): "remission" = lack gastrointestinal symptoms and normal CRP; "mild" = ambulatory patient, eating and drinking, no significant weight loss, lack of fever, obstruction, mass or tenderness and CRP increased above the upper limit of normal; "moderate" = intermittent vomiting or weight loss, ineffective treatment for mild disease, tenderness or mass but no overt obstruction, raised CRP; "severe" = severe weight loss or obstruction or abscess, persistence of symptoms despite intensive treatment and increased CRP at time points corresponding to each MRE. As noted above, if the gastroenterologist did not feel they

had enough clinical information to assign this classification confidently, the time point was excluded from the analysis.

Definition of clinical response

Patients were defined as clinical responders if their clinical status improved by at least one rank along the scale of disease activity (e.g. by moving from moderate to mild), between baseline and first follow up MRI. A secondary category of strong clinical responders based on a drop of at least two ranks along the scale of activity was also assigned. The first follow up MRI was used to classify response when more than one follow up scan was available.

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics, version 22. Mean MEGS scores of the two readers were used in the primary analysis; MEGS values were non-normally distributed. Wilcoxon signed rank test was utilised for paired MEGS data comparisons between baseline and first follow-up according to response category (overall clinical responder or non-responder [primary analysis] and strong clinical responder versus those with a clinical response of just one activity rank [secondary analysis]). The primary analysis was repeated used the simplified model suggested during the initial validation of MEGS (1.8.wall thickness + 0.08.mural T2 signal + 0.19. length -0.192).

Kruskal-Wallis ANOVA with post-hoc analysis was used for multiple group comparison of MEGS score according to disease activity across all time points (remission, mild, moderate and severe).

Spearman rank correlation between MEGS values and clinical activity across all time points (remission, mild, moderate and severe) was also performed.

Interobserver agreement between the readers' MEGS scores was assessed using the 95% Bland-Altman limits of agreement (BA LoA) including all MRI scans. For all analyses, $P<0.05$ was taken to represent statistical significance.

Results

The final study cohort consisted of 36 patients; 26 clinical responders and 10 non-responders. Among the clinical responders, there were 14 strong clinical responders and 12 patients with a fall of disease activity of 1 rank. Patient demographics and clinical characteristics (split by clinical responders and non-responders) are presented in **Table 2.** Seventeen patients received 3 MRE's (baseline and two follow-up); the remaining 19 received 2 MRE's, giving 89 MRE overall. Temporal intervals between baseline MRE and first follow-up ranged between 3 and 62 months (median 28 months); the same range applied for overall follow-up time. Mean MEGS values ranged between 0 and 155.

Change in MEGS according to clinical response

In clinical responders median MEGS decreased significantly from 28 (IQR = 17 to 54) at baseline to 6 (IQR = 2 to 28) at first follow-up (P<0.001). Conversely, in those with no clinical response there was no significant change in MEGS score from baseline (median = 26; IQR = 13 to44) and first follow-up (median = 18; IQR = 90 to 26; P=0.28) (**Figure 3**).

The median difference between MEGS at baseline and at first follow-up was 24 (IQR = 10 to 28) in clinical responders and 10 (IQR = 6 to 24) in non-responders. Overall 15/26 (58%) of responders had a decrease in MEGS of at least 60% compared to 3/10 (30%) of non-responders, giving a sensitivity of 58% and specificity of 70% for clinical response.

In the 17 patients with a second follow-up MRI, median MEGS at the second follow-up were 5 (IQR = 0 to 21) in clinical responders (n = 12) and 10 (IQR = 9 to 26) in non-responders (n = 5).

Strong clinical responders had a larger drop in MEGS (median = 29 [IQR 11 to 54] pre-treatment versus median = 6.4 [IQR 1 to 28] post-treatment, P=0.002) compared to responders with a decrease

in clinical activity of 1 rank (median = 31 [IQR 17 to 52] pre-treatment versus median = 14 [IQR 6 to 29] post-treatment, P=0.002).

Using the simplified model (1.8.wall thickness + 0.08.mural T2 signal + 0.19. length -0.192), clinical responders median model score decreased significantly from 3.9 (IQR = 2.1 to 4.3) at baseline to 2.6 (IQR = 0 to 5.7) at first follow-up (P=0.01). Conversely, in those with no clinical response there was no significant model score change from baseline (median = 4.9; IQR = 3.9 to 5.9) and first follow-up (median = 3.9; IQR = 0 to 5.7; P=0.07).

Correlation between MEGS and clinical status

A significant, moderate positive correlation was found between MEGS and clinical activity according to the physician global assessment (r=0.53; P<0.001).

Including all time points, median MEGS differed significantly according to disease activity grade (Kruskal-Wallis ANOVA, P<0.001). Pairwise post-hoc analysis revealed significant median MEGS differences between patients in clinical remission and those with moderate (P<0.001) and severe activity (P<0.001, **Figure 4**).

Inter-observer agreement

The mean difference between the two observers' MEGS was -0.57 (BA LoA -19.70 to 18.57). Agreement was superior in patients with quiescent disease or mild activity (mean difference = -0.58; BA LoA -12.50 to 11.34) (**Figure 5**).

Discussion

We found that changes in a quantitative MRI score of Crohn's disease burden and activity (MEGS) reflected, to some extent, overall clinical response to anti-TNFα therapy. Furthermore, at a single

time point MEGS differed significantly between patients in clinical remission and those with clinically moderate or severe disease activity. The drop in MEGS was larger in those with a strong clinical response; a simplified model based on T2 signal, wall thickness and disease length also significantly fell in clinical responders.

Accurate assessment of disease activity is fundamental to management of Crohn's disease. Medical therapies are aimed at suppressing the immune system to reduce inflammatory activity but have significant side effects as a consequence, including sepsis. It is thus vital that the use of such medications is monitored in a safe, objective and reproducible way (2).

MRI is used increasingly as a first line test for diagnosis and classification of CD and influences patient care significantly, over and above clinical assessment and endoscopy (18, 19). More recently MRI scores of disease activity have been advocated to monitor therapeutic response (7-9).

The MEGS score represents an evolution of the MRI Crohn's activity score (7) and has prospectively showed reasonable correlation with faecal calprotectin as a marker of global disease activity (12). In the current study MEGS again demonstrated a moderate positive correlation with an independent reference standard, this time a clinician-derived activity grade. In the case of CD activity assessment there is no "perfect" standard of reference against which to compare new methods. Indeed, had MEGS shown very high correlation with existing standards of reference, this would imply it does not provide any additional information over and above these methods. Instead, a moderate correlation could suggest additional utility of MEGS over standard clinical tools.

Tielbeek et al. (11) employed a modified activity score based on the MRI Crohn's activity score (12), a score similar to MEGS, in patients receiving serial MRI during anti-TNF-α therapy, and also used a physician global assessment to define therapeutic response. Like in our study, they too found that MRI scores improved significantly in clinical responders but overall did not significantly change in non-responders. Tielbeek et al. found improvement of MRI activity scores in around 30% of clinical

non-responders. Similarly, in the current cohort around 30% of non-responders had a sizeable fall in MEGS of 60% or more. It is clear therefore that MRI and clinical response do not fully overlap. Nevertheless the concordance of results between two studies in different patient cohorts is reassuring and suggests MRI is a reasonable marker of treatment outcome.

Ordás et al. have also recently reported that the MaRIA score is a good marker of therapeutic response, as early as 12 weeks (9). Although similar to the MRI Crohn's activity score, the MaRIA score includes an assessment of mucosal ulceration and has been well validated against the CD endoscopic activity score (8).

Any proposed MRI activity score should demonstrate good reproducibility. Tielbeek et al. recently reported good inter-observer agreement for both the MRI Crohn's activity score and MaRIA score (20). In the current study, MEGS also demonstrated good interobserver agreement, with mean differences between observers close to zero. Agreement was strongest in patients with quiescent or mild disease. We did use experienced readers, as a clear learning curve for MRI interpretation has been shown, with a likely need to review 100 cases to achieve acceptable accuracy (21).

Whilst it is clear MRI is a powerful tool to assess treatment response, the optimum scoring system is not yet defined. The MaRIA score is well validated (8, 9) and increasingly implemented. It includes mucosal ulceration as a very useful marker of activity which is best assessed in well distended segments (most validation work has used a colonic water enema). Evaluation of ulceration is not included in the MEGS score which has been developed in datasets without specific colonic preparation. MaRIA does not include evaluation of extra-enteric changes such as abscess or fistulae, which are likely important in assessing disease activity. Nevertheless, there are common facets to both scores such as wall thickness, T2 mural signal and contrast enhancement. The use of diffusion weighted imaging also shows promise (22), although to date has not been fully validated and it is unclear if it adds to conventional MRI sequences. Finally, it is important to assess chronic bowel damage as well as disease activity. The Lémann score (23) has been developed to assess chronic

bowel damage, and includes some imaging findings. Furthermore recent data suggest increased bowel wall enhancement at 7 minutes may be a good marker of chronic fibrosis (24) and it is possible it could be added to activity scores in the future.

Our study does have limitations. Its retrospective nature means that no systematic clinical, laboratory or imaging data collection was possible; the scoring clinician had to express an overall subjective judgement at each time point on the basis of available information, which varied in nature. Importantly, when the clinician felt they were unable to assign a disease activity category, that time point was excluded from the analysis. While a physician's global assessment of disease activity has some limitations and is arguably subjective, alternative reference standards are also imperfect. Much of the data validating MRI activity scores have used endoscopic or histopathological reference standards (7, 8, 25-28). By estimating only segmental mucosal activity within reach of the endoscope tip, such scoring systems fail to capture the potential of MRI to assess true disease burden by visualising the entire intra-abdominal gastrointestinal tract, including both mural and extra-enteric tissues. A physician global assessment mirrors what happens in day to day clinical practice, whenever therapeutic decisions are made. Reassuringly, the drop in MEGS was larger in patients with clinical response of at least two activity ranks compared to those with a drop of just one rank. We used both 1.5 T and 3 T platforms but, as MEGS includes only qualitative data, platforms have not been found to affect scoring (12). The temporal interval between baseline and follow-up imaging also varied considerably, in some cases being as long as 5 years, well beyond a realistic time interval for response assessment. However the majority of follow-up intervals were around 1 year, which is the recommended time for detailed re-assessment of patients taking anti-TNF-α therapy (13). It should be noted that, given the length of follow-up, some patients may have developed more fibrotic disease and, although the physician global assessment focused on symptoms of activity, it is possible fibrotic disease influenced the evaluation of activity status.

We were unable to control for the indication for follow-up MRI. Indications typically include worsening clinical symptoms, patient or clinician desire to reduce medication in clinically stable disease, or routine 1-year assessment. It is possible that some non-responders had actually initially responded before losing response, triggering the MRI, but we achieved a good split between responders and non-responders in our cohort. Indeed all patients in the current cohort were started on anti-TNF-α therapy for presumed active disease. Overall, perhaps as would be expected, all those with Montreal classification of B1 responded, whilst around two thirds of those with B2 and B3 responded (suggesting active disease) and around a third did not. Furthermore the need for patients to have MRI at the same time as starting therapy potentially creates a spectrum bias in our cohort. Although all patients were started on anti-TNF-α medication, many were on existing treatment for Crohn's disease. Overall, however, there was little difference in this concomitant medication between responders and non-responders.

The proposed scoring system is time consuming for radiologists, which may limit clinical uptake. A simplified score dropped significantly in responders but not in non-responders and thus shows promise as a quicker simpler evaluation more suited to clinical practice, with perhaps the full MEGS score reserved for those with extra-enteric complications and research studies. Inter-observer agreement may be less for inexperienced radiologists and training would likely be needed (21).

Finally, as noted above, MEGS focuses particularly on inflammatory activity and does not specifically assess fibro-stenotic lesions or previous surgery. Alternative approaches such as Lémann index (23) may address this "chronic bowel damage", and could be scored in parallel as required.

In summary, MEGS significantly reduces in clinical responders to anti-TNF-α therapy, but not in non-responders. It demonstrates good inter observer agreement and moderate correlation with clinical disease activity.

Acknowledgements

BLINDED

References

1. Mayberry JF, Lobo A, Ford AC, Thomas A. NICE clinical guideline (CG152): the management of Crohn's disease in adults, children and young people. Aliment Pharmacol Ther. 2013;37(2):195-203.
2. D'Haens GR, Sartor RB, Silverberg MS, Petersson J, Rutgeerts P. Future directions in inflammatory bowel disease management. J Crohns Colitis. 2014.
3. De Cruz P, Kamm MA, Prideaux L, Allen PB, Moore G. Mucosal healing in Crohn's disease: a systematic review. Inflamm Bowel Dis. 2013;19(2):429-44.
4. D'Haens G, Baert F, van Assche G, et al. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. Lancet. 2008;371(9613):660-7.
5. Ben-Horin S, Chowers Y. Review article: loss of response to anti-TNF treatments in Crohn's disease. Aliment Pharmacol Ther. 2011;33(9):987-95.
6. 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(7):556-85.
7. Steward MJ, Punwani S, Proctor I, et al. Non-perforating small bowel Crohn's disease assessed by MRI enterography: derivation and histopathological validation of an MR-based activity index. Eur J Radiol. 2012;81(9):2080-8.
8. Rimola J, Ordás I, Rodriguez S, et al. Magnetic resonance imaging for evaluation of Crohn's disease: validation of parameters of severity and quantitative index of activity. Inflamm Bowel Dis. 2011;17(8):1759-68.
9. Ordás I, Rimola J, Rodríguez S, et al. Accuracy of magnetic resonance enterography in assessing response to therapy and mucosal healing in patients with Crohn's disease. Gastroenterology. 2014;146(2):374-82.e1.
10. Van Assche G, Herrmann KA, Louis E, et al. Effects of infliximab therapy on transmural lesions as assessed by magnetic resonance enteroclysis in patients with ileal Crohn's disease. J Crohns Colitis. 2013;7(12):950-7.
11. Tielbeek JA, Löwenberg M, Bipat S, et al. Serial magnetic resonance imaging for monitoring medical therapy effects in Crohn's disease. Inflamm Bowel Dis. 2013;19(9):1943-50.
12. Makanyanga JC, Pendsé D, Dikaios N, et al. Evaluation of Crohn's disease activity: initial validation of a magnetic resonance enterography global score (MEGS) against faecal calprotectin. Eur Radiol. 2014;24(2):277-87.
13. Biological therapies audit sub group. National clinical audit report of biological therapies. Adult national report. UK IBD audit. Royal College of Physicians2013.
14. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut. 2006;55(6):749-53.
15. Lauenstein TC, Schneemann H, Vogt FM, Herborn CU, Ruhm SG, Debatin JF. Optimization of oral contrast agents for MR imaging of the small bowel. Radiology. 2003;228(1):279-83.
16. Ajaj WM, Lauenstein TC, Pelster G, et al. Magnetic resonance colonography for the detection of inflammatory diseases of the large bowel: quantifying the inflammatory activity. Gut. 2005;54(2):257-63.
17. Van Assche G, Dignass A, Panes J, et al. The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. J Crohns Colitis. 2010;4(1):7-27.
18. Messaris E, Chandolias N, Grand D, Pricolo V. Role of magnetic resonance enterography in the management of Crohn disease. Arch Surg. 2010;145(5):471-5.
19. Hafeez R, Punwani S, Boulos P, et al. Diagnostic and therapeutic impact of MR enterography in Crohn's disease. Clin Radiol. 2011;66(12):1148-58.

20. Tielbeek JA, Makanyanga JC, Bipat S, et al. Grading Crohn disease activity with MRI: interobserver variability of MRI features, MRI scoring of severity, and correlation with Crohn disease endoscopic index of severity. AJR Am J Roentgenol. 2013;201(6):1220-8.
21. Tielbeek JA, Bipat S, Boellaard TN, Nio CY, Stoker J. Training readers to improve their accuracy in grading Crohn's disease activity on MRI. Eur Radiol. 2014;24(5):1059-67.
22. Hordonneau C, Buisson A, Scanzi J, et al. Diffusion-weighted magnetic resonance imaging in ileocolonic Crohn's disease: validation of quantitative index of activity. Am J Gastroenterol. 2014;109(1):89-98.
23. Pariente B, Mary JY, Danese S, et al. Development of the Lémann index to assess digestive tract damage in patients with Crohn's disease. Gastroenterology. 2015;148(1):52-63.e3.
24. Rimola J, Planell N, Rodríguez S, et al. Characterization of inflammation and fibrosis in Crohn's disease lesions by magnetic resonance imaging. Am J Gastroenterol. 2015;110(3):432-40.
25. Zappa M, Stefanescu C, Cazals-Hatem D, et al. Which magnetic resonance imaging findings accurately evaluate inflammation in small bowel Crohn's disease? A retrospective comparison with surgical pathologic analysis. Inflamm Bowel Dis. 2011;17(4):984-93.
26. Rimola J, Rodriguez S, García-Bosch O, et al. Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease. Gut. 2009;58(8):1113-20.
27. Taylor SA, Punwani S, Rodriguez-Justo M, et al. Mural Crohn disease: correlation of dynamic contrast-enhanced MR imaging findings with angiogenesis and inflammation at histologic examination--pilot study. Radiology. 2009;251(2):369-79.
28. Ziech ML, Bipat S, Roelofs JJ, et al. Retrospective comparison of magnetic resonance imaging features and histopathology in Crohn's disease patients. Eur J Radiol. 2011;80(3):e299-305.

Tables

Table 1. Magnetic resonance enterography global score (MEGS). The score is applied to each bowel segment (jejunum, ileum, terminal ileum, caecum, ascending, transverse, descending, sigmoid colon and rectum), summated and added to the score for extramural features. *Measured using electronic calipers; **compared with normal small bowel; ***compared with nearest vessel.

Score	0	1	2	3
Mural thickness* small bowel	<3 mm	>3–5 mm	>5–7 mm	>7 mm
Mural T2 signal**	Equivalent to normal bowel wall	Minor increase in signal: bowel wall appears dark grey on fat-saturated images	Moderate increase in signal: bowel wall appears light grey on fat-saturated images	Marked increase in signal: bowel wall contains areas of white high signal approaching that of luminal content
Peri-mural T2 signal (mesenteric oedema)	Equivalent to normal mesentery	Increase in mesenteric signal but no fluid	Small fluid rim (≤2 mm)	Larger fluid rim (>2 mm)
T1 Enhancement***	Equivalent to normal bowel wall	Minor enhancement: bowel wall signal greater than normal small bowel but significantly less than nearby vascular structures	Moderate enhancement: bowel wall signal increased but somewhat less than nearby vascular structures	Marked enhancement: bowel wall signal approaches that of nearby vascular structures
Mural enhancement pattern	N/A or homogeneous	Mucosal	Layered	
Haustral loss (colon only)	None	<1/3 segment	1/3 to 2/3 segment	>2/3 segment
Multiplication factor per segment				
Length of disease segment		0–5 cm × 1	5–15 cm × 1.5	>15 cm × 2
Additional score for extramural features				
Score	0	5		
Lymph nodes (≥1 cm measured in shortest diameter)	Absent	Present		
Comb sign (linear densities on the mesenteric side of affected bowel segments)	Absent	Present		
Abscess	Absent	Present		
Fistulae	Absent	Present		

Table 2. Patient demographic and clinical characteristics divided between clinical responders and non-responders. * = according to the Montreal classification.

	Clinical responders (n=26)	Non-responders (n=10)
Males (%)	13 (50%)	6 (60%)
Median age (range)	24 years (21-47)	35 years (18-50)
Median age at diagnosis (range)	19 years (15-29)	27 years (16-45)
Median disease duration at baseline MRE (range)	4 years (0-30)	5 years (2-14)
Median duration of MRE follow-up	34 months (18-62 months)	35 months (7-55 months)
Disease location *		
L1 (%)	2 (7.7%)	1 (10%)
L2 (%)	2 (7.7%)	0 (0%)
L3 (%)	22 (84.6%)	9 (90%)
L4 (%)	6 (23.1%)	1 (10%)
P (%)	7 (26.9%)	2 (20%)
Disease behaviour*		
B1 (%)	9 (34.6%)	0 (0%)
B2 (%)	10 (38.5%)	7 (70%)
B3 (%)	7 (26.9%)	3 (30%)
Previous bowel resections		
YES (%)	11 (42.3%)	7 (70%)
NO (%)	15 (57.7%)	3 (30%)
Anti-TNF agent		
Infliximab (%)	13 (50%)	2 (20%)
Adalimumab (%)	13 (50%)	8 (80%)
Pre-existing immunosuppression at the time of anti-TNF start		
None (%)	7 (26.9%)	3 (30%)
Azathioprine (%)	14 (53.8%)	5 (50%)
Methotrexate (%)	3 (11.5%)	2 (20%)
Mercaptopurine (%)	1 (3.8%)	0 (0%)
Missing data (%)	1 (3.8%)	0 (0%)

Pre-existing steroids at the time of anti-TNF start	7 (26.9%)	1 (10%)

Figures legends

Figure 1. Sample MRE images from a clinical responder: axial SSTSE and coronal fat suppressed SSTSE sequences. Baseline MRE (left) demonstrates marked wall thickening of the terminal ileum (black arrow; > 7 mm; score = 3) extending for approximately 10 cm (multiplication factor = 1.5), mural oedema with areas of fluid-signal in the bowel wall (score = 3), perimural oedema with a large rim of free fluid (white arrowheads; score = 3) and mesenteric vessel hyperaemia (comb sign; score = 5). On follow-up (right) changes have markedly improved with minimal wall thickening (black arrow; 3-5 mm; score = 1) and mural oedema (white arrowhead; minimally increased T2 signal; score = 1).

Figure 2. Sample MRE images from a clinical non-responder at baseline: axial SSTSE, coronal fat suppressed SSTSE and coronal dynamic contrast-enhanced sequences. Long segments of diseased ileum are present in the upper quadrants (> 30 cm; multiplication factor = 2), showing marked wall thickening, mural and perimural oedema, hyperenhancement of intensity similar to the mesenteric and iliac arteries (score = 3) and fairly homogenous enhancement pattern (score = 0).

Figure 3. Box-and-whisker plots of MEGS values at baseline and first follow-up in clinical responders and non-responders. Median MEGS significantly decreased between baseline and first follow up in clinical responders and strong responders (drop of two clinical activity ranks), but not in non-responders.

Figure 4. Box-and-whisker plots of MEGS values for each rank of clinically assessed disease activity. Median MEGS differed significantly between ranks. Post-hoc analysis revealed significant differences (*) between patients in remission and those with moderate and severe activity.

Figure 5. Inter-observer agreement. Bland-Altman plots of MEGS values with relative inter-observer differences plotted against the mean. Agreement was superior for patients in remission or with mild activity (B) versus all patients (A).

Appendix

Appendix 1. MRI parameters at 1.5 Tesla. SSTSE = single shot turbo spin echo; True FISP = true fast imaging with steady-state precision. Dynamic contrast-enhanced MRI commenced at the start of contrast medium injection and the final time point was used to assess contrast enhancement.

	Coronal/axial SSTSE	**Coronal/axial true FISP with and without fat saturation**	**Baseline volume interpolated gradient ECHO**	**Dynamic contrast-enhanced MRI**
Field of view (mm)	Variable	Variable	Variable	Variable
No. slices	20/26	25/34	40	40
Stacks	1/3	1/3	1	1
Repetition time (ms)	1,200/800	3.98/4.25	3.07	2.73
Echo time (ms)	86/86	1.72/2.13	1.08	0.9
Image matrix	256/256	256/256	256	256

	Coronal/axial SSTSE	Coronal/axial true FISP with and without fat saturation	Baseline volume interpolated gradient ECHO	Dynamic contrast-enhanced MRI
Slice thickness (mm)	4/4	4/4	3.5	3.5
Averages	1	1	1	1
Flip angle			15 °	15 °

Appendix 2. MRI parameters at 3 Tesla. SSTSE = single shot turbo spin echo; True FISP = true fast imaging with steady-state precision. Dynamic contrast-enhanced MRI commenced at the start of contrast medium injection and the final time point was used to assess contrast enhancement.

	Coronal/axial SSTSE	Coronal/axial true FISP with and without fat saturation	Baseline volume interpolated gradient ECHO	Dynamic contrast-enhanced MRI
Field of view (mm)	Variable	Variable	Variable	Variable
No. slices	20/26	25/34	40	40
Stacks	1/3	1/3	1	1
Repetition time (ms)	1,200/800	3.98/4.25	3.07	2.73
Echo time (ms)	86/86	1.72/2.13	1.08	0.9
Image matrix	256/256	256/256	256	256
Slice thickness (mm)	4/4	4/4	3.5	3.5
Averages	1	1	1	1
Flip angle			15 °	15 °