# Magnetic Resonance Enterography and Histology in Patients With Fibrostenotic Crohn's Disease: A Multicenter Study

Alexandre Coimbra, PhD[1], Jordi Rimola, MD, PhD[2], Miriam Cuatrecasas, MD, PhD[2], Gert De Hertogh, MD, PhD[3], Gert Van Assche, MD, PhD[3], Ragna Vanslembrouck, MD[3], Henning Glerup, MD, PhD[4], Agnete Hedemann Nielsen, MD[4], Rikke Hagemann-Madsen, MD[4], Yoram Bouhnik, MD, PhD[5], Magaly Zappa, MD[5], Dominique Cazals-Hatem, MD PhD[5], Geert D'Haens, MD, PhD[6], Jaap Stoker, PhD[6], Sybren Meijer, MD, PhD[6], Gerhard Rogler, MD, PhD[7], Andreas Boss, MD, PhD[8], Achim Weber, MD[9], Rui Zhao, PhD[1], Mary E. Keir, PhD[1], Alexis Scherl, MD, PhD[1], Alex de Crespigny, PhD[1], Timothy T. Lu, MD, PhD[1] and Julián Panés, MD[2]

INTRODUCTION: **Magnetic resonance enterography (MRE) is useful for detecting bowel strictures, whereas a number of imaging biomarkers may reflect severity of fibrosis burden in Crohn's disease (CD). This study aimed to verify the association of MRE metrics with histologic fibrosis independent of inflammation.**

METHODS: **This prospective European multicenter study performed MRE imaging on 60 patients with CD with bowel strictures before surgical resection. Locations of 61 histological samples were annotated on MRE examinations, followed by central readings using the Chiorean score and measurement of delayed gain of enhancement (DGE), magnetization transfer ratio, T2-weighted MRI sequences (T2R), apparent diffusion coefficient (ADC), and the magnetic resonance index of activity (MaRIA). Correlations of histology and MRE metrics were assessed. Least Absolute Shrinkage and Selection Operator and receiver operator characteristic (ROC) curve analyses were used to select composite MRE scores predictive of histology and to estimate their predictive value.**

RESULTS: **ADC and MaRIA correlated with fibrosis ($R = -0.71$, $P < 0.0001$, and 0.59, $P < 0.001$) and more moderately with inflammation ($R = -0.35$, $P < 0.01$, and $R = 0.53$, $P < 0.001$). Lower or no correlations of fibrosis or inflammation were found with DGE, magnetization transfer ratio, or T2R. Least Absolute Shrinkage and Selection Operator and ROC identified a composite score of MaRIA, ADC, and DGE as a very good predictor of histologic fibrosis (ROC area under the curve = 0.910). MaRIA alone was the best predictor of histologic inflammation with excellent performance in identifying active histologic inflammation (ROC area under the curve = 0.966).**

DISCUSSION: **MRE-based scores for histologic fibrosis and inflammation may assist in the characterization of CD stenosis and enable development of fibrosis-targeted therapies and clinical treatment of stenotic patients.**

**SUPPLEMENTARY MATERIAL** accompanies this paper at http://links.lww.com/CTG/A833

*Clinical and Translational Gastroenterology 2022;13:e00505. https://doi.org/10.14309/ctg.0000000000000505*

## INTRODUCTION

Assessment of fibrosis is critical for the accurate and comprehensive evaluation of patients with Crohn's disease (CD) and, if reliable, key in informing disease management. Although patients with predominantly inflammatory strictures may benefit from anti-inflammatory treatment, those with predominantly fibrotic strictures will see limited benefit from such treatment and most likely require endoscopic balloon dilation or surgery. Inflammatory disease can

[1]Early Clinical Development, Genentech, Inc., South San Francisco, California [2]Hospital Clinic, University of Barcelona, Barcelona, Spain; [3]University Hospitals Leuven and University of Leuven, Belgium; [4]Silkeborg Hospital, Silkeborg, Denmark; Lillebaelt Hospital, Vejle, Denmark; [5]Hôpital Beaujon, Paris, France; [6]Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, Netherlands; [7]Department of Gastroenterology and Hepatology, University Hospital of Zurich and University of Zurich, Zurich, Switzerland; [8]Institute of Diagnostic and Interventional Radiology, University Hospital of Zurich and University of Zurich, Zurich, Switzerland; [9]Department of Pathology and Molecular Pathology, University Hospital of Zurich and University of Zurich, Zurich, Switzerland.
**Correspondence:** Alexandre Coimbra, PhD. E-mail: coimbraa@gene.com.
**Received September 29, 2021; accepted May 6, 2022; published online June 13, 2022**

currently be assessed by endoscopy or imaging techniques such as magnetic resonance enterography (MRE) and computerized tomography enterography (CTE). Several studies have evaluated MRE accuracy in detecting the presence and severity of inflammation (1,2). Fewer studies have assessed both inflammation and fibrosis of the bowel wall (3). Identification of reliable tools to measure the presence and extent of intestinal fibrosis will enable comprehensive evaluation of disease for physicians and patients to make better treatment decisions. Such tools will also facilitate development of new antifibrotic treatments.

Chiorean et al. (4) investigated CTE features to differentiate inflammatory from fibrotic lesions in CD with surgical pathology specimens as a reference standard. They identified the presence of stenosis and a 3-point ordinal CTE-based score of fibrostenosis as main indicators of histologic fibrosis in the specimens. However, the risk of cumulative radiation exposure and limited multiparametric analysis associated with CTE makes the development of techniques such as MRE a more desirable, safer alternative for detecting bowel strictures and assessing severity of both inflammation and fibrosis. Furthermore, continuous variable scores may be more sensitive and possibly enable assessment of fibrosis even in the absence of strictures.

Metrics related to tissue uptake of intravenously administered gadolinium-based contrast agents, including quantitative metrics such as delayed gain of enhancement (DGE), have been reported to be associated with fibrosis (5–8). The magnetization transfer ratio (MTR), derived from the magnetization transfer technique reflecting burden of macromolecules such as collagen accumulating in the bowel wall, has been reported to be associated with bowel fibrosis both in animal models and in patients (9–11). The apparent diffusion coefficient (ADC) has also shown promise for detecting fibrosis in CD. Lower ADC values correlate significantly with Chiorean scores for histologic fibrosis (7) presumably because the presence of fibrosis reduces extracellular space, thus restricting diffusion of water molecules in the bowel wall tissue (12).

The Magnetic Resonance Index of Activity (MaRIA), a weighted sum of 4 MRI features including bowel wall thickness, relative contrast enhancement, and presence of edema or ulcer, is a composite score that was originally designed to maximize the correlation between MRE and endoscopic scores of mucosal inflammation (1,13). However, some components of the MaRIA score, such as wall thickness and presence of edema, are associated with fibrosis (14). Similarly, the ratio of bowel wall signal intensity relative to normal muscle on T2-weighted MRI sequences (T2R) infers the presence of edema as an indicator of active inflammation (1,2).

This study aimed to identify MRE sequences and metrics associated with histologic fibrosis, independent of inflammation, in CD. The primary objective was to verify concordance between DGE and the histopathology fibrosis score. MTR and ADC were also evaluated against the histopathology score as secondary and exploratory objectives. A final technical objective was to evaluate the feasibility of a standardized MRE-fibrosis protocol in a multicenter setting to obtain high-quality data.

## METHODS

### Study design and patient population

This prospective, multicenter imaging study was conducted at 6 sites in Europe and enrolled patients with a previous diagnosis of CD undergoing elective bowel surgery for disease management (Figure 1). Disease duration, phenotype, location, concomitant medications, past surgical history, and smoking status were documented. At screening, a blood chemistry profile was used to estimate glomerular filtration rate.

Patients underwent MRE examination with intravenous and oral contrast during the study. The MRE results were used to guide the surgical resection.

The study was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practice. Approval from the institutional review boards and ethics committees was obtained before its start. Patient consent was obtained before enrollment.

### Imaging acquisition and analysis

Sites were required to perform MRE examinations $< 8$ weeks before surgery with no treatment changes between imaging examination and surgery (see Supplemental Figure 1, Supplemental Methods, http://links.lww.com/CTG/A833). Site-specific scanning parameters were optimized and agreed on *a priori* (see Supplemental Methods, Supplementary Digital Content 1, http://links.lww.com/CTG/A833).

Following surgery, local pathologists and radiologists for primary on-site evaluation identified the most stenosed region of the resected specimens and annotated the corresponding location of histological samples on the MRE scans. One central reader with 14 years of experience in gastrointestinal radiology then reviewed all MRE sequences, delineated regions of interest in the appropriate sequences, and recorded measurements for derivation of the following quantitative metrics: DGE, MTR, T2R, ADC, and MaRIA. Supplemental Materials, Supplementary Digital Content 1, http://links.lww.com/CTG/A833, describe the MRE sequences and definition of these variables.

To assess technical feasibility of MRI procedures in the context of a multicenter trial, the central reader also evaluated the quality of MRE images following an algorithm previously described (15). In short, each sequence in every examination was subjectively assessed for quality and rated as either Excellent, Good, Fair, or Unevaluable.

### Histopathology analysis

All surgical bowel resection specimens were subject to gross examination and photographic documentation by the local site pathologist. All tissue sections were formalin fixed and paraffin embedded using usual site protocols and were shipped to a central histopathology laboratory (University Hospitals, Leuven) for sectioning and hematoxylin and eosin and Masson trichrome staining.

Surgical resection margins were included *post hoc* in the analyses to increase the dynamic range of scores and provide a baseline comparison of histologic features present in the intestinal segment. These samples were normal or with little evidence of fibrosis or inflammation (score 0 or 1, respectively), depending on the status of the surgical resection margin.

Histopathologic assessments were conducted independently by 2 expert pathologists (GdH, 13 years of experience, and MC 23, years of experience) in a blinded fashion. Inflammation and fibrosis were evaluated using a Modified Chiorean Score (see Supplemental Table 1, Supplementary Digital Content 1, http://links.lww.com/CTG/A833). Briefly, the Modified Chiorean score uses 4, instead of 3, grades. The original Chiorean grade 0 is now subdivided into a new grade 0 for normal histology with "no

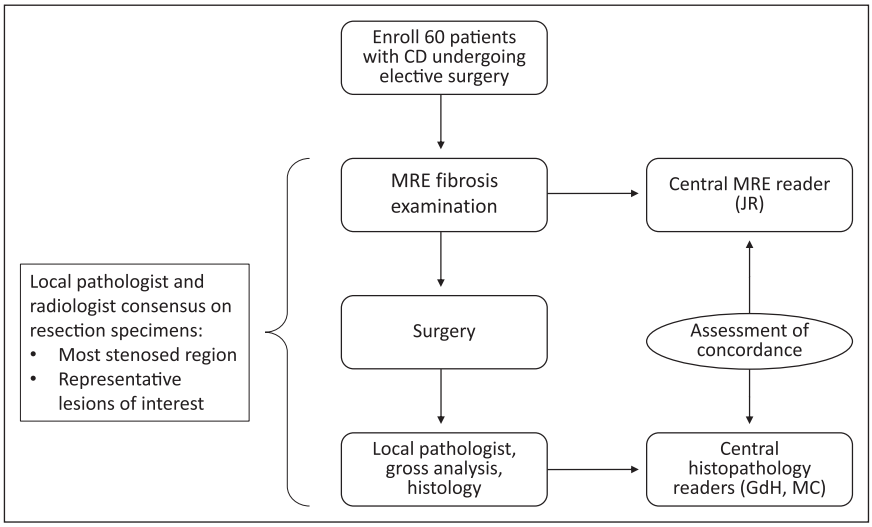


**Figure 1.** Study design. Patients were enrolled and underwent an MRE scan before surgery in which pathological specimens were collected and examined first by a local pathologist and then centrally. MRE scans were reviewed by a central reader. Concordance was assessed and data analyzed statistically. CD, Crohn's disease; MRE, magnetic resonance enterography.

fibrosis" and grade 1 for "minimal fibrosis limited to submucosa (<25% thickness)." The original grades 1 and 2 now equate to grades 2 and 3. Inflammation was similarly graded, 0–3. The severity of both inflammation and fibrosis is determined by both the extent (prevalence) of the features and depth of penetration with respect to the intestinal anatomy (i.e., confined to mucosa or extending deeper to submucosa and/or muscularis propria). All samples receiving discordant scores were reevaluated jointly by both pathologists and adjudicated to reach a single consensus score.

### Statistics

It was estimated that 60 patients would sufficiently power the study to verify the association between DGE and the histological fibrosis score. For this analysis, 80% power (2-sided, $\alpha = 0.05$) could be obtained with a minimum of $n = 40$ histological samples with associated MRE images of sufficient quality. This estimate also controlled for a moderate association between the histological fibrosis score and the histological inflammation score and for moderate variability in the levels of DGE across sites (see Supplemental Materials, Supplementary Digital Content 1, http://links.lww.com/CTG/A833 for further details).

The association between histological and MRE metrics was evaluated using Pearson correlation coefficient, R, and associated *P* values for testing the null hypothesis against a nonzero correlation between pairs of MRE metrics and histological scores. Because of a potential confounding, dependent association between histologic inflammation and fibrosis, linear partial Pearson correlation coefficients, Rho, were also measured, where correlation between MRE metrics and histologic fibrosis was controlled for histologic inflammation, and correlation with histologic inflammation was controlled for histologic fibrosis.

The least absolute shrinkage and selection operator (LASSO) was used to select linear models based on the 5 MRE features tested (ADC, MaRIA, DGE, MTR, and T2R) that best predict histologic fibrosis and inflammation. This method uses a shrinkage penalty or regularization term in its cost function to enable elimination (shrinkage) of variables that either have low or no predictive value or are highly collinear with other predictive variables. The regularization parameter $\lambda$ determines the strength of the shrinkage penalty and was optimized for each regression run using a randomized 8-fold cross-validation. Selected models were used to compute MRE-based predictive scores of histologic fibrosis and inflammation. The receiver operating characteristic (ROC) curve method with the area under the curve (AUC) value was used to quantify the ability of the resulting MRE-based regression models to discriminate between histological samples with low-mild (0–1) from moderate-severe (2,3) scores.

Cohen's kappa coefficient was used to measure interrater reliability between the 2 pathology experts assessing the histological scores. Statistical analyses were performed in MATLAB (version R2018a, MathWorks Inc., Natick, MA) and R (version 3.5, https://www.r-project.org/about.html).

## RESULTS

### Patients

Sixty eligible patients were enrolled in the study. Patient demographics, enrollment, and follow-up flow are summarized in Table 1 and Supplemental Figure 2, Supplementary Digital Content 1, http://links.lww.com/CTG/A833. Of the 60 enrolled patients, 52 completed the study. Of those, 50 patients had evaluable MRE examination and histological data available for correlation assessments. Two patients had unevaluable histological samples because of missing bowel wall layers.

### Imaging

Imaging was performed a median of 2 weeks (4-week interquartile range) before surgery with no treatment changes between imaging and surgery. Pathologists and radiologists annotated the location of 61 histological samples on the associated MRE sequences, including 10 samples from the margins and 51 samples from the most stenosed areas of resected specimens. These samples originated from the terminal ileum (n = 52), proximal

**Table 1. Summary of subject demographics**

| Characteristic | |
|---|---|
| Female, n (%) | 37 (61.7) |
| Age at enrollment (yrs), median (IQR) | 27 (29–50) |
| Disease duration (yrs), median (IQR) | 9.5 (2.5–20.5) |
| Location of strictures | |
| Ileal, n (%) | 49 (78.3) |
| Colonic, n (%) | 2 (3.3) |
| Ileocolonic, n (%) | 1 (1.7) |
| Unknown[a], n (%) | 8 (13.3) |
| Previous surgery, n (%) | 23 (38.3) |
| Concomitant medication | |
| 5-ASA, n (%) | 3 (5) |
| 6-MP, n (%) | 3 (5) |
| Steroids, n (%) | 16 (26.7) |
| Antibiotics, n (%) | 8 (13.3) |
| Immunosuppressants, n (%) | 7 (11.7) |
| Anti-TNF antibodies, n (%) | 7 (11.7) |
| Anti-integrins, n (%) | 2 (3.3) |
| Other biologics, n (%) | 2 (3.3) |
| Other, n (%) | 7 (11.7) |

5-ASA, 5-aminosalicylic acid; 6-MP, mercaptopurine; IQR, interquartile range; TNF, tumor necrosis factor.
[a]Unknown because of patient discontinuation before determination of stricture location.

ileum (n = 7), or transverse colon (n = 2). Representative contrast-enhanced and ADC sequences are shown in Figure 2.

The central reader monitored image data quality, judging >95% of sequences to be evaluable (sequences with a quality score of Fair or better) at each site and across sites, overall. In 5 of 6 sites, >80% of sequences were assessed as Good or Excellent quality. At one site, approximately 59% of data were assessed as Good or Excellent quality (see Supplemental Table 2, Supplementary Digital Content 1, http://links.lww.com/CTG/A833).

### Histopathology

Histology score distribution for all 61 evaluable samples from 50 patients (Figure 3) shows that most samples (30/61) presented with the same score for inflammation and fibrosis. Most remaining samples were more fibrotic than inflamed (26/61); only 5 samples showed inflammation as the prevailing component. In particular, there were only 2 cases with low fibrosis (score $\leq$ 1) and higher inflammation, as would be expected from targeted sampling of radiographically confirmed, grossly observed strictures. Overall, the histological scores for fibrosis and inflammation were well correlated ($R = 0.66$, $P < 0.0001$). In addition, interrater reliability between the 2 expert readers was high with Cohen's kappa coefficients of 0.77 and 0.89 for fibrosis and inflammation scores, respectively.

### Association between imaging and histopathology

Weak or no association was found between histologic fibrosis scores and DGE, MTR, and T2R, with correlation coefficients ranging from 0.21 to 0.27 (Figure 4). Stronger correlation was found between histologic fibrosis and ADC ($R = -0.71$, $P < 0.0001$) and the MaRIA score (0.59, $P < 0.001$), respectively. No association of histologic inflammation was found with DGE, MTR, or T2R. ADC and MaRIA were correlated with inflammation ($R = -0.35$, $P < 0.01$; and $R = 0.53$, $P < 0.001$, respectively). There was no significant variability in the MRE measurements associated with the site of origin of measurement (see Supplemental Figure 3, Supplementary Digital Content 1, http://links.lww.com/CTG/A833).

Controlling for associated histologic inflammation scores, partial Pearson correlation coefficients between MRE metrics and histologic fibrosis were significant for histologic score associations with ADC and MaRIA scores (Rho = $-0.68$, $P < 0.001$; 0.37, $P < 0.005$, respectively). Partial correlation between MRE metrics and histologic inflammation when controlling for histologic fibrosis did not show statistical significance. The partial correlation coefficient between MaRIA and histologic inflammation when controlling for fibrosis was Rho = 0.24 ($P = 0.078$).

Excluding normal areas of resected samples from the analysis reduced the dynamic range of both histological scores and MRE measurements, affecting the estimation of the Pearson correlation coefficient. Under these conditions, only ADC and MaRIA show a significant association with the histological fibrosis score ($R = -0.42$, $P < 0.005$; 0.31, $P < 0.05$, respectively), and only MaRIA shows association with the histological inflammation score ($R = 0.34$, $P < 0.05$; see Supplemental Figure 4, Supplementary Digital Content 1, http://links.lww.com/CTG/A833). In these severe stenotic areas, controlling for the effect of inflammation, only ADC showed a significant association with histologic fibrosis (Rho = $-0.49$, $P < 0.001$.) Similarly, only ADC showed a significant association, albeit weaker, with histologic inflammation when controlling for fibrosis (Rho = 0.29, $P < 0.05$).

Of the 5 MRE features considered, the LASSO approach identified a multivariate linear score including ADC, MaRIA, and DGE as the best predictor of the fibrosis histological score. This composite score had a Pearson correlation coefficient with histologic fibrosis of $R = 0.75$ ($P < 0.0001$) and a partial correlation controlling for histologic inflammation of Rho = 0.68 ($P < 0.0001$). The score is described as follows:

$$Fibrosis_{MRE} = -1.61 \cdot ADC + 0.02 \cdot MaRIA + 0.72 \cdot DGE$$

The LASSO selection identified a univariate linear model including only MaRIA as the best predictor of the histological score of inflammation. Figure 5 depicts scatter and box plots of histologic fibrosis and inflammation scores and associated MRE predictive scores $Fibrosis_{MRE}$ and MaRIA, respectively. Horizontal dashed lines depict relevant predictive score thresholds obtained with ROC analysis.

ROC curve analysis indicated very good performance of the predictive MRE score of fibrosis, $Fibrosis_{MRE}$, in detecting moderate to severe histological fibrosis ($ROC_{AUC} = 0.910$; Figure 6, Table 2). The predictive performance of $Fibrosis_{MRE}$ was improved over the performance of MaRIA and ADC as univariate predictors (Table 2). The coefficients for the $Fibrosis_{MRE}$ model and its predictive performance were verified in a *post hoc* cross-validation experiment described in Supplemental Materials.

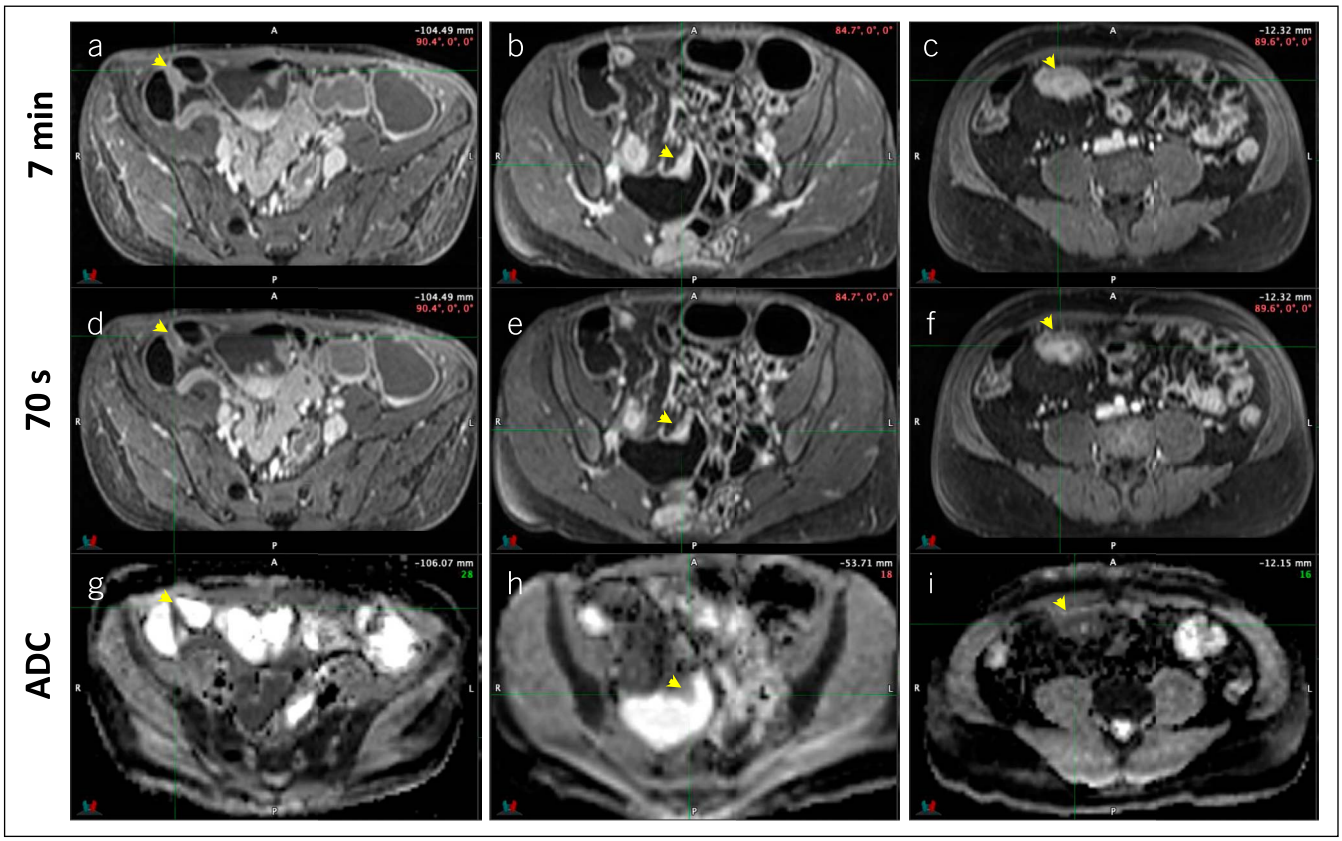


**Figure 2.** Representative contrast-enhanced sequences acquired at 7 minutes (**a**, **b**, and **c**) and 70 seconds (**d**, **e**, and **f**) post-Gd administration and apparent diffusion coefficient (ADC) sequences (**g**, **h**, and **i**) from 3 patients with different histological grades of inflammation and fibrosis. Panels **a**, **d**, and **g** are from a patient with low inflammation (grade 1) and low fibrosis (grade 1). Panels **b**, **e**, and **h** are from a patient with low inflammation (grade 1) and moderate fibrosis (grade 2). Panels **c**, **f**, and **i** are from a patient with moderate inflammation (grade 2) and high fibrosis (grade 3). Arrows indicate the region from where magnetic resonance enterography measurements and associated histological samples were obtained.

The predictive MRE score for inflammation, MaRIA, had fair performance in predicting moderate to severe histological inflammation ($ROC_{AUC}$ = 0.763) and excellent performance in detecting active histological inflammation ($ROC_{AUC}$ = 0.966). Note that the optimal predictive threshold for active histologic inflammation, MaRIA ≥ 7.12, is consistent with previously published diagnostic thresholds for MaRIA to detect active endoscopic activity (1,13).

## DISCUSSION

Overall, in this prospective multicenter European study, MaRIA and ADC scores exhibited moderate univariate correlations with histologic fibrosis and inflammation scores, which were preserved even when controlling for associated histologic inflammation. A stronger correlation was observed between ADC and fibrosis than between ADC and inflammation. However, DGE and MTR correlated only weakly at best with histologic fibrosis scores. In all cases, correlations were increased by the *post hoc* inclusion of normal margins, suggesting that the difference between nonfibrotic and fibrotic tissue is more readily detectable than increasing fibrosis within stenotic regions.

From the 5 MRE metrics investigated in this study, LASSO selection of multivariate predictive scores identified a score composed of MaRIA, ADC, and DGE, the $Fibrosis_{MRE}$, as the best predictor of histologic fibrosis in a multisite cohort. The same approach identified the univariate MaRIA score as the best predictor of histologic inflammation. ROC curve analysis indicates very good and fair accuracy of $Fibrosis_{MRE}$, and MaRIA, respectively, to detect moderate to severe histologic fibrosis and inflammation (Chiorean scores 2–3).

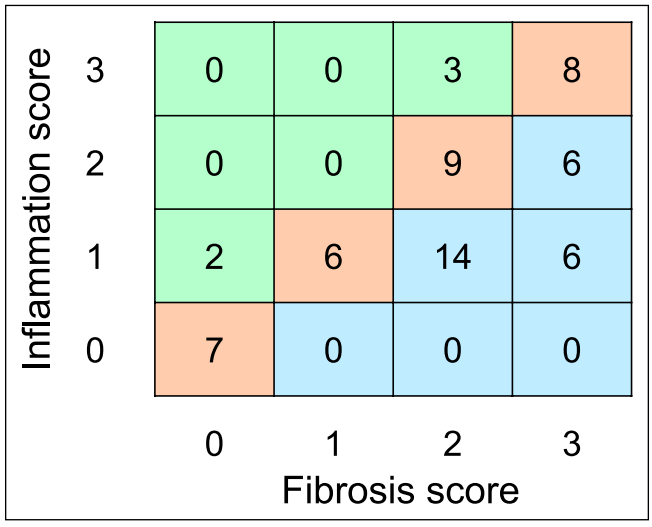


**Figure 3.** Frequency matrix of histological sample observations with each combination of inflammation and fibrosis scores. Overall, the histological scores for fibrosis and inflammation were well correlated (R = 0.66, $P < 0.0001$).

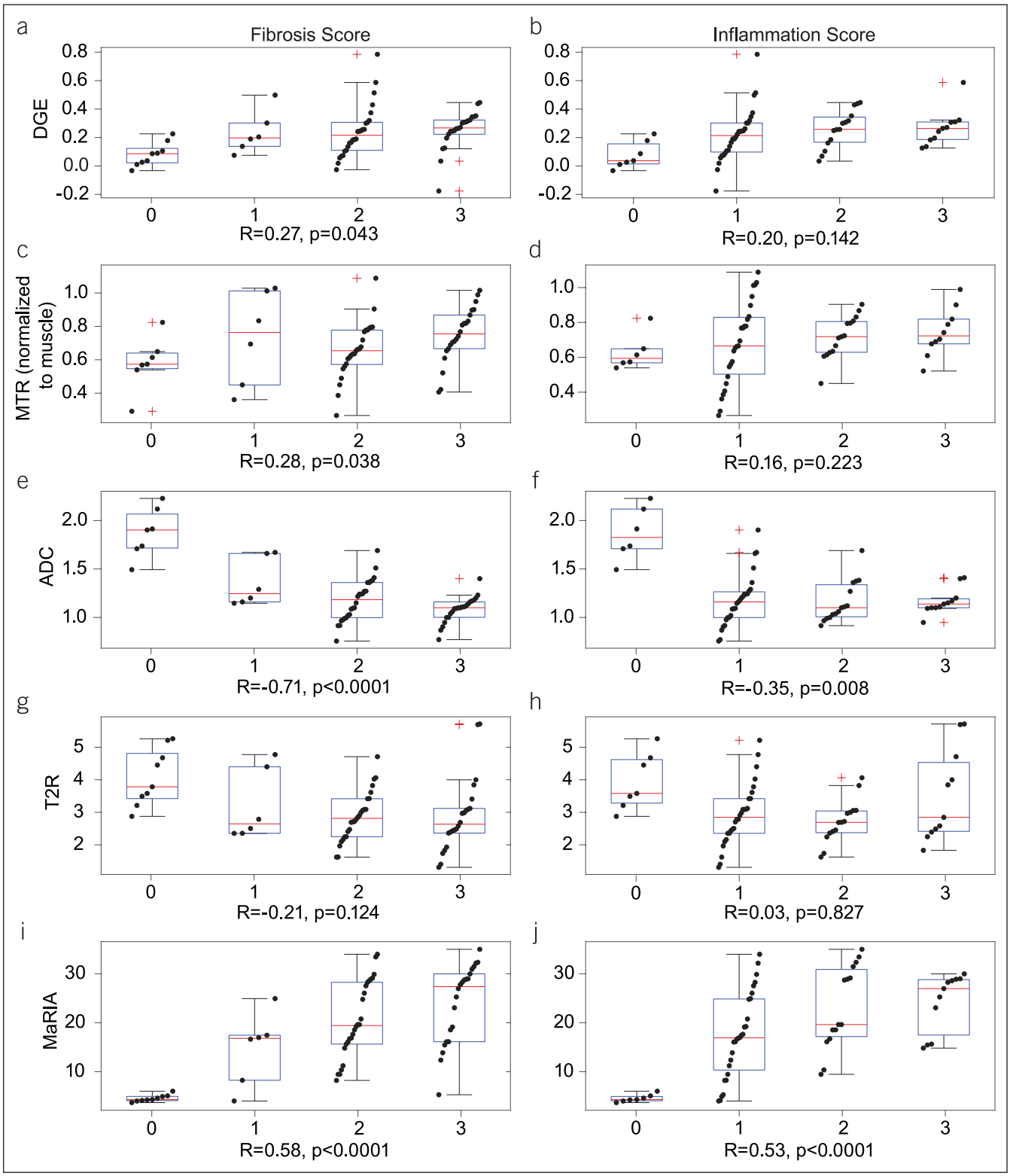


**Figure 4.** Boxplots and overlaid scatter plots of each MRE metric of interest depicted against histological scores for fibrosis and inflammation for DGE (**a** and **b**), MTR (**c** and **d**), ADC (**e** and **f**), T2R (**g** and **h**), and MaRIA (**i** and **j**) to show an association between imaging and histopathology. Boxplots indicate median, quartiles, and 95% CI. Red crosses indicate outliers (beyond 95% CI). Correlation coefficients, R, and respective *P* values are indicated at the bottom of each comparison panel. ADC, apparent diffusion coefficient; CI, confidence interval; DGE, delayed gain of enhancement; MaRIA, magnetic resonance index of activity; MRE, magnetic resonance enterography; MTR, magnetization transfer ratio; T2R, T2-weighted magnetic resonance imaging (MRI) sequences.

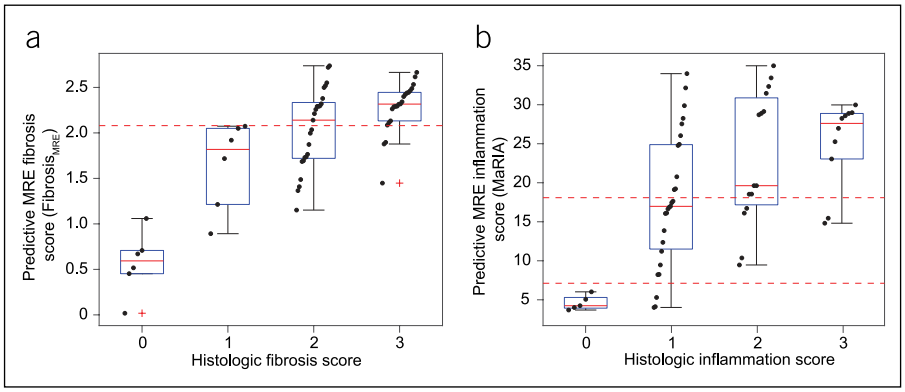


**Figure 5.** Scatter and box plots of histologic fibrosis (**a**) and inflammation (**b**) scores and associated MRE predictive scores Fibrosis$_{MRE}$ and MaRIA, respectively. Boxes are IQR, whiskers depict extreme data points, and red crosses in boxplots indicate outlier data points (beyond 99.3% of estimated normal distribution). Horizontal red dashed lines depict relevant predictive score thresholds obtained with ROC curve analysis. IQR, interquartile range; MaRIA, magnetic resonance index of activity; MRE, magnetic resonance enterography; ROC, receiver operator characteristic.

In addition, the MaRIA score showed excellent performance in predicting active histologic inflammation (Chiorean scores ≥1) with a detection cutoff of MaRIA ≥ 7. This is consistent with previous reports on the accuracy of MaRIA in predicting endoscopically assessed mucosal inflammation (1,13,15).

The quality of images obtained with the protocol used in this study was reliably high across all sites (most sequences were good or excellent), with few unevaluable sequences. Together with the lack of significant variability in the measurements that could be attributed to a site, these data indicate that the harmonized protocol used in this study is technically feasible and may be used prospectively.

The lack of univariate associations between MRE features and histologic fibrosis from this study contrast with previously published findings showing a strong correlation between DGE and MTR with fibrosis (7–9,12). This may be due to a spatial mismatch between the histologic tissue sample and the region of interest for imaging measurements, inconsistencies in the interval between imaging and surgery, the effects of histologic preparation, or lack of granularity in or robustness of histological scoring. This may reflect challenges presented by multisite collection of data likely to be encountered in clinical trial design and execution not found in earlier single-center studies.

An important confounding factor may also be muscular hypertrophy (12,16). It is likely that muscular hypertrophy has a particular MRE measurement signature distinct from fibrosis, where one might expect to observe lower DGE (due to higher blood perfusion), higher ADC and T2R (due to more fluid), but rather similar MTR (similar concentration of macromolecules) in the presence of the former compared with the latter histopathologic feature. Meanwhile, although the Chiorean scoring system used in this study does consider muscle hypertrophy as a factor in scoring fibrosis, it does not provide much distinct granularity between fibrosis and muscular hypertrophy. Unfortunately, muscular hypertrophy was neither controlled for nor explicitly quantified with any tool better than the Chiorean scoring system. Therefore, it is not possible to address the impact of muscle hypertrophy, nor derive an associated MRE-based composite score directly from our data, as was done for prediction of the histologic fibrosis and inflammation scores.

Regarding the association between MRE features and histologic inflammation, our analysis concludes that across 5 MRE-based metrics investigated, univariate MaRIA is the best predictor of histologic

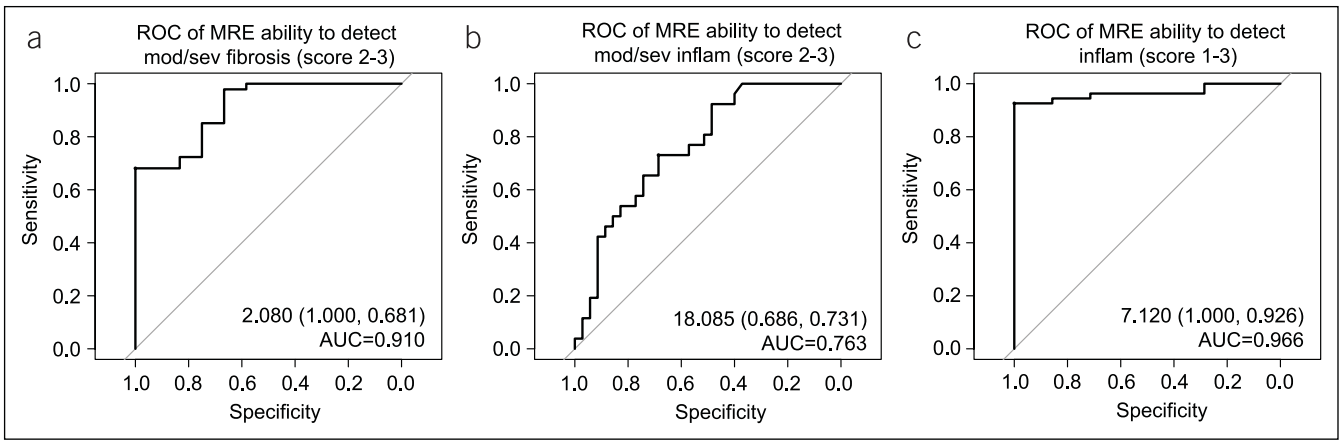


**Figure 6.** ROC curves for (**a**) ability of the predictive MRE score for fibrosis, Fibrosis$_{MRE}$, to detect moderate to severe histologic fibrosis (scores 2–3); (**b**) ability of the predictive MRE score for inflammation, MaRIA, to detect moderate to severe histological inflammation (scores 2–3), and (**c**) ability of the MaRIA score to detect active histological inflammation (histologic inflammation scores >0). The values inside the individual panels indicate the predictive score cutoff (sensitivity, specificity) and area under the curve. AUC, area under curve; MaRIA, magnetic resonance index of activity; MRE, magnetic resonance enterography; ROC, receiver operator characteristic.

**Table 2. Diagnostic accuracy of predictive MRE scores for histologic fibrosis and inflammation, $Fibrosis_{MRE}$ and MaRIA, respectively, for detection of moderate to severe fibrosis and inflammation and active inflammation**

| | $ROC_{AUC}$ | Cutoff | Sensitivity | Specificity |
|---|---|---|---|---|
| Fibrosis moderate/severe ($Fibrosis_{MRE}$) | 0.910 | ≥2.080 | 0.681 | 1.000 |
| Fibrosis moderate/severe (MaRIA) | 0.896 | ≥8.857 | 0.958 | 0.733 |
| Fibrosis moderate/severe (ADC) | 0.894 | ≥1.452 | 0.958 | 0.692 |
| Inflammation mod/severe (MaRIA) | 0.763 | ≥18.085 | 0.731 | 0.686 |
| Inflammation active (MaRIA) | 0.966 | ≥7.120 | 0.926 | 1.000 |

AUC, area under curve; ADC, apparent diffusion coefficient; MRE, magnetic resonance enterography; MaRIA, magnetic resonance index of activity; ROC, receiver operator characteristic. $Fibrosis_{MRE}$ was a very good predictor of moderate to severe histological fibrosis. MaRIA had excellent performance in detecting active histological inflammation.

inflammation. MaRIA is a composite score developed to maximize its correlation with an endoscopic score of mucosal inflammation. The Pearson correlation coefficient observed between MaRIA and histologic inflammation showed only a moderate correlation compared with higher correlations between MaRIA and endoscopic scores reported in a number of studies (e.g., (1,13,15,17)). This could be due to the lower dynamic range of the modified Chiorean score and its 4-rank scale, contrasting with the wider dynamic range and continuous variable nature of the endoscopic scores. Our study sample was limited to either highly stenotic areas or margins of resected stenotic specimens with either no or mild activity, whereas other previous studies investigated an association between MaRIA and endoscopic inflammation samples with a wider variety of inflammation severity.

Finally, although we were able to show good interrater variability of the modified Chiorean scoring system, an important limitation of any attempt to validate a cross-sectional imaging tool against histological assessment of fibrosis is the lack of a fully validated histopathologic scoring system for either inflammation or fibrosis in CD (18).

In conclusion, our observations suggest that inflammation and fibrosis coexist and are highly correlated in most stenotic lesions in CD. However, the predictive $Fibrosis_{MRE}$ and MaRIA scores may assist in the characterization of fibrosis severity independent of inflammation. Given its high specificity, a high (>2.08) $Fibrosis_{MRE}$ score validation of $Fibrosis_{MRE}$ in an independent cohort of patients, ideally from multiple clinical sites, is needed. Further evaluation, development, and validation of histologic fibrosis scores and other alternative MRE metrics informative of inflammation and fibrosis will inform future MRE protocol development in CD in general and enable development of targeted therapies and clinical treatment of stenotic patients.

## CONFLICTS OF INTEREST

**Guarantor of the article:** Alexandre Coimbra, PhD.

**Specific author contributions:** A.C. was involved in planning and conducting the study, interpreting data, and drafting the manuscript. J.R. was involved in planning and conducting the study, collecting data, and drafting the manuscript. M.C. was involved in planning and conducting the study, collecting data, and drafting the manuscript. G.de H. was involved in planning and conducting the study, collecting and interpreting data, and drafting the manuscript. G. v A. was involved in study design, study conduct, and manuscript review. R.V. was involved in planning and conducting the study, collecting data, and drafting the manuscript. H.G. was involved in planning and conducting the study, collecting and/or interpreting data, and drafting the manuscript. I have included patients in the study. A.H.N. was involved in collecting and interpreting data and drafting the manuscript. R.H-M. was involved in collecting data and drafting the manuscript. Y.B. was involved in planning and conducting the study, collecting data, and drafting the manuscript. M.Z. was involved in collecting data and drafting the manuscript. D.C-H. was involved in planning and collecting histological data. G.d'H. was involved in study design, data collection and interpretation, and critical review of the manuscript. J.S. was involved in planning and conducting the study, collecting data, and drafting the manuscript. S.M. was involved in planning and collecting histological data and drafting the manuscript. G.R. was involved in planning and conducting the study, collecting data, and drafting the manuscript. A.B. was involved in planning and conducting the study, collecting data, and drafting the manuscript. A.W. was involved in collecting/interpreting data and drafting the manuscript. R.Z. was involved in planning the study, collecting data, and drafting the manuscript. M.K. was involved in planning the study, collecting data, and drafting the manuscript. A.S. was involved in planning the study, collecting the data, and drafting the manuscript. A.deC. was involved in planning the study and drafting the manuscript. T.T.L. was involved in planning and conducting the study, collecting and interpreting data, and drafting and revising the manuscript. J.P. was involved in planning and conducting the study, collecting data, and drafting the manuscript. All authors approved the final draft of the submitted manuscript.

**Financial support:** Genentech, Inc., funded this study. Genentech, Inc., participated in the study design, the collection, analysis, and interpretation of the data and in the writing of the report.

**Potential competing interests:** A.C.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. J.R.: Research grants from AbbVie, Genentech; speaker fees for lectures or consultancy from Janssen, Gilead, Takeda, Bioclinica, Origo biopharma, Lument, Alimentiv, and TiGenix. M.C.: Nothing to declare. G.de H.: Nothing to declare. G. v A.: No COI in the last 18 months. R.V.: Nothing to declare. H.G.: Nothing to declare. A.H.N.: Nothing to declare. R.H-M.: Nothing to declare. Y.B.: Relationships with AbbVie, Amgen, Biogaran, Biogen, Boehringer Ingelheim, Celltrion, Ferring, Fresenius Kabi, Gilead, Hospira, Iterative Scopes, Janssen, Lilly, Mayoli Spindler, Merck, MSD, Norgine, Pfizer, Roche, Sandoz, Sanofi, Shire, Takeda, Tillotts, and UCB. M.Z.: Speaker honoraria from AbbVie. D.C.-H.: Nothing to declare. G.d'H.: Nothing to declare. J.S.: Research agreement with Takeda on a nonrelated topic. S.M.: Nothing to declare. G.R.: Consulted for AbbVie, Augurix, BMS, Boehringer, Calypso, Celgene, FALK, Ferring, Fisher, Genentech, Gilead, Janssen, MSD, Novartis, Pfizer, Phadia, Roche, UCB, Takeda, Tillots, Vifor, Vital Solutions, and Zeller; speaker's honoraria from AstraZeneca, AbbVie, FALK, Janssen, MSD, Pfizer, Phadia, Takeda, Tillots, UCB, Vifor, and Zeller; and educational grants and research grants from

AbbVie, Ardeypharm, Augurix, Calypso, FALK, Flamentera, MSD, Novartis, Pfizer, Roche, Takeda, Tillots, UCB, and Zeller. A.B.: Nothing to declare. A.W.: Nothing to declare. R.Z.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. M.K.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. A.S.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. A.deC.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. T.T.L.: Employee of Genentech, Inc., and shareholder of F. Hoffmann-La Roche Ltd. J.P.: Research grants from AbbVie and Pfizer; speaker's fees from AbbVie, Ferring, Janssen, Merck, Pfizer, Shire, Takeda, and Theravance; and has been a consultant for AbbVie, Arena, Boehringer Ingelheim, Celgene, Celltrion, Ferring, Genentech, GlaxoSmithKline, GoodGut, Janssen, Merck, Nestlé, Origo, Pandion, Pfizer, Progenity, Robarts Clinical Trials, Roche, Takeda, Theravance, and Wassermann. The study was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practice. Approval from the institutional review boards and ethics committees was obtained before its start. Patient consent was obtained before enrollment.

## Study Highlights

### WHAT IS KNOWN

- ✓ Magnetic resonance enterography (MRE) is an accurate tool to identify and measure the degree of bowel inflammation.
- ✓ Fibrosis and inflammation usually coexist in an intestinal lesion.
- ✓ Some single-center pilot studies have previously identified a number of MRE-derived metrics to identify fibrosis in the bowel.

### WHAT IS NEW HERE

- ✓ The development of an index based on MRE for characterization of fibrosis severity independent of inflammation will help improve patient treatment decisions and the development of fibrosis-targeted therapies.

## ACKNOWLEDGMENTS

We thank the patients and their families who took part in the study as well as the staff, research coordinators, and investigators at each participating institution. Writing and editing assistance was provided by Genentech, Inc.

## REFERENCES

1. 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;58:1113–20.
2. Church PC, Turner D, Feldman BM, et al. Systematic review with meta-analysis: Magnetic resonance enterography signs for the detection of inflammation and intestinal damage in Crohn's disease. Aliment Pharmacol Ther 2015;41:153–66.
3. Bettenworth D, Bokemeyer A, Baker M, et al. Stenosis therapy and anti-fibrotic research (STAR) Consortium. Assessment of Crohn's disease-associated small bowel strictures and fibrosis on cross-sectional imaging: A systematic review. Gut 2019;68:1115–26.
4. Chiorean MV, Sandrasegaran K, Saxena R, et al. Correlation of CT enteroclysis with surgical pathology in Crohn's disease. Am J Gastroenterol 2007;102:2541–50.
5. Punwani S, Rodriguez-Justo M, Bainbridge A, et al. Mural inflammation in Crohn disease: Location-matched histologic validation of MR imaging features. Radiology 2009;252:712–20.
6. Rimola J, Rodríguez S, Cabanas ML, et al. MRI of Crohn's disease: From imaging to pathology. Abdom Imaging 2012;37:387–96.
7. Tielbeek JA, Ziech ML, Li Z, et al. Evaluation of conventional, dynamic contrast enhanced and diffusion weighted MRI for quantitative Crohn's disease assessment with histopathology of surgical specimens. Eur Radiol 2014;24:619–29.
8. 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;110:432–40.
9. Pazahr S, Blume I, Frei P, et al. Magnetization transfer for the assessment of bowel fibrosis in patients with Crohn's disease: Initial experience. MAGMA 2013;26:291–301.
10. Li XH, Mao R, Huang SY, et al. Characterization of degree of intestinal fibrosis in patients with Crohn disease by using magnetization transfer MR imaging. Radiology 2018;287:494–503.
11. Fang ZN, Li XH, Lin JJ, et al. Magnetisation transfer imaging adds information to conventional MRIs to differentiate inflammatory from fibrotic components of small intestinal strictures in Crohn's disease. Eur Radiol 2020;30:1938–47.
12. Wagner M, Ko HM, Chatterji M, et al. Magnetic resonance imaging predicts histopathological composition of ileal Crohn's disease. J Crohns Colitis 2018;12:718–29.
13. 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;17:1759–68.
14. 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;17:984–93.
15. Coimbra AJ, Rimola J, O'Byrne S, et al. Magnetic resonance enterography is feasible and reliable in multicenter clinical trials in patients with Crohn's disease, and may help select subjects with active inflammation. Aliment Pharmacol Ther 2016;43:61–72.
16. Chen W, Lu C, Hirota C, et al. Smooth muscle hyperplasia/hypertrophy is the most prominent histological change in Crohn's fibrostenosing bowel strictures: A semiquantitative analysis by using a novel histological grading scheme. J Crohns Colitis 2017;11:92–104.
17. Takenaka K, Ohtsuka K, Kitazume Y, et al. Correlation of the endoscopic and magnetic resonance scoring systems in the deep small intestine in Crohn's disease. Inflamm Bowel Dis 2015;21:1832–8.
18. Gordon IO, Bettenworth D, Bokemeyer A, et al. Stenosis therapy and anti-fibrotic research (STAR) Consortium. Histopathology scoring systems of stenosis associated with small bowel Crohn's disease: A systematic review. Gastroenterology 2020:137–50.