**Magnetic resonance enterography and histology in patients with fibrostenotic Crohn’s disease: a multicenter study**

**ONLINE SUPPLEMENTAL MATERIALS**

**Supplemental Methods**

**Protocol for patient preparation and scanning**

Patients fasted for at least 4 hours and were prepped for administration of IV or oral contrast agents. Patients with colonic strictures also received a colon distension enema when positioned on the scanner. Following administration of anti-peristaltics (scopolamine butylbromide 10 mg or glucagon 0.5 mg IV), balanced gradient echo (GRE), fast spin echo T2-weighted with and without fat suppression, and 2D fast low-angle shot GRE magnetization transfer sequences with and without off-resonance magnetization transfer preparation pre-pulse were performed. A second (booster) dose of anti-peristaltic was given before performing a fast 3D GRE (volumetric interpolated breath-hold examination) T1-weighted sequence. After Gd contrast administration, a series of identical fast 3D sequences were performed at 70 seconds, 5 minutes, and 7 minutes following Gd administration. Lastly, 2D echo planar diffusion-weighted imaging sequences were performed. ADC maps were automatically generated.

**MRE quantitative definitions**

DGE was derived from pre-Gd, 70 sec. and 7 min post-Gd sequences and was defined as follows:

where *SI70s*, and *SI7min* are the signal intensities.

MTR was derived from the magnetization transfer sequences and is defined as follows:

where *MTROI* and *MTpsoas* are the measured MT in the stenotic ROI and psoas muscle, respectively; *MT*on is the signal intensity at the wall ROI in the series with pre-pulse on and *MT*off is the corresponding ROI signal intensity in the series with pre-pulse off.

T2R was derived from the fast T2-weighted sequences, and is defined as follows:

where *SIROI*, and *SIpsoas*, are the signal measured from T2-weighted sequences in the stenotic ROI, and psoas muscle, respectively.

ADC at stenotic ROI was measured from the ADC maps computed by commercial software by scanner manufacturers.

MaRIA was derived from the fast T2-weighted and pre- and post-Gd sequences (70s), and is defined as follows:

where RCE is of the relative contrast enhancement; wall thickness is measured in mm; presence of edema = 1; absence of edema = 0; and presence of ulcerations = 1; absence of ulcerations = 0.

**Sample size calculations**

This study was powered at 80% (two-sided, alpha= 0.05) with 40 good quality (evaluable) datasets as assessed by simulations based on prior data obtained from a single site suggesting high value of DGE in predicting histologic fibrosis score (n=44) (8). Simulations used a multiple linear regression (MLR) approach to estimate sample size controlling for moderate association between DGE and histologic inflammation score.

Rationale: This supposes there is a set of candidate predictor variables X = {x1, x2, … xp}. Using MLR the aim is to identify a subset of variables from X that are statistically significant predictors of the dependent variable Y. This problem is studied by simulation.

Simulations setup: Y = b1\*x1 + b2\*x2 + b3\*x3 + b4\*x4 + … + b8\*x8 + error

- All predictor variables and Y are zero-mean with unit variance.

- cor(x1,x2)=cor(x1,x3)=cor(x2,x3)=0.5

- b1=0.5, b2=0.3, b3=0.2

- b4=b5=…=b8=0; w/ correlation between variables in {x4, …, x8} equal to zero

- correlation between variables in {x1, x2, x3} and variables in {x4, …, x8} equals 0.3

- parameters chosen to mimic correlations in MRE inflammation variables.

Simple linear regression approach: The alternative simple linear regression approach for primary analysis provides 90% power (two-sided, alpha = 0.05) with a minimum of N=40 samples, assuming a hypothesized slope of 0.3, variation in the explanatory variable equal to 1, and a residual standard deviation of 0.56. This sample size estimate was calculated using the single-arm multiple linear regression (MLR) design in EAST version 6.4.1 (https://www.cytel.com/software/east/).

**Post hoc cross-validation of the FibrosisMRE predictive model**

Five sets of pseudo-randomized cross-validation splits of the data were created with respective derivation and validation subsets in a 4:1 ratio. Care was taken to balance the number of samples with low-mild and moderate-severe scores across data splits. Using these cross-validation splits, we repeated the LASSO regression runs 5,000 times for estimation of variability of the linear model coefficients with the “derivation” data subsets, and used the respective “validation” subsets to measure ROCAUC, accuracy, sensitivity and specificity of the derived models. This generated 5,000 sets of regression coefficient estimates and respective ROCAUC, accuracy, sensitivity and specificity estimates. The results are summarized in Supplemental Table 3 and show that the distribution of the estimated coefficients are in a range that includes the coefficients derived with the whole data set, and the medians of said distributions and associated performance metrics (ROCAUC, accuracy, sensitivity and specificity) approximated the coefficients and performance metrics estimated with the whole data set.

**Data availability**

Qualified researchers may request access to individual patient level data through the clinical study data request platform (https://vivli.org/). Further details on Roche's criteria for eligible studies are available here (https://vivli.org/members/ourmembers/). For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see here (https://go.gene.com/datasharing)

**Investigator Experience**

J.R. (Radiologist): 14 years

M.C. (Pathologist): 23 years

G.d.H. (Pathologist): 13 years

R.V. (Radiologist): 22 years

A.H.N. (Radiologist): 24 years

R.H-M. (Pathologist): 22 years

D.C-H. (Pathologist): 28 years

J.S. (Radiologist): 26 years

A.B. (Radiologist): 18 years

A.W. (Pathologist): 19 years

A.S. (Pathologist): 13 years

Supplemental Table 1. Modified Chiorean Score differences competed with the Original Chiorean Score.

|  |  |  |
| --- | --- | --- |
|  | **Original Chiorean Score** | **Modified Chiorean Score** |
| **Fibrosis Feature** | **Fibrosis Score** | **Fibrosis Score** |
| No fibrosis | Grade 0 | Grade 0 (normal) |
| Minimal fibrosis limited to submucosa (<25% thickness) | Grade 0 | Grade 1 (minimal) |
| Mild stricture (>15mm) with non-dilated lumen  Submucosal fibrosis and muscular hyperplasia >25% with preserved layers | Grade 1 | Grade 2 (mild /moderate) |
| Massive transmural fibrosis; effacement of normal layers; severe stricture | Grade 2 | Grade 3 (severe) |
| **Inflammation Feature** | **Inflammation Score** | **Inflammation Score** |
| No mucosal inflammation | Grade 0 (normal) | Grade 0 (normal) |
| Aphthous ulcers affected surface <50%; cryptitis <50%; inflammation limited to mucosa | Grade 1 (mild) | Grade 1 (mild) |
| Large, superficial ulcers (0.5-2 cm); ulcered surface <50%; affected surface 50%-100%; cryptitis >50%; crypt abscess; submucosal inflammation | Grade 2 (moderate) | Grade 2 (moderate) |
| Deep ulcers or size > 2 cm; circumferential ulcers; transmural inflammation; deep fissures | Grade 3 (severe) | Grade 3 (severe) |

Supplemental Table 2. Prevalence of quality scores attributed to all sequences acquired per site.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Quality Score (%)** | | | |
|  | **Unevaluable** | **Fair** | **Good** | **Excellent** |
| Site A | 1 | 0 | 28 | 71 |
| Site B | 0 | 4 | 48 | 47 |
| Site C | 3 | 38 | 24 | 35 |
| Site D | 0 | 17 | 33 | 50 |
| Site E | 0 | 3 | 31 | 66 |
| Site F | 2 | 12 | 56 | 29 |
| All sites | 1 | 12 | 34 | 53 |

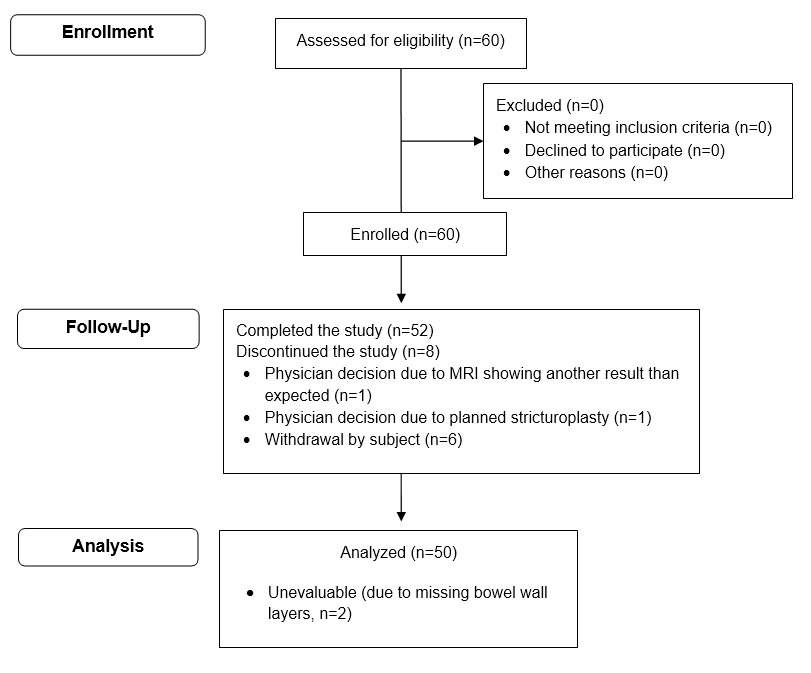
Supplemental Table 3. Summary of 5,000 cross-validation runs of predictive MRE score for histologic fibrosis, FibrosisMRE, for detection of moderate to severe fibrosis.

|  |  |
| --- | --- |
| **Derived coefficients and performance metrics** | **Median (IQR)** |
| ADC | -1.546 (0.383) |
| MaRIA | 0.014 (0.012) |
| DGE | 0.645 (0.489) |
| ROCAUC | 0.880 (0.018) |
| Accuracy | 0.667 (0.144) |
| Sensitivity | 0.556 (0.300) |
| Specificity | 1.000 (0.250) |

Supplemental Figure 1. MRE image acquisition protocol. GRE: gradient echo; T2w: T2-weighted; 3D: three-dimensional; T1w: T1 weighted; Gd: gadolinium; MTR: magnetization transfer ratio; DGE: delayed gadolinium enhancement; ADC: apparent diffusion coefficient .

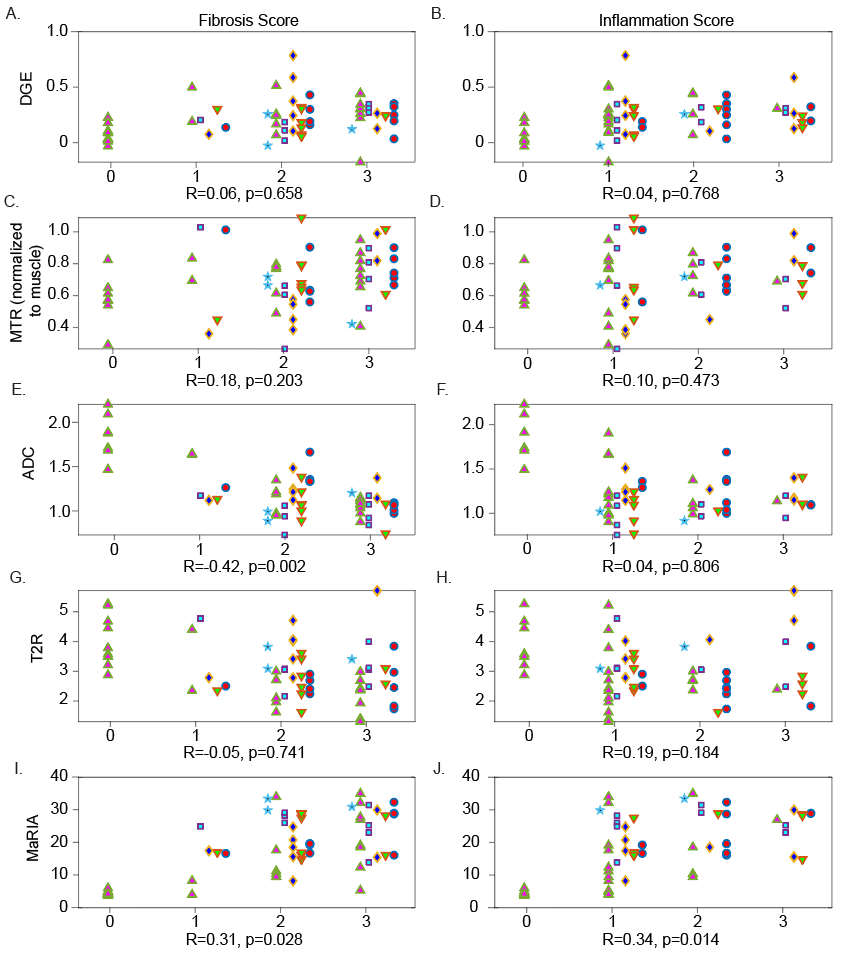


Supplemental Figure 2. Patient flow diagram.



Supplemental Figure 3. Scatter plots depicting association between MRE metrics and fibrosis or inflammation scoresindicating site of origin for each data point for DGE (A, B), MTR (C, D), ADC (E, F), T2R (G, H), and MaRIA (I, J). DGE: delayed gain of enhancement; MTR: magnetization transfer ratio; ADC: apparent diffusion coefficient;

T2R: T2-weighted MRI sequences; MaRIA:magnetic resonance index of activity. Each symbol in this plot represents data from a unique participating site. The distribution of MRE measurements within each metric is comparable across sites indicating no significant variability in MRE measurements associated with site of origin of measurement.



Supplemental Figure 4. This figure includes only data from the most stenosed region of resected specimens, and excludes data from resected specimen margins. Compare this figure to Figure 4 in the main text which includes both most stenosed and marginal areas. Boxplots and overlaid scatter plots of each MRE metric of interest depicted against histological scores for fibrosis and inflammation for for DGE (A, B), MTR (C, D), ADC (E, F), T2R (G, H), and MaRIA (I, J). Boxplots indicate median, quartiles, and 95% CI. Correlation coefficients (R) and respective p-values are indicated at the bottom of each comparison panel. Plots, and associated correlation results presented in this figure exclude normal samples obtained from site E. DGE: delayed gain of enhancement; MTR: magnetization transfer ratio; ADC: apparent diffusion coefficient; T2R: T2-weighted MRI sequences; MaRIA:magnetic resonance index of activity.

