



HAL
open science

Development and validation of a score to assess transmural healing and response in patients with Crohn's disease

Anthony Buisson, Jérémy Junda, Jeanne Vignette, Emma Lecoq, Guillaume Bouguen, Félix Goutorbe, Julien Scanzi, Dilek Coban, Marie Dodel, Maëva Bazoge, et al.

► To cite this version:

Anthony Buisson, Jérémy Junda, Jeanne Vignette, Emma Lecoq, Guillaume Bouguen, et al.. Development and validation of a score to assess transmural healing and response in patients with Crohn's disease. *Clinical Gastroenterology and Hepatology*, 2024, 22 (11), pp.2271-2279.e11. 10.1016/j.cgh.2024.06.007 . hal-04646830

HAL Id: hal-04646830

<https://hal.science/hal-04646830v1>

Submitted on 26 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

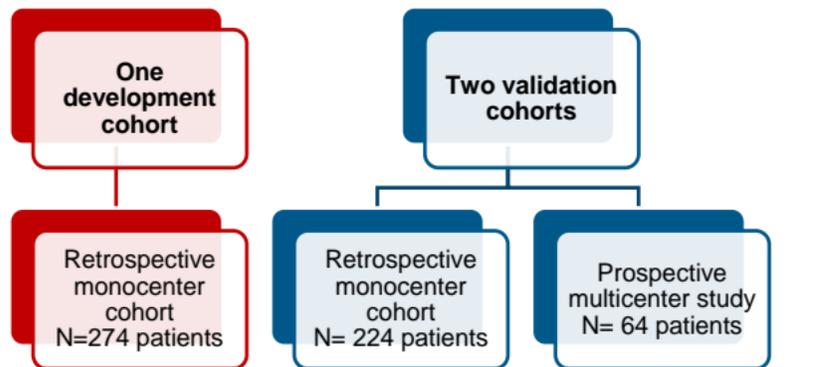
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

DEVISE-CD project

DEvolution and **V**alidation of a dedicated tool to assess tran**S** mural response after therapeutic interv**E**ntion in patients with Crohn's **D**isease

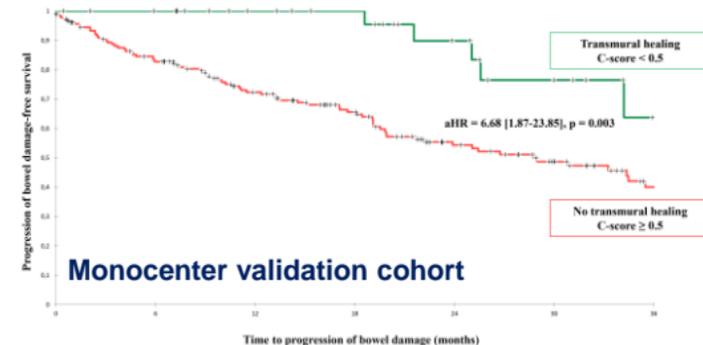
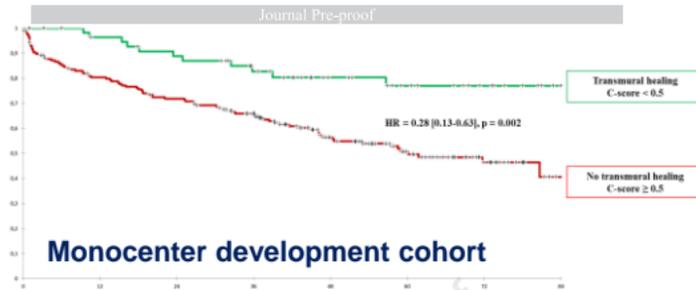


1 Development and validation of a new MRI score
C-score

(modified Clermont score to assess transmural response)

Items	Calculation	Results
Bowel wall thickening compared to the normal	Bowel wall thickness (mm) - 3	
Presence of any ulcer	+ 2	
Presence of enlarged lymph nodes	+ 1.5	
TOTAL		

Buisson A. et al. Clin Gastroenterol Hepatol 2024



TH, TR50 and TR25 are associated with lower risk of bowel damage progression



TH, TR50 and TR25 are reachable with current medications



New definitions of transmural healing and response

Transmural healing:

C-score < 0.5 in each segment

Transmural response (TR50):

Decrease of at least 50% of C-score compared to baseline, in all the segments that were considered as active at baseline

Transmural response (TR25):

Decrease of at least 25% of C-score compared to baseline, in all the segments that were considered as active at baseline

Rate of patients achieving transmural endpoints after 12 weeks of anti-TNF therapy
(*Multicenter validation cohort*)

TH = 28.3%
TR50 = 41.3%
TR25 = 47.8%

Clinical Gastroenterology and Hepatology

Development and validation of a score to assess transmural healing and response in patients with Crohn's disease

Running head: Development and validation of C-score

Anthony Buisson^{1,2}, Jérémy Junda³, Jeanne Vignette¹, Emma Lecoq¹, Guillaume Bouguen⁴, Felix Goutorbe⁵, J Scanzi^{1,6}, Dilek Coban¹, Marie Dodel¹, Maëva Bazoge¹, Bruno Pereira⁷, Constance Hordonneau³

Affiliations :

¹ Université Clermont Auvergne, Inserm, 3iHP, CHU Clermont-Ferrand, Service d'Hépatogastroentérologie, Clermont-Ferrand, France

² Université Clermont Auvergne, Inserm U1071, M2iSH, USC-INRA 2018, F-63000 Clermont-Ferrand, France

³ Université Clermont Auvergne, CHU Clermont-Ferrand, DRCI, Service de radiologie, Clermont-Ferrand, France

⁴ CHU Rennes, Univ Rennes, INSERM, CIC1414, Institut NUMECAN (Nutrition Metabolisms and Cancer), F-35000 Rennes, France

⁵ Centre Hospitalier de la côte basque, Service d'Hépatogastro Entérologie, Bayonne, France

⁶ CH Thiers, Service d'Hépatogastro Entérologie, Thiers, France

⁷ Université Clermont Auvergne, CHU Clermont-Ferrand, DRCI, Unité de Biostatistiques, Clermont-Ferrand, France

Corresponding author:

Prof. Anthony BUISSON, MD, PhD

University Hospital Estaing, Gastroenterology Department,

1 place Aubrac, 63100 Clermont-Ferrand, France

Phone: +33 4 73 750 523; Fax: +33 4 73 750 524; e-mail: a_buisson@hotmail.fr

Grant support: none.

Disclosures:

Anthony BUISSON: Consulting fees from Abbvie, Amgen, Arena, Biogen, Celltrion Healthcare, CTMA, Galapagos, Janssen, MSD, Nexbiome, Pfizer, Roche, Takeda and Tillotts. Lecture fees from Abbvie, Amgen, Biogen, Celltrion Healthcare Galapagos, Janssen, Mayoli-Spindler, MSD, Nordic Pharma, Norgine, Pfizer, Roche, Takeda, Tillotts and Vifor Pharma. Research grant from Abbvie, Celltrion Healthcare Janssen, Lessaffre, Lilly, Pfizer and Takeda

Guillaume BOUGUEN: received lecture fees from Abbvie, Ferring, MSD, Takeda and Pfizer and consultant fees from Takeda, Janssen.

Felix GOUTORBE : none

Julien SCANZI : none

Jérémy JUNDA, Jeanne VIGNETTE, Emma LECOQ, Dilek COBAN, Marie DODEL, Maëva BAZOGE, Bruno PEREIRA, Constance HORDONNEAU : none

Acknowledgments: We thank CHU Clermont-Ferrand (DRCI) for its recurrent support.

Guarantor or the article: Anthony BUISSON

Author contributions

Anthony BUISSON: Conception and design of the study; analysis and interpretation of data; Drafting of the manuscript; Approval of the final version of the manuscript.

Jérémy JUNDA: Generation, collection, assembly of data; analysis and interpretation of data; Drafting of the manuscript; Approval of the final version of the manuscript.

Jeanne VIGNETTE, Emma LECOQ, Dilek COBAN, Marie DODEL, Maëva BAZOGE,

Guillaume BOUGUEN, Felix GOUTORBE, Julien SCANZI: Generation, collection, assembly of data; Substantial revision of the manuscript; Approval of the final version of the manuscript.

Bruno PEREIRA: analysis and interpretation of data; statistical analyses; Substantial revision of the manuscript; Approval of the final version of the manuscript.

Constance HORDONNEAU: Conception and design of the study; Generation, collection, assembly of data; analysis and interpretation of data; Substantial revision of the manuscript; Approval of the final version of the manuscript.

ABSTRACT

Background and Aims

As transmural healing (TH) could be the best therapeutic target in Crohn's disease (CD), we aimed to build and validate a score to assess TH and transmural response (TR), and to confirm their association with favorable CD outcomes.

Methods

DEVISE-CD project encompassed two retrospective cohorts (274 and 224 CD patients for development and validation phase, retrospectively) and one multicenter prospective validation cohort (N=46 patients). A step-by-step process was used to build the modified Clermont score (C-score). The primary endpoints were time to bowel damage progression, and steroid-free clinical remission with fecal calprotectin < 250 (CFREM) at one year for retrospective and prospective validation cohorts, respectively.

Results

Edema, ulcer, contrast enhancement, diffusion-weighted hyperintensity, fat wrapping, bowel thickening (>3 mm), and enlarged lymph nodes were associated to higher risk of bowel damage progression ($p < 0.01$). Edema, diffusion-weighted hyperintensity, post-gadolinium contrast enhancement, and bowel thickening were highly coexistent (>95%) and collinear ($p < 0.0001$). Bowel thickness had the highest sensitivity to change after treatment (SMD=0.30±1.0)($p=0.001$). C-score was calculated as $0.2 \times (\text{bowel thickness} - 3\text{mm}) + 1.5 \times \text{enlarged lymph nodes} + 2 \times \text{ulcer}$. TH (C-score<0.5) (HR=0.28[0.13-0.63], $p=0.002$; aHR=0.15[0.04-0.53], $p=0.003$), TR50 (50%-decrease of C-score)(HR=0.30[0.15-0.63], $p=0.001$; aHR = 0.36[0.14-0.88], $p=0.025$) or TR25(25%-decrease of C-score)(HR=0.37[0.19-0.71], $p=0.003$; aHR=0.46[0.23-0.94], $p=0.034$) prevented bowel damage progression in development and validation cohorts, respectively.

In the prospective validation cohort, achieving TH (OR=4.6[1.3–15.6], $p=0.016$), TR50 (OR=6.9[1.8–26.0], $p=0.008$) or TR25 (OR=6.0[1.6–22.3], $p=0.008$) after 12 weeks of anti-TNF therapy led to higher rate of CFREM at one year.

Conclusion

C-score is a validated, reliable and easy-to-use tool to assess TH and TR in CD patients.

Keywords: Crohn's disease; transmural healing; MRI; C-score; modified Clermont score; transmural response; inflammatory bowel disease

- **BACKGROUND**

- Despite its strong association with favorable outcomes, transmural healing is not recommended as therapeutic target in Crohn's disease (CD) especially owing to the lack of consensual definition.

- **FINDINGS**

- We developed and validated an MRI score to define and assess transmural healing (TH) and response (TR) in patients with CD that are associated with favorable outcomes.

- **IMPLICATION FOR CARE**

- C-score and its definition of transmural healing (TH) and transmural response (TR25 and TR50) are now a validated, reliable and easy-to-use tool to assess therapeutic efficacy.

INTRODUCTION

Crohn's disease (CD) is a chronic inflammatory bowel disease (IBD) that can highly alter patients' quality of life and lead to bowel damage due to its transmural pattern. Magnetic Resonance Imaging (MRI) could be an attractive alternative due to its better patients' acceptability than endoscopy¹. MRI enables by itself a concomitant evaluation of both small bowel and colon, a transmural assessment of bowel inflammation, and the detection of CD-related complications (stricture, fistula). The concept of transmural healing has emerged as a promising therapeutic target as it has been associated with longer time spent in steroid-free clinical remission, decreased risk of hospitalization, slower progression of bowel damage and reduced risk of subsequent surgery²⁻⁶. Furthermore, recent works suggested that transmural healing could lead to better outcomes, such as prevention of bowel damage progression, than endoscopic remission^{2,5-7}. However, despite this accumulation of arguments, the recent international guidelines STRIDE 2 did not recommend its use in daily practice⁸.

In 2014, Sandborn and colleagues considered that the ideal treatment goal in any chronic disease was one that is clearly defined, achievable with medical or surgical therapy, predictive of long-term outcomes, affordable, non-invasive and relevant across disease subtypes, with a low test-to-test variability⁹. The two main scientific missing characteristics to retain transmural healing as the best therapeutic target in CD are the lack of consensual definition and insufficient evidence of being reachable with the current medications.

In the DEVISE-CD project, we aimed to 1) build and validate a dedicated score to assess transmural response (TR) and transmural healing (TH) in patients with CD, 2) to confirm that these definitions of TR and TH are associated with improved outcomes and 3) to describe the proportion of CD patients that can achieve these specific endpoints with anti-TNF therapy.

PATIENTS AND METHODS

Ethical considerations

The study was approved by local Ethics Committee (IRB00013412, “CHU de Clermont Ferrand IRB #1”, IRB number 2022-CF065 and CPP Sud-Est VI [#2014/CE 72]).

Selected population

For the development cohort, we retrospectively screened all consecutive CD patients who underwent MRI in our IBD unit between January 2012 and June 2018 from a dedicated database. We included all the patients who had at least two MRIs: 1) with objective signs of inflammation on 1st MRI (at least one sign among edema, bowel thickening or ulcer), 2) 2nd MRI indicated for re-evaluation of disease activity and 3) follow-up >6 months or occurrence of surgery or bowel damage progression after the 2nd MRI in case of shorter follow-up (**Supplementary Figure S1**). The patients who had bowel resection between the two MRI scans were excluded. In case of repeated MRIs (more than 2 MRIs) the closest was selected for the study. The same inclusion criteria were applied for the monocenter validation cohort in patients who underwent MRI between June 2018 and January 2022 from the same MRI database. We used IRMA study⁴, a multicenter prospective study as another validation cohort.

Magnetic resonance imaging

MRI protocol was detailed in **Supplementary Appendix 1** and **Supplementary Table 1**. It was similar for the three cohorts. MRI were reviewed independently by two experienced IBD radiologists (C.H. and J.J) blinded from clinical data. The bowel segmentation for MRI review is detailed in **Supplementary Appendix 1**.

Description of the step-by-step process to build the modified Clermont score for assessing transmural therapeutic efficacy in Crohn’s disease

The process is described in **Figure 1 and supplementary Appendix 2**.

Endpoints

Bowel damage progression was considered in the case of bowel resection, new stricture or fistula, or worsening of pre-existing stricture (increase in pre-stenotic dilation on imaging or occurrence of major obstructive symptoms) or fistula (new abscess or new fistula tract) as previously defined in several studies^{4,5}. For the multicenter validation cohort, namely the IRMA study, corticosteroid-free remission (CFREM) at week 52 (W52)⁴.

Statistical considerations

Statistical analyses

Sample size calculation is detailed **in Supplementary Appendix 3**. To assess persistence of MRI lesions, or C-score as predictor of bowel damage progression (censored data), estimates were constructed using the Kaplan-Meier method. Log-rank test was used for univariate analyses and Cox proportional hazards regression in multivariable analysis. The proportional-hazard hypothesis was verified using Schoenfeld's test and plotting residuals. Results were expressed as hazard-ratio (HR) and 95% confidence interval.

To evaluate the coexistence of MRI lesions at baseline MRI (multicollinearity), categorical data have been analyzed estimating the concordance rate (accuracy) and the Kappa agreement coefficient (κ). The relationships between quantitative variables were explored with Pearson or Spearman correlation coefficients, according to the statistical distribution. Finally, for relations between categorical and continuous variables, effect-sizes were estimated from marginal models with generalized estimating equations carried out to take into account several measures (*i.e.* segments) from a same patient. The results were expressed as effect size (ES) with 95% confidence intervals.

The sensitivity to change was evaluated estimating standardized mean difference (SMD) (calculated as the difference between values of two MRI divided by the standard-deviation of value at first MRI) and compared by paired tests (Student paired test or Wilcoxon test). For categorical data, Cramer Index (V) was estimated with 95% confidence interval. V indexes were compared using 95%CI and Stuart-Maxwell test.

The MRI items suitable to build the modified Clermont score for assessing transmural therapeutic efficacy in Crohn's disease (C-score) was selected according to clinical relevance and to aforementioned analyses, completed by LASSO approach. Furthermore, multicollinearity was tested with variance inflation factor index.

In multicenter prospective validation cohort, TH, TR25 and TR50 were tested as predictor of CFREM using mixed models with center as random-effect. More precisely, generalized linear mixed model was used. Results were expressed as odds-ratio (OR) and 95% confidence interval. Of note, C-score is calculated in each active segment at baseline and its capability to predict outcomes is based on the achievement of this specific endpoint in each analyzed segment.

Statistical analyses were performed using Stata software (version 15, StataCorp LP, College Station, US). The tests were two-sided, with a type I error set at 5%.

RESULTS

1. Development cohort

Population study

Overall, 443 patients undergoing 889 MRI were screened for the development cohort. Among them, 274 patients were included (**Supplementary Figure S1**). Their characteristics at the time of the second MRI are detailed in **Supplementary Table S2**. The median interval between the two MRIs was 9.2 months [6.0 - 14.1]. The median follow-up was 14.5 months [4.0-31.5]. Among the 274 patients, 53 (19.3%) required bowel resection over time and progression of bowel damage occurred in 106 patients (38.6%).

Identification of persistence of MRI lesions after treatment as predictor of bowel damage progression

On the first MRI, we identified 413 active segments in 274 patients. Among them, 63.1% (173/274) had only one active segment. The number of patients with each MRI lesions in at least one segment on the first and the second MRI is detailed in **Supplementary Table S3**. After univariate and multivariable analyses adjusted on potential confounders (**Supplementary Table S3**), the detection on the second MRI of all MRI lesions except for length of the lesions were associated with the risk of bowel damage progression (**Supplementary Table S3**).

Coexistence and concordance of MRI lesions at baseline MRI

Edema (> 95%), bowel thickening (bowel wall > 3 mm) (> 99%), increased contrast enhancement (> 97%) and diffusion-weighted hyperintensity (> 94%) were the more frequent lesions observed in case of active segments (N=413) (**Figure 2A**). They were concomitantly observed two-by-two in more than 94 % of the segments giving similar information, *i.e.*

detection of active segment (**Supplementary Table S4**). The quantitative parameters of diffusion-weighted (ADC) or injected sequences (RCE) as well as bowel thickness (continuous variable) were also highly concordant with edema ($ES > 0.8$ for each) provided close information (**Supplementary Table S4, Supplementary Figure S2**).

Ulcer was observed in only two thirds of the active segments (between 65.6% and 68.2%) (**Figure 2**) with negligible to mild concordance with lesions reflecting bowel wall inflammation (edema, bowel thickening, increased contrast enhancement and diffusion-weighted hyperintensity) meaning that detection of ulcer provides additional information (**Supplementary Table S4**). There was no concordance between the presence of peri-enteric lesions, *i.e.* fat wrapping and enlarged lymph nodes, and lesions reflecting bowel wall inflammation (**Supplementary Table S4**). The concordance of ulcer and peri-enteric lesions was mild (κ -coefficient ranging from 0.19 to 0.24) (**Supplementary Table S4**). There was no concordance between the presence of fat wrapping and enlarged lymph nodes ($\kappa = 0.01$) (**Supplementary Table S4**). Results on concomitant healing of MRI lesions at the time of the second MRI are detailed in **Supplementary Appendix 3, Supplementary Table S4 and Supplementary Figure S4**).

Sensitivity to change after therapeutic intervention

Among 413 active segments (including 98 colonic segments: right colon = 38, transverse colon = 21, left/sigmoid colon = 25 and rectum = 14) on baseline MRI, the sensitivity to change after treatment was mild for bowel thickness ($SMD=0.30\pm 1.01$) and RCE ($SMD=0.23\pm 1.37$), but negligible for ADC ($SMD=0.19\pm 0.09$), edema ($SMD=0.16[0.08-0.21]$), contrast enhancement ($SMD=0.19[0.12-0.24]$), diffusion-weighted hyperintensity ($SMD=0.16[0.09-0.21]$), ulcer ($SMD=0.17[0.09-0.22]$), fat wrapping ($SMD=0.03[0.00-0.08]$) and enlarged lymph nodes ($SMD=0.04[0.00-0.09]$) (**Figure 3**). We did not find any significant

different between small bowel and colonic segments regarding sensitivity to change after treatment (**Supplementary Appendix 4**).

Selection of MRI items to build the modified Clermont score for assessing transmural therapeutic efficacy in Crohn's disease (C-score)

To identify the segments where the score should be calculated (active segments), 4 items were eligible: edema, bowel thickening (bowel wall > 3 mm), increased contrast enhancement, and diffusion-weighted hyperintensity. We chose edema as it was the most “easy to use” for radiologist, did not need any specific sequence (injected or diffusion-weighted) and was lacking in case of fibrotic lesions (contrary to bowel thickening).

Among the quantitative parameters available to assess the variation after treatment with very high agreement, bowel thickness has a higher sensitivity to change than RCE or ADC and is more convenient for radiologist. After multivariable analysis, bowel thickening compared to the normal value (bowel thickness in mm – 3) was associated with a greater risk of bowel damage progression (aHR=1.19[1.06–1.32], p=0.003) (**Supplementary table S5**).

Among the other items, which are giving independent information, *i.e.* ulcer, enlarged lymph nodes and fat wrapping, a multivariable model showed that ulcer (HR=2.31[1.11–4.76], p=0.024) and enlarged lymph nodes (HR=1.62[1.01–2.61], p=0.048) are associated with the risk of bowel damage progression (**Supplementary table S5**). The selected items have been rounded for best convenience for users. How to calculate C-score is detailed in **Table 2**.

C-score as predictor of favorable outcomes

Based on 274 patients, the C-score after treatment was associated with higher risk of bowel damage progression (aHR=1.43[1.21–1.70], p < 0.001). TH defined as C-score < 0.5

(HR=0.28[0.13-0.63], $p=0.002$ (**Figure 4A**), TR defined as either TR50 (at least 50%-improvement of C-score in each active segment at baseline) (HR=0.30[0.15-0.63], $p=0.001$) (**Figure 4B**) or TR25 (at least 25%-improvement of C-score in each active segment at baseline) (HR=0.37[0.19-0.71], $p=0.003$) (**Figure 4C**) were associated with lower risk of bowel damage progression. The rationale for choosing these cut-off values is given in Supplementary Appendix 5. In addition, TH (HR=0.07[0.01-0.55], $p=0.011$), TR50 (HR=0.32[0.12-0.82], $p=0.017$) and TR25 (HR=0.38[0.17-0.87], $p=0.022$) were associated with reduced risk of bowel resection (**Supplementary Figure S6, S7 and S8**, respectively). We performed a sensitivity analysis focusing on the subgroup of patient with inflammatory phenotype at baseline (no stricture and no fistula) and confirmed that TH according to C-score was associated to lower risk of bowel damage progression (HR = 0.29 [0.09-0.96], $p= 0.043$). TH was also associated with reduced risk of bowel damage progression among patients with complicated phenotype (stricture or fistula) at baseline (HR = 0.40 [0.19-0.84], $p= 0.016$).

2. Monocenter validation cohort

Overall, 224 patients were included (**Table 1**). The rate of TH (C-score < 0.5), TR 50 and TR25 were 16.6% (37/224), 21.9% (49/224), and 34.4% (77/224), respectively. The median interval between the two MRIs was 12.5 months [8.9–18.6]. The median follow-up was 19.6 months [7.9 - 31.3]. In multivariable analysis, TH was associated with lower risk of progression of bowel damage (aHR=0.15[0.04-0.53], $p=0.003$) and surgery (aHR=0.07[0.01-0.57], $p=0.014$). TR50 (aHR=0.36[0.14-0.88], $p=0.025$; aHR=0.28[0.08-0.97], $p=0.045$, respectively) and TR25 (aHR=0.46[0.23-0.94], $p=0.034$; aHR=0.31[0.11-0.85], $p=0.024$) were associated with a reduced risk of progression of bowel damage (**Figures 4D, 4E, and 4F**) and surgery (**Supplementary Figure S9, S10 and S11**).

Due to the limited number of patients with isolated colonic CD, we pooled data from the two

cohorts to investigate whether the association between TH (defined using C-score) and the risk of bowel damage progression is also generalizable to pure colonic CD. We found that TH was associated with lower risk of bowel damage ($p = 0.042$) in this subgroup (**Supplementary Figure S12**).

3. Multicentre validation cohort

Among the 46 patients enrolled in the prospective validation cohort (**Table 1**), The rate of TH, TR50 and TR25 were 28.3% (13/46), 41.3% (19/46), and 47.8% (22/46), respectively. Achieving TH (OR=4.6[1.3–15.6], $p=0.016$), TR50 (OR=6.9[1.8–26.0], $p=0.008$) or TR25 (OR = 6.0[1.6 – 22.3], $p=0.008$) after 12 weeks of anti-TNF therapy led to higher rate of steroid-free remission at week 52. The performances of TH, TR50 and TR25 according to C-score are illustrated in **Figure 5**.

DISCUSSION

In the DEVISE-CD project, we developed and validated an MRI score dedicated to the evaluation of TH and TR in daily practice and clinical trials in patients with CD using an original and pragmatic approach. We also provided clear and easy-to-use definitions of TH and TR (TR50 and TR25) that are associated with reduced risk of bowel damage progression and surgery, as well as higher rate of mid-term clinical and biological remission. Furthermore, we showed that these endpoints are achievable in a significant part of CD patients treated with anti-TNF therapy.

While the association between TH and favorable outcomes is now well established, TH is not yet recommended in daily practice. One of the main drawbacks of TH is the lack of consensual definition. Rather than using one or several available MRI scores or indexes to define TH in CD¹⁰⁻¹⁴, we decided to conduct the DEVISE-CD project because available MRI scores have not been developed to assess TH. Our team, like the others, were mistaken in developing MRI scores directly or indirectly in comparison to endoscopy^{10,11}. Endoscopy is not suitable to evaluate TH because it focuses on the mucosa. Consequently, the most popular MRI scores such as MaRIA¹⁰ or Clermont score¹¹ are reliable to assess mucosal healing¹⁵, but still less sensitive to detect superficial mucosal lesions than colonoscopy. MaRIA was obtained by multivariate regression model predicting the presence of ulcers in endoscopy. The original Clermont score was derived from the MaRIA and subsequently validated compared to endoscopy. It has been demonstrated that TH and endoscopic remission are not strongly concordant^{2,5,7}. If IBD physicians want to exploit the full potential of MRI to define TH, a dedicated tool is necessary. TH and endoscopic remission have close performances to predict outcomes related to disease activity (symptomatic relapse, drug discontinuation or therapeutic escalation), but TH is associated with reduced risk of long-term pejorative course such as surgery or bowel damage progression^{2,5,7}. Preventing bowel damage is now recognized as a

major long-term objective in the management of patients with CD. As there was no obvious gold standard to develop our new MRI score dedicated to assessing TH, we chose the prediction of bowel damage progression as reference, which seemed to us the most relevant option from a physician's point of view.

During the development stage, we identified that the persistence all individual CD-related MRI lesions were associated with the risk of bowel damage progression except the length of active lesions. This finding is meaningful because endoscopic scores (CDEIS and SES-CD) are taking into account the affected area. Our data confirm that considering the most severe lesion is more impactful than the affected length to predict the risk of bowel damage progression. In addition, we did not include stricture and fistula into our model as these two complications are a marker of bowel damage and not a marker of inflammatory activity¹⁶. Bowel damage and disease activity are two different concepts. Our goal was to develop an MRI score assessing transmural activity and predict long-term risk of bowel damage progression.

Edema, bowel thickening (bowel wall > 3 mm), increased contrast enhancement, and diffusion-weighted hyperintensity were highly concordant (coexistence > 94 %) meaning that these four lesions give the same information, *i.e.* detection of active segments. Bowel thickness (continuous variable), RCE, and ADC, which are reflecting the level of bowel wall inflammation, provided also close information. Due to high coexistence between MRI lesions translating into statistical wording as substantial collinearity, usual multivariable models were not suitable. Thus, we based our selection of items on clinical relevance. For disease activity, we did not retain increased contrast enhancement or diffusion-weighted hyperintensity due to the need of additional sequences^{10,11}. As mild bowel thickening could be due to fibrotic complications, we selected edema to detect active segments that should be further assessed to calculate the dedicated score. To quantify the degree or level of inflammation, bowel thickness (quantitative value) was the most sensitive to change after therapeutic intervention. In addition,

we did not select RCE or ADC owing to the need of additional sequences, the additional time-consuming nature of assessing these items, which did not take part of routine practice, and the questioning about the inter-reader reliability due to different MRI scans¹⁷.

Thanks to the DEVISE-CD project, the modified Clermont score for assessing TH, so-called C-score, is calculated in active segments (defined by the presence of edema) and includes three items: bowel thickness as quantitative value, presence of ulcer and enlarged lymph nodes. These lesions are commonly written down on radiological report, confirming the feasibility of using the C-score in daily practice and clinical trials. While bowel thickness is reflecting the level of inflammation and ulcer is a mucosal marker of severity, enlarged lymph nodes illustrates the presence of peri-enteric inflammation. Of note, enlarged lymph nodes was significantly associated with endoscopic ulcer or endoscopic severity in works from the Barcelona team^{12,18}, but it is still unclear why it was not retained in final score. Contrary to existing available MRI scores built using endoscopy^{10-13,15} or surgical specimen¹⁴ as reference, C-score has been specifically developed to assess TH and predict favorable CD outcomes, does not consist exclusively of binary items (major problem for assessing response)^{12,13}, does not include overlapping items¹³, and does not require burdensome and time-consuming evaluations of quantitative value^{10,11}.

The DEVISE-CD project proposes three clearly defined endpoints such as TH (C-score < 0.5), TR50 and TR25. Using these definitions, we confirmed that TH or TR are associated with higher rate of CFREM, lower risk of surgery and reduced risk of bowel damage progression, which is in line with the current literature on TH to predict favorable course in patients with CD²⁻⁷. Of note, the rate of TH, TR50 and TR25 after three months of anti-TNF therapy was achieved by a significant proportion of patients with CD (28.3%, 41.3% and 47.8%, respectively) which is in line with available data^{6,19}.

The retrospective design of two cohorts is a limitation of our study, even though the evaluation of bowel damage, requiring a significant follow-up, could have compromised the feasibility of two different prospective cohorts. The heterogeneity of medications is also a potential limitation even if this study was not designed to assess a specific medication, but rather to assess outcomes, regardless of the medication used. In addition, the low number of isolated colonic CD makes difficult to draw any firm conclusion in this specific subgroup. In contrast, several strengths could be underlined such as an original and pragmatic approach, the use of three independent cohorts (two large sample size retrospective cohorts and one multicenter prospective cohort) and meaningful results for IBD physicians applicable in daily practice.

In conclusion, C-score is a validated, reliable and easy-to-use tool to assess TH and TR in patients with CD. TH or TR according to C-score are non-invasive and clearly defined endpoints that are achievable with medical therapy, and predictive of favorable long-term outcomes. TH or TR defined by C-score should be considered as therapeutic goal in daily practice and clinical trials.

FIGURES LEGENDS

Figure 1: Flow chart describing the process of C-score building.

Figure 2: Coexistence of MRI lesions at baseline in 274 patients with Crohn's disease.

Figure 3: Sensitivity to change after treatment of each MRI lesions.

Figure 4: Bowel damage progression-free survival according to achievement of transmural healing or transmural response (TR50 and TR25) in patients with Crohn's disease enrolled in development (4A-C) and validation cohorts (4D-F).

Figure 5: Performances of transmural healing or transmural response (TR50 and TR25) after anti-TNF induction therapy to predict CFREM at week 52.

LEGENDS OF SUPPLEMENTARY FIGURES

Supplementary Figure S1: Design of the study

Supplementary Figure S2: Flow chart illustrating the selection of the patients.

Supplementary Figure S3: Concordance of MRI lesions at baseline. *Concordance between qualitative et quantitative parameters as sample size.*

Supplementary Figure S4: Concomitant healing between normalization of each MRI lesion between pre- and post-treatment MRI in 274 patients with Crohn's disease.

Supplementary Figure S5: Concordance between normalization of each MRI lesion between pre- and post-treatment MRI in 274 patients with Crohn's disease. *Concordance between qualitative et quantitative parameters as sample size.*

Supplementary Figure S6: Bowel resection-free survival according to achievement of transmural healing (C-score < 0.5) in the development cohort of patients with Crohn's disease.

Supplementary Figure S7: Bowel resection-free survival according to achievement of transmural response (TR50) in the development cohort of patients with Crohn's disease.

Supplementary Figure S8: Bowel resection-free survival according to achievement of transmural response (TR25) in the development cohort of patients with Crohn's disease.

Supplementary Figure S9: Bowel resection-free survival according to achievement of transmural healing (C-score < 0.5) in validation cohort of patients with Crohn's disease.

Supplementary Figure S10: Bowel resection-free survival according to achievement of transmural response (TR50) in validation cohort of patients with Crohn's disease.

Supplementary Figure S11: Bowel resection-free survival according to achievement of transmural response (TR25) in validation cohort of patients with Crohn's disease.

Supplementary Figure S12: Bowel damage progression-free survival according to achievement of transmural healing (C-score < 0.5) in pooled cohort of patients with isolated colonic Crohn's disease.

Journal Pre-proof

TABLES**Table 1: Recommendation for calculating the modified Clermont score for assessing transmural therapeutic efficacy in Crohn's disease (C-score)**

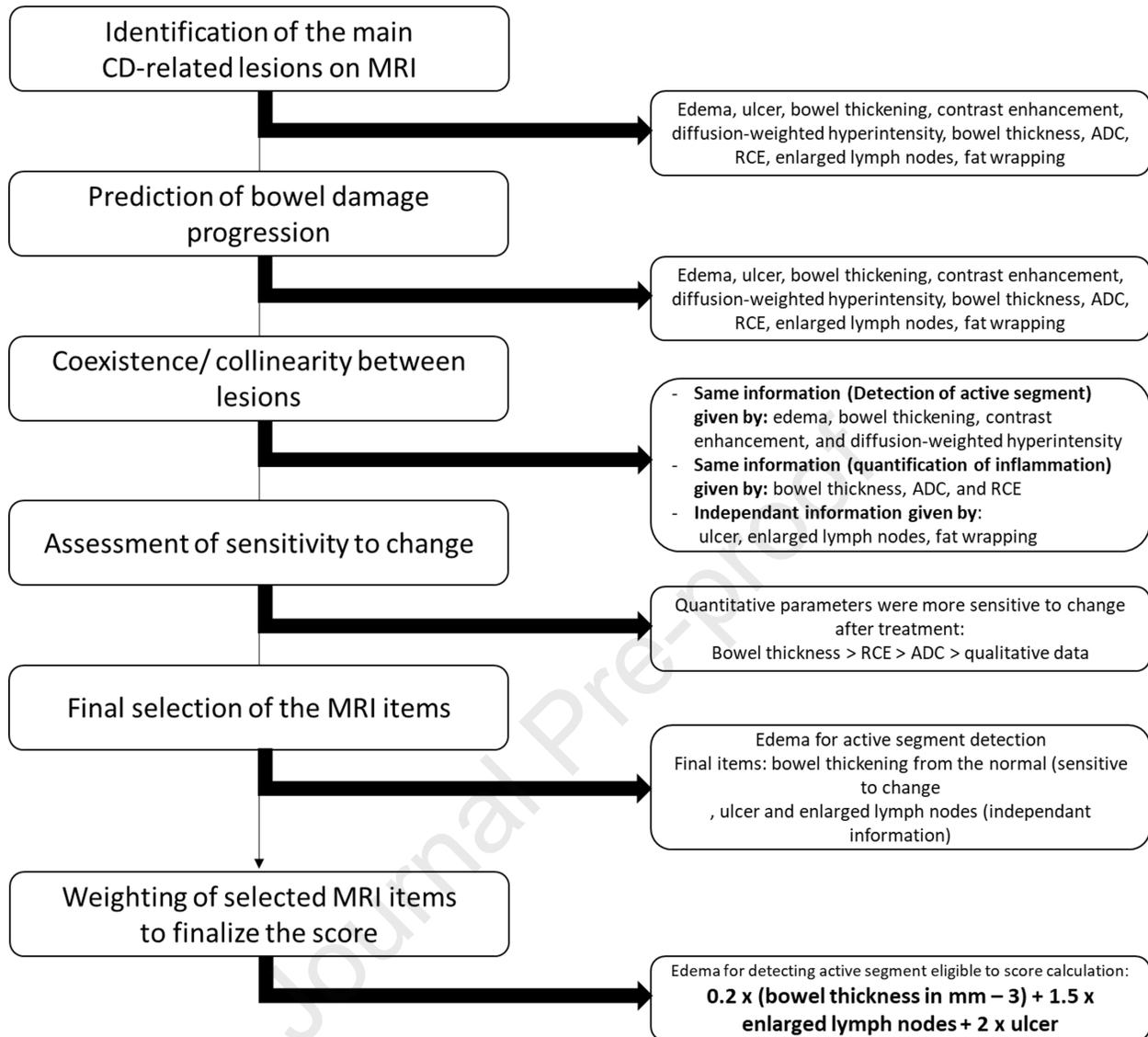
Items	Calculation	Results
Bowel wall thickening compared to the normal	Bowel wall thickness (mm) – 3	
Presence of any ulcer	+ 2	
Presence of enlarged lymph nodes	+ 1.5	
TOTAL		

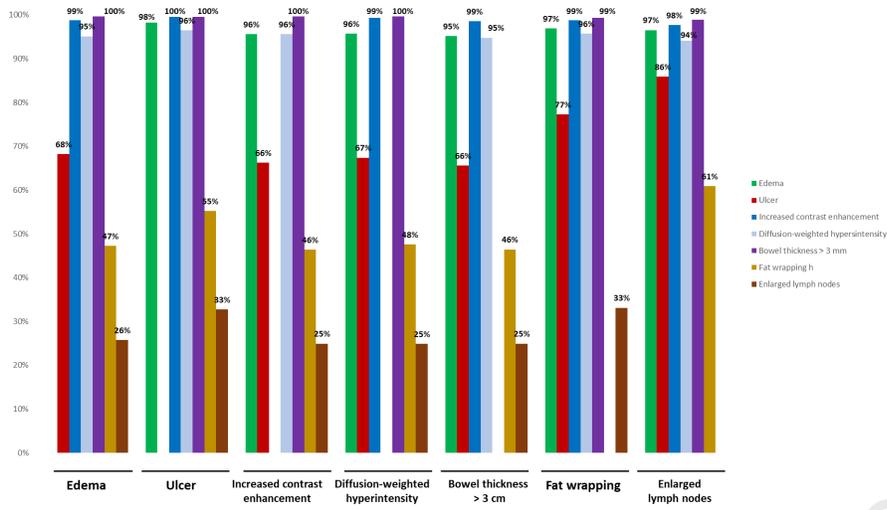
The C-score is calculated on the most severe lesion within a segment and should be calculated on each active segment for a same patient.

REFERENCES

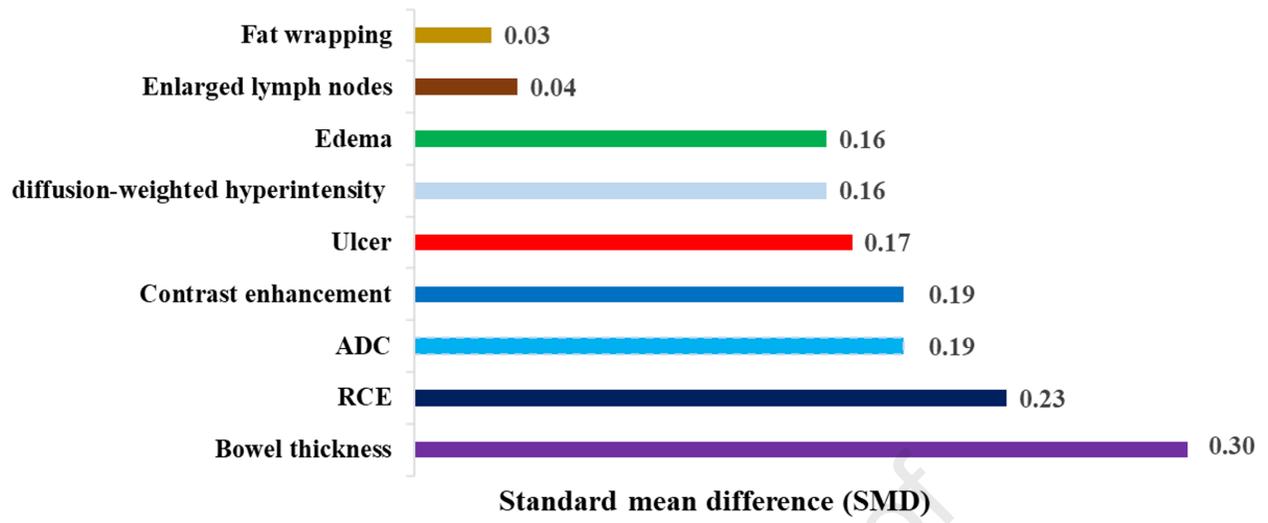
1. Buisson A, Gonzalez F, Poullenot F, et al. Comparative Acceptability and Perceived Clinical Utility of Monitoring Tools: A Nationwide Survey of Patients with Inflammatory Bowel Disease. *Inflamm Bowel Dis* 2017;23:1425–1433.
2. Fernandes SR, Rodrigues RV, Bernardo S, et al. Transmural Healing Is Associated with Improved Long-term Outcomes of Patients with Crohn's Disease. *Inflamm Bowel Dis* 2017.
3. Buisson A, Hordonneau C, Goutorbe F, et al. Bowel wall healing assessed using magnetic resonance imaging predicts sustained clinical remission and decreased risk of surgery in Crohn's disease. *J Gastroenterol* 2019;54:312–320.
4. Messadeg L, Hordonneau C, Bouguen G, et al. Early Transmural Response Assessed Using Magnetic Resonance Imaging Could Predict Sustained Clinical Remission and Prevent Bowel Damage in Patients with Crohn's Disease Treated with Anti-Tumour Necrosis Factor Therapy. *J Crohns Colitis* 2020;14:1524–1534.
5. Lafeuille P, Hordonneau C, Vignette J, et al. Transmural healing and MRI healing are associated with lower risk of bowel damage progression than endoscopic mucosal healing in Crohn's disease. *Aliment Pharmacol Ther* 2021;53:577–586.
6. Takenaka K, Kawamoto A, Kitazume Y, et al. Transmural Remission Characterized by High Biologic Concentrations Demonstrates Better Prognosis in Crohn's Disease. *J Crohns Colitis* 2023;17:855–862.
7. Fernandes SR, Serrazina J, Botto IA, et al. Transmural remission improves clinical outcomes up to 5 years in Crohn's disease. *United Eur Gastroenterol J* 2023;11:51–59.
8. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: An Update on the STRIDE Initiative of the International Organization for the Study of IBD (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. *Gastroenterology* 2021;160:1570–1583.
9. Sandborn WJ, Hanauer S, Van Assche G, et al. Treating beyond symptoms with a view to improving patient outcomes in inflammatory bowel diseases. *J Crohns Colitis* 2014;8:927–935.
10. 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;58:1113–1120.
11. 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;109:89–98.
12. Ordás I, Rimola J, Alfaro I, et al. Development and Validation of a Simplified Magnetic Resonance Index of Activity for Crohn's Disease. *Gastroenterology* 2019;157:432-439.e1.
13. Oussalah A, Laurent V, Bruot O, et al. Diffusion-weighted magnetic resonance without bowel preparation for detecting colonic inflammation in inflammatory bowel disease. *Gut* 2010;59:1056–1065.

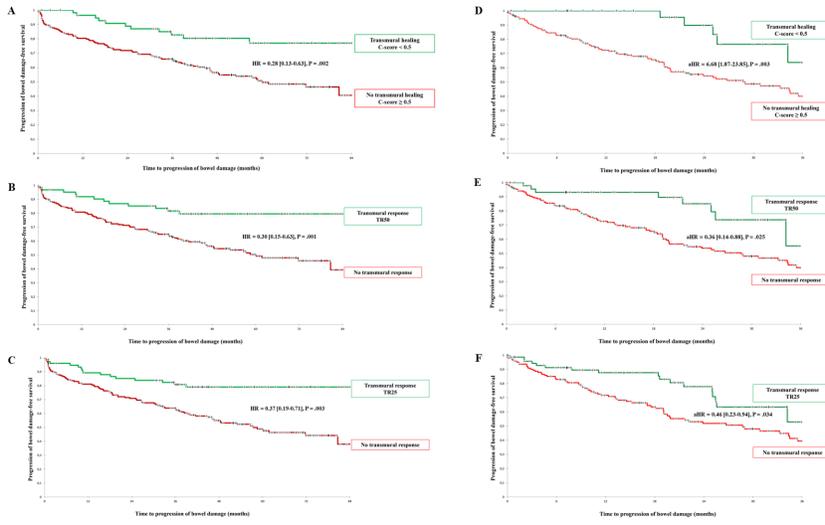
14. Steward MJ, Punwani S, Proctor I, et al. Non-perforating small bowel Crohn's disease assessed by MRI enterography: derivation and histopathological validation of an MR-based activity index. *Eur J Radiol* 2012;81:2080–2088.
15. Buisson A, Hordonneau C, Goutte M, et al. Diffusion-weighted magnetic resonance imaging is effective to detect ileocolonic ulcerations in Crohn's disease. *Aliment Pharmacol Ther* 2015;42:452–460.
16. Pariente B, Torres J, Burisch J, et al. Validation and Update of the Lémann Index to Measure Cumulative Structural Bowel Damage in Crohn's Disease. *Gastroenterology* 2021;161:853-864.e13.
17. Buisson A, Hordonneau C, Goutte M, et al. What is the Role of Diffusion-weighted Imaging in Ileocolonic Crohn's Disease? *Inflamm Bowel Dis* 2015;21:E9.
18. 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;146:374-382.e1.
19. Fernandes SR, Bernardo S, Saraiva S, et al. Proactive Infliximab Monitoring Improves the Rates of Transmural Remission in Crohn's Disease: A Propensity Score-Matched Analysis. *Inflamm Bowel Dis* 2023:izad272.



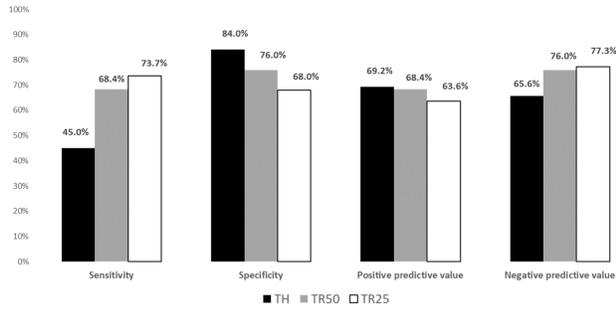


Journal Pre-proof



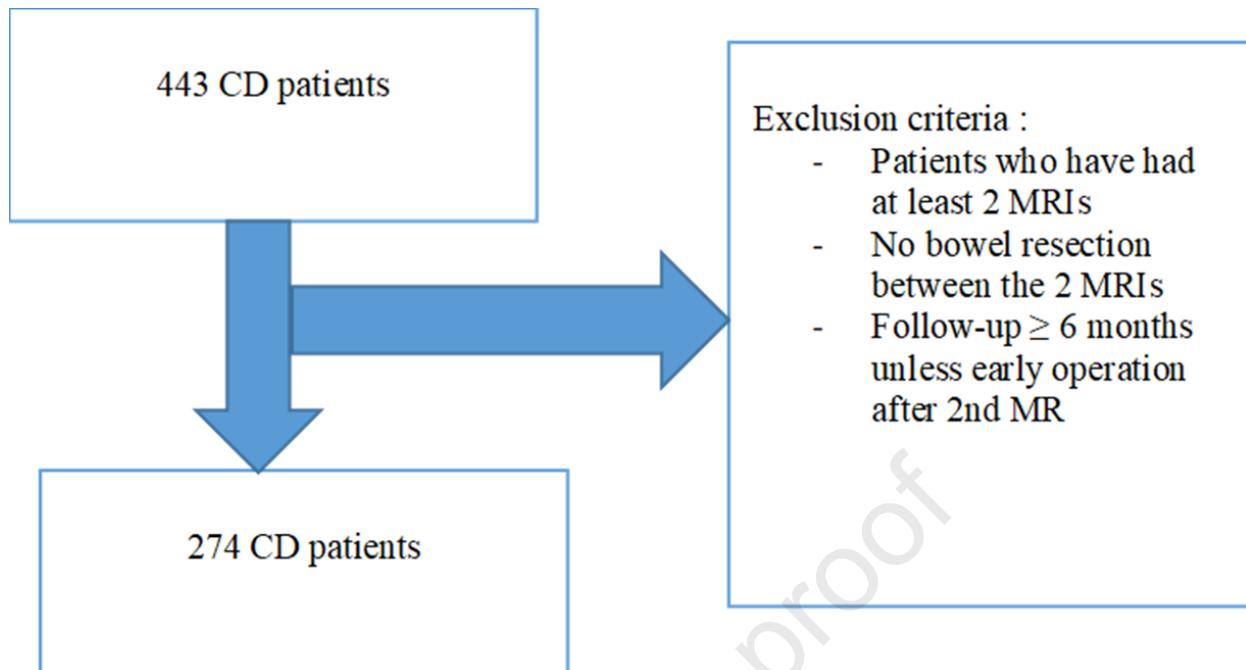


Journal Pre-proof

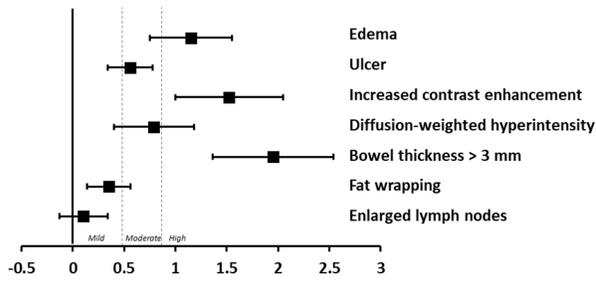


Endpoint using C-score	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Transmural healing (TH)	45.0 [25.9 – 65.8]	84.0 [64.6 – 94.1]	69.2 [44.1 – 94.3]	65.6 [49.2 – 82.1]
Transmural response (TR50)	68.4 [45.8 – 84.7]	76.0 [56.1 – 88.7]	68.4 [47.5 – 89.3]	76.0 [59.3 – 92.7]
Transmural response (TR25)	73.7 [50.8 – 88.4]	68.0 [48.2 – 82.8]	63.6 [43.5 – 83.7]	77.3 [59.8 – 94.8]

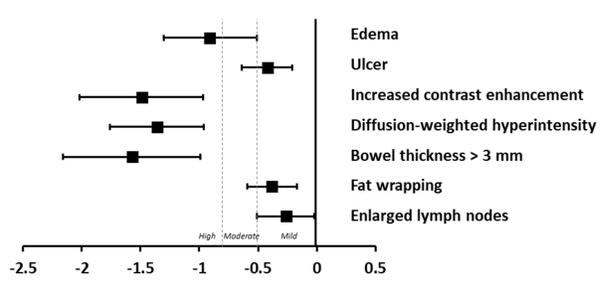
Journal Pre-proof



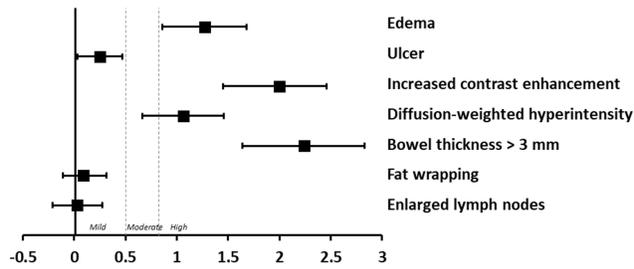
Bowel thickness (continuous variable)



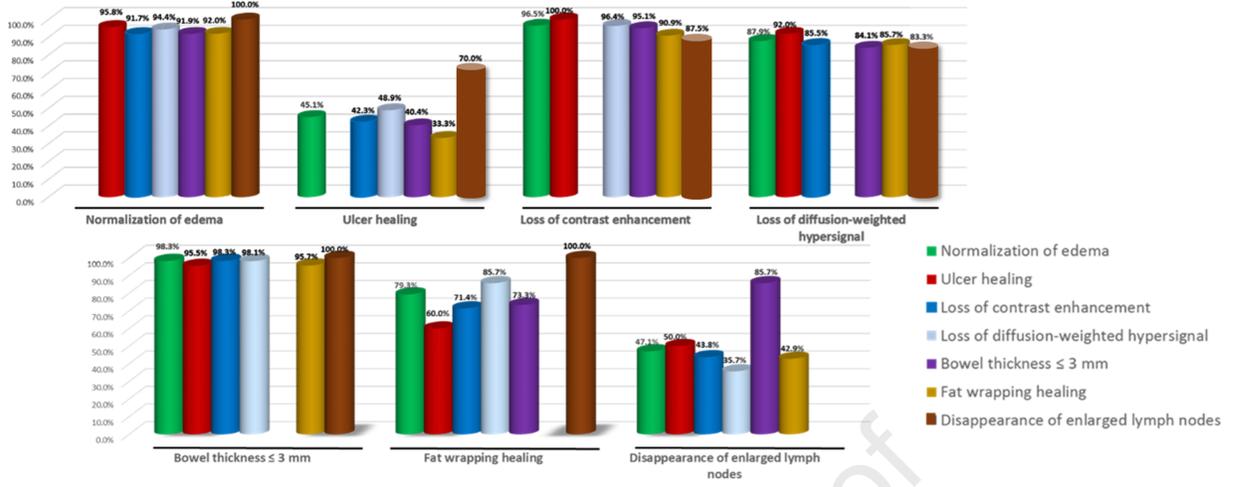
Apparent Diffusion Coefficient (ADC)



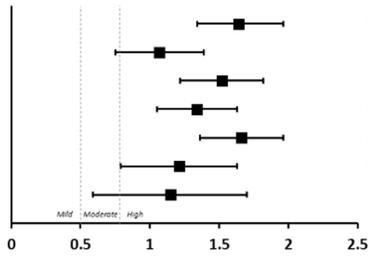
Relative contrast enhancement (RCE)



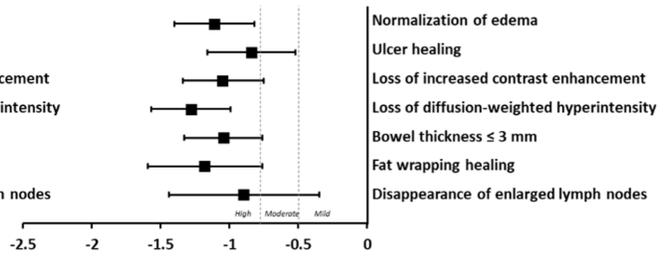
Journal Pre-proof



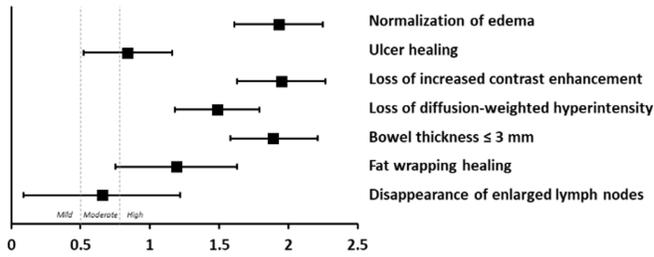
Variation of bowel thickness



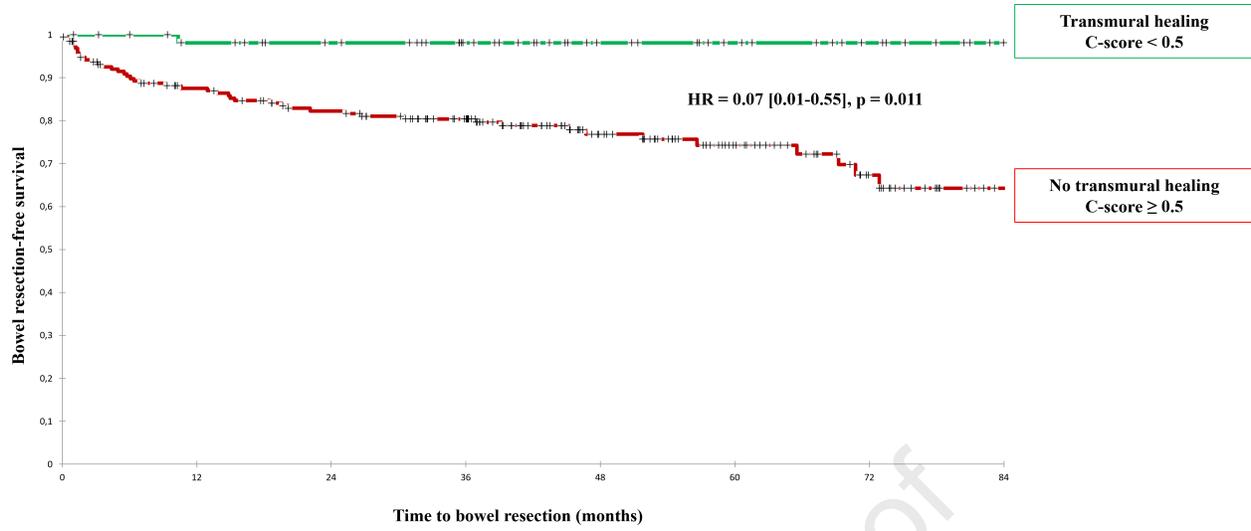
Variation of Apparent Diffusion Coefficient (ADC)

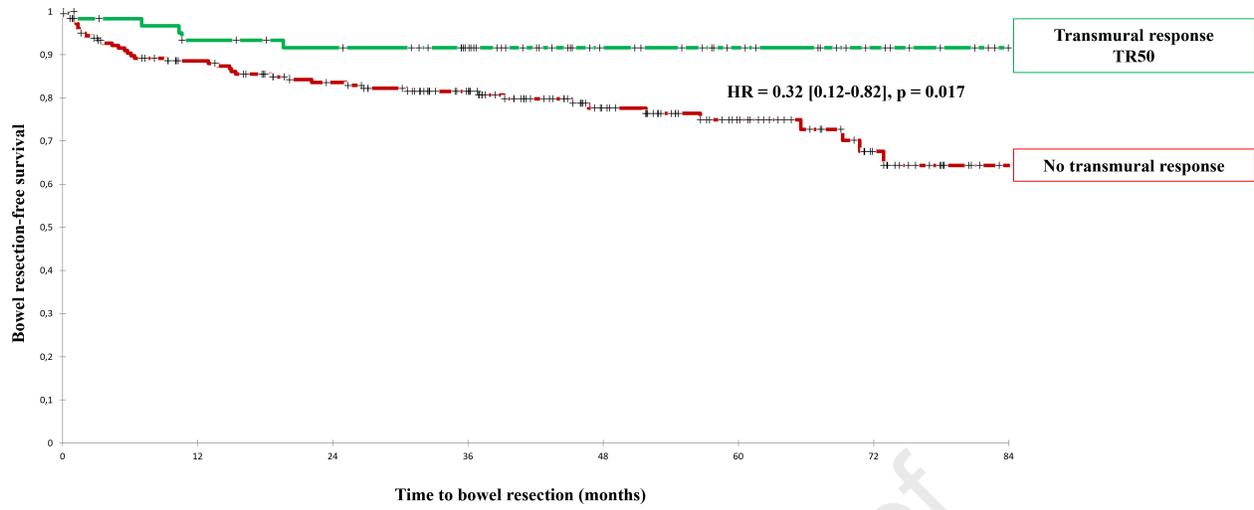


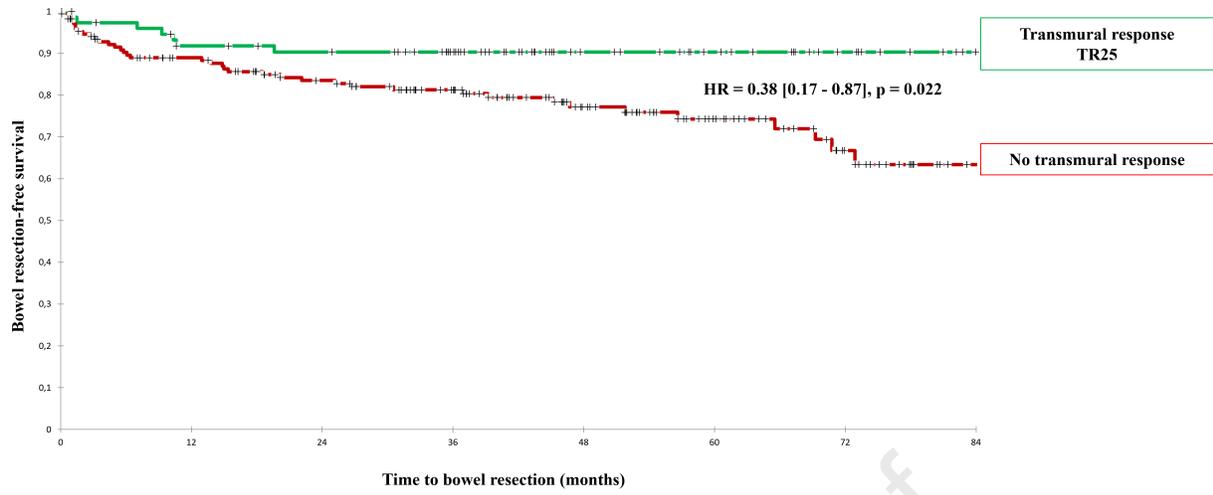
Variation of relative contrast enhancement (RCE)

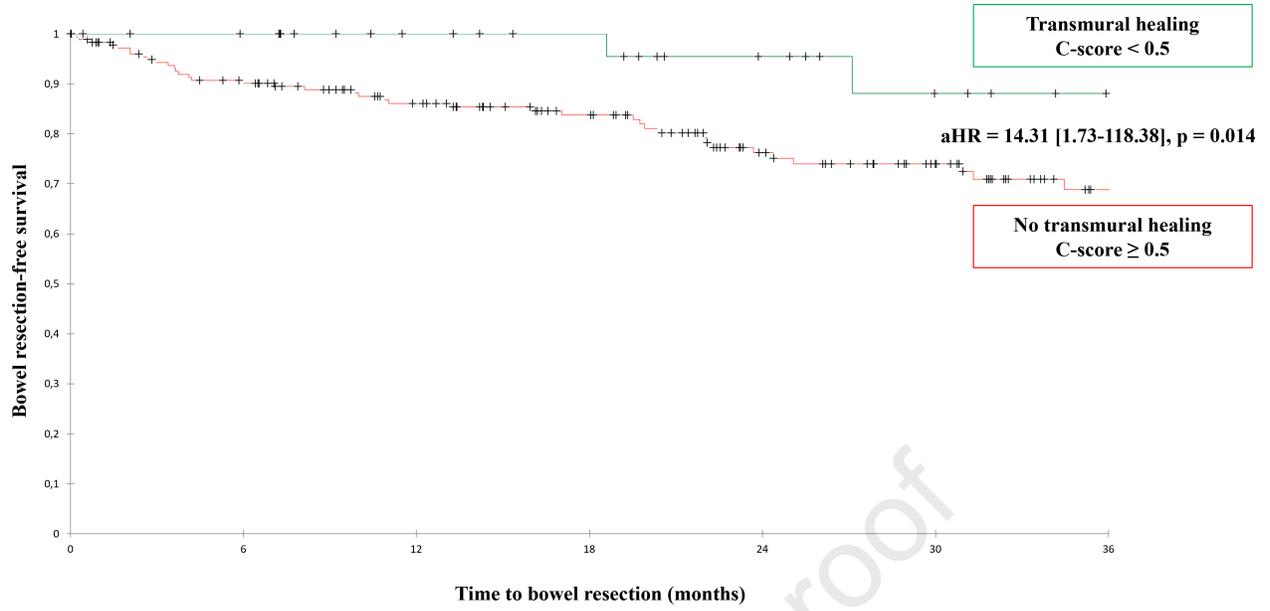


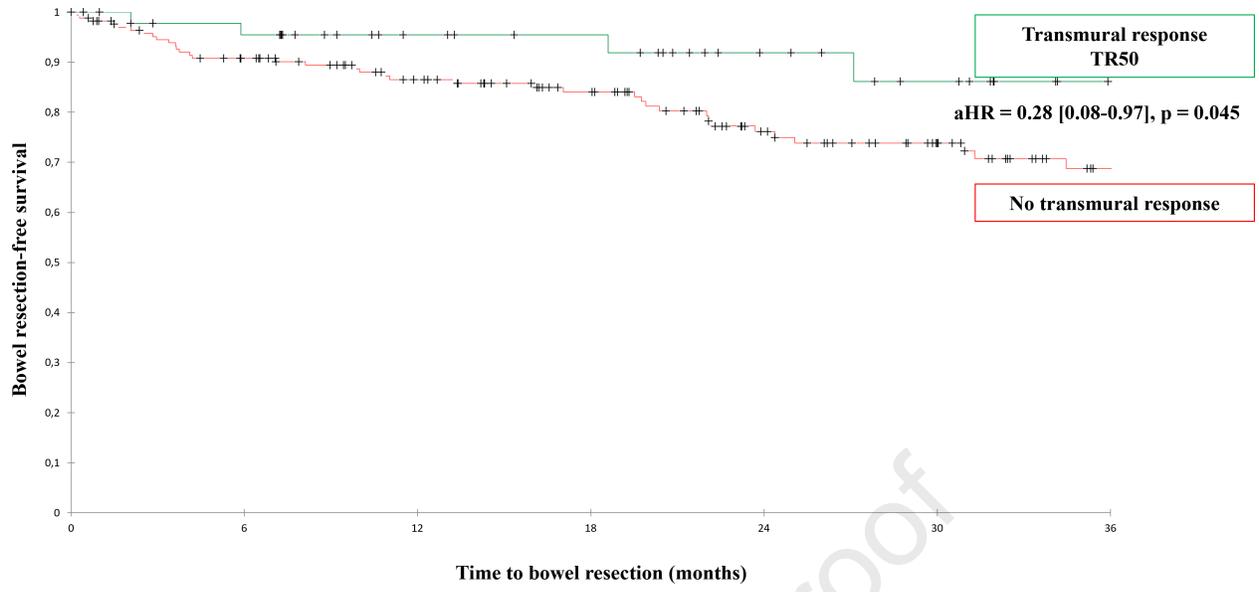
Journal Pre-proof

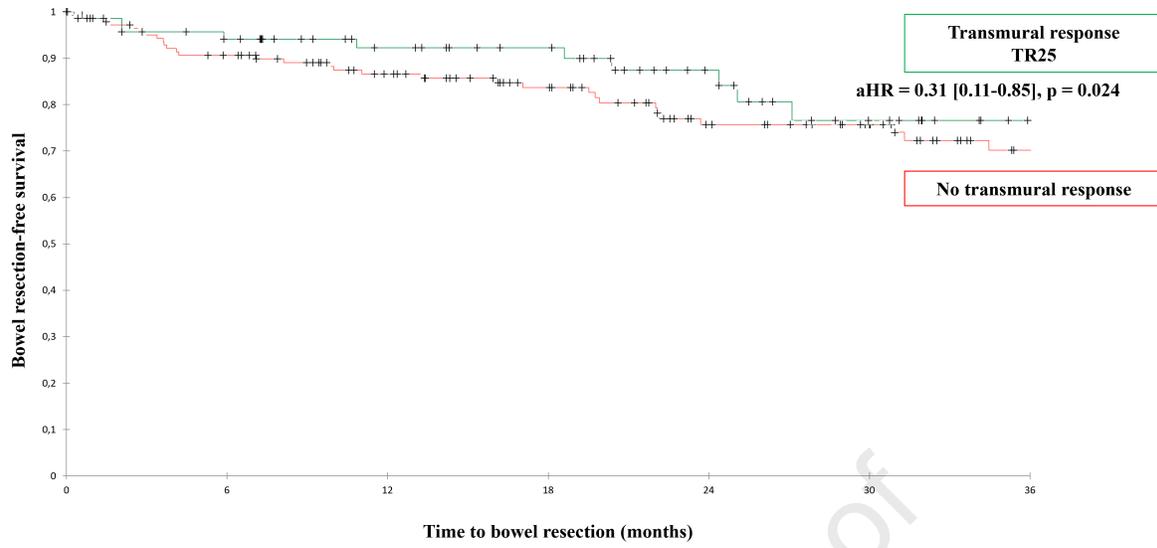


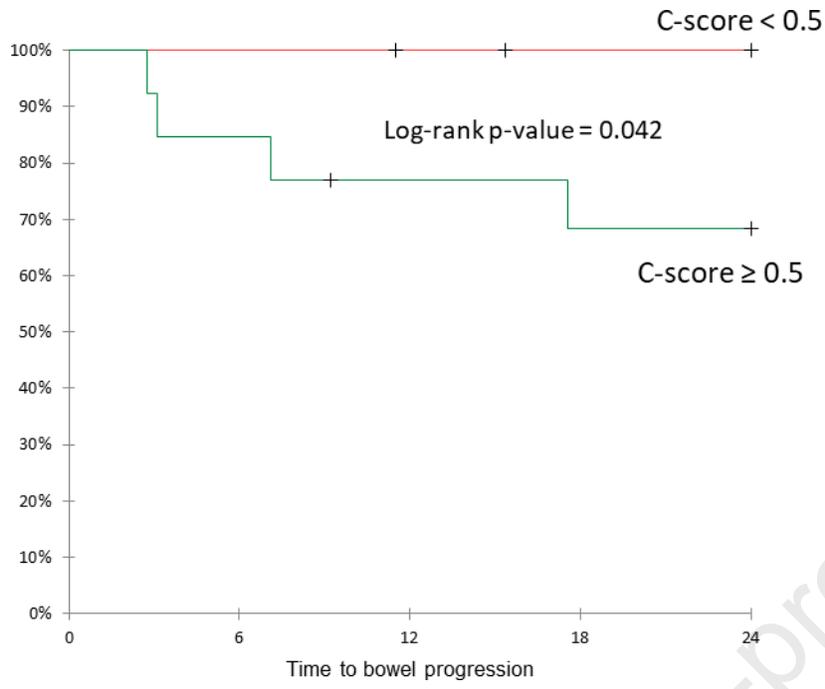












Journal Pre-proof

Supplementary Appendix 1

All the patients underwent MRI with a 1.5 Tesla General Electric Optima MR 450w (GE HealthCare, Fairfield, CT) with evaluation of small bowel, colon and rectum. Diffusion-weighted and injected sequences were performed in all patients. For the development cohort each examination was interpreted retrospectively and independently by two radiologists: a junior (J.J.) and an experienced IBD radiologist (C.H.) who were not aware of the clinical or biological data. MRI was performed with no bowel cleansing the day before the examination and no rectal enema during the procedure. An oral ingestion of polyethylene glycol (0.5 to 1L from 30 to 45 minutes before the MRI) was applied to obtain adequate distension of ileal loops. Regarding per-segment analyses, we used a division into nine segments (proximal jejunum, distal jejunum, proximal ileum, distal ileum, terminal ileum, right colon, transverse colon, left colon/sigmoid, and rectum). Image analysis was performed using a dedicated post-processing workstation (Advantage Window WorkStation, GE HealthCare Fairfield, CT).

Supplementary Appendix 2 : Description of the step-by-step process to build the modified Clermont score for assessing transmural therapeutic efficacy in Crohn's disease

From the available MRI indexes or scores and based on clinical relevance, we firstly identified the main CD-related MRI lesions. We identified the following main CD-related MRI lesions: edema, ulcers, bowel thickening (> 3 mm), bowel thickness (quantitative value), diffusion hyperintensity, contrast enhancement, relative contrast enhancement (quantitative value), apparent diffusion coefficient (quantitative value), the length of the lesions, mesenteric fat wrapping, and mesenteric lymph nodes enlargement. As we aimed to develop a score to assess disease activity and not bowel damage by itself, we excluded fistula and stricture from the future score. Secondly, we investigated whether the persistence of each lesion was associated with the progression of bowel damage. Thirdly, we assessed the coexistence and the concordance of the lesions between them to avoid selecting several items giving the same information. Fourthly, we evaluated the sensitivity to change of each item after therapeutic intervention. Considering all the previous steps, we used a pragmatic approach to select the most relevant items based on the best compromise between clinical relevance, convenience for radiologist, and statistical data such as prediction of bowel damage progression, coexistence between lesions (clinical correspondence to collinearity), and sensitivity to change after treatment.

Supplementary Appendix 3

Sample size calculation

For the development cohort, sample size was determined according to the recommendations from Green¹³ and Harrell and colleagues¹⁴, who have defined rules for minimal number of patients for a specific number of variables in multivariable analyses. Based on previous studies^{7,8}, we could expect an occurrence of bowel damage progression (between 36 and 44%). Thus, at least 250 patients allowed to construct a score with 4 to 6 parameters in multivariable analysis.

For the validation cohort, sample size was calculated in order to highlight difference of bowel damage progression between patients with or without TH or TR. For a two-sided type I error at 5%, an absolute difference of 20% (40% vs. 20%) can be shown with 200 patients for a satisfactory statistical power at 90%.

In the multicenter validation cohort, at least 40 patients were necessary for a type I error at 5% and a statistical power greater than 80% to detect a true absolute difference of 50% to predict CFREM at W52 using our definition of TH or TR at W12.

Supplementary Appendix 4

Concomitant healing of MRI lesions at the time of the second MRI

We also investigated the proportion of concomitant healing of MRI lesions (**Supplementary Figure S3**) as well as the concordance of lesions healing (disappearance of qualitative variable or variation of quantitative parameter between pre- and post-treatment MRI) (**Supplementary Table S3 and Supplementary Figure S4**). Normalization of edema, bowel thickness (bowel wall ≤ 3 mm), or increased contrast enhancement and loss diffusion-weighted hyperintensity were associated with normalization of all the investigated parameters (qualitative items) in $> 91\%$, $> 95\%$, $> 87\%$ and $> 83\%$, respectively (**Supplementary Table S3**). Ulcer healing was associated with normalization of inflammatory sign of bowel wall from 40.4% to 48.9% of the cases, fat wrapping healing in 33.3% and disappearance of enlarged lymph nodes in 70.0% of the segments (**Supplementary Figure S3**). The proportion of lesions healing ranged from 60.0% to 100.0% in case of fat wrapping healing and from 35.7% to 85.7% in segments with disappearance of enlarged lymph nodes (**Supplementary Figure S3**).

The concordances side-by-side of normalization of edema, normalization of bowel thickness, loss of increased contrast enhancement and loss of diffusion-weighted hyperintensity were almost perfect with κ -coefficient ranging from 0.82 to 0.95 (**Supplementary Table S3**). The concordance was moderate between ulcer healing and normalization of edema ($\kappa = 0.54$), normalization of bowel thickness (bowel wall ≤ 3 mm) ($\kappa = 0.50$), loss of increased contrast enhancement ($\kappa = 0.53$), loss diffusion-weighted hyperintensity ($\kappa = 0.57$) and disappearance of enlarged lymph nodes ($\kappa = 0.50$) but was mild between ulcer healing and fat wrapping ($\kappa = 0.36$) (**Supplementary Table S3**). There were high correlations between ADC, RCE and bowel thickness as continuous variable and all the other investigated lesions ($ES > 0.8$) (**Supplementary Table S3, Supplementary Figure S4**).

Supplementary Appendix 4

We performed additional analyses on 98 colonic segments (right colon = 38, transverse colon = 21, left/sigmoid colon = 25 and rectum = 14) the sensitivity to change after treatment was negligible for bowel thickness (SMD=0.16±0.93) and RCE (SMD=0.15±38.0), but negligible for ADC (SMD=0.13± 1.0). For qualitative parameters, the sensitivity to change after treatment was negligible for edema (SMD=0.10[0.02-0.18]), contrast enhancement (SMD=0.14[0.01-0.24]), diffusion-weighted hyperintensity (SMD=0.12[0.03-0.21]), ulcer (SMD=0.23[0.05-0.32]), fat wrapping (SMD=0.04[0.00-0.08]) and enlarged lymph nodes (SMD=0.11[0.00-0.22]).

Supplementary Appendix 5

The definition of a cut-off value for a score is always a compromise between performances and relevance. For example, in CD, the current definitions of endoscopic response (Δ SES-CD > 25 % or 50 %) were defined from post-hoc analyses of the SONIC trial. The evidence level for choosing these thresholds is low but as it seemed meaningful for physicians to improve lesions > 50 %. In the same line, we attempted to find a MRI score with two main objectives: being intuitive for physicians with high level of performances (prediction of bowel damage progression). We performed a ROC curve to confirm that our choices for defining transmural healing (C-score < 0.5) or transmural response (TR50 and TR25) are relevant. Of note, our performance metrics are about prediction and not diagnosis and are thus lower than commonly in diagnosis studies. The variation of C-score was associated with the risk of bowel damage progression (AUC = 0.65 [0.58-0.72]). A reduction of 25 % was the lowest improvement that reaches a positive predictive value > 80 % [69-88] to predict the absence of bowel damage progression. After that, PPV plateaued between reduction of 25 and 50 % with PPV > 80 %. Regarding the definition of TH, C-score after treatment was associated with decreased risk of bowel damage progression (AUC = 0.66 [0.60-0.73]) and is the highest value with PPV > 80 % [67-89]. Of course, C-score = 0 had a PPV of 93.1% [77.2-99.2%] but is more difficult to achieve. Consequently, based on a pragmatic approach including statistical models and clinical relevance in mirror of endoscopic definitions, we decided to use TH (C-score < 0.5), TR25 and TR50.

Supplementary table S1: Technical characteristics of sequences.

Sequences	Plane	FOV (cm)	TE (ms)	TR (ms)	Flip angle (degree)	Slice thickness (mm)	Acquisition duration (s)
T2 SSFSE	Axial	34	120	820	90	7	30
T2 SSFSE	Coronal	41	120	710	90	5	20
2D FIESTA FS	Axial	34	2	4.8	85	6.5	40
DWI b0-b800	Axial	34	70	3.200	90	6	120
T1 FS LAVA	Coronal	42	1.9	4	12	3.2	240
T1 FS LAVA	Axial	40	1.9	4	12	3.8	45

DWI: diffusion-weighted imaging; FIESTA: fast imaging employing steady-state acquisition with fat suppression; FOV: field of view; FS: fat sat; MRI: magnetic resonance imaging; SSFSE: single-shot fast spin-echo; TE: echo time; TR: repetition time; 2D: two dimensional

Supplementary Table S2: Characteristics at the time of the second MRI of the 274 patients with CD included in this study.

	Development cohort	Monocenter validation cohort	Multicenter validation cohort
Characteristics at the time of the 2nd MRI	N= 274 patients	N=224 patients	N=46 patients
Active smokers, n (%)	100 (36.5%)	47 (21.0%)	17 (37.0%)
Age at the time of inclusion (mean+/-sd)	33.1 ± 15.8	27.8 ± 13.8	34.9 ± 14.1
CD duration, median [IQR]	7.0 [2.0-13.0]	11.6 ± 9.9	9.0 ± 9.5
Prior intestinal resection, n (%)	86 (31.4%)	79 (35.3%)	12 (26.1%)
Montreal Classification, n (%)			
CD location			
L1, n (%)	141 (51.5%)	136 (60.7%)	22 (47.8%)
L2, n (%)	15 (5.5%)	10 (4.5%)	6 (13.0%)
L3, n (%)	118 (43.1%)	78 (34.8%)	18 (39.1%)
L4, n (%)	20 (7.4%)	23 (10.3%)	0 (0.0%)
Perianal lesions, n (%)	71 (25.9%)	62 (27.7%)	13 (28.3%)
Phenotype			
B1, n (%)	92 (33.6%)	90 (39.7%)	13 (28.3%)
B2, n (%)	97 (35.4%)	77 (36.4%)	18 (39.1%)
B3, n (%)	85 (31.0%)	57 (23.8%)	15 (32.6%)
Medications at the time of 2 nd MRI			
Steroids, n (%)	17 (6.3%)	7 (3.2%)	10 (21.7%)
Immunosuppressants, n (%)	122 (45.2%)	96 (42.9%)	27 (58.7%)
Anti-TNF, n (%)	178 (65.7%)	148 (66.1%)	46 (100.0%)
Infliximab, n (%)	90 (50.6%)	67 (29.9%)	22 (47.8%)
Adalimumab, n (%)	87 (48.9%)	80 (35.7%)	24 (52.2%)
Golimumab, n (%)	1 (2.6%)	1 (0.4%)	0 (0.0%)
Ustekinumab, n (%)	7 (2.6%)	40 (17.9%)	0 (0.0%)
Vedolizumab, n (%)	0 (0.0%)	5 (2.2%)	0 (0.0%)

n : number; sd: standard deviation; IQR: interquartile range; ASA: aminosalicylates; TNF: tumor necrosis factor; CD: Crohn's disease; MRI: magnetic resonance imaging

Supplementary Table S3: Evaluation of the persistence of each MRI lesion on the second MRI as predictor of bowel damage progression among 274 patients with Crohn's disease (Development cohort).

Type of MRI lesions	1 st MRI		2 nd MRI		Univariate analysis		Multivariable analysis*	
	Number of patients with the lesion	Number of patients with the lesion	Healing of the lesion	Hazard ratio [95% CI]	p-value	Hazard ratio [95% CI]	p-value	
Qualitative parameters								
Edema	259	215	17.0%	3.39 [1.59- 9.65]	0.003	4.78 [1.74- 13.13]	0.002	
Ulcer	203	154	24.1%	1.74 [1.14- 2.67]	0.011	1.90 [1.21- 2.97]	0.005	
Contrast enhancement	268	224	16.4%	4.05 [1.49- 11.03]	0.006	4.13 [1.50- 11.36]	0.006	
Diffusion-weighted hyperintensity	257	211	17.9%	5.61 [2.06- 15.27]	0.001	6.97 [2.50- 19.41]	< 0.001	
Fat wrapping	142	120	15.5%	1.55 [1.04- 2.30]	0.030	1.87 [1.22- 2.87]	0.004	
Bowel thickening (> 3 mm)	274	232	15.3%	3.82 [1.67- 8.74]	0.001	3.73 [1.61- 8.66]	0.002	
Enlarged lymph nodes	83	67	19.3%	2.11 [1.40- 3.18]	< 0.001	2.31 [1.48- 3.63]	< 0.001	
Quantitative parameters								
Bowel thickness	-	-	-	1.14 [1.07- 1.24]	< 0.001	1.15 [1.06- 1.25]	< 0.001	
Apparent diffusion coefficient (ADC)	-	-	-	0.27 [0.13- 0.59]	0.001	0.27 [0.13- 0.60]	0.001	
Relative contrast enhancement (RCE)	-	-	-	1.00 [1.00- 1.00]	0.003	1.00 [1.00- 1.00]	0.005	
Length of active lesions	-	-	-	1.02 [0.99 - 1.05]	0.10	-	-	

Supplementary Table S4: Concordance between lesions on baseline MRI (blue scale) and between variation of MRI lesions between pre- and post-treatment MRIs (red scale).

Variation between pre- and post-treatment MRIs	Baseline MRI										
	Qualitative variables							Quantitative variables			
	Ulcers	Edema	Qualitative contrast enhancement	Qualitative diffusion-weighted hyperintensity	Bowel thickness > 3 mm	Fat wrapping	Enlarged lymph nodes	ADC	RCE	Bowel thickness (mm)	
Ulcers		0.20 ± 0.035 (69.5%)	0.13 ± 0.028 (67.5%)	0.16 ± 0.036 (67.8%)	0.10 ± 0.024 (66.5%)	0.24 ± 0.049 (61.3%)	0.19 ± 0.038 (54.1%)	ES = -0.42 1.40 vs 1.54 p = 0.0008	ES = 0.25 168 vs 151 p = 0.036	ES = 0.58 7.3 vs 6.1 p < 0.0001	
Edema	0.54 ± 0.061 (86.9%)		0.53 ± 0.050 (94.7%)	0.42 ± 0.053 (91.5%)	0.53 ± 0.047 (95.0%)	0.07 ± 0.025 (50.0%)	0.03 ± 0.030 (30.8%)	ES = -0.90 1.43 vs 1.73 p = 0.0025	ES = 1.27 170 vs 85 p = 0.0002	ES = 1.15 7.0 vs 4.6 p < 0.0001	
Qualitative contrast enhancement	0.53 ± 0.060 (86.3%)	0.93 ± 0.056 (97.8%)		0.60 ± 0.050 (95.2%)	0.77 ± 0.052 (98.3%)	0.05 ± 0.019 (48.2%)	0.01 ± 0.013 (27.8%)	ES = -1.49 1.43 vs 1.90 p < 0.0001	ES = 2.00 168 vs 39 p < 0.0001	ES = 1.52 7.0 vs 3.8 p < 0.0001	
Qualitative diffusion-weighted hyperintensity	0.57 ± 0.063 (87.8%)	0.89 ± 0.057 (98.2%)	0.88 ± 0.056 (96.5%)		0.51 ± 0.047 (94.7%)	0.05 ± 0.026 (49.7%)	0.01 ± 0.014 (28.4%)	ES = -1.36 1.42 vs 1.84 p < 0.0001	ES = 1.06 168 vs 97 p = 0.0004	ES = 0.79 7.0 vs 5.3 p = 0.0004	
Bowel thickness > 3 mm	0.50 ± 0.058 (85.8%)	0.94 ± 0.055 (96.7%)	0.95 ± 0.054 (98.8%)	0.88 ± 0.055 (96.5%)		0.04 ± 0.017 (47.9%)	0.01 ± 0.011 (27.1%)	ES = -1.58 1.43 vs 1.96 p = 0.0001	ES = 2.24 167 vs 24 p < 0.0001	ES = 1.95 7.0 vs 3.0 p < 0.0001	
Fat wrapping	0.36 ± 0.085 (87.1%)	0.82 ± 0.0805 (94.8%)	0.76 ± 0.080 (93.5%)	0.80 ± 0.081 (94.1%)	0.79 ± 0.078 (94.4%)		0.01 ± 0.014 (29.4%)	ES = -0.38 1.39 vs 1.52 p = 0.0003	ES = 0.09 166 vs 159 p = 0.36	ES = 0.35 7.3 vs 6.5 p = 0.0008	
Enlarged lymph nodes	0.50 ± 0.115 (86.1%)	0.58 ± 0.010 (89.0%)	0.52 ± 0.010 (87.7%)	0.44 ± 0.101 (87.1%)	0.58 ± 0.010 (89.2%)	0.56 ± 0.012 (92.4%)		ES = -0.26 1.39 vs 1.48 p = 0.11	ES = 0.03 164 vs 161 p = 0.78	ES = 1.0 7.0 vs 6.8 p = 0.29	
ADC	ES = 0.84 0.25 vs 0.00 p < 0.0001	ES = 1.11 0.37 vs 0.01 p < 0.0001	ES = 1.04 0.37 vs 0.02 p < 0.0001	ES = 1.28 0.41 vs 0.02 p < 0.0001	ES = 1.04 0.38 vs 0.02 p < 0.0001	ES = 1.18 0.41 vs 0.02 p < 0.0001	ES = 0.90 0.23 vs 0.00 p = 0.013		q = -0.04	q = -0.24	
RCE	ES = 0.84 70 vs 2 p < 0.0001	ES = 1.93 145 vs 6 p < 0.0001	ES = 1.95 151 vs 5 p < 0.0001	ES = 1.49 128 vs 7 p < 0.0001	ES = 1.90 149 vs 5 p < 0.0001	ES = 1.19 106 vs 10 p = 0.002	ES = 0.66 51 vs 0 p = 0.07	q = -0.28		q = 0.13	
Bowel thickness (mm)	ES = 1.08 1.9 vs 0.1 p < 0.0001	ES = 1.65 3.2 vs 0.3 p < 0.0001	ES = 1.52 3.1 vs 0.3 p < 0.0001	ES = 1.34 2.8 vs 0.3 p < 0.0001	ES = 1.66 3.2 vs 0.3 p < 0.0001	ES = 1.22 2.5 vs 0.2 p = 0.003	ES = 1.15 2.1 vs 0.2 p = 0.003	q = -0.35	q = 0.38		

For the relationship between two qualitative parameters, the results are expressed as $\kappa \pm$ standard error and agreement (in round brackets). For the relationship between one qualitative and one quantitative parameters, the results are expressed as effect size (ES), mean values with or without the presence of the lesion and p-values. For the relationship between two quantitative parameters, the results are expressed as coefficient of correlation (q).

Concordance (κ)	0.00 – 0.19	0.20 – 0.39	0.4 – 0.59	0.60 – 0.79	≥ 0.80	Concordance (κ)	0.00 – 0.19	0.20 – 0.39	0.4 – 0.59	0.60 – 0.79	≥ 0.80
Effect size (ES)	0.00 – 0.19	0.20 – 0.49	0.50 – 0.79	0.80 – 1.59	≥ 1.60	Effect size (ES)	0.00 – 0.19	0.20 – 0.49	0.50 – 0.79	0.80 – 1.59	≥ 1.60

Supplementary Table S5: Multivariable model including MRI items to predict bowel damage progression (weighting of selected items to finalize the C-score).

	HR [CI 95%]	p-value
Bowel thickening compared to the normal value (3 mm)	1.19 [1.06-1.32]	0.003
Ulcer	2.31 [1.11-4.76]	0.024
Enlarged lymph nodes	1.62 [1.01-2.61]	0.048

- **BACKGROUND**

- Despite its strong association with favorable outcomes, transmural healing is not recommended as therapeutic target in Crohn's disease (CD) especially owing to the lack of consensual definition.

- **FINDINGS**

- We developed and validated an MRI score to define and assess transmural healing (TH) and response (TR) in patients with CD that are associated with favorable outcomes.

- **IMPLICATION FOR CARE**

- C-score and its definition of transmural healing (TH) and transmural response (TR25 and TR50) are now a validated, reliable and easy-to-use tool to assess therapeutic efficacy.