



Validation of the simplified magnetic resonance index of activity by using DWI without gadolinium enhancement to evaluate bowel inflammation in Crohn's disease

Heejin Bae¹ · Nieun Seo¹ · Eun Ae Kang² · Jae Hee Cheon² · Joon Seok Lim¹ · Myeong-Jin Kim¹

Received: 4 January 2023 / Revised: 4 January 2023 / Accepted: 3 February 2023 / Published online: 11 March 2023
© The Author(s), under exclusive licence to European Society of Radiology 2023

Abstract

Objectives To validate the modified simplified magnetic resonance index of activity (sMARIA) score using DWI on non-contrast magnetic resonance enterography (MRE) to evaluate active inflammation in patients with Crohn's disease (CD), compared to the original sMARIA scoring system, with and without contrast enhancement.

Methods This retrospective study included 275 bowel segments from 55 CD patients who underwent ileocolonoscopy and MRE within a 2-week period. Two blinded radiologists evaluated original sMARIA on both conventional MRE (CE-sMARIA) and non-contrast MRE (T2-sMARIA). Modified sMARIA was then evaluated using non-contrast MRE, replacing ulcerations with DWI grades. Three scoring systems were compared for diagnostic accuracy of active inflammation, correlation with simple endoscopic score (SES)-CD, and interobserver reproducibility.

Results The AUC of modified sMARIA for detecting active inflammation (0.863, 95% confidence interval [0.803–0.923]) was significantly higher than T2-sMARIA (0.827 [0.773–0.881], $p=0.017$), and comparable to CE-sMARIA (0.908 [0.857–0.959], $p=0.122$). CE-sMARIA, T2-sMARIA, and modified sMARIA all showed moderate correlation with SES-CD ($r=0.795$, 0.722, and 0.777, respectively). Interobserver reproducibility of diffusion restriction (κ , 0.686 [0.602–0.770]) was significantly better than ulcers on conventional MRE (κ , 0.382 [0.212–0.552]; $p=0.001$) and T2-weighted image (κ , 0.312 [0.034–0.590]; $p=0.012$).

Conclusions Modified sMARIA using DWI can improve the diagnostic performance of sMARIA on non-contrast MRE, showing comparable performance to sMARIA using contrast-enhanced MRE.

Key Points

- DWI can improve the diagnostic performance of non-contrast magnetic resonance enterography (MRE) for assessing active inflammation in patients with Crohn's disease.
- Modified simplified magnetic resonance index of activity (sMARIA) using DWI grades in place of ulcers showed comparable diagnostic performance to sMARIA using conventional MRE with contrast-enhanced sequences.

Keywords Crohn's disease · Magnetic resonance enterography · Diagnosis · Diffusion-weighted imaging

Abbreviations

CD	Crohn's disease	MARIA	Magnetic resonance index of activity
CE	Contrast-enhanced	MRE	Magnetic resonance enterography
CI	Confidence interval	SD	Standard deviation
ICC	Intraclass correlation coefficient	SES	Simple endoscopic score
		sMARIA	Simplified magnetic resonance index of activity
		T1WI	T1-weighted imaging
		T2WI	T2-weighted imaging

✉ Nieun Seo
SLDMSDL@yuhs.ac

¹ Department of Radiology and Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-Ro, Seodaemun-Gu, Seoul, Republic of Korea 03722

² Department of Internal Medicine, Institute of Gastroenterology, Yonsei University College of Medicine, Seoul, Republic of Korea

Introduction

Crohn's disease (CD), a chronic gastrointestinal inflammatory disorder with increasing incidence and prevalence, is characterized by episodes of relapse and remission [1–3].

Therefore, objective evaluation of disease severity and monitoring treatment response are crucial for successful disease management. Various clinical, endoscopic, and imaging scoring systems have been developed for the evaluation of disease activity in CD patients and have been implemented in both clinical practice and research [4]. Of several image-based scoring systems, the magnetic resonance index of activity (MARIA) system was most frequently used and validated [5]. However, measurement of quantitative parameters, such as relative contrast enhancement for MARIA, is not only time-consuming but also requires contrast-enhanced (CE) sequences. To overcome these weaknesses, the simplified MARIA (sMARIA) was recently developed to include a binary assessment of imaging features, and it showed excellent correlation with both the CD endoscopic index of severity and original MARIA scores [6]. Moreover, imaging components included in the sMARIA calculation could be evaluated using only non-contrast magnetic resonance enterography (MRE), even though sMARIA was originally developed from CE-MRE [6].

In many institutions, routine MRE protocols for patients with CD include CE sequences. Although gadolinium-based contrast agents are generally considered safe, there is increased unease about gadolinium retention in the brain and other organs, including bone, skin, and liver due to repeated gadolinium-enhanced MRI [7, 8]. Despite the uncertainty of its clinical importance and dose thresholds, growing evidence of gadolinium retention has raised concerns about the safety of gadolinium-based contrast media in the long term. The US Food and Drug Administration suggests that healthcare professionals acknowledge the retention characteristics of gadolinium contrasts and minimize repeated gadolinium-enhanced MRI, particularly in patients who may need multiple lifetime doses, or in children [9]. Additionally, CE-MRE is difficult to perform in patients with contraindications to gadolinium-based contrast media, such as impaired renal function, hypersensitivity to contrast media, or pregnancy [10, 11]. These issues need to be considered deeply in patients with CD because they are relatively young and receive repetitive imaging evaluations throughout the course of their disease. Previous studies have found that sMARIA evaluated on non-contrast T2-weighted imaging (T2WI) is effective for evaluating disease activity and response to treatment in CD patients [12, 13]. However, there are also concerns that T2WI alone is not sufficient for evaluating the inflammatory activity of CD, especially because the accurate detection and assessment of ulcers on T2WI are challenging [14]. Therefore, to improve the diagnostic performance of sMARIA evaluated on non-contrast MRE, imaging features more reliable than ulcers need to be recruited as components of sMARIA.

Similar to T2WI, DWI does not require contrast enhancement, and restricted mural diffusion is also known to be

correlated with bowel inflammatory severity in CD [15, 16]. A previous study showed combined assessment using T2WI and DWI was noninferior to CE-MRE for evaluating inflammation in CD patients [17]. In addition, Kim et al. [14] introduced modified MARIA scores by substituting ulcers with DWI grades, and the modified MARIA scores showed improved interobserver reproducibility while maintaining overall diagnostic performance compared to original MARIA scores on CE-MRE. Therefore, DWI can be a potential imaging feature for improving the diagnostic performance of sMARIA evaluation on non-contrast MRE.

The purpose of this study was to investigate the diagnostic performance of a modified sMARIA using diffusion restriction instead of ulcers on non-contrast MRE, compared to sMARIA on T2WI alone and conventional CE-MRE to evaluate inflammatory activity in patients with CD.

Materials and methods

Study population

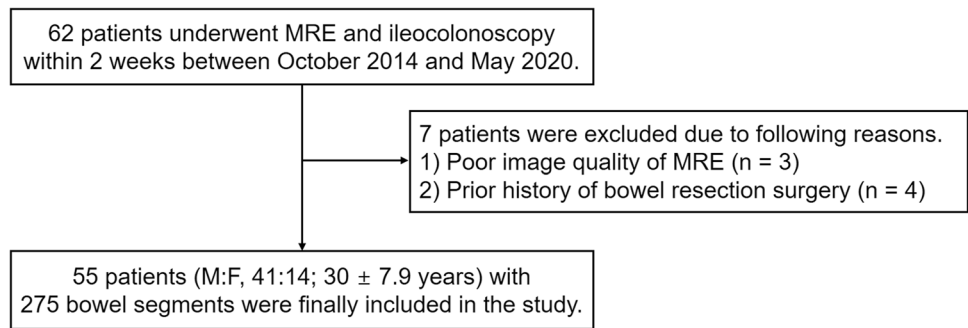
The institutional review board at our institution approved this retrospective study and the requirement for informed consent was waived. A flow diagram of patient selection is shown in Fig. 1. We retrospectively recruited 62 patients who underwent MRE within 2 weeks of receiving an ileocolonoscopy for known or suspected CD between October 2014 and May 2020. After the exclusion of 7 patients due to poor image quality ($n=3$) and prior history of bowel resection ($n=4$), a total of 55 patients (41 men; mean age \pm standard deviation [SD], 30 ± 7.9 years) were included in the final study population. The following demographic and medical information were obtained from electronic medical records: age, sex, body mass index, CD activity index, and C-reactive protein level.

Image acquisition

To achieve adequate bowel distension, patients ingested 1250 mL of polyethylene glycol solution (Coolprep; Taejoon Pharmaceutical) 40 min before the examination. An intravenous injection of 10 mg scopolamine-N-butylbromide (Buscopan; Boehringer Ingelheim) was administered before initiating the scan. The same dose was administered additionally before acquiring the coronal T1-weighted sequence, to reduce bowel peristalsis.

MRE scans were performed using a 3.0-T (Ingenia CX or Achieva, Philips Healthcare; Discovery MR 750, GE Medical Systems) MR scanner. Routine MR sequences were as follows: coronal T2-weighted half-Fourier sequence, without fat suppression; coronal balanced gradient-echo sequence with fat suppression; coronal DWI (with b factors of 0 and 800 s/mm^2); ADC map; coronal T1-weighted spoiled

Fig. 1 A flow diagram of patient selection. MRE magnetic resonance enterography



gradient-echo sequences with fat suppression conducted before and after the contrast agent injection, including enteric and portal phases; and axial delayed CE T1-weighted spoiled gradient-echo sequence with fat suppression. For CE T1-weighted images (CE T1WI), a volume of 0.2 mL/kg of gadolinium-based contrast agent (Prohance; Bracco Diagnostics Inc.) was injected intravenously at a fixed rate of 2 mL/s, followed by a saline bolus injection. Axial T2-weighted half-Fourier sequence with fat suppression was included in the routine sequences from 2017.

Image analysis

Two board-certified abdominal radiologists (with 3 and 9 years of experience in MRE imaging, respectively), who were unaware of clinical information and ileocolonoscopy results, independently reviewed the images in two review sessions with a washout period of more than 4 weeks. In this study, conventional MRE (T2WI and CE T1WI) and non-contrast MRE sequences (T2WI and DWI) were analyzed in separate sessions for each patient. In order to avoid bias caused by the order of image review, the study population was randomly divided into two groups. During the first reading session, conventional MRE images of one group and non-contrast MRE images of the other group were reviewed, and vice versa in the second review session.

Radiologists assessed the absence or presence of the qualitative parameters of sMARIA for each segment of the terminal ileum, ascending colon, transverse colon, left colon including descending and sigmoid colons, and rectum: mural thickening (> 3 mm), mural edema, fat stranding, and ulcers, on both conventional and non-contrast MRE sequences [6]. For DWI, reviewers evaluated mural diffusion restriction of each bowel segment using the following definition [17]: 0, no diffusion restriction; 1, increased DWI signal intensity slightly lower than that of a lymph node; and 2, increased DWI signal intensity similar to or higher than that of a lymph node. To exclude T2 shine-through effects, ADC images were also provided during DWI interpretation.

Using a previously validated formula [(1 × bowel wall thickness > 3 mm) + (1 × edema) + (1 × fat stranding) + (2 × ulcers)]

[6], the sMARIA for each bowel segment was calculated from conventional MRE sequences (referred to as CE-sMARIA) and T2WI only (referred to as T2-sMARIA). Since a past study showed the potential of the modified MARIA scoring system using DWI grades as a substitute for ulcers with high interobserver reproducibility [14], we similarly modified the sMARIA scoring system by replacing ulcers with DWI grades as: [(1 × bowel wall thickness > 3 mm) + (1 × edema) + (1 × fat stranding) + (1 × DWI grades)] from non-contrast MRE sessions (referred to as modified sMARIA).

Additionally, for each patient, the type (sinus tract, fistula, or abscess) and location of penetrating disease were analyzed. The level of confidence for penetrating disease was rated using a 5-point numerical rating scale: 1, definitely absent; 2, probably absent; 3, indeterminate; 4, probably present; and 5, definitely present. Disagreements between the two reviewers were resolved through discussion, and reviewers reached a consensus.

Ileocolonoscopy

Ileocolonoscopy images were retrospectively reviewed by a board-certified gastroenterologist and were used as the reference standard for CD disease activity. For each bowel segment, the Simple Endoscopic Score for Crohn's disease (SES-CD) was calculated [18], and the presence of inflammatory mucosal lesions, including aphthoid lesions, erythema, and both superficial and deep ulcers, was evaluated. According to endoscopic results, bowel segments were classified into three categories: 1, no active inflammation; 2, inflammatory lesions such as erythema, edema, or aphthae without ulcers; 3, severe inflammation with the presence of superficial or deep ulcers. Both mild and severe inflammation were considered active inflammation in this study.

Statistical analysis

All statistical analyses were performed on a per-segment basis. First, the diagnostic performances of the three scoring systems, CE-sMARIA, T2-sMARIA, and modified

sMARIA, for detecting active inflammation and severe inflammation were evaluated using the receiver operating characteristic curve analysis. The AUC was compared between two scoring systems using a method by Obuchowski, considering intracluster correlation [19].

Second, the correlation between segmental SES-CD and sMARIA was estimated using Spearman's correlation. The correlation coefficient, ρ , ranges from -1 to $+1$ with the absolute value representing the strength of the correlation (0, no correlation; 0.2, weak correlation; 0.5, moderate correlation; 0.8, strong correlation; 1, perfect correlation) [20].

Third, to assess the diagnostic performance of MRE features (MRE-defined ulcer and diffusion restriction) to predict endoscopic ulcers, sensitivity, specificity, and accuracy were calculated and compared using a generalized estimating equations model. As superficial ulcerations less than 5 mm in depth have less clinical significance and worse reliability compared to deep ulcers, the diagnostic performance of MRE was analyzed for ulcerations ≥ 0.5 cm. Additionally, the detection sensitivity of MRE features for endoscopic ulcerations was compared according to ulcer size (< 0.5 cm, 0.5–2 cm, and ≥ 2 cm). DWI grade of 2 was defined as positive diffusion restriction for the analysis.

Fourth, the interobserver reproducibility of the segmental sMARIA scores between the two reviewers was evaluated using the intraclass correlation coefficient (ICC) with a linear mixed model to consider intracluster correlation: $ICC \leq 0.20$, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect reliability [21]. The interobserver reproducibility of each component of sMARIA (mural thickening, mural edema, fat stranding, ulcers) and DWI grade (0–2) was analyzed using weighted κ statistics: $\kappa \leq 0.20$, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; and 0.81–1.00, excellent agreement [22, 23].

Statistical analyses were performed using R software (version 4.0.4; R Foundation for Statistical Computing) and SAS (version 9.4, SAS Institute). A p value less than 0.05 was considered statistically significant. Bonferroni's correction was applied to adjust p values for multiple comparison.

Results

Patient characteristics

A total of 55 patients with 275 bowel segments were included in the analyses and patient characteristics are summarized in Table 1. The median global SES-CD was 8 (IQR, 4–13) for 35 patients who had at least one bowel segment with active inflammation. Out of 275 bowel segments, ileocolonoscopy identified active inflammation in 75 bowel segments (27.3%), including 38 segments with severe

inflammation. Consensus results of conventional MRE demonstrated eight patients (8/55, 14.5%) had intraabdominal penetrating complications including sinus tract ($n = 1$) and fistula ($n = 7$). On the other hand, non-contrast MRE overlooked two of these patients (one sinus tract and one fistula).

Validation of sMARIA for evaluating active bowel inflammation

For detection of active inflammation, CE-sMARIA showed the highest AUC (0.908, 95% confidence interval (CI)

Table 1 Characteristics of the study population

Parameter	Value
Patients ($n = 55$)	
Age (years), mean \pm SD	30 \pm 7.9
Male (%)	41 (74.5)
Body mass index (kg/m ²), mean \pm SD	21.0 \pm 2.8
CRP (mg/L), mean \pm SD	8.1 \pm 13.5
CD activity index	89 (52.5–176.5)
SES-CD, global ^a	8 (4–13)
Duration between MRE and colonoscopy (days)	0 (0–1.5)
Disease location ^b (%)	
Terminal ileum	13 (23.6)
Colonic	3 (5.5)
Ileocolonic	19 (34.5)
No active inflammation	20 (36.4)
Penetrating disease (%)	
Conventional MRE (T2WI+CE T1WI)	8 (14.5)
Non-contrast MRE (T2WI+DWI)	6 (10.9)
Bowel segments ($n = 275$) ^b	
Segments with mild inflammation (%)	37 (13.5)
Segments with severe inflammation ^c (%)	38 (13.8)
SES-CD, segmental ^a	4 (3–7)
sMARIA, segmental ^a	
CE-sMARIA	3 (2–4.5)
T2-sMARIA	2 (0–3)
Modified sMARIA	3 (0.5–4.5)

Unless otherwise specified, continuous values are expressed as median (interquartile range)

SD, standard deviation; CRP, C-reactive protein; CD, Crohn's disease; SES, simple endoscopic score; MRE, magnetic resonance enterography; T2WI, T2-weighted imaging; CE, contrast-enhanced; T1WI, T1-weighted imaging; sMARIA, simplified magnetic resonance index of activity

^aValues of SES-CD and sMARIA are summarized for 75 bowel segments with active inflammation

^bActive inflammation was determined by ileocolonoscopy

^cSevere active inflammation was defined by the presence of endoscopic ulcerations

[0.857–0.959]) followed by modified sMARIA (0.863 [0.803–0.923]), and T2-sMARIA (0.827 [0.773–0.881]; Table 2). The AUC of T2-sMARIA was significantly lower than CE-sMARIA ($p=0.001$) and modified sMARIA ($p=0.017$). On the other hand, the AUCs were not significantly different between CE-sMARIA and modified sMARIA ($p=0.122$). The cutoff values were 1 for diagnosing active inflammation, and 2 for diagnosing severe inflammation, and were kept the same for the three scoring systems (CE-sMARIA, T2-sMARIA, and modified sMARIA).

For detecting severe inflammation with endoscopic ulcerations, modified sMARIA showed the highest AUC (0.858 [0.757–0.958]), followed by CE-sMARIA (0.835 [0.748–0.922]) and T2-sMARIA (0.806 [0.717–0.895]). The AUC of modified sMARIA was significantly higher than that of T2-sMARIA ($p=0.036$). There was no significant difference between the AUCs of CE-sMARIA and T2-sMARIA ($p=0.971$), or between CE-sMARIA and modified sMARIA ($p>0.999$).

Segmental SES-CD and sMARIA scores showed moderate correlation across all scoring systems, with correlation coefficients of 0.795 (95% CI, 0.747–0.835) for CE-sMARIA, 0.722 (0.660–0.774) for T2-sMARIA, and 0.777 (0.726–0.820) for modified sMARIA. Representative cases evaluating CE-sMARIA, T2-sMARIA, and modified sMARIA are depicted in Figs. 2 and 3.

Diagnostic performances of MRE findings to predict endoscopic ulcers

In this study, ulcers from the original sMARIA system were replaced with DWI grades in the formula used to calculate the modified sMARIA on non-contrast MRE. Using ileocolonoscopy results as the reference standard, the diagnostic performance of MRE ulcer detection and diffusion restriction (DWI grade 2) for identifying endoscopic ulcers were compared. For all parameters, ulcers on conventional MRE (T2WI + CE T1WI) and diffusion

restriction did not show significant differences in endoscopic ulcer prediction ($ps>0.05$, Table 3). The sensitivity of ulcers on T2WI (18.4%) was significantly lower than those of ulcers on conventional MRE (42.1%, $p=0.008$) and diffusion restriction (57.9%, $p<0.001$). On the other hand, the specificity of ulcers on T2WI (99.2%) was significantly higher than those of ulcers on conventional MRE (94.9%, $p=0.004$) and diffusion restriction (95.8%, $p=0.031$). The accuracy was not significantly different between ulcers on conventional MRE, ulcers on T2WI, and diffusion restriction ($ps>0.05$).

Regardless of ulcer size, ulcers evaluated on T2WI showed the lowest sensitivity for endoscopic ulcerations (Table 4). Ulcers on conventional MRE showed better sensitivity than ulcers on T2WI for the detection of endoscopic ulcers < 0.5 cm (37.5% vs. 8.3%, $p=0.024$), and 0.5–2 cm (45% vs. 5%, $p=0.015$). On the other hand, sensitivity was not significantly different between ulcers on conventional MRE and diffusion restriction for ulcers < 2 cm ($p>0.999$). For ulcers ≥ 2 cm, diffusion restriction showed significantly higher sensitivity (72.2%) than ulcers on conventional MRE (38.9%, $p=0.042$) and ulcers on T2WI (33.3%, $p=0.024$).

Interobserver reproducibility

Interobserver agreement for sMARIA between the two reviewers ranged from substantial to almost perfect for CE-sMARIA (ICC, 0.769 [95% CI, 0.714–0.816]), T2-sMARIA (0.819 [0.774–0.857]), and modified sMARIA (0.834 [0.793–0.869]). Interobserver agreement for each parameter of sMARIA is summarized in Table 5. All parameters except for ulcers showed similarly good interobserver agreement between conventional MRE (κ , 0.660–0.746) and non-contrast MRE (κ , 0.603–0.797). MRE-detected ulcers showed fair agreement on both conventional (κ , 0.382) and non-contrast MRE (κ , 0.312). Diffusion restriction (DWI grade 2) showed significantly better agreement (κ , 0.686) than ulcer detection on conventional MRE ($p=0.001$) and T2WI ($p=0.012$).

Table 2 Per-segment diagnostic performances of MRE-based indices for predicting active and severe inflammation

	AUC (95% CI)			Adjusted p value ^a		
	CE-sMARIA	T2-sMARIA	Modified sMARIA ^b	CE vs. T2	CE vs. modified	T2 vs. modified
Active inflammation	0.908 (0.857–0.959)	0.827 (0.773–0.881)	0.863 (0.803–0.923)	0.001	0.122	0.017
Severe inflammation	0.835 (0.748–0.922)	0.806 (0.717–0.895)	0.858 (0.757–0.958)	0.971	>0.999	0.036

MRE, magnetic resonance enterography; CI, confidence interval; CE, contrast-enhanced; sMARIA, simplified magnetic resonance index of activity

^aFor multiple comparisons, p values were adjusted using Bonferroni's correction

^bModified sMARIA score was calculated by replacing ulcers with DWI grades on non-contrast MRE

Fig. 2 A 20-year-old man with active Crohn's disease inflammation in the terminal ileum. **a** Coronal contrast-enhanced T1-weighted image shows mural thickening (arrow) and ulcerations (arrowheads) of the terminal ileum. **b** Coronal T2-weighted image without fat saturation shows mural thickening of the terminal ileum, but without definite edema or ulceration. **c** Coronal DWI image shows multifocal diffusion restriction (grade 2) in the terminal ileum (arrowheads), resulting in a CE-sMARIA and modified sMARIA of 3, and a T2-sMARIA of 1. **d** Endoscopic image of the terminal ileum shows focal erythema, aphthous and superficial ulcers (arrows)

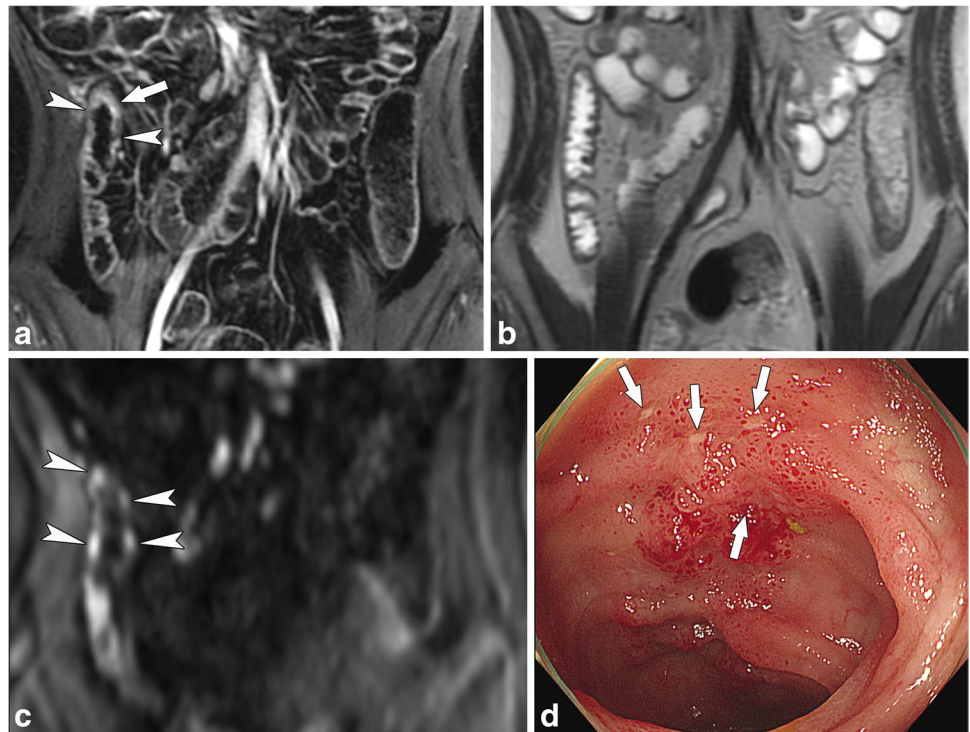


Fig. 3 A 20-year-old woman with active Crohn's disease inflammation in the descending colon. **a** Coronal contrast-enhanced T1-weighted image shows mural thickening (arrow) and ulcerations (arrowheads) of the descending colon. **b** Coronal T2-weighted image without fat saturation shows mural thickening and edema of the descending colon, but without definite ulceration. **c** Coronal DWI image shows diffusion restriction (grade 2) in the corresponding segment, resulting in a CE-sMARIA and modified sMARIA of 4, and a T2-sMARIA of 2. **d** Endoscopic image of the descending colon shows large longitudinal ulcerations and cobblestone appearance

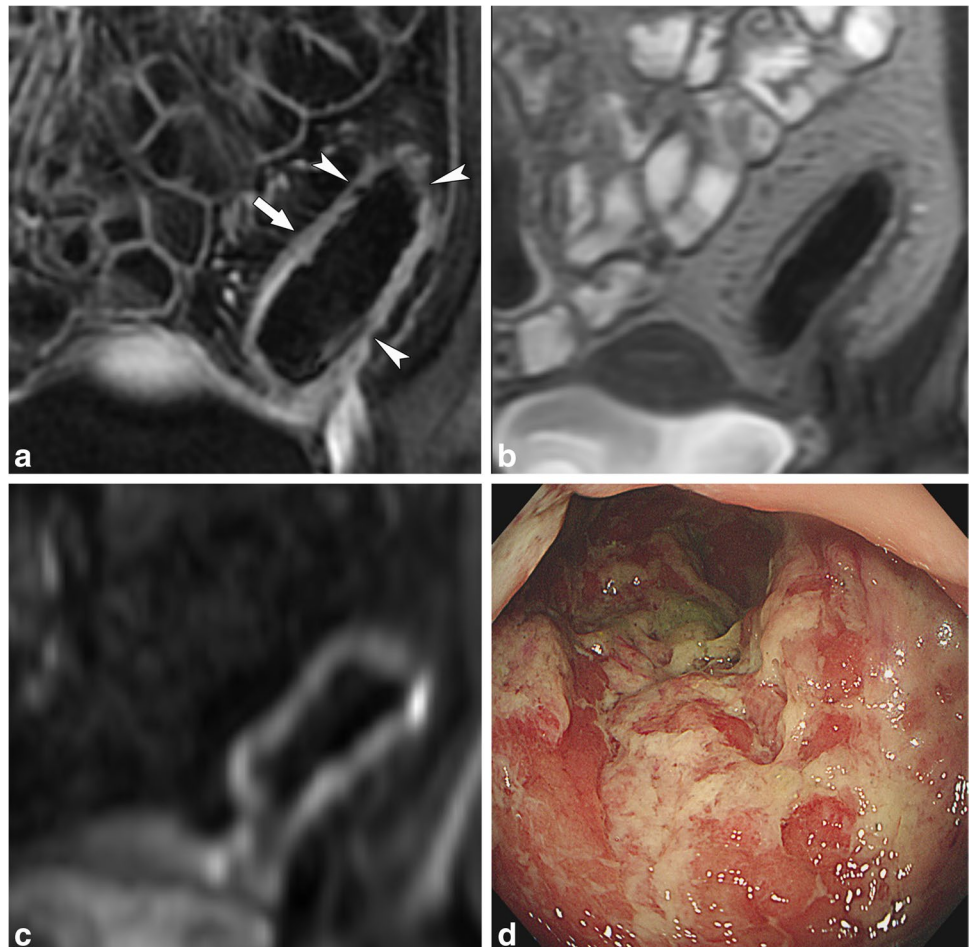


Table 3 Diagnostic performances of MRE findings for identifying endoscopic ulcerations

Parameters	Ulcer on CE	Ulcer on T2	Diffusion restriction ^b	Adjusted <i>p</i> value ^a		
				CE vs. T2	CE vs. DWI	T2 vs. DWI
Sensitivity	42.1 (16/38) [26.41–57.80]	18.4 (7/38) [6.10–30.75]	57.9 (22/38) [42.20–73.59]	0.008	0.290	< 0.001
Specificity	94.9 (225/237) [92.15–97.73]	99.2 (235/237) [97.99–100]	95.8 (227/237) [93.22–98.34]	0.004	> 0.999	0.031
Accuracy	87.6 (241/275) [83.75–91.53]	88.0 (242/275) [84.16–91.84]	90.6 (249/275) [87.09–94.00]	> 0.999	0.302	0.480

Endoscopic ulcerations ≥ 0.5 cm were included in the analyses

Values in square brackets indicate 95% confidence interval. For *p* values, CE and T2 indicate ulcers observed on conventional MRE and T2-weighted imaging, respectively, and DWI implies diffusion restriction

MRE, magnetic resonance enterography; CE, contrast-enhanced

^aFor multiple comparisons, *p* values were adjusted using Bonferroni's correction

^bStrong diffusion restriction (DWI grade of 2) is considered positive

Discussion

This study validated that sMARIA scoring is effective for detecting active inflammation and grading inflammatory severity of bowel inflammation in CD patients. In contrast to the original MARIA score requiring CE sequences, sMARIA can theoretically be evaluated solely on non-contrast scans, such as T2WI. Despite encouraging results from previous studies regarding sMARIA scoring without gadolinium-enhanced sequences [12, 13], our study revealed that T2-sMARIA had significantly inferior diagnostic performance compared to CE-sMARIA for assessing active inflammation. Additionally, T2-sMARIA showed a worse correlation with SES-CD than CE-sMARIA. In this study, we modified the original sMARIA scoring system by replacing ulcers with DWI grades, which could be evaluated on non-contrast sequences (T2WI and DWI). The

modified sMARIA resulted in significantly higher AUCs than T2-sMARIA for evaluating active and severe inflammation and it showed a similar diagnostic performance to the original CE-sMARIA. The modified sMARIA also demonstrated almost perfect interobserver agreement. Therefore, with the incorporation of DWI, the diagnostic performance of sMARIA without gadolinium enhancement could be significantly improved, resulting in a similar performance to sMARIA on conventional MRE.

Ulcers are one of the most important findings associated with active inflammation in CD on both endoscopy and MRE. Ulcer depicted on MRE is significantly correlated with CD endoscopic index of severity scores; thus, it is included in most MRE-based activity scoring systems, including sMARIA [6, 24, 25]. In this study, we evaluated the diagnostic performance of MRE-defined ulcers and diffusion restriction for the prediction of ulcer detection by endoscopy. Despite

Table 4 Detection sensitivity of MRE findings for identifying endoscopic ulcerations according to ulcer size

Parameters	Ulcer on CE	Ulcer on T2	Diffusion restriction ^b	Adjusted <i>p</i> value ^a		
				CE vs. T2	CE vs. DWI	T2 vs. DWI
Ulcer size < 0.5 cm	37.5 (9/24) [17.15–71.19]	8.3 (2/24) [1.01–30.10]	29.2 (7/24) [11.73–60.09]	0.024	> 0.999	0.177
0.5 cm \leq Ulcer size < 2 cm	45.0 (9/20) [20.58–85.42]	5.0 (1/20) [0.13–27.86]	45.0 (9/20) [20.58–85.42]	0.015	> 0.999	0.015
Ulcer size ≥ 2 cm	38.9 (7/18) [15.64–80.13]	33.3 (6/18) [12.23–72.55]	72.2 (13/18) [38.46–123.5]	> 0.999	0.042	0.024

Values in square brackets indicate 95% confidence interval. For *p* values, CE and T2 indicate ulcers observed on conventional MRE and T2-weighted imaging, respectively, and DWI implies diffusion restriction

MRE, magnetic resonance enterography; CE, contrast-enhanced

^aFor multiple comparisons, *p* values were adjusted using Bonferroni's correction

^bStrong diffusion restriction (DWI grade of 2) is considered positive

Table 5 Interobserver agreement for interpretation of MRE findings

	κ statistics (95% CI)	
	Conventional MRE	Non-contrast MRE
Mural thickening	0.720 (0.632–0.809)	0.797 (0.713–0.881)
Mural edema	0.746 (0.667–0.825)	0.718 (0.616–0.821)
Fat stranding	0.660 (0.526–0.793)	0.603 (0.507–0.699)
Ulcers	0.382 (0.212–0.552)	0.312 (0.034–0.590)
Diffusion restriction	NA	0.686 (0.602–0.770)

MRE, magnetic resonance enterography; CI, confidence interval; NA, not applicable

the high specificity of MRE-defined ulcers, the sensitivity of ulcer detection (≥ 0.5 cm) was 42.1% on conventional MRE (T2WI and CE T1WI), and was only 18.4% on T2WI alone. Moreover, the interobserver agreement of ulcer detection was unsatisfactory, with κ values of 0.382 for conventional MRE, and 0.312 for T2WI. Previous studies have also reported that ulcer detection rates on MRE are variable, ranging from 0 to 71%, and have lower interobserver reproducibility than other imaging parameters [5, 14, 26]. The low interobserver reproducibility may be because recognition of CD ulcers on MRE is dependent not only on ulcer size and depth, but also on their orientation with respect to the image plane [14]. Additionally, as inflamed mucosa appears prominently with contrast enhancement, a lack of CE sequences may further hinder the detection of subtle mucosal breaks.

Therefore, in order to improve the accuracy of sMARIA on non-contrast MRE, modification of the ulcer detection requirement is essential. As evidenced by the results of the modified sMARIA described in this study, diffusion restriction can be a good substitute for the ulcer detection requirement. In this study, diffusion restriction showed 57.9% of sensitivity to predict endoscopic ulcers (≥ 0.5 cm), which was significantly higher than ulcer detection on T2WI and comparable with ulcer detection on conventional MRE. Of note, for the detection of large ulcers (≥ 2 cm), diffusion restriction showed significantly higher sensitivity (72.2%) than ulcers on conventional MRE (38.9%) and ulcers on T2WI (33.3%). Interobserver agreement of diffusion restriction ($\kappa = 0.686$) was also significantly better than ulcer detection on both conventional MRE and T2WI. DWI is currently included as part of routine MRE sequences in many institutions and has shown potential to detect active CD inflammation and to evaluate disease severity quantitatively [15, 27]. The Clermont score, which included an ADC value in place of relative contrast enhancement from the original MARIA scores, effectively detected endoscopic ulcers with a sensitivity of 79% and a specificity of 73% [25, 28, 29]. Kim et al. [14] also modified original MARIA scores by substituting ulcer detection with diffusion restriction and reported that modified MARIA scores showed better interobserver reproducibility when

assessing inflammatory severity, and comparable diagnostic performance to original MARIA scores using CE-MRE. However, there are still several pitfalls to using DWI, including inconsistent accuracy for diagnosing active CD inflammation, and relatively low specificity with a high false positive rate, particularly in less distended bowel or colonic segments [15, 16, 30]. Therefore, technical optimization and careful interpretation with other sequences, such as T2WI, are warranted.

There are several limitations to this study. First, a small number of patients were included in this study, because we selected a 2-week time interval between ileocolonoscopy and MRE for precise comparison. Despite the limited number of patients, all statistical analyses were performed on a per-segment basis with a total of 275 bowel segments, and statistical methods were chosen that took into account intracluster correlation within a patient. Although our study was able to validate the modified scoring system using DWI for “segmental” sMARIA, special caution should be exercised when interpreting our results at the patient level. Future larger prospective studies are necessary to generalize this modified scoring system to “global” sMARIA before adopting it in daily clinical practice. Second, we retrospectively enrolled CD patients with recent ileocolonoscopy results, and a relatively high percentage of patients (36.4%) without active bowel inflammation were included. In this study, approximately half of the bowel segments with active inflammation showed only mild inflammation on endoscopy. In contrast, most previous studies used segments with moderate to severe disease. Nevertheless, this study has shown that sMARIA and modified sMARIA are also effective at evaluating either mild or inactive disease states. Additionally, retrospective evaluation of endoscopic images has its own limitations. Third, axial T2WI was not included in the routine MRE protocols during the early study period. Since the detection of ulcers is affected by the orientation of ulcers with respect to the viewing plane, the lack of axial T2WI might decrease the diagnostic performance of T2WI for ulcer evaluation. Fourth, we did not evaluate DWI alone, only with T2WI; thus, T2WI findings might affect the DWI evaluation. However, our study design might reflect clinical practices because DWI alone has poor anatomical details, and MRE interpretation in daily clinical practice includes the use of combined sequences. Fifth, we did not perform sophisticated regression analysis when developing the modified sMARIA score. Instead, we practically weighted coefficient 1 for DWI grades to maintain similarity to the original sMARIA scale. Further large-scale prospective studies are necessary to identify the optimal combination of DWI to improve upon our results. Sixth, non-contrast MRE might have limited value for evaluating penetrating complications in CD patients. Out of eight patients in this study that had penetrating complications detected on conventional MRE, two had penetrating complications overlooked on non-contrast MRE. As previous studies have also shown that CE sequences might enhance sensitivity

to detect penetrating complications [17, 31], caution is warranted before omitting contrast enhancement if penetrating complications are suspected. Finally, since the primary aim of this study was to introduce and validate sMARIA using non-contrast MRE with DWI for diagnosing active inflammation in CD patients, further studies regarding the accuracy of modified sMARIA in monitoring treatment response or drug efficacy are needed for the comprehensive implementation of this scoring system in daily practice.

In conclusion, our study proposed a modified sMARIA that uses DWI instead of ulcers in its calculation and externally validated the sMARIA scoring system to diagnose active inflammation in CD patients. Modified sMARIA using DWI can potentially improve the diagnostic performance of non-contrast MRE and achieve comparable performance to sMARIA using CE-MRE. However, further prospective studies are warranted in a larger population to generalize our results.

Funding This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) [grant number NRF-2018R1C1B6002747], a faculty research grant from Yonsei University College of Medicine [grant number 6-2021-0061], and Bracco diagnostics.

Declarations

Guarantor The scientific guarantor of this publication is Dr. Nieuw Seo.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry Hye Jung Shin (from the Biostatistics Collaboration Unit, Department of Biomedical Systems Informatics, Yonsei University College of Medicine) kindly provided statistical advice for this manuscript.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval from Severance Hospital was obtained.

Methodology

- retrospective
- diagnostic or prognostic study
- performed at one institution

References

1. Cushing K, Higgins PDR (2021) Management of Crohn disease: a review. *JAMA* 325:69–80
2. Lichtenstein GR, Loftus EV, Isaacs KL, Regueiro MD, Gerson LB, Sands BE (2018) ACG clinical guideline: management of Crohn's disease in adults. *Am J Gastroenterol* 113:481–517
3. Maaser C, Sturm A, Vavricka SR et al (2018) ECCO-ESGAR Guideline for diagnostic assessment in IBD Part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 13:144–164K
4. Sturm A, Maaser C, Calabrese E et al (2018) ECCO-ESGAR Guideline for diagnostic assessment in IBD part 2: IBD scores and general principles and technical aspects. *J Crohns Colitis* 13:273–284
5. Rimola J, Ordás I, Rodriguez S et al (2010) Magnetic resonance imaging for evaluation of Crohn's disease: validation of parameters of severity and quantitative index of activity. *Inflamm Bowel Dis* 17:1759–1768
6. Ordás I, Rimola J, Alfaro I et al (2019) Development and validation of a simplified magnetic resonance index of activity for Crohn's disease. *Gastroenterology* 157:432–439.e431
7. Maximova N, Gregori M, Zennaro F, Sonzogni A, Simeone R, Zanon D (2016) Hepatic gadolinium deposition and reversibility after contrast agent-enhanced MR imaging of pediatric hematopoietic stem cell transplant recipients. *Radiology* 281:418–426
8. Kobayashi M, Levendovszky SR, Hippe DS et al (2021) Comparison of human tissue gadolinium retention and elimination between gadoteridol and gadobenate. *Radiology* 300:559–569
9. FDA Drug Safety Communication: FDA warns that gadolinium-based contrast agents (GBCAs) are retained in the body; requires new class warnings, U.S. Food and Drug Administration (2018) Available via <https://www.fda.gov/Drugs/DrugSafety/ucm589213.htm>. Accessed 20 Nov 2022
10. Bruining DH, Zimmermann EM, Loftus EV et al (2018) Consensus recommendations for evaluation, interpretation, and utilization of computed tomography and magnetic resonance enterography in patients with small bowel Crohn's disease. *Gastroenterology* 154:1172–1194
11. Moy MP, Sauk J, Gee MS (2016) The role of MR enterography in assessing Crohn's disease activity and treatment response. *Gastroenterol Res Pract* 2016:8168695
12. Capozzi N, Ordás I, Fernandez-Clotet A et al (2020) Validation of the simplified magnetic resonance index of activity [sMARIA] without gadolinium-enhanced sequences for Crohn's disease. *J Crohns Colitis* 14:1074–1081
13. Fernández-Clotet A, Sapena V, Capozzi N et al (2022) Avoiding contrast-enhanced sequences does not compromise the precision of the simplified MaRIA for the assessment of non-penetrating Crohn's disease activity. *Eur Radiol* 32:3334–3345
14. Kim JS, Jang HY, Park SH et al (2017) MR enterography assessment of bowel inflammation severity in Crohn disease using the MR index of activity score: modifying roles of DWI and effects of contrast phases. *AJR Am J Roentgenol* 208:1022–1029
15. Choi SH, Kim KW, Lee JY, Kim K-J, Park SH (2015) Diffusion-weighted magnetic resonance enterography for evaluating bowel inflammation in Crohn's disease: a systematic review and meta-analysis. *Inflamm Bowel Dis* 22:669–679
16. Park SH (2016) DWI at MR enterography for evaluating bowel inflammation in Crohn disease. *AJR Am J Roentgenol* 207:40–48
17. Seo N, Park SH, Kim KJ et al (2016) MR Enterography for the evaluation of small-bowel inflammation in Crohn disease by using diffusion-weighted imaging without intravenous contrast material: a prospective noninferiority study. *Radiology* 278:762–772
18. Daperno M, D'Haens G, Van Assche G et al (2004) Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointest Endosc* 60:505–512
19. Obuchowski NA (1997) Nonparametric analysis of clustered ROC curve data. *Biometrics* 53:567–578
20. Zou KH, Tuncali K, Silverman SG (2003) Correlation and simple linear regression. *Radiology* 227:617–622
21. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
22. Yang Z, Zhou M (2014) Kappa statistic for clustered matched-pair data. *Stat Med* 33:2612–2633

23. Vanbelle S (2017) Comparing dependent kappa coefficients obtained on multilevel data. *Biom J* 59:1016–1034
24. Rimola J, Rodriguez S, Garcia-Bosch O et al (2009) Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease. *Gut* 58:1113–1120
25. Buisson A, Joubert A, Montoriol P-F et al (2013) Diffusion-weighted magnetic resonance imaging for detecting and assessing ileal inflammation in Crohn's disease. *Aliment Pharmacol Ther* 37:537–545
26. Tielbeek JA, Makanyanga JC, Bipat S et al (2013) Grading Crohn disease activity with MRI: interobserver variability of MRI features, MRI scoring of severity, and correlation with Crohn disease endoscopic index of severity. *AJR Am J Roentgenol* 201:1220–1228
27. Dohan A, Taylor S, Hoeffel C et al (2016) Diffusion-weighted MRI in Crohn's disease: current status and recommendations. *J Magn Reson Imaging* 44:1381–1396
28. Buisson A, Hordonneau C, Goutte M, Boyer L, Pereira B, Bommelaer G (2015) Diffusion-weighted magnetic resonance imaging is effective to detect ileocolonic ulcerations in Crohn's disease. *Aliment Pharmacol Ther* 42:452–460
29. Hordonneau C, Buisson A, Scanzi J et al (2014) Diffusion-weighted magnetic resonance imaging in ileocolonic Crohn's disease: validation of quantitative index of activity. *Am J Gastroenterol* 109:89–98
30. Kim K-J, Lee Y, Park SH et al (2014) Diffusion-weighted MR enterography for evaluating Crohn's disease: how does it add diagnostically to conventional MR enterography? *Inflamm Bowel Dis* 21:101–109
31. Barat M, Hoeffel C, Bouquot M et al (2019) Preoperative evaluation of small bowel complications in Crohn's disease: comparison of diffusion-weighted and contrast-enhanced MR imaging. *Eur Radiol* 29:2034–2044

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.