Retrospective Study

Quantitative diffusion-weighted magnetic resonance enterography in ileal Crohn's disease: A systematic analysis of intra and interobserver reproducibility

Hao Yu, Ya-Qi Shen, Fang-Qin Tan, Zi-Ling Zhou, Zhen Li, Dao-Yu Hu, John N Morelli

ORCID number: Hao Yu (0000-0002-4495-897X); Ya-Qi Shen (0000-0003-0589-8975); Fang-Qin Tan (0000-0003-7383-8573); Zi-Ling Zhou (0000-0002-9793-4382); Zhen Li (0000-0002-9289-6785); Dao-Yu Hu (0000-0001-9281-8522); John N Morelli (0000-0002-6451-6984).

Author contributions: Hu DY and Shen YQ designed the research; Yu H and Tan FQ performed the MR examinations; Yu H and Zhou ZL analyzed the data; Yu H wrote the paper; Morelli JN, Li Z, Shen YQ, and Hu DY revised the paper.

Supported by National Natural Science Foundation of China, No. 81571642, No. 81701657, and No. 81771801; and the Fundamental Research Funds for the Central Universities, No. 2017KFYXJ126.

Institutional review board statement: This study was approved by the Ethics Committee of the Tongji Hospital of Huazhong University of Science and Technology.

Informed consent statement: Informed consent for this study was waived due to its retrospective nature.

Conflict-of-interest statement: All authors declare no conflict of interest related to this study.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article which was

Abstract

BACKGROUND

Magnetic resonance enterography (MRE) is increasingly attractive as a noninvasive and radiation-free tool for assessing Crohn’s disease (CD). Diffusion-weighted imaging (DWI) is recommended as an optional MRE sequence for CD by the European Society of Gastrointestinal and Abdominal Radiology, and has shown a superb potential as a quantitative modality for bowel inflammation evaluation. However, the measurement reproducibility of quantitative DWI analysis in MRE has not been ascertained so far. To facilitate the application of quantitative diffusion-weighted MRE in the clinical routine, systematic investigations of the intra and interobserver reproducibility of DWI quantitative parameters should be performed.

AIM

To evaluate the intra and interobserver reproducibility of quantitative analysis for diffusion-weighted MRE (DW-MRE) in ileal CD.

METHODS

Forty-four subjects (21 with CD and 23 control subjects) who underwent ileocolonoscopy and DW-MRE (b = 800 s/mm²) within one week were included. Two radiologists independently measured apparent diffusion coefficients (ADC) of the terminal ileum and signal intensity ratio (SR) of the terminal ileum to ipsilateral psoas muscle on DWI images (b = 800 s/mm²). Between- and within-reader agreements were assessed using intraclass correlation coefficients (ICC), coefficients of variation (CoV), and 95% limits of agreement of Bland-Altman plots (BA- LA LoA). Diagnostic performances of ADC and SR for identifying
INTRODUCTION

Crohn’s disease (CD) is a life-long, relapsing, chronic inflammatory disease of the gastrointestinal tract that most frequently involves the terminal ileum[1]. Serial evaluation including monitoring of disease and assessment of treatment efficacy is important throughout the patient’s life[2-4]. For non-invasive assessment, magnetic resonance imaging (MRI) enterography (MRE) is increasingly attractive as an ionizing radiation-free supplement to endoscopy in assessment of CD patients[5-8]. MRE presents several advantages as it not only assesses abnormalities of the bowel lumen[2,6], but also identifies extra-intestinal complications (abscesses or fistulae)[8,9], offers an assessment of the whole gastrointestinal tract[5], and has the merit owing to its high soft tissue contrast[9,10]. For non-invasive assessment, magnetic resonance enterography (MRE) is increasingly attractive as a radiation-free supplement to endoscopy in assessment of CD patients[5-8]. MRE presents several advantages as it not only assesses abnormalities of the bowel lumen[2,6], but also identifies extra-intestinal complications (abscesses or fistulae)[8,9], offers an assessment of the whole gastrointestinal tract[5], and has the merit owing to its high soft tissue contrast[9,10]. Since the conventional MRE is based upon the diffusion-weighted magnetic resonance imaging (DWI) could qualitatively and quantitatively evaluate random motion of water molecules at the microstructural level in biological tissues, and is recommended upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

CONCLUSION
Quantitative parameters derived from DW-MRE have good to excellent intra and interobserver agreements with high diagnostic accuracy, and can serve as robust and efficient quantitative biomarkers for CD evaluation.

Key words: Magnetic resonance imaging; Diffusion-weighted imaging; Crohn’s disease; Ileum; Reproducibility

©The Author(s) 2019. Published by Baishideng Publishing Group Inc. All rights reserved.

Core tip: We evaluated the measured reproducibility of quantitative apparent diffusion coefficients (ADC) and signal contrast ratio (SR) of the terminal ileum to ipsilateral psoas muscle on diffusion-weighted imaging (DWI; b = 800 s/mm²). The intra and interobserver reproducibility of ADC and SR values was good to excellent for Crohn’s disease (CD) and normal control subjects. Quantitative parameters of ADC and SR derived from DWI had outstanding diagnostic efficacy for CD. Therefore, both ADC and SR could serve as robust and efficient quantitative biomarkers for the evaluation of CD.

Citation: Yu H, Shen YQ, Tan FQ, Zhou ZL, Li Z, Hu DY, Morelli JN. Quantitative diffusion-weighted magnetic resonance enterography in ileal Crohn’s disease: A systematic analysis of intra and interobserver reproducibility. World J Gastroenterol 2019; 25(27): 3619-3633

URL: https://www.wjgnet.com/1007-9327/full/v25/i27/3619.htm
DOI: https://dx.doi.org/10.3748/wjg.v25.i27.3619

selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
segments in CD patients, and is suitable for the patients who cannot receive intravenous contrast agents. Furthermore, the latest consensus recommendations pointed out that restricted diffusion on DWI was a valuable and supportive sign correlating with the inflammatory lesions demonstrated on endoscopy. The apparent diffusion coefficient (ADC) derived from DWI has been reported as a potential quantitative biomarker for monitoring the course of disease and therapeutic efficiency in patients with CD.

However, ADC values calculated from DWI could be degraded by variations in field homogeneity, eddy currents, and motion, particularly in hollow organs like the gastrointestinal tract or gallbladder. And systematic investigations on measurement reproducibility of quantitative DWI analysis in MRE have not been ascertained so far. The characterization of reader agreement and variability with quantitative parameters is imperative before such parameters are to be incorporated into clinical practice. As ADC values could be affected by the image mismatch calculated from DWI with at least two b values, we proposed an abbreviated proxy for diffusion restriction, a signal intensity ratio (SR), between the terminal ileum and ipsilateral psoas muscle on DWI with an appropriate b value (800 s/mm²), and explored the application of SR and ADC values in CD patients.

The purpose of this study was to investigate the level of inter and intraobserver agreement for ADC and SR measurements from DWI in the terminal ileum, and then to assess the diagnosis efficiency of quantitative ADC and SR for the evaluation of CD.

**MATERIALS AND METHODS**

**Patients**

The study was approved by our institutional review board, and informed consent was waived for the retrospective study design. From November 2015 to May 2018, a total of 297 patients underwent MRE examinations for suspected gastrointestinal disorders at our gastrointestinal medical center of the referral university-based hospital. Sixty-one patients were included based on the follow criteria: (1) Complete clinical data available; (2) DWI with a b-value of 800 s/mm² included in the MRE protocol; (3) Colonoscopy was successfully conducted to assess the terminal ileum within one week of the MR enterography; and (4) No bowel surgery performed prior to MRI examination. Among the 61 patients, 12 diagnosed with other diseases and 5 with inadequate MRE images due to serious artifacts were excluded. Finally, the remaining 44 patients, including 21 with CD involving the terminal ileum and 23 with a normal terminal ileum as control subjects, were included for the further analysis. A flowchart of the study population is shown in Figure 1. The diagnosis of CD was confirmed by clinical information and a combination of endoscopic, histological, and radiological investigations according to the European Crohn’s and Collitis Organization criterion.

**MRI data acquisition**

MRE was performed in a supine position using a 3.0-T GE scanner (Discovery MR750; GE Healthcare, Milwaukee, WI, United States) with a 32-channel torso phased array body coil. Per protocol, all patients had fasted for at least 6 h before MRE. To achieve adequate distension of the small bowel, about 1500 mL of 2.5% isosmotic mannitol solution was administered orally to the patients 45 min prior to MR scanning. Anisodamine 10 mg was administered intramuscularly 5 min before the MR examination to reduce bowel peristalsis. Conventional MRE was acquired including T2-weighted single shot fast spin echo (SSFSE; TR/TE: a respiratory cycle/68 ms, matrix: 288 × 288, slice thickness: 4 mm for coronal/5 mm for axial), the fast imaging employing steady state acquisition (FIESTA; TR/TE: 3.2/1.2 ms, matrix: 288 × 288, slice thickness: 4 mm for coronal/5 mm for axial, flip angle: 45 degrees), and coronal and axial T1 weighted three-dimensional fast spoiled gradient echo (liver acquisition with volume acceleration sequence, LAVA; TR/TE: 3.8/1.7 ms, matrix: 260 × 210, slice thickness: 4 mm, flip angle: 15 degrees).

The axial DWI sequence was obtained using single shot echo planar imaging (SS-EPI; TR/TE: a respiratory cycle/minimum, field of view: 38.0 cm × 30.4 cm, matrix: 160 × 128, slice thickness: 6 mm) with b values of 0 and 800 s/mm²; the number of excitations (NEX) was 6. The total acquisition time of the whole examination for every patient was approximately 25 min depending on the patient’s respiratory rate.

**Image analysis**

All images were transferred to a workstation (Advantage Workstation 4.6, GE Healthcare) where ADC maps were generated by using a mono-exponential fit. Two
gastrointestinal radiologists (R1 with 5 years of experience and R2 with 3 years of experience), both blinded to clinical details as well as endoscopy results, measured the signal intensity (SI) of the terminal ileum (defined as the final 10 cm of the small bowel) and ipsilateral psoas muscle on DWI images with $b = 800 \text{ s/mm}^2$. Separately and simultaneously, ADC values for the terminal ileum were recorded. For the SI and ADC measurements in the terminal ileum, the images were magnified, and three oval regions of interest (ROI), as large as possible ($12-30 \text{ mm}^2$) to encompass the bowel wall, were placed on the area of brightest signal on the DWI along with hypointensity on ADC maps (Figure 2), as having been described in previous studies [9,23]. The means of the three measurements (SI and ADC) were used for the further analysis. Then, SI of the psoas muscle on DWI with $b = 800 \text{ s/mm}^2$ was measured by placing a larger oval ROI ($60-100 \text{ mm}^2$). The SR of the terminal ileum to ipsilateral psoas muscle was calculated according to the prior studies [24,25] by the equation: $\text{SR} = \frac{\text{SI}_{\text{ileum}}}{\text{SI}_{\text{muscle}}}$, where $\text{SI}_{\text{ileum}}$ and $\text{SI}_{\text{muscle}}$ represent the measured SI of the terminal ileum and ipsilateral psoas muscle, respectively. After a washout period of 4 wk, the same two radiologists repeated the above described measurements to assess intraobserver agreement.

**Reference standard: Endoscopy assessment**

The terminal ileum of all the subjects was assessed by ileocolonoscopy as the reference standard. Prior endoscopy, bowel preparation was performed via oral ingestion of about 4 L of polyethylene glycol solution according to the standard protocol. Endoscopy was successfully conducted to assess the terminal ileum in all of the included patients. Endoscopic analysis of the terminal ileum was performed by an endoscopist with 10 years of experience who was blinded to MRE results and clinical data. The colonoscopy findings of the terminal ileum, such as overt ulcers, aphthoid lesions, and erythema or edema, were considered as inflammatory conditions [17,22].

**Statistical analysis**

Statistical calculations were performed using the Statistical Package for the Social Sciences Software Package (SPSS Statistics, version 20.0; SPSS Inc., Chicago, IL, United States) and MedCalc (release 15.2.2; MedCalc Software, Mariakerke, Belgium). The normal distribution of quantitative parameters was ascertained by the Kolmogorov-Smirnov test. Differences in ADC and SR between the two readers and two sessions within the CD and control groups were compared using a paired $t$-test. Inter and intraobserver agreements for SR and ADC measurements were assessed using the intraclass correlation coefficient (ICC), coefficients of variation (CoV) from duplicate
Figure 2 A 48-year-old man with Crohn’s disease involving the terminal ileum. A: Axial diffusion-weighted imaging with \( b = 800 \) s/mm\(^2\) demonstrated high signal intensity of the terminal ileum (arrow). B: The terminal ileum was hypointense (arrow) on the corresponding apparent diffusion coefficient map. C: Ileocolonoscopy showed mucosal ulcers. D: Histopathology demonstrated neutrophil infiltration.

measurements, and the Bland-Altman analysis with 95% limits of agreement (LoA). ICC was interpreted as follows: 0.00-0.20, poor; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, good; 0.81-1.00, excellent. Diagnostic efficiency of the DWI parameters was determined by receiver operating characteristic (ROC) curve analyses. The comparison of the areas under ROC curve (AUROC) was performed between SR and ADC using the method of Delong et al. Spearman’s correlation test was used to identify correlation between SR and ADC values in CD patients. \( P \)-values less than 0.05 were considered statistically significant.

RESULTS

Patient characteristics
A total of 44 subjects (age: Mean ± SD, 44.9 ± 15.7 years; 22 males and 22 females) were included in this study, and all of the DWI images were eligible for evaluation. Twenty-three control subjects had a normal terminal ileum in colonoscopy examination. The remaining 21 CD patients with the terminal ileum involved presented 12 overt ulcers, 5 aphthoid lesions, and 4 erythema or edema conditions according to endoscopy. The patients’ data of this cohort are summarized in Table 1.

Comparisons of quantitative ADC and SR between two reading sessions or two radiologists
There were no significant differences (paired \( t \)-test, \( P > 0.05 \)) in ADC or SR values between the two sessions or between the two radiologists either in the CD or control group (Tables 2 and 3). The intra and interobserver agreements of quantitative DWI parameters measurements are shown in Tables 2 and 3 and Figures 3 and 4.

Intraobserver agreement analysis
The intraobserver agreement for SR was excellent in both the CD and control groups with similar results obtained by the two readers (R1: ICC = 0.989, CoV = 3.51%; R2: ICC = 0.979, CoV = 4.64% for CD and R1: ICC = 0.974, CoV = 5.93%; R2: ICC = 0.944, CoV = 8.69% for the normal). Similarly, the ADC measurements in the CD group showed excellent intraobserver reproducibility (R1: ICC = 0.952, CoV = 6.28%; R2: ICC = 0.977, CoV = 4.71%). However, intraobserver agreement was inferior when measuring the ADC on the normal terminal ileum (R1: ICC = 0.736, CoV = 10.47%; R2:
Table 1 Clinical characteristics of included subjects

| Characteristic | Value |
|---------------|-------|
| CD patients (n) | 21 |
| Age (yr) (mean ± SD) | 43.4 ± 16.1 |
| Gender (males: Females) | 11:10 |
| CRP (mg/dL) [median (range)] | 21.0 (0.9-94.0) |
| ESR (mm/h) [median (range)] | 19 (7-51) |
| Days between MRE and endoscopy median [median (range)] | 4 (0-7) |
| Endoscopic findings of terminal ileum (n) | |
| Overt ulcers | 12 |
| Aphthoid lesions | 5 |
| Erythema or edema conditions | 4 |
| Control group (n) | 23 |
| Age (yr) (mean ± SD) | 46.3 ± 15.4 |
| Gender (males:females) | 11:12 |
| Days between MRE and endoscopy [median (range)] | 3 (1-7) |
| Total (n) | 44 |
| Age (yr) (mean ± SD) | 44.9 ± 15.7 |
| Gender (males:females) | 22:22 |
| Days between MRE and endoscopy [median (range)] | 4 (0-7) |

CD: Crohn’s disease; SD: Standard deviation; CRP: C-reactive protein; ESR: Erythrocyte sedimentation rate; MRE: Magnetic resonance enterography.

ICC = 0.711, CoV = 10.64%). Bland-Altman plots for intraobserver agreement showed rather narrow 95% LoA ranges except for the ADC measurements in the control group; no systematic bias between sessions was present (Figure 3).

Interobserver agreement analysis
Similar to the analysis of intraobserver reproducibility, all ADC measurements in the CD group and SR in both the CD and normal groups demonstrated excellent interobserver agreement (Table 3). The lowest agreement and highest variability between readers were observed in ADC measurements of the control group (first session: ICC = 0.641, CoV = 11.43%; second session: ICC = 0.736, CoV = 11.07%). Similarly, only the interreader ADC measurements for the control group demonstrated broad 95% LoA ranges (-33.93% to 25.96% for first session; -32.03% to 29.62% for second session) as illustrated in Figure 4.

Spearman’s correlation analysis between SR and ADC values in the CD patients
The correlation between the quantitative DWI parameters of SR and ADC is shown in Figure 5. SR was moderately and negatively correlated with ADC values in CD patients for both readers (R1: r = - 0.438, P = 0.047; R2: r = - 0.485, P = 0.026).

Diagnostic efficiency of ADC and SR for differentiating inflammatory bowel segments from the normal
Due to the good to excellent intraobserver agreement of ADC and SR measurements in two sessions of the two readers, the mean ADC and SR values were calculated based on an average from the two sessions for further analysis. ADC and SR from both readers achieved a high diagnostic accuracy for differentiating inflammatory segments from the normal bowel with high sensitivities and specificities (Table 4 and Figure 6). The areas under the ROC curve for SR (R1: 0.981; R2: 0.988) were slightly better than those of ADC (R1: 0.923; R2: 0.929), but these differences were not statistically significant (P-values: 0.117 for R1; 0.067 for R2). The cutoff values maximizing sensitivity and specificity for ADC and SR parameters were approximately 1.80 × 10⁻³ mm²/s and 2.00, respectively.

DISCUSSION
DW-MRE has been increasingly utilized for the noninvasive evaluation of CD over the last few years[5-10,14,26]. Meanwhile, quantitative analysis of DW-MRE is gradually
Yu H et al. Reproducibility of quantitative DW-MRE in CD

Table 2  Intraobserver agreement and variability of apparent diffusion coefficient and signal ratio measurements between the two readers over two sessions

| Group | CD (n = 21) | Normal (n = 23) |
|-------|-------------|----------------|
|       | 1st session | 2nd session | P-value | 1st session | 2nd session | P-value | 1st session | 2nd session | P-value |
| ADC   | 1.50 ± 0.40 | 1.54 ± 0.45 | 0.201    | 1.52 ± 0.45 | 1.56 ± 0.47 | 0.125    | 2.55 ± 0.53 | 2.53 ± 0.53 | 0.772   |
| ICC   | 0.952       | 0.977        |          | 0.736       |             |          | 0.711       |             |         |
| CoV   | 6.28        | 4.71         |          | 10.47       |             |          | 10.64       |             |         |
| Bias (%) | 1.77      | 1.89         |          | -1.07       |             |          | -1.71       |             |         |
| LoA (%) | -14.10 to | -10.65 to 14.42 |          | -30.75 to 28.62 |             |          | -28.30 to 24.86 |             |         |
| SR    | 4.17 ± 1.43 | 4.20 ± 1.36 | 0.505    | 4.21 ± 1.34 | 4.23 ± 1.42 | 0.794    | 1.45 ± 0.54 | 1.44 ± 0.53 | 0.697   |
| ICC   | 0.989       | 0.979        |          | 0.974       |             |          | 0.944       |             |         |
| CoV   | 3.51        | 4.64         |          | 5.93        |             |          | 8.69        |             |         |
| Bias (%) | 1.39       | -0.07        |          | -0.58       |             |          | 0.45        |             |         |
| LoA (%) | -9.22 to 12.01 | -13.81 to 13.67 |          | -19.16 to 17.99 |             |          | -23.99 to 24.88 |             |         |

Apparent diffusion coefficient (ADC) and signal ratio parameters are expressed as the mean ± standard deviation, ADC units: ×10⁻² mm²/s. ADC: Apparent diffusion coefficient; SR: Signal ratio; CD: Crohn’s disease; ICC: Intraclass correlation coefficient; CoV: Coefficients of variation; LoA: Bland-Altman analysis with 95% limits of agreement.

drawing attention as an objective biomarker for bowel inflammation[9,17,26-29]. ADC values along with the Clermont score have been reported as having excellent accuracy for identifying inflamed bowel segments and evaluating therapeutic response in CD patients[19,20,26]. However, the measurement agreement for DW-MRE quantitative parameters has not been ascertained to date[19]. To facilitate the application of quantitative DW-MRE into the clinical routine, a thorough study of the intra and interobserver reproducibility of DWI quantitative parameters should be performed, similar to what has been done in other organs[13,20].

In the current study, intra and interobserver agreements of ADC measurements in CD patients were excellent with high ICC, low CoV, and a limited 95% LoA range (Tables 2 and 3). ADC values of the terminal ileum in the CD group were robust both between and within individual readers, suggesting that different radiologists could obtain comparable ADC values in CD patients over time. In previous studies[9,23,29], interobserver agreement of measured ADC values in CD patients was also superb with high ICC, similar to our results of ADC analysis between radiologists. As the same with ADC values, excellent intra and interobserver agreements of SR measurement were found in inflamed terminal ileum. As expected, SR demonstrated moderately negative correlations with ADC in CD patients in both readers (Figure 5), indicating that SR could also represent the diffusion condition on DWI just like the ADC value. Therefore, both ADC and SR could act as the robust and quantitative methods for the evaluation of CD patients.

Although the comprehensive Clermont score has been introduced as a valuable tool for the ideal CD assessment in academic studies[12,23,26], the calculation of Clermont score is based on quantitative ADC values and wall thickness along with qualitative signs of bowel ulcer and mural edema. As a completed quantitative analysis of DWI in this study, Clermont score was not applied for its qualitative signs on conventional MRE. And the Kappa values of inter-reader agreement were only fair for the presence of ulcer (0.34) and moderate for the presence of mural edema (0.47) on conventional MRE[19]. Furthermore, ADC values were reported to have similarly high diagnostic performance for CD patients compared with Clermont score in prior studies[22,23]. In addition, ADC measurement from the post-processing ADC maps and SR of the terminal ileum to ipsilateral psoas muscle on b = 800 s/mm² images are relatively convenient to implement in the daily clinical practice.

However, the measured ADC in normal terminal ileum demonstrated lower agreement between sessions and readers (ICC: 0.641-0.736; CoV: 10.47%-11.43%; BA- LA LoA: ±26.59% to ±30.83%). This might be due to more pronounced partial volume effects of the thin normal intestinal tissues during the ROI placement compared to the thickened inflammatory segments[9]. Additionally, post-processing ADC maps are generated by applying a mono-exponential fit with at least two b values, a technique prone to image degradation due to bowel peristalsis[23]. The SR in the subjects with a normal terminal ileum still achieved high intra and interobserver agreements (ICC:
Yu H et al. Reproducibility of quantitative DW-MRE in CD

Table 3 Interobserver agreement and variability of apparent diffusion coefficient and signal ratio measurements between the two readers over two sessions

| Group | CD (n = 21) | Normal (n = 23) |
|-------|-----------|----------------|
|       | 1st sessi-on reader 1 | 1st sessi-on reader 2 | P-value | 1st sessi-on reader 1 | 1st sessi-on reader 2 | P-value |
| ADC   | 1.50 ± 0.40 | 1.52 ± 0.45 | 0.402 | 1.54 ± 0.45 | 1.56 ± 0.47 | 0.379 | 2.55 ± 0.53 | 2.57 ± 0.50 | 0.815 | 2.53 ± 0.53 | 2.64 ± 0.57 | 0.208 |
| ICC   | 0.969 | 0.984 | 0.641 | 0.736 |
| CoV   | 4.71 | 3.73 | 11.43 | 11.07 |
| Bias (%) | -0.72 | -0.82 | -1.20 | -3.98 |
| LoA (%) | -14.29 to 12.85 | -12.09 to 10.44 | -32.03 to 29.62 | -33.93 to 25.96 |
| SR    | 4.17 ± 1.43 | 4.21 ± 1.34 | 0.473 | 4.20 ± 1.36 | 4.23 ± 1.42 | 0.569 | 1.45 ± 0.54 | 1.44 ± 0.52 | 0.760 | 1.44 ± 0.53 | 1.45 ± 0.56 | 0.618 |
| ICC   | 0.981 | 0.989 | 0.950 | 0.971 |
| CoV   | 4.53 | 3.72 | 8.12 | 6.27 |
| Bias (%) | -1.85 | -0.39 | 0.43 | -0.59 |
| LoA (%) | -17.30 to 13.60 | -12.38 to 11.61 | -22.17 to 23.04 | -19.47 to 18.29 |

Apparent diffusion coefficient (ADC) and signal ratio parameters are expressed as the mean ± standard deviation, ADC units: ×10⁻³ mm²/s. ADC: Apparent diffusion coefficient; SR: Signal ratio; CD: Crohn’s disease; ICC: Intraclass correlation coefficient; CoV: Coefficients of variation; LoA: Bland-Altman analysis with 95% limits of agreement.

0.944-0.974) and was less variable than ADC between sessions and readers. The main reasons could be that SR is calculated from a single b value DWI, limiting the propensity for motion degradation, and acquisition of a single b value DWI image using SS-EPI is rapid, sub-second using the latest acquisition techniques and hardware⁸,¹⁹. Hence, SR could be more robust for quantitative DWI assessment in normal or inflamed ileal segments of CD patients. And the measurement of SR avoids the post-processing maps compared to ADC, which would be easier and more applicable in clinic.

In line with prior quantitative analysis of DWI on CD patients⁹,¹¹,¹²,¹³, the ADC values of inflammatory lesions were significant lower than those of normal segments in this study, and achieved optimal AUROC (R1: 0.923 and R2: 0.929) with high sensitivities and specificities by a threshold value of 1.8 × 10⁻³ mm²/s. Furthermore, diagnostic efficiency of SR (AUROC: 0.981 for R1 and 0.988 for R2) was slightly superior to the ADC values. According to our results and Table 4, SR values greater than 2 (meaning more than twice signal intensity of the terminal ileum to ipsilateral psoas muscle) could be used to differentiate inflamed terminal ileum from the normal. Quantitative parameters of ADC and SR from DWI both had excellent diagnostic accuracy for CD. Similarly, the relative signal intensity (rSI) of lesion to muscle demonstrated higher diagnostic performance than ADC in assessing complete tumor response on post neoadjuvant chemoradiotherapy DWI in rectal cancer¹⁴. The SR, as an alternative modality of ADC, would be a reliable and effective quantitative parameter of DW-MRE to identify the normal and inflamed bowel segments based on this initial study. The further exploration of SR in the differentiation of CD stages and medical therapy assessment during CD course is worthwhile in the future study.

However, our study had some limitations due to the retrospective nature and relatively small sample size included. Further prospective trials with larger sample sizes would help to confirm these results. Since DW-MRE examinations of this primary study were acquired from a single center using a same 3.0-T MR scanner, a multicenter study with various MR systems could further increase the applicability of the data. Additionally, ROIs placement on the bowel wall was relatively challenging and subjective for the observers; improvements in automated approaches to ROI placement would be potentially of great value. Finally, the quantitative endoscopic score was not acquired due to this retrospective study.

In conclusion, the intra and interobserver agreements of quantitative ADC and SR
Figure 3  Bland-Altman plots of intraobserver agreement between 2 two sessions of two readers. Bland-Altman 95% limits of agreement with diffusion weighted imaging parameters including apparent diffusion coefficient and signal ratio from the Crohn’s disease group and normal group. The solid line represents the mean relative difference as a percentage and the dashed lines represent the upper and lower 95% limits of agreement. ADC: Apparent diffusion coefficient; SR: Signal ratio; CD: Crohn’s disease; SD: Standard deviation; R1S1: First session of reader 1; R1S2: Second session of reader 1; R1S2: First session of reader 2; R2S2: Second session of reader 2.
Figure 4  Bland-Altman plots of interobserver agreement between two readers at two sessions. Bland-Altman 95% limits of agreement with Crohn’s disease parameters including apparent diffusion coefficient and signal ratio from the Crohn’s disease group and normal group. The solid line represents the mean relative difference as a percentage and the dashed lines represent the upper and lower 95% limits of agreement. ADC: Apparent diffusion coefficient; SR: Signal ratio; CD: Crohn’s disease; SD: Standard deviation; R1S1: First session of reader 1; R1S2: Second session of reader 1; R2S1: First session of reader 2; R2S2: Second session of reader 2.
analysis on DW-MRE images were found to be good or excellent in this initial study. Both the ADC and SR demonstrated high diagnostic yield for distinguishing inflamed terminal ileum in CD patients from the normal. Therefore, ADC value or SR could serve as robust and efficient biomarkers to aid in the noninvasive evaluation of CD.

ACKNOWLEDGEMENTS

The authors thank Jian Han from our gastrointestinal medical center for analysis of endoscopy examinations, whose important contributions to this study are indispensable to its success.
Table 4  Diagnostic performance of apparent diffusion coefficient and signal ratio for differentiating inflammatory lesions from normal terminal ileum

|                     | CD (n = 21) | Normal (n = 23) | P-value   | Cut-off value | AUROC | Sensitivity (%) | Specificity (%) |
|---------------------|-------------|-----------------|-----------|---------------|-------|----------------|-----------------|
| First reader        |             |                 |           |               |       |                |                 |
| ADC value           | 1.52 ± 0.42 | 2.54 ± 0.49     | < 0.001*  | 1.80          | 0.923 | 85.7           | 95.7            |
| SR value            | 4.19 ± 1.39 | 1.44 ± 0.53     | < 0.001*  | 2.24          | 0.981 | 90.5           | 95.7            |
| Second reader       |             |                 |           |               |       |                |                 |
| ADC value           | 1.54 ± 0.46 | 2.61 ± 0.48     | < 0.001*  | 1.81          | 0.929 | 85.7           | 95.7            |
| SR value            | 4.22 ± 1.37 | 1.45 ± 0.53     | < 0.001*  | 2.07          | 0.988 | 95.2           | 95.7            |

*P < 0.05 was considered statistically significant. Apparent diffusion coefficient (ADC) and signal ratio parameters are expressed as the mean ± SD, ADC units: ×10^{-3} mm²/s. ADC: Apparent diffusion coefficient; SR: Signal ratio; CD: Crohn’s disease; AUROC: Areas under receiver operating characteristic curve.

Figure 6  Receiver operating characteristic curve for quantitative apparent diffusion coefficient and signal ratio of the two readers. ADC: Apparent diffusion coefficient; SR: Signal ratio; R1: Reader 1; R2: Reader 2.

ARTICLE HIGHLIGHTS

Research background
The diagnosis and monitoring of Crohn’s disease (CD) are important for the clinical therapy during its longstanding course. As avoiding the use of intravenous contrast agents, diffusion weighted (DW) imaging (DWI) is increasingly considered as a promising magnetic resonance (MR) enterography (MRE) technique to the qualitative and quantitative evaluation of CD in recent years, and is recommended as an optional MR sequence for CD. The quantitative DWI analysis may be more objective for CD assessment. However, the measurement reproducibility of quantitative DWI analysis in MRE has not been ascertained so far. Thus, the investigation of intra and interobserver agreements of quantitative DWI parameters could be valuable to facilitate the application of quantitative DW-MRE into the clinical routine.

Research motivation
DWI has shown a superb potential as a quantitative modality for CD evaluation. The characterization of reader agreement and variability with quantitative DWI parameters is imperative before such parameters are to be incorporated into clinical practice. Therefore, it is necessary to investigate the level of measurement reproducibility of quantitative DWI parameters.

Research objectives
This study aimed to explore intra and interobserver reproducibility of quantitative analysis for DW-MRE in CD, and then to assess the diagnosis efficiency of quantitative parameters for the evaluation of CD.

Research methods
Forty-four subjects (21 diagnosed with CD and 23 normal control subjects) who underwent ileocolonoscopy and DW-MRE (b = 800 s/mm²) within one week were included. Two radiologists, blinded to clinical details as well as endoscopy results, measured the signal intensity (SI) of the terminal ileum and ipsilateral psoas muscle on DWI with b = 800 s/mm². Separately and simultaneously, apparent diffusion coefficients (ADC) of the terminal ileum were
recorded. Then, SI of ipsilateral psoas muscle on DWI with b = 800 s/mm² was also measured by placing a larger oval ROI. Signal intensity ratio (SR) of the terminal ileum to ipsilateral psoas muscle was calculated according to the equation: SR = SI terminal / SI muscle. After a washout period of 4 weeks, the same two radiologists repeated the above described measurements to assess interobserver agreement. Between- and within-reader agreements were assessed using intraclass correlation coefficients (ICC), coefficients of variation (CoV), and 95% limits of agreement of Bland-Altman plots (BA-LA LoA). Diagnostic performances of ADC and SR for identifying inflamed terminal ileum from the normal were evaluated by receiver operating characteristic (ROC) curve analysis.

Research results

There were no significant differences in quantitative DWI parameters (ADC and SR) between the two sessions or between the two radiologists either in the CD or control group (paired t-test, P > 0.05). The intra and interobserver reproducibility of ADC (ICC: 0.952-0.984; CoV: 3.73%-6.28%; BA-LA LoA: ±11.27% to ±15.88%) and SR (ICC: 0.969-0.989; CoV: 3.51%-4.64%; BA-LA LoA: ±10.62% to ±15.45%) was excellent for CD. Agreement of ADC measurements was slightly less in control subjects (ICC: 0.641-0.736; CoV: 10.47%-11.43%; BA-LA LoA: ±26.59% to ±30.83%). SR of the normal terminal ileum demonstrated high intra and interobserver reproducibility (ICC: 0.944-0.974; CoV: 3.5%-6.28%; BA-LA LoA: ±18.58% to ±24.45%). The areas under the ROC curve for SR (reader 1: 0.981; reader 2: 0.988) were slightly better than those of ADC (reader 1: 0.923; reader 2: 0.929), but these differences were not statistically significant (P-values: 0.117 for reader 1; 0.067 for reader 2). The cutoff values maximizing sensitivity and specificity for ADC and SR parameters were approximately 1.80 × 10⁻³ mm²/s and 2.00, respectively. In addition, SR was moderately and negatively correlated with ADC values in CD patients for both readers (reader 1: r = -0.438, P = 0.047; reader 2: r = -0.485, P = 0.026).

Research conclusions

From this study, the intra and interobserver agreements of quantitative ADC and SR analysis from DW-MRE images were found to be good or excellent. Both the ADC and SR demonstrated high diagnostic yield for distinguishing inflamed terminal ileum in CD patients from the normal. Additionally, SR demonstrated moderately negative correlations with ADC in CD patients, and could also represent the diffusion condition of the terminal ileum. Therefore, ADC value or SR could serve as robust and efficient biomarkers to aid in the noninvasive evaluation of CD.

Research perspectives

Quantitative ADC and SR values derived from DW-MRE achieved good to excellent intra and interobserver agreements with high diagnostic accuracy based on this initial study. The further exploration of these quantitative parameters in the differentiation of CD stages and medical therapy assessment during CD course is worthwhile in the future study.

REFERENCES

1. Torres J, Mehndiratta S, Colombel JF, Peyrin-Biroulet L. Crohn’s disease. Lancet 2017; 389: 1741-1755 [PMID: 27014655 DOI: 10.1016/S0140-6736(16)31711-1]
2. Panes J, Jairath V, Levesque BG. Advances in Use of Endoscopy, Radiology, and Biomarkers to Monitor Inflammatory Bowel Diseases. Gastroenterology 2017; 152: 362-373.e3 [PMID: 27751880 DOI: 10.1053/j.gastro.2016.10.005]
3. Kilcoyne A, Kaplan JL, Gee MS. Inflammatory bowel disease imaging: Current practice and future directions. World J Gastroenterol 2016; 22: 917-932 [PMID: 26811637 DOI: 10.3748/wjg.v22.i3.917]
4. Papay P, Iglaviovic A, Karmiris K, Amarante H, Millheller P, Feagan B, D’Haens G, Marteau P, Reinisch W, Sturm A, Steinwuz F, Egan L, Panés J, Louis E, Colombel JF, Panaccione R. Optimising monitoring in the management of Crohn’s disease: a physician’s perspective. J Crohns Colitis 2013; 7: 653-669 [PMID: 23626762 DOI: 10.1016/j.crohns.2013.02.005]
5. Dohan A, Taylor S, Hoeffff C, Barret M, Allez M, Dautry R, Zappa M, Savoye-Collet C, Dray X, Boudiaf M, Reinhold C, Soyer P. Diffusion-weighted MRI in Crohn’s disease: Current status and recommendations. J Magn Reson Imaging 2016; 44: 1381-1396 [PMID: 27249184 DOI: 10.1002/jmri.25322]
6. Seo N, Park SH, Kim KJ, Kang BK, Lee Y, Yang SK, Ye BD, Park SH, Kim SY, Baek S, Han K, Ha HK. 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 2016; 278: 762-772 [PMID: 26348103 DOI: 10.1148/radiol.2015150809]
7. Baillef P, Caudot G, Goutte M, Goutoube F, Briciu H, Hoeff C, Allamnt C, Reymond M, Obriri-Guilhen H, Maguin B, Bommelaer G, Pereira B, Hordonneau C, Buisson A. Faecal calprotectin and magnetic resonance imaging in detecting Crohn’s disease: histological and endoscopic endoscopic recurrences. World J Gastroenterol 2018; 24: 641-650 [PMID: 29434453 DOI: 10.3748/wjg.v24.i5.641]
8. Park SH. DWI at MR Enterography for Evaluating Bowel Inflammation in Crohn Disease. AJR Am J Roentgenol 2016; 207: 40-48 [PMID: 26993362 DOI: 10.2214/AJR.15.13862]
9. Li XH, Sun CH, Mao P, Huang SY, Zhang ZW, Yang XF, Huang L, Lin J, Zhang J, Ben-Horin S, Feng CH, Men H, Li ZP. Diffusion-weighted MRI Enables to Accurately Grade Inflammatory Activity in Patients of Ileocolonic Crohn’s Disease: Results from an Observational Study. J Crohns Colitis 2018; 12: 24-243 [PMID: 28079618 DOI: 10.1097/MIB.0000000000001001]
10. Tielbeek JA, Ziech ML, Li Z, Lavini C, Bipat S, Bemelman WA, Roelofs JJ, Ponsioen CY, Vos FM, Steker J. Evaluation of conventional, dynamic contrast enhanced and diffusion weighted MRI for quantitative Crohn’s disease assessment with histopathology of surgical specimens. Eur Radiol 2014; 24: 619-629 [PMID: 24037299 DOI: 10.1007/s00330-013-3015-7]
11. Zou X, Luo Y, Li Z, Hu Y, Li H, Tang H, Shen Y, Hu D, Kamel IR. Volumetric Apparent Diffusion
Coefficient Histogram Analysis in Differentiating Intrahepatic Mass-Forming Cholangiocarcinoma From Hepatocellular Carcinoma. J Magn Reson Imaging 2019; 49: 975-983 [PMID: 30276726 DOI: 10.1002/jmri.26253]

Buisson A, Hordonneau C, Coutue M, Boyer L, Pereira B, Bommelaer G. Diffusion-weighted magnetic resonance imaging is effective to detect ileocolonic ulcerations in Crohn's disease. Aliment Pharmacol Ther 2015; 42: 452-460 [PMID: 26088810 DOI: 10.1111/apt.13287]

Taylor SA, Avni F, Cronin CG, Hoefel C, Kim SH, Laghi A, Napoliitano M, Petit P, Rimola J, Tolan DJ, Yorkzad MR, Zappar M, Bhattragar G, Puslaert CAJ, Stoker J. The first joint ESGAR/ESPR consensus statement on the technical performance of cross-sectional small bowel magnetic resonance imaging. Eur Radiol 2017; 27: 2570-2582 [PMID: 27775721 DOI: 10.1007/s00330-016-4615-9]

Sirin S, Kathemann S, Schweiger B, Hahnemenn ML, Forsting M, Lauenstein TC, Kinner S. Magnetic resonance colonography including diffusion-weighted imaging in children and adolescents with inflammatory bowel disease: do we really need intravenous contrast? Invest Radiol 2015; 50: 32-39 [PMID: 25215934 DOI: 10.1097/RLI.0000000000000097]

Bruining DH, Zimmermann EM, Loftus EV, Sandborn WJ, Sauer CG, Strong SA; Society of Abdominal Radiology Crohn’s Disease-Focused Panel. Consensus Recommendations for Evaluation, Interpretation, and Utilization of Computed Tomography and Magnetic Resonance Enterography in Patients With Small Bowel Crohn's Disease. Radiology 2018; 286: 776-799 [PMID: 29319414 DOI: 10.1148/radiol.2018171737]

Bruining DH, Zimmermann EM, Loftus EV, Sandborn WJ, Sauer CG, Strong SA; Society of Abdominal Radiology Crohn’s Disease-Focused Panel. Consensus Recommendations for Evaluation, Interpretation, and Utilization of Computed Tomography and Magnetic Resonance Enterography in Patients With Small Bowel Crohn's Disease. Gastroenterology 2018; 154: 1172-1194 [PMID: 29329905 DOI: 10.1053/j.gastro.2017.11.274]

Hub J, Kim KJ, Park SH, Park SH, Yang SK, Ye BD, Park SH, Han K, Kim AY. Diffusion-Weighted MR Enterography to Monitor Bowel Inflammation after Medical Therapy in Crohn's Disease: A Prospective Longitudinal Study. Korean J Radiol 2017; 18: 162-172 [PMID: 28696720 DOI: 10.3348/kjr.2017.18.1.162]

Kim SJ, Lee JM, Kim H, Yoon JH, Han JK, Choi BI. Role of diffusion-weighted magnetic resonance imaging in the diagnosis of gallbladder cancer. J Magn Reson Imaging 2013; 38: 127-137 [PMID: 23281048 DOI: 10.1002/jmri.23956]

Priola AM, Priola SM, Gned D, Girrado MT, Brunda M, Righi L, Veltri A. Diffusion-weighted quantitative MRI of pleural abnormalities: Intra- and interobserver variability in the apparent diffusion coefficient measurements. J Magn Reson Imaging 2017; 46: 769-782 [PMID: 28117923 DOI: 10.1002/jmri.25633]

Clauser P, Marcon M, Maier M, Zuiani C, Bazzocchi M, Baltzer PA. Is there a systematic bias of apparent diffusion coefficient (ADC) measurements of the breast if measured on different workstations? An inter- and intra-reader agreement study. Eur Radiol 2016; 26: 2291-2296 [PMID: 24463404 DOI: 10.1007/s00330-015-4051-2]

Krajewski KM, Nishino M, Franchetti Y, Ramiaya NH, Van den Abbeele AD, Choueiri TK. Intraobserver and interobserver variability in computed tomography size and attenuation measurements in patients with renal cell carcinoma receiving antiangiogenic therapy: implications for alternative response criteria. Cancer 2014; 120: 711-721 [PMID: 24264883 DOI: 10.1002/cncr.28493]

Van Asche G, Dignass A, Panes J, Beaugerie L, Karagiannis J, Allez M, Ochsenkühn T, Orchard T, Rogler G, Louis E, Kupcinets L, Manzaris G, Travis S, Stange E; European Crohn’s and Colitis Organisation (ECCO). The second European evidence-based Consensus on the diagnosis and management of Crohn’s disease: Definitions and diagnosis. J Crohns Colitis 2010; 4: 7-27 [PMID: 21112488 DOI: 10.1016/j.jcjo.2009.12.003]

Hordonneau C, Buisson A, Scanzi J, Goutorbe F, Pereira B, Borderon C, Da Ines D, Montoriol PF, Garcia JM, Boyer L, Bommelaer G, Petillon V. Diffusion-weighted magnetic resonance imaging in ileocolonic Crohn’s disease: validation of quantitative index of activity. Am J Gastroenterol 2014; 109: 89-88 [PMID: 24257212 DOI: 10.1038/ajg.2013.357]

Cai PQ, Wu YP, An X, Qiu X, Kong LH, Liu GC, Xie CM, Pan ZZ, Wu PH, Ding PR. Simple measurements on diffusion-weighted MR imaging for assessment of complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. Eur Radiol 2014; 24: 2962-2970 [PMID: 25088851 DOI: 10.1007/s00330-014-3251-5]

Mottola JC, Sahni VA, Erturk SM, Swanson R, Banks PA, Morteke KJ. Diffusion-weighted MRI of focal cystic pancreatic lesions at 3.0-Tesla: preliminary results. Abdom Imaging 2012; 37: 110-117 [PMID: 21512724 DOI: 10.1007/s00261-011-9373-6]

Buisson A, Joubert A, Montoriol PF, Da Ines D, Hordonneau C, Pereira B, Garcia JM, Bommelaer G, Petillon V. Diffusion-weighted magnetic resonance imaging for detecting and assessing ileal inflammation in Crohn's disease. Aliment Pharmacol Ther 2013; 37: 537-545 [PMID: 23289713 DOI: 10.1111/app.12201]

Choil SH, Kim KW, Lee JY, Kim KJ, Park SH. Diffusion-weighted Magnetic Resonance Enterography for Evaluating Bowel Inflammation in Crohn’s Disease: A Systematic Review and Meta-analysis. Inflamm Bowel Dis 2016; 22: 669-679 [PMID: 26457380 DOI: 10.1097/MIB.0000000000000567]

Sato H, Tamura C, Narimatsu K, Shimizu M, Takajyo T, Yamashita M, Inoue Y, Ozaki H, Furushashi H, Moruta K, Yasutake Y, Yoshikawa K, Watanabe C, Komoto S, Tornta K, Nagao S, Miura S, Shinmoto H, Hokari R. Magnetic resonance enterocolonography in detecting erosion and redness in intestinal mucosa of patients with Crohn's disease. J Gastroenterol Hepatol 2015; 30: 667-673 [PMID: 25387770 DOI: 10.1111/jgh.12851]

Stanesca-Sigmund N, Nimsch Y, Wunderlich AP, Wagner M, Meier R, Juchems MS, Beer M, Schmidt SA. Quantification of inflammatory activity in patients with Crohn's disease using diffusion weighted imaging (DWI) in MR enteroclysis and MR enterography. Acta Radiol 2017; 58: 264-271 [PMID: 27178931 DOI: 10.1177/0284185116648030]

Rees MA, Dillman JR, Anton CG, Rattan MS, Smith EA, Towbin AJ, Zhang B, Trout AT. Inter-radiologist agreement using Society of Abdominal Radiology-American Gastroenterological Association (SAR-AGA) consensus nomenclature for reporting CT and MR enterography in children and young adults with small bowel Crohn disease. Abdom Radiol (NT) 2019; 44: 391-397 [PMID: 30120514 DOI: 10.1007/s00261-018-1743-5]

Kurugol S, Freeman M, Alacan O, Domachevsky L, Perez-Rossello JM, Callahan MJ, Warfield SK.
Motion-robust parameter estimation in abdominal diffusion-weighted MRI by simultaneous image registration and model estimation. *Med Image Anal* 2017; 39: 124-132 [PMID: 28494271 DOI: 10.1016/j.medfae.2017.04.006]
