

Long-Term Performance of Readers Trained in Grading Crohn Disease Activity Using MRI

Carl A. J. Puylaert, MSc, Jeroen A. W. Tielbeek, MD, PhD, Shandra Bipat, PhD, Thierry N. Boellaard, MD, PhD, C. Yung Nio, MD, Jaap Stoker, MD, PhD

Rationale and Objectives: We aim to evaluate the long-term performance of readers who had participated in previous magnetic resonance imaging (MRI) reader training in grading Crohn disease activity.

Materials and Methods: Fourteen readers (8 women; 12 radiologists, 2 residents; mean age 40; range 31–59), who had participated in a previous MRI reader training, participated in a follow-up evaluation after a mean interval of 29 months (range 25–34 months). Follow-up evaluation comprised 25 MRI cases of suspected or known Crohn disease patients with direct feedback; cases were identical to the evaluation set used in the initial reader training (of which readers were unaware). Grading accuracy, overstaging, and understaging were compared between training and follow-up using a consensus score by two experienced abdominal radiologists as the reference standard.

Results: In the follow-up evaluation, overall grading accuracy was 73% (95% confidence interval [CI]: 62%–81%), which was comparable to reader training grading accuracy (72%, 95% CI: 61%–80%) ($P = .66$). Overstaging decreased significantly from 19% (95% CI: 12%–27%) to 13% (95% CI: 8%–21%) between training and follow-up ($P = .03$), whereas understaging increased significantly from 9% (95% CI: 4%–21%) to 14% (95% CI: 7%–26%) ($P < .01$).

Conclusions: Readers have consistent long-term accuracy for grading Crohn disease activity after case-based reader training with direct feedback.

Key Words: Crohn disease; magnetic resonance imaging; education; medical; graduate; teaching; follow-up studies.

© 2016 The Association of University Radiologists. Published by Elsevier Inc. All rights reserved.

INTRODUCTION

Magnetic resonance imaging (MRI) is the preferred technique for assessment of disease activity in Crohn disease because of its capability to evaluate intra- and extramural disease of both small bowel and colon, without the use of ionizing radiation. A previous study showed that inexperienced radiologists and residents could be successfully trained to grade Crohn disease activity using case-based training (1). In that study, readers graded 100 MRI cases of patients with suspected or known Crohn disease, with the readers receiving direct feedback after each case. Significant improvement of grading accuracy was seen (66% to 75%, $P = .003$), whereas understaging decreased significantly (15% to 7%, $P < .001$).

Although case-based reader training can provide short-term improvement, to our knowledge, no studies have investigated long-term performance, as reader evaluation was

always performed at the time of training. Furthermore, there is no knowledge of how daily practice impacts reader's level of experience after training.

The primary objective of this study was to compare readers' results in grading Crohn disease activity between reader training and a follow-up evaluation of the same set of MRI cases. The secondary objective was to determine the influence of readers' interim experience on grading results at follow-up evaluation.

MATERIALS AND METHODS

Thirty-one readers from different hospitals, who had previously participated in a reader training at a single tertiary center (1), were invited for a follow-up evaluation. Before the start of initial reader training, readers had no ($n = 14$) or limited ($n = 17$) MRI experience (<25 magnetic resonance enterography/enteroclysis examinations) for evaluating Crohn disease. During the initial training readers had graded 100 MRI cases of patients with suspected or known Crohn disease, with the readers receiving direct feedback after each case (ie, expert radiologic report and endoscopic or surgical findings). Before reading those 100 cases, the readers had graded 25 MRI cases without direct feedback, serving as a baseline. These 25

Acad Radiol 2016; ■■■■-■■■

From the Department of Radiology, Academic Medical Center, University of Amsterdam, G1-229, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Received May 12, 2016; accepted August 1, 2016. **Address correspondence to:** C.A.J.P. e-mail: c.a.puylaert@amc.uva.nl

© 2016 The Association of University Radiologists. Published by Elsevier Inc. All rights reserved.
<http://dx.doi.org/10.1016/j.acra.2016.08.006>

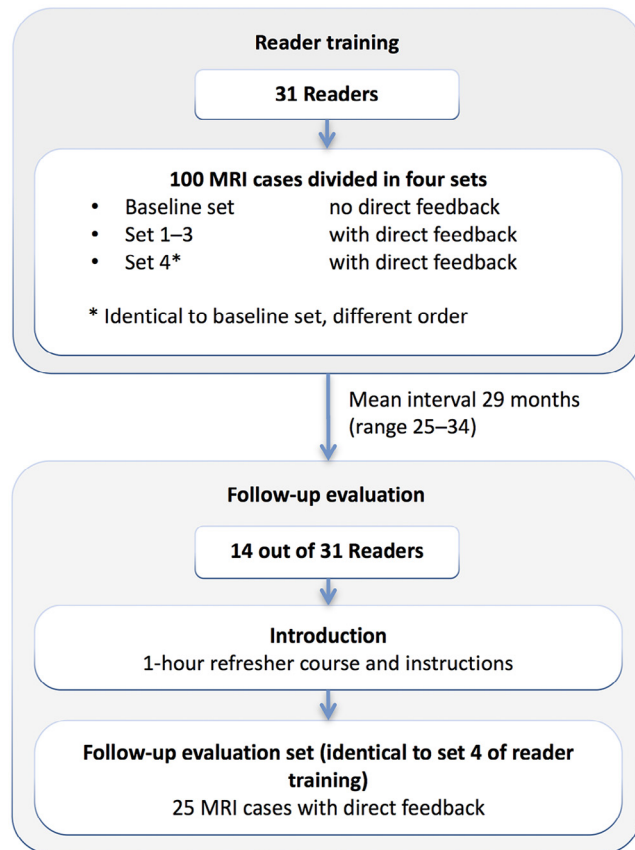


Figure 1. Flowchart of reader training and follow-up evaluation.

baseline cases were reassessed as the last 25 cases of the aforementioned 100 cases but in a differently randomized order; this time the readers received feedback case by case. This setup facilitated comparing grading accuracy, understaging, and overstaging before and after the case-by-case reader training.

Fourteen of these 31 readers from 13 hospitals attended the follow-up evaluation and were included in the present study. The mean interval between reader training and the follow-up evaluation was 29 months (range: 25–34). Readers were asked to describe their interim experience with magnetic resonance enterography/enteroclysis and abdominal MRI examinations between reader training and follow-up evaluation (described as the average number of examinations seen per month).

Follow-Up Evaluation

During follow-up evaluation, readers evaluated 25 MRI cases (16 enterography and 9 enteroclysis), which were identical to the baseline set and the reassessed last 25 cases from the reader training, with direct feedback after each case (Fig 1). The readers were not informed that these were identical cases from the previous training session. Patient and MRI characteristics for these cases are presented in Table 1. All examinations were performed at a 1.5 T unit (Avanto, Siemens Healthcare, Erlangen, Germany) or a 3.0 T unit (Philips Medical

TABLE 1. Patient and MRI Characteristics

Parameter		
Gender	Male or female	13/12
Age at time of imaging	Years (mean/SD)	37 (13)
Previous surgery	<i>n</i> (%)	11 (44%)
MRI technique	1.5 T	20
	3.0 T	5
Preparation	Enterography	16
	Enteroclysis	9
Disease activity of the patients per examination (reference standard)	None	7
	Mild	6
	Moderate	5
	Severe	7

MRI, magnetic resonance imaging; SD, standard deviation.

Systems, Best, The Netherlands) in the supine position, and the scan protocol included at least a coronal and axial T2-weighted sequence, a coronal fat-saturated T1-weighted unenhanced sequence, and coronal and axial fat-saturated postcontrast T1-weighted sequences (1). Follow-up evaluation started with a 1-hour refresher course on how to use the grading system (including two instruction cases) and technical instructions for our picture archiving and communication system. Subsequently, readers individually evaluated all 25 MRI cases on using the same online scoring form used in reader training to grade disease activity (<http://mrentero.webklik.nl>). This scoring system is a modification of a validated qualitative scoring system, with the addition of pattern of enhancement, comb sign, disease length, and assessment of extra-enteric complications (2) (Table 2). As an extension to this scoring system, an arbitrary classification was used to classify patients in four categories: none, mild, moderate, or severe disease (Table 3). This modified scoring system was used, as clinically important imaging findings (eg, fistula, abscess) are not integrated in the published scoring systems. The following MRI features were recorded for each patient for the most severe lesion: mural thickness, mural T2 signal, T1 enhancement, mural enhancement pattern, total length of disease, and presence of comb sign (vascular enlargement of the vasa recta). The following complications were described to be present: infiltrate (tethering and kinking of bowel loops), abscess, fistula, and severe stenosis (defined as 80% lumen reduction with prestenotic dilatation and a moderate-to-severe increase in mural T2 signal) (Fig 2). Additionally, readers were asked to grade their confidence for severity grading (0–10). The scoring system and classification system were identical to the initial reader training. Surgical history was provided for each case if applicable. Cases were reviewed in the same order as during the initial case-by-case reader training, and direct feedback was given after each case.

Reference Standard

All MRI examinations had been scored by two abdominal radiologists (CYN and JS), with 18 and 21 years of experience

Download English Version:

<https://daneshyari.com/en/article/5725768>

Download Persian Version:

<https://daneshyari.com/article/5725768>

[Daneshyari.com](https://daneshyari.com)