Impact of 3 Tesla MRI on interobserver agreement in clinically isolated syndrome

A MAGNIMS multicentre study

For the MAGNIMS Study Group; Barkhof, Frederik; Ciccarelli, Olga; Yousry, Tarek; Fredriksen, Jette Lautrup; Rovira, Alex; Sastre-Garriga, Jaume; Vrenken, Hugo

Published in:
Multiple Sclerosis Journal

DOI:
10.1177/1352458517751647

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

MULTIPLE SCLEROSIS JOURNAL | MSJ

Original Research Paper

Impact of 3 Tesla MRI on interobserver agreement in clinically isolated syndrome: A MAGNIMS multicentre study

Marloes HJ Hagens, Jessica Burggraaff, Iris D Kilsdonk, Serena Ruggieri, Sara Collorone, Rosa Cortese, Niamh Cawley, Emilia Sbardella, Michaela Andelova, Michael Amann, Johanna M Lieb, Patrizia Pantano, Birgit I Lissenberg-Witte, Joep Killestein, Celia Oreja-Guevara, Jens Wuerfel, Olga Ciccarelli, Claudio Gasperini, Carsten Lukas, Alex Rovira, Frederik Barkhof and Mike P Wattjes; For the MAGNIMS Study Group

Multiple Sclerosis Journal

2019, Vol. 25(3) 352–360

DOI: 10.1177/1352458518751647

Abstract

Background: Compared to 1.5 T, 3 T magnetic resonance imaging (MRI) increases signal-to-noise ratio leading to improved image quality. However, its clinical relevance in clinically isolated syndrome suggestive of multiple sclerosis remains uncertain.

Objectives: The purpose of this study was to investigate how 3 T MRI affects the agreement between raters on lesion detection and diagnosis.

Methods: We selected 30 patients and 10 healthy controls from our ongoing prospective multicentre cohort. All subjects received baseline 1.5 and 3 T brain and spinal cord MRI. Patients also received follow-up brain MRI at 3–6 months. Four experienced neuroradiologists and four less-experienced raters scored the number of lesions per anatomical region and determined dissemination in space and time (McDonald 2010).

Results: In controls, the mean number of lesions per rater was 0.16 at 1.5 T and 0.38 at 3 T ($p = 0.005$). For patients, this was 4.18 and 4.40, respectively ($p = 0.657$). Inter-rater agreement on involvement per anatomical region and dissemination in space and time was moderate to good for both field strengths. 3 T slightly improved agreement between experienced raters, but slightly decreased agreement between less-experienced raters.

Conclusion: Overall, the interobserver agreement was moderate to good. 3 T appears to improve the reading for experienced readers, underlining the benefit of additional training.

Keywords: Multiple sclerosis, clinically isolated syndrome, magnetic resonance imaging, interobserver variation, multicentre study

Date received: 3 July 2017; revised: 20 October 2017; accepted: 29 November 2017

Introduction

Magnetic resonance imaging (MRI) plays a pivotal role in the diagnosis and monitoring of multiple sclerosis (MS).[1,2] After clinically isolated syndrome (CIS), which is commonly the first manifestation of MS, 56%–82% of patients with brain MRI abnormalities will develop clinically definite MS within the next 20 years.[3,4] For patients with a normal brain MRI, this is much lower, approximately 20%.[3,4] An early accurate diagnosis is highly relevant in clinical decision making, such as initiation of disease-modifying therapy in early stage of the disease. Moreover, precise lesion detection is important in identifying patients with an increased risk of long-term disability, mainly patients with a high lesion load, gadolinium enhancing lesions and infratentorial lesions.[5–7] In addition, adequate monitoring of CIS and MS patients requires an accurate detection of new lesions.[1,2]

Correspondence to:
MHJ Hagens
Department of Neurology, MS Center Amsterdam, VU University Medical Center, De Boelelaan 1117 (ZH 0A 65), 1081 HV Amsterdam, The Netherlands.
m.hagens1@vumc.nl

Marloes Hendrika Johanna Hagens
Jessica Burggraaff
Joep Killestein
Department of Neurology, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands

Iris D Kilsdonk
Department of Radiology and Nuclear Medicine, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands/Department of Radiology and Nuclear Medicine, Onze Lieve Vrouwen Gasthuis, Amsterdam, The Netherlands

Serena Ruggieri
Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy/Department of Neurosciences, San Camillo-Forlanini Hospital, Rome, Italy

Sara Collorone
Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy/NMR Research Unit, Queen Square Multiple Sclerosis Centre, UCL Institute of Neurology, London, UK

Rosa Cortese
NMR Research Unit, Queen Square Multiple Sclerosis Centre, UCL Institute of Neurology, London, UK; Department of Basic Medical Sciences, Neurosciences and Sense Organs, University of Bari, Bari, Italy

The current McDonald 2010 diagnostic criteria for MS do not define MRI acquisition parameters such as magnetic field strength, spatial resolution and the selection of pulse sequences.[8] Mainly due to the improved signal-to noise ratio leading to an improvement of image quality, brain imaging at higher magnetic field strengths offers new possibilities with respect to the diagnosis and follow-up of neuroinflammatory disease.[9–11] Current expert panel guidelines recommend 3 T brain imaging,[1,2] as the improved signal-to-noise ratio results in an increased lesion detection in anatomical regions relevant for dissemination in space (DIS), especially in the (juxta)cortical, periventricular and infratentorial region.[12,13] However, the clinical relevance of high field strength MRI is uncertain. In particular, the question remains, whether the use of 3 T leads to an earlier diagnosis of MS. A previous prospective single-centre and single-vendor study with 40 CIS patients demonstrated an increased lesion detection on brain scans, but as such this did not lead to an earlier diagnosis of MS according to the McDonald 2005 and Swanton criteria.[14,15] Moreover, when retrospectively applying the 2010 revised McDonald criteria to this dataset, this outcome did not change.[16]

The purpose of this prospective multicentre, multivendor and multi-rater study in patients presenting with a CIS was to evaluate the effect of 3 T MRI on interobserver agreement on lesions detection and subsequently fulfilment of the criteria for DIS and dissemination in time (DIT). Additionally, we evaluated the effect of the raters' experience on the interobserver agreement for both the lesion detection and the McDonald diagnostic criteria.

Materials and methods

This study is part of a MAGNIMS (Magnetic Resonance Imaging in MS, http://www.magnims.eu) prospective multicentre, multi-vendor project conducted at the following MS Centres: VU University Medical Center Amsterdam, University Hospital of Basel, St. Josef Hospital Bochum, UCL Institute of Neurology London, Hospital Clínico San Carlos Madrid and Sapienza University of Rome.

At each centre, the study design was approved by the local institutional review board. Written informed consent was obtained from all participants.

For the CIS patients, two visits were used for this analysis: the baseline visit and the first follow-up 3 to 6 months later (Figure 1). As at this interval, no change on MRI scans is to be expected for healthy controls; only baseline visits were scheduled for the control group.

Recruitment of subjects

Patients with CIS suggestive of MS, as defined by the International Panel on MS diagnosis,[8] were recruited from the outpatient clinics of the six participating centres between July 2013 and September 2015. Patients were recruited within 6 months after the first clinical episode suggestive of demyelination. All subjects were aged 18 to 59 years at baseline. Exclusion criteria were a history of vascular, malignant or other immunological disease and MRI-related contra-indications, such as claustrophobia and a previous allergic reaction to a gadolinium-based contrast agent.

Thirty patients and ten healthy controls were selected for this project. Subjects were randomly selected per site, for the patients based on availability of completed follow-up visits.

Neurological examination

At baseline, a medical history was taken and the Expanded Disability Status Scale (EDSS) was assessed by a trained physician. At follow-up visits, possible new symptoms leading to diagnosis of clinically definite MS were registered and the EDSS assessment was repeated.

MRI acquisition

All patients received baseline MRI scans of the brain and spinal cord at both 1.5 and 3 T separated by 24–72 hours (see Figure 1 for the illustration of the scanning protocol and study design). For both magnetic field strengths, a multisequence scanner optimized acquisition protocol was used (detailed information is given in Supplementary Table 1). In summary, brain imaging included isotropic three-dimensional (3D) T1 and 3D fluid-attenuated inversion recovery (FLAIR), as well as axial 3 mm two-dimensional (2D) T2, proton density (PD) and post-contrast T1 spin-echo (SE) sequences. From the 3D sequences, 3 mm axial reconstructions were made following the same repositioning compared to the 2D sequences. Spinal cord imaging included post-contrast sagittal 3 mm T1 SE and PD/T2. According to the MAGNIMS guidelines on MS diagnosis and monitoring, axial spinal cord imaging was not included due to the substantial increase in scan duration.

In healthy controls, the same protocol was used without the administration of intravenous contrast. For the patients' follow-up, the brain MRI protocol was repeated without the administration of intravenous contrast.

Niamh Cawley
Olga Ciccarelli
NMR Research Unit, Queen Square Multiple Sclerosis Centre, UCL Institute of Neurology, London, UK

Emilia Sbardella
Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy

Michaela Andelova
Department of Neurology, University Hospital Basel, Basel, Switzerland

Michael Amann
Department of Neurology, University Hospital Basel, Basel, Switzerland/Medical Image Analysis Center (MIAC), Basel, Switzerland/Division of Neuroradiology, Department of Radiology, University Hospital Basel, Basel, Switzerland

Johanna M Lieb
Division of Neuroradiology, Department of Radiology, University Hospital Basel, Basel, Switzerland

Patrizia Pantano
Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy/Istituto Neurologico Mediterraneo, Neuromed, Pozzilli, Italy

Birgit I Lissenberg-Witte
Department of Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, The Netherlands

Celia Oreja-Guevara
Department of Neurology, Hospital Clínico San Carlos, Madrid, Spain

Jens Wuerfel
Medical Image Analysis Center (MIAC), Basel, Switzerland/NeuroCure, Charité – Berlin University of Medicine, Berlin, Germany/Department of Biomedical Engineering, University Hospital Basel, Basel, Switzerland

Claudio Gasperini
Department of Neurosciences, San Camillo-Forlanini Hospital, Rome, Italy

Carsten Lukas
Department of Diagnostic and Interventional Radiology and Nuclear Medicine, St. Josef Hospital, Ruhr University, Bochum, Germany

Alex Rovira
Department of Radiology, Hospital Universitari Vall d'Hebron, Barcelona, Spain

Frederik Barkhof
Department of Radiology and Nuclear Medicine, MS Center Amsterdam, VU University Medical Center, Amsterdam, The

Netherlands/Institutes of Neurology and Healthcare Engineering, UCL Institute of Neurology, London, UK

Mike P Wattjes
Department of Radiology and Nuclear Medicine, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands

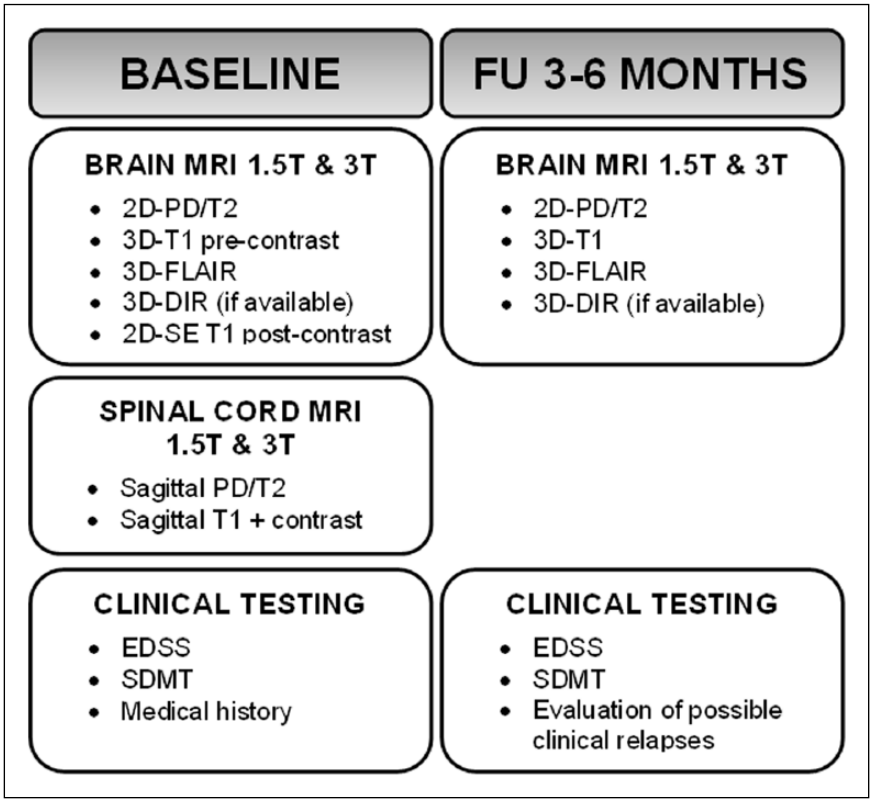

Figure 1. Study protocol.
DIR: double inversion recovery; EDSS: Expanded Disability Status Scale; FLAIR: fluid-attenuated inversion recovery; FU: follow-up; PD: proton density; SDMT: standard digit modalities test; SE: spin-echo; T: Tesla.

Imaging analysis

All scans were centrally collected and checked for completeness. The scans were rated independently by eight raters during a central reading session: four experienced raters (C.L., neuroradiologist for 8 years; A.R., neuroradiologist for 26 years; M.P.W., neuroradiologist for 9 years; F.B., neuroradiologist for 19 years) and four MS researchers or radiology residents considered as less-experienced raters (I.D.K., S.R., S.C., R.C.). For this central reading, the full scan protocol, as described in Figure 1, was available. For each subject, the 1.5 and 3 T scans were presented separately with approximately a 20-hour time interval. The order of presentation was randomized between sessions, but the same for all the eight raters. Localization of symptoms at onset was presented for each patient, as per McDonald 2010 criteria symptomatic brainstem or spinal cord lesions are excluded from demonstration of DIS.[8] Besides location of onset, the raters were blinded for clinical information such as age, gender and centre.

For all baseline scans, the number of inflammatory lesions larger than 3 mm in size were scored and categorized according to the anatomical region (periventricular, juxtacortical, infratentorial and spinal cord). In CIS patients but not in healthy control subjects (no contrast administered), the number of enhancing lesions per region was reported. Subsequently, the presence of DIS and DIT according to the McDonald 2010 criteria was determined. For follow-up scans, new lesions per region were scored and again fulfilment of the criteria for DIS and DIT was determined.

Statistical analysis

The difference in lesion detection between 1.5 and 3 T was tested using generalized estimating equations (GEEs) with a logit link function and an exchangeable correlation structure. Repeated measures for each subject were defined as the scores of the different observers.

Inter-rater agreement on number of lesions detected per region was calculated with Conger's kappa. Agreement on involvement per anatomical region, independent of the number of lesions scored in

Table 1. Demographics of clinically isolated syndrome patients and healthy controls.

Characteristics	Patients ($n=30$)	Controls ($n=10$)
Age, mean ± SD (years)	34.5 ± 7.0	38.7 ± 9.3
Gender, male/female (n)	11/19	2/8
EDSS, median and range	2.0 (0–6)	
Location presenting symptoms (n):		
Optic nerve	12	
Cerebral hemisphere	3	
Infratentorial	4	
Spinal cord	11	

EDSS: Expanded Disability Status Scale; SD: standard deviation.

that region, was calculated with Cohen's kappa. This statistical analysis was also used to determine agreement on the fulfilment of the criteria for DIS, DIT and MS. Values of 0.41 to 0.60 were considered as moderate agreement, 0.61 to 0.80 as substantial agreement and >0.81 as good agreement.[17] Calculations were performed using SPSS 22.0 (Windows) and 'R' version 3.1.1.

Results

Patient characteristics

Detailed demographic information of the study subjects is given in Table 1. The mean age for patients was 34.5 ± 7.0 years, 64% was female. The median EDSS at baseline was 2.0 (range 0–6). Most CIS patients presented with an optic neuritis ($n=12$) or spinal cord syndrome ($n=11$). Patients were scanned with a median of 90 days (interquartile range (IQR) = 29–123) after onset of the symptoms.

In healthy controls, the mean age was 38.7 ± 9.3 years, 80% were female.

Lesion detection and diagnosis

In healthy controls, no spinal cord lesions were scored. The mean total number of brain lesions scored per rater per subject was 0.38 at 3 T (median 0, IQR = 0–0.8) and 0.16 at 1.5 T (median 0, IQR = 0–0) ($p=0.005$). In the patient group, the mean overall number of lesions at baseline was 4.40 at 3 T (median 3, IQR = 1–7) and 4.14 at 1.5 T (median 3, IQR = 1–6) ($p=0.732$), see Figure 2. Only very few enhancing juxtacortical and infratentorial lesions at baseline and new infratentorial lesions at follow-up were identified leading to the exclusion of these regions at these time points from further analyses.

The mean number of cases per rater diagnosed as MS based on radiological criteria was at baseline 1.63 at 1.5 T (median 2, IQR = 1–2) and 2.25 at 3 T (median 2, IQR = 2–2), and at follow-up 4.63 at 1.5 T (median 5, IQR = 3.25–5.75) and 6.38 at 3 T (median 6, IQR = 6–6). Full statistical analysis will be presented based on a consensus score after completion of our ongoing cohort study.

Inter-rater agreement on lesion detection

Inter-rater agreement on involvement per anatomical region for all the raters was moderate to good on both 1.5 and 3 T, with kappa scores (κ) varying from 0.49 to 0.84, see Figure 3. The agreement was highest for baseline infratentorial lesions (3 T: κ 0.84, 1.5 T: κ 0.76) and lowest for baseline juxtacortical lesions (3 T: κ 0.53, 1.5 T: κ 0.49). Agreement on presence of spinal cord lesions was lower at 1.5 T compared to 3 T (3 T: κ 0.76, 1.5 T: κ 0.66). Agreement on enhancing lesions was substantial for periventricular lesions (3 T: κ 0.70, 1.5 T: κ 0.80) and moderate for spinal cord lesions (3 T: κ 0.57, 1.5 T: κ 0.59). Overall, agreement on involvement of regions was higher at baseline compared to follow-up.

As can be expected, inter-rater agreement dropped for the category 'exact number of lesions scored per region', see Figure 3. Agreement on enhancing lesions was not affected, as there was no more than one enhancing lesion in any anatomical region.

When looking at the kappa scores for involvement per anatomical region for the groups by experience, agreement on involvement per anatomical region was overall higher at 3 T for the experienced raters and overall higher at 1.5 T for the less-experienced raters, see Figure 4.

Inter-rater agreement on diagnosis

In CIS patients, the inter-rater agreement for DIS, DIT and diagnosis of MS at baseline was also moderate to good, with κ scores varying from 0.51 to 1.00, see Figure 5. The remarkable κ of 1.00 for DIT at 1.5 T at baseline for both experienced and less-experienced raters is due to full agreement on non-symptomatic

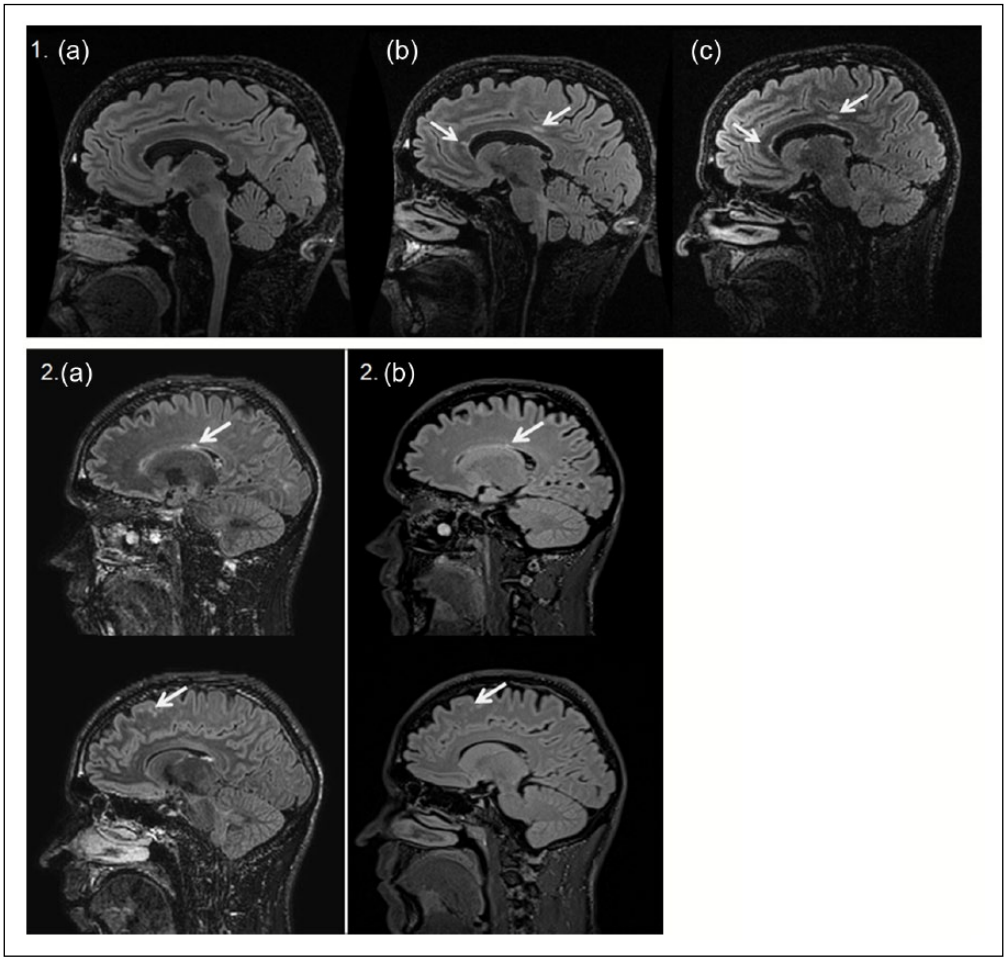

Figure 2. 1.5 and 3 T MRI scans of two CIS patients. 1. 3DFLAIR brain scans of one CIS patient presenting with optic neuritis: (a) baseline scan on 3 T with no brain lesions, (b) follow-up scan on 3 T showing two new T2 lesions in the corpus callosum, (c) follow-up scan on 1.5 T on which only one of the new lesion can be identified. 2. Baseline (a) 3 and (b) 1.5 T 3DFLAIR brain scans of one CIS patient presenting with a spinal cord syndrome. All raters identified additional periventricular and juxtacortical lesions on 3 T MRI leading to dissemination in space, while only three experienced raters on 1.5 T.

enhancing lesions, and therefore DIT, in two patients. At 3 T part of the raters identified a non-symptomatic enhancing lesion in another six patients, leading to a drop in inter-rater agreement on DIT and the diagnosis of MS at 3 T.

At follow-up, 3 T slightly improved the inter-rater agreement for the experienced raters on DIS, DIT and MS, while the agreement between less-experienced raters slightly decreased on all criteria. Overall, the inter-rater agreement on the diagnosis of MS at follow-up was substantial (κ 0.61–0.80) at both field strengths

Discussion

The McDonald criteria for the diagnosis of MS do not define important MRI acquisition parameters such as field strength.[8] However, recent MAGNIMS guidelines recommend the use of 3 T brain MRI based on an improved signal-to-noise ratio resulting in higher lesion detection.[1,2] Nonetheless, to date the extent of and the clinical relevance of a higher detection rate using higher magnetic field strength with respect to diagnostic and prognostic purposes remains unclear. This multicentre, multi-vendor and multi-rater study provides important information on the lesion detection rates and interobserver variation with respect to MS lesion detection for diagnostic purposes in patients presenting with CIS suggestive of MS. Overall, inter-rater agreement on involvement per anatomical region was moderate to good, which was not substantially influenced by field strength. With respect to the lesion location, the agreement was the lowest for juxtacortical lesions at baseline. When comparing this to the

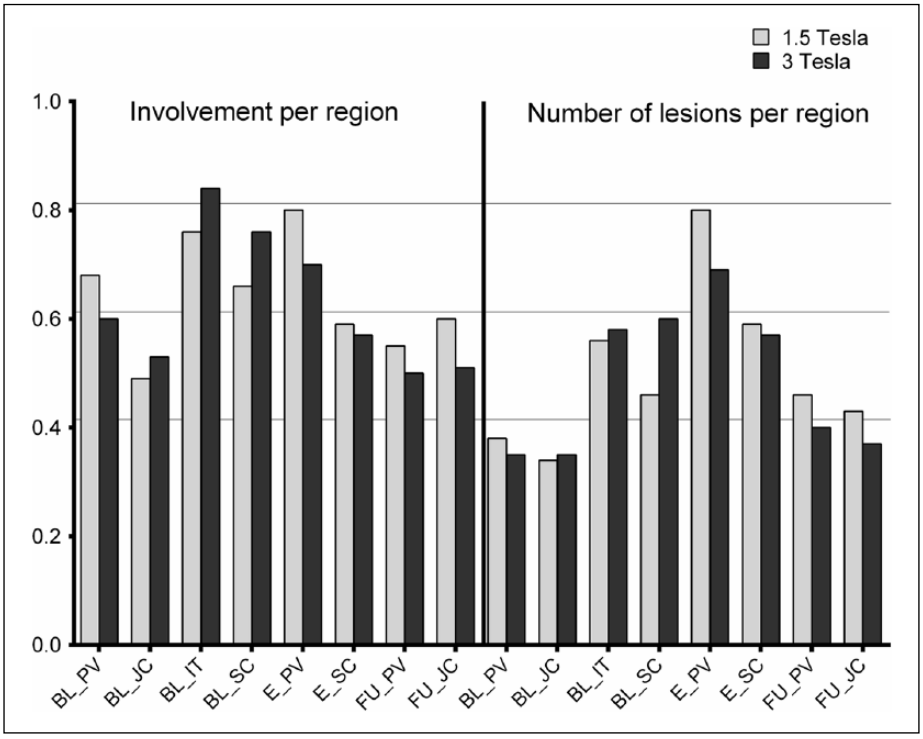

Figure 3. Agreement on lesions per anatomical region per field strength. Agreement between the eight raters on the involvement of an anatomical region, calculated with Cohen's kappa scores, and on the exact number of lesions per anatomical regions, calculated with weighted Conger's kappa scores. The horizontal lines indicate the cut-off values of 0.41 for moderate agreement, 0.61 for substantial agreement and 0.81 for good agreement.
BL: baseline; E: enhancement; FU: follow-up; IT: infratentorial; JC: juxtacortical; PV: periventricular; SC: spinal cord.

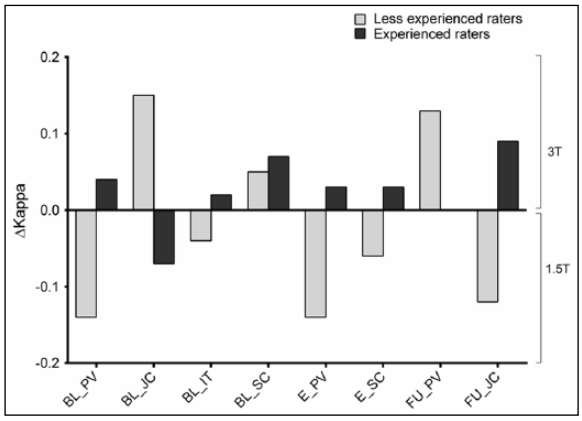

Figure 4. Effect of experience on agreement on involvement per anatomical region per field strength. Calculated by subtracting Cohen's kappa for 3 T by Cohen's kappa for 1.5 T.
BL: baseline; E: enhancement; FU: follow-up; IT: infratentorial; JC: Juxtacortical; PV: periventricular; SC: spinal cord; T: Tesla.

agreement on the exact number of lesions per region, the largest decrease in agreement was understandably in the periventricular region, as this is the region where most lesions were identified.

In contrast to a previous single-centre and single-vendor study,[15] we used 3D brain imaging with 3-mm-thick axial reconstructions on both field strengths. Moreover, we also studied spinal cord imaging at both field strengths. Previous studies have shown that the identification of a spinal cord lesion does not only facilitate the fulfilment of the MRI criteria for diagnosis of MS, but is also predictive for conversion to clinically definite MS in CIS patients.[18,19] However, spinal cord MRI is challenging – especially at 3 T – due to various possible artefacts due to patient motion, swallowing, respiration and pulsation of the cerebrospinal fluid and blood vessels.[20] In addition, it has not conclusively been demonstrated that 3 T leads to higher lesion detection levels compared to lower field strength.[21] Contrary to this, agreement on spinal cord lesions was highest at 3 T for both the experienced and less-experienced raters.

When demonstrating the effect of the experience of the raters on the variability of lesion detection, overall the inter-rater agreement for the less-experienced raters is higher for the 1.5 T scans, while the more experienced raters agree more at 3 T. This

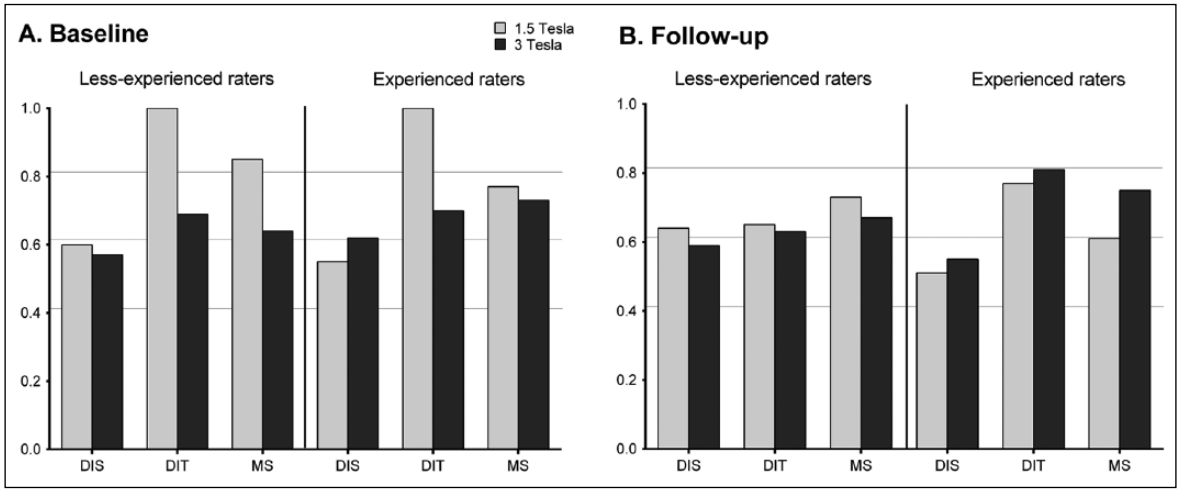

Figure 5. Agreement on the diagnosis per field strength dependent on experience of the raters. Calculated using Cohen's kappa scores. The horizontal lines indicate the cut-off values 0.41 for moderate agreement, 0.61 for substantial agreement and 0.81 for good agreement.
DIS: dissemination in space; DIT: dissemination in time; MS: diagnosis of multiple sclerosis.

could be explained by an effect of training. Most probably, a correct interpretation of high field strength MRI requires more experience as smaller details become visible, including more incidental lesions in healthy controls.

Even though all eight raters were well familiar with the McDonald 2010 criteria, applying these criteria consistently to all the scans appeared to be more challenging than anticipated. A good working knowledge of these complex criteria was not without doubt even for the experienced neuroradiologists. The difficulty of applying the diagnostic criteria for MS has previously also been demonstrated when using the McDonald 2001 criteria.[22] For the 2010 revision of the diagnostic criteria, most questions arose on how to exclude the symptomatic brainstem and spinal cord lesions in the criteria for DIS. In the current criteria, symptomatic lesions localized in the brainstem or spinal cord are to be excluded from lesion count. However, it is unclear as to whether only the one symptomatic lesion or all the lesions in the symptomatic area should be excluded when scoring DIS. Moreover, it can be quite difficult, if not impossible, to identify the particular lesion causing the clinical symptoms. These doubts ask for a simplification of the McDonald 2010 criteria, as recently proposed by the MAGNIMS study group.[23] This is supported by recent studies indicating that including the symptomatic lesion in the criteria for DIS, does not lead to a decrease in specificity and even increases the sensitivity of these diagnostic criteria.[24,25]

As a future perspective, the introduction of ultra-high-field MRI creates new possibilities and challenges. Given the strong effect of tissue relaxation times, in particular on clinically recommended sequences (such as FLAIR, conventional T2 and optionally double inversion recovery), and the different appearances of cortical grey matter and white matter structures, the reading of 7 T images in the context of MS is likely to be even more challenging.[26–32] 7 T is now exclusively used in research and its future role in clinical practice remains uncertain. Possibly, the effect of training will be even stronger for ultra-high-field MRI.

In conclusion, this study demonstrates a moderate to good interobserver agreement on lesion detection, DIS and DIT, which was not substantially influenced by field strength. Furthermore, interobserver agreement at 3 T was lower for less-experienced raters compared to experienced raters, indicating correct interpretation of high field strength MRI may require more training.

Acknowledgements

MAGNIMS steering committee members: F Barkhof (MS Centre Amsterdam, VU University Medical Centre, Amsterdam, The Netherlands and Institutes of Neurology and Healthcare Engineering, UCL Institute of Neurology, London, UK), O Ciccarelli and T Yousry (Queen Square Multiple Sclerosis Centre, UCL Institute of Neurology, London, UK), N De Stefano (University of Siena, Siena, Italy), C Enzinger (Department of Neurology, Medical University of

Graz, Graz, Austria), M Filippi and M A Rocca (San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy), J L Frederiksen (Glostrup Hospital and University of Copenhagen, Copenhagen, Denmark), C Gasperini (San Camillo-Forlanini Hospital, Rome, Italy), L Kappos (University of Basel, Basel, Switzerland), J Palace (University of Oxford Hospitals Trust, Oxford, UK), A Rovira and J Sastre-Garriga (Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain), H Vrenken (MS Centre Amsterdam, VU Medical Centre, Amsterdam, The Netherlands).

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: M.H.J.H., J.B., I.D.K., S.R., R.C., N.C., E.S., M.An., M.Am., J.M.L. and B.I.L-W. have nothing to disclose. S.C. received meeting compensations from Novartis. P.P. has received founding for travel from Novartis, Genzyme and Bracco and speaker honoraria from Biogen. J.K. has accepted speaker and consultancy fees from Merck-Serono, Teva, Biogen, Genzyme, Roche and Novartis. C.O-G received honoraria as speaker from Biogen Idec, Bayer Schering, Merck-Serono, Teva, Genzyme and Novartis. J.W. is CEO of MIAC AG, has received research grants from the German Ministries of Science and Economy, the European Union and from Novartis. He received speaker honoraria from Bayer, Biogen, Genzyme Sanofi, Novartis and Teva, and he served for advisory boards for Novartis, Biogen, Genzyme and Roche. O.C. is an Associate Editor of Neurology, and she serves as consultant for Novartis, Roche, Genzyme and Teva, and payments are made to the institution. C.G. received fees as speaker for Bayer Schering Pharma, Sanofi-Aventis, Genzyme, Biogen, Teva, Novartis and Merck-Serono and received a grant for research by Teva. C.L. holds an endowed professorship supported by the Novartis foundation, has received consulting and speaker's honoraria from Biogen Idec, Bayer Schering, Novartis, Sanofi, Genzyme and TEVA and has received research scientific grant support from Merck-Serono and Novartis. A.R. serves on scientific advisory boards for Biogen Idec, Novartis, Genzyme and OLEA Medical, and on the editorial board of the American Journal of Neuroradiology and Neuroradiology, has received speaker honoraria from Bayer, Genzyme, Sanofi-Aventis, Bracco, Merck-Serono, Teva Pharmaceutical Industries Ltd, Stendhal, Novartis and Biogen Idec and has research agreements with Siemens AG. F.B. serves on the editorial boards of Brain, Neurology, Neuroradiology, Multiple Sclerosis Journal and Radiology and serves as a consultant for Bayer Schering Pharma, Sanofi-Aventis, Genzyme, Biogen, Teva, Novartis, Roche, Synthon BV and Jansen Research. M.P.W. serves on the editorial boards of Neuroradiology, Journal of Neuroimaging, European Radiology, Frontiers of Neurology and serves as a consultant for Roche, Novartis and Biogen.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This research has been supported by a programme grant (14-358e) from the Dutch MS Research Foundation (Voorschoten, The Netherlands). The study in London was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

References

1. Wattjes MP, Rovira A, Miller D, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis – Establishing disease prognosis and monitoring patients. *Nat Rev Neurol* 2015; 11: 597–606.
2. Rovira A, Wattjes MP, Tintore M, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis-clinical implementation in the diagnostic process. *Nat Rev Neurol* 2015; 11: 471–482.
3. Beck RW, Trobe JD, Moke PS, et al. High- and low-risk profiles for the development of multiple sclerosis within 10 years after optic neuritis: Experience of the optic neuritis treatment trial. *Arch Ophthalmol* 2003; 121: 944–949.
4. Fisniku LK, Brex PA, Altmann DR, et al. Disability and T2 MRI lesions: A 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* 2008; 131: 808–817.
5. Minneboo A, Barkhof F, Polman CH, et al. Infratentorial lesions predict long-term disability in patients with initial findings suggestive of multiple sclerosis. *Arch Neurol* 2004; 61: 217–221.
6. Tintore M, Rovira A, Rio J, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain* 2015; 138: 1863–1874.
7. Tintore M, Rovira A, Arrambide G, et al. Brainstem lesions in clinically isolated syndromes. *Neurology* 2010; 75: 1933–1938.
8. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010

revisions to the McDonald criteria. *Ann Neurol* 2011; 69: 292–302.

9. Wattjes MP and Barkhof F. High field MRI in the diagnosis of multiple sclerosis: High field-high yield? *Neuroradiology* 2009; 51: 279–292.
10. Kilsdonk ID, de Graaf WL, Barkhof F, et al. Inflammation high-field magnetic resonance imaging. *Neuroimaging Clin N Am* 2012; 22: 135–157, ix.
11. Filippi M, Evangelou N, Kangarlu A, et al. Ultra-high-field MR imaging in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2014; 85: 60–66.
12. Wattjes MP, Lutterbey GG, Harzheim M, et al. Higher sensitivity in the detection of inflammatory brain lesions in patients with clinically isolated syndromes suggestive of multiple sclerosis using high field MRI: An intraindividual comparison of 1.5 T with 3.0 T. *Eur Radiol* 2006; 16: 2067–2073.
13. Simon B, Schmidt S, Lukas C, et al. Improved in vivo detection of cortical lesions in multiple sclerosis using double inversion recovery MR imaging at 3 Tesla. *Eur Radiol* 2010; 20: 1675–1683.
14. Wattjes MP, Harzheim M, Kuhl CK, et al. Does high-field MR imaging have an influence on the classification of patients with clinically isolated syndromes according to current diagnostic MR imaging criteria for multiple sclerosis? *AJNR Am J Neuroradiol* 2006; 27: 1794–1798.
15. Wattjes MP, Harzheim M, Lutterbey GG, et al. Does high field MRI allow an earlier diagnosis of multiple sclerosis? *J Neurol* 2008; 255: 1159–1163.
16. Kilsdonk ID, Barkhof F and Wattjes MP. 2010 revisions to McDonald criteria for diagnosis of multiple sclerosis: Impact of 3-Tesla magnetic resonance imaging. *Ann Neurol* 2011; 70: 182–183.
17. Landis JR and Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
18. Sombekke MH, Wattjes MP, Balk LJ, et al. Spinal cord lesions in patients with clinically isolated syndrome: A powerful tool in diagnosis and prognosis. *Neurology* 2013; 80: 69–75.
19. Bot JC, Barkhof F, Polman CH, et al. Spinal cord abnormalities in recently diagnosed MS patients: Added value of spinal MRI examination. *Neurology* 2004; 62: 226–233.
20. Bot JC and Barkhof F. Spinal-cord MRI in multiple sclerosis: Conventional and nonconventional MR techniques. *Neuroimaging Clin N Am* 2009; 19: 81–99.
21. Stankiewicz JM, Neema M, Alsop DC, et al. Spinal cord lesions and clinical status in multiple sclerosis: A 1.5 T and 3 T MRI study. *J Neurol Sci* 2009; 279: 99–105.
22. Korteweg T, Uitdehaag BM, Knol DL, et al. Interobserver agreement on the radiological criteria of the International Panel on the diagnosis of multiple sclerosis. *Eur Radiol* 2007; 17: 67–71.
23. Filippi M, Rocca MA, Ciccarelli O, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. *Lancet Neurol* 2016; 15: 292–303.
24. Brownlee WJ, Swanton JK, Miszkiel KA, et al. Should the symptomatic region be included in dissemination in space in MRI criteria for MS? *Neurology* 2016; 87: 680–683.
25. Tintore M, Otero-Romero S, Rio J, et al. Contribution of the symptomatic lesion in establishing MS diagnosis and prognosis. *Neurology* 2016; 87: 1368–1374.
26. De Graaf WL, Kilsdonk ID, Lopez-Soriano A, et al. Clinical application of multi-contrast 7-T MR imaging in multiple sclerosis: Increased lesion detection compared to 3 T confined to grey matter. *Eur Radiol* 2013; 23: 528–540.
27. De Graaf WL, Zwanenburg JJ, Visser F, et al. Lesion detection at seven Tesla in multiple sclerosis using magnetisation prepared 3D-FLAIR and 3D-DIR. *Eur Radiol* 2012; 22: 221–231.
28. Kilsdonk ID, Jonkman LE, Klaver R, et al. Increased cortical grey matter lesion detection in multiple sclerosis with 7 T MRI: A post-mortem verification study. *Brain* 2016; 139: 1472–1481.
29. Mistry N, Tallantyre EC, Dixon JE, et al. Focal multiple sclerosis lesions abound in 'normal appearing white matter'. *Mult Scler* 2011; 17: 1313–1323.
30. Tallantyre EC, Dixon JE, Donaldson I, et al. Ultra-high-field imaging distinguishes MS lesions from asymptomatic white matter lesions. *Neurology* 2011; 76: 534–539.
31. Kilsdonk ID, Wattjes MP, Lopez-Soriano A, et al. Improved differentiation between MS and vascular brain lesions using FLAIR* at 7 Tesla. *Eur Radiol* 2014; 24: 841–849.
32. Sinnecker T, Kuchling J, Dusek P, et al. Ultrahigh field MRI in clinical neuroimmunology: A potential contribution to improved diagnostics and personalised disease management. *EPMA J* 2015; 6: 16.