doi:10.1093/brain/awaa233 BRAIN 2020: 143; 2988–2997 | 2988

BRAIN
A JOURNAL OF NEUROLOGY

Histopathology-validated recommendations for cortical lesion imaging in multiple sclerosis

Piet M. Bouman,[1] Martijn D. Steenwijk,[1] Petra J. W. Pouwels,[2] Menno M. Schoonheim,[1] Frederik Barkhof,[2,3] Laura E. Jonkman[1] and Jeroen J. G. Geurts[1]

Cortical demyelinating lesions are clinically important in multiple sclerosis, but notoriously difficult to visualize with MRI. At clinical field strengths, double inversion recovery MRI is most sensitive, but still only detects 18% of all histopathologically validated cortical lesions. More recently, phase-sensitive inversion recovery was suggested to have a higher sensitivity than double inversion recovery, although this claim was not histopathologically validated. Therefore, this retrospective study aimed to provide clarity on this matter by identifying which MRI sequence best detects histopathologically-validated cortical lesions at clinical field strength, by comparing sensitivity and specificity of the thus far most commonly used MRI sequences, which are T_2, fluid-attenuated inversion recovery (FLAIR), double inversion recovery and phase-sensitive inversion recovery. Post-mortem MRI was performed on non-fixed coronal hemispheric brain slices of 23 patients with progressive multiple sclerosis directly after autopsy, at 3 T, using T_1 and proton-density/T_2-weighted, as well as FLAIR, double inversion recovery and phase-sensitive inversion recovery sequences. A total of 93 cortical tissue blocks were sampled from these slices. Blinded to histopathology, all MRI sequences were consensus scored for cortical lesions. Subsequently, tissue samples were stained for proteolipid protein (myelin) and scored for cortical lesion types I–IV (mixed grey matter/white matter, intracortical, subpial and cortex-spanning lesions, respectively). MRI scores were compared to histopathological scores to calculate sensitivity and specificity per sequence. Next, a retrospective (unblinded) scoring was performed to explore maximum scoring potential per sequence. Histopathologically, 224 cortical lesions were detected, of which the majority were subpial. In a mixed model, sensitivity of T_1, proton-density/T_2, FLAIR, double inversion recovery and phase-sensitive inversion recovery was 8.9%, 5.4%, 5.4%, 22.8% and 23.7%, respectively (20, 12, 12, 51 and 53 cortical lesions). Specificity of the prospective scoring was 80.0%, 75.0%, 80.0%, 91.1% and 88.3%. Sensitivity and specificity did not significantly differ between double inversion recovery and phase-sensitive inversion recovery, while phase-sensitive inversion recovery identified more lesions than double inversion recovery upon retrospective analysis (126 versus 95; $P < 0.001$). We conclude that, at 3 T, double inversion recovery and phase-sensitive inversion recovery sequences outperform conventional sequences T_1, proton-density/T_2 and FLAIR. While their overall sensitivity does not exceed 25%, double inversion recovery and phase-sensitive inversion recovery are highly pathologically specific when using existing scoring criteria and their use is recommended for optimal cortical lesion assessment in multiple sclerosis.

1 Department of Anatomy and Neurosciences, MS Center Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands
2 Department of Radiology and Nuclear Medicine, MS Center Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands
3 UCL Institutes of Neurology and Healthcare Engineering, UCL, London, UK

Correspondence to: Piet M. Bouman
Department of Anatomy and Neurosciences, Amsterdam University Medical Centre
Location VUmc, PO Box 7075, 1007 MB Amsterdam, The Netherlands
E-mail: p.bouman@amsterdamumc.nl

Keywords: multiple sclerosis; cortical lesions; post-mortem imaging; double inversion recovery; phase-sensitive inversion recovery

Received December 23, 2019. Revised April 10, 2020. Accepted June 1, 2020. Advance access publication August 21, 2020

Abbreviations: DIR = double inversion recovery; FLAIR = fluid-attenuated inversion recovery; PD = proton-density; PSIR = phase-sensitive inversion recovery

Introduction

Multiple sclerosis is a chronic inflammatory, demyelinating and neurodegenerative disease of the CNS, which is accompanied by a great variety of clinical and cognitive deficits (Compston and Coles, 2008). Although initially considered a white matter disease, cortical grey matter lesions were found to be abundant in multiple sclerosis (Kidd *et al.*, 1999; Geurts *et al.*, 2005*a*; Kutzelnigg *et al.*, 2005). Consequently, as white matter abnormalities alone insufficiently explained clinical symptomatology (Barkhof, 2002; Geurts and Barkhof, 2008; Calabrese *et al.*, 2010; Klaver *et al.*, 2013), research focus has shifted to (imaging of) cortical lesions over the past years. Thus, cortical lesions have been associated with disease conversion and progression, and have enabled a better understanding of cognitive decline in multiple sclerosis (Calabrese *et al.*, 2009; Roosendaal *et al.*, 2009; Nelson *et al.*, 2011; Geurts *et al.*, 2012). As cortical lesions seem highly specific for multiple sclerosis, they have become an integral part of the diagnostic criteria for multiple sclerosis (Thompson *et al.*, 2018). However, despite their obvious clinical relevance, cortical lesions are notoriously hard to detect using conventional MRI techniques.

Early histopathological validation studies at clinical field strengths (1.5 T and 3 T) suggested superior sensitivity of double inversion recovery (DIR) when compared to proton density (PD)/T_2-weighted imaging and 3D fluid-attenuated inversion recovery (FLAIR) imaging (Geurts *et al.*, 2005*b*). Reasons for this relative superiority may be that contrast between cortical lesions and normal cortex is subtly enhanced in DIR, following suppression of both white matter and CSF signals in the double inversion scheme. Official scoring criteria for DIR were developed and published (Geurts *et al.*, 2011) and, using these criteria, DIR was shown to have a very high pathological specificity, though only a maximum sensitivity of 18% (Seewann *et al.*, 2012). This given, combined with the fact that DIR is a time-consuming technique in acquisition, with a low signal-to-noise ratio, and a host of artefacts (Geurts *et al.*, 2011), its routine use has remained limited. In light of this, more recent studies developed a phase-sensitive inversion recovery (PSIR) technique, which in a study comparing the two sequences (Sethi *et al.*, 2012), was suggested to have a higher sensitivity than DIR. However, as histopathological validation of PSIR has been lacking, this claim was never verified.

The aim of the present study was to assess sensitivity and specificity of DIR and PSIR, set against conventional clinical sequences (T_1, PD/T_2, FLAIR). As such, we aim to provide an evidence-based recommendation as to which currently available MRI technique(s) should be 'method of choice' for imaging cortical lesions in multiple sclerosis.

Materials and methods

Patients and autopsy procedure

Imaging data and tissue were obtained through the standardized MS Center Amsterdam rapid autopsy protocol (Popescu *et al.*, 2015). Inclusion criteria were availability of T_1, PD/T_2, FLAIR, DIR and PSIR MRI and tissue samples from the corresponding patients. Prior to death, all patients had registered with the Netherlands Brain Bank, thereby giving consent for use of their tissue and medical records for research purposes. Permission for the autopsy protocol was further granted by the institutional ethics review board. A summary of the included patient characteristics is displayed in Table 1. An overview of the methods is displayed in Fig. 1.

MRI

MRI was performed on a 3 T whole-body scanner (GE MR750 Discovery), with an eight-channel phased-array head coil. After extracting the brain from cranium, it was cut into 1-cm thick coronal slices, which were then immediately scanned with MRI, prior to fixation of the tissue. Brain slices were scanned in a custom-made brain-slice holder that allows for imaging of five brain slices at the same time. These five brain slices were *a priori* selected for MRI in a standardized manner (Popescu *et al.*, 2015). The imaging protocol included 2D-T_1-weighted, 2D-PD T_2-weighted and 3D-FLAIR. Furthermore, 3D-DIR and 2D-PSIR sequences were acquired. Sequence parameters are displayed in Table 2.

Prospective MRI scoring

Cortical lesions were manually rated on each sequence independently, while randomizing patients and sequence, using Medical Image Processing, Analysis and Visualization software (MIPAV; version 8.0.2, Centre for Information Technology, National Institutes of Health, Bethesda, MD, USA). All regions were scored for cortical lesions, blinded to histopathology. Scoring was performed by P.M.B. in consensus with J.J.G.G. (who has 15 years of experience) according to an iterative consensus scoring scheme. Scoring was based on consensus guidelines developed by the MAGNIMS consortium for DIR and scoring criteria developed by Sethi and colleagues for PSIR (Geurts *et al.*, 2011; Sethi *et al.*, 2012). Furthermore, prospective inter-rater variability was determined between P.M.B. and L.E.J. (who has 7 years of experience in defining lesions on post-mortem MRI). In addition, prospective intra-rater variability was determined. Inter- and intra-rater measures were determined using a randomly selected subset of the images from five of the included patients. All inter- and intra-rater scoring was performed in a random order for both patients and sequences as well.

Table 1 Demographics of included patients

Patient	Sex	Age, years	PMD, h: min	DD, years	Disease type	Cause of death
1	F	58	4:00	25	PPMS	Euthanasia
2	F	77	4:00	26	PPMS	Natural
3	F	82	3:40	60	SPMS	Euthanasia
4	F	75	5:15	25	Unknown	Aspirational pneumonia
5	F	65	4:15	16	PPMS	CVA
6	F	50	4:25	25	PPMS	Euthanasia
7	F	40	2:30	8	SPMS	Morphine induction
8	M	59	6:20	28	SPMS	Urosepsis
9	F	49	3:55	34	SPMS	Pneumonia
10	F	49	6:20	28	SPMS	Urosepsis
11	M	66	5:15	37	SPMS	Pneumonia
12	M	48	3:15	15	SPMS	Dehydration
13	F	58	6:25	20	SPMS	Euthanasia
14	F	66	7:00	23	PPMS	Pneumonia
15	F	75	6:00	36	SPMS	Subdural haematoma
16	M	61	4:50	28	SPMS	Euthanasia
17	M	57	6:00	25	PPMS	Urosepsis
18	F	70	3:52	32	SPMS	Euthanasia
19	M	66	5:00	25	PPMS	Euthanasia
20	M	60	5:09	17	PPMS	Euthanasia
21	M	54	4:15	21	SPMS	Euthanasia
22	F	61	4:30	2	Unknown	Euthanasia
23	F	74	3:50	50	SPMS	Euthanasia
Mean (±SD)		62.3 (±10.1)	4:30 (±1:03)	26.5 (±14.1)		

CVA = cerebrovascular accident; DD = disease duration from diagnosis; PMD = post-mortem delay at time of arrival in the hospital; PPMS = primary progressive multiple sclerosis; SPMS = secondary progressive multiple sclerosis.

Histopathological staining and validation of MRI scores

After MRI scanning, the 1-cm thick coronal brain slices were cut in half to obtain 5-mm tissue slices from the centre of the MRI planes. Subsequently, a standardized set of tissue blocks was dissected from the slices. From the remaining tissue, a number of regions of interest was dissected, guided by MRI to increase the harvest of cortical lesions, which can be difficult to detect by the naked eye. Immediately after dissection of the selected tissue samples, they were formalin-fixed for paraffin embedding (FFPE), and stored in the Amsterdam MS Center Biobank until further use. The tissue used in this study had been fixated in formalin for 48 h before embedding. FFPE tissue samples were cut to 10-µm thick sections. To undo crosslinking of proteins due to embedding in paraffin, after deparaffinization, antigen retrieval was performed by heating sections submerged in Tris-EDTA buffer (10 mM; pH 9.0) to 90–95°C for 30 min in a steam cooker. Endogenous peroxidase was blocked using 1% hydrogen peroxide in Tris-buffered saline (TBS; pH 7.6). To block non-specific binding, sections were incubated with 3% bovine serum albumin in TBS-Triton X. Subsequently, sections were incubated with primary antibody proteolipid protein (PLP; Bio-Rad) overnight at 4°C. The next day, sections were incubated with biotin labelled donkey anti-mouse (Jackson IgG) 1:400 diluted in TBS-Triton X for 2 h. Next, sections were incubated with ABC (Vector) diluted 1:400 in TBS- Triton X for 1 h. Colour development was performed using 3,3′-diaminobenzidine (DAB) for 10 min. Sections were counterstained with thionin (Brand). Finally, sections were dehydrated and mounted with Entellan® (Merck).

Histopathological lesions were then scored according to the criteria developed by Bø *et al.* (2003). Lesions were scored based on their position in the cortex: mixed grey-white matter (type I; leukocortical), purely intracortical (type II), subpial (type III) or cortex-spanning (type IV). Scored MRI abnormalities that were not multiple sclerosis lesions upon histopathological validation were marked as false positives. Subsequently, regions of interest were drawn on PLP sections for all cortical lesions that were visible on DIR and/or PSIR using ImageJ v1.52q (http://rsbweb.nih.gov/ij). Regions of interest area was then calculated to measure lesion surface area in square millimetres.

Matching and retrospective MRI scoring

After prospective MRI and histopathological scoring, tissue samples were matched to their corresponding MRI images, using as many anatomical cortical and white matter landmarks as possible. Then, the prospective MRI scores were unblinded and a retrospective unblinded scoring (P.M.B.) was performed: i.e. lesions were scored on MRI while knowing the lesion location and type. Retrospective inter-rater variability was determined between P.M.B. and L.E.J. for a randomly selected subset of the images from five of the included patients. Additionally, retrospective intra-rater variability was determined.

Image contrasts

Contrast ratios were calculated in a random subsample of 10 patients, for different sequences, based on signal-intensity

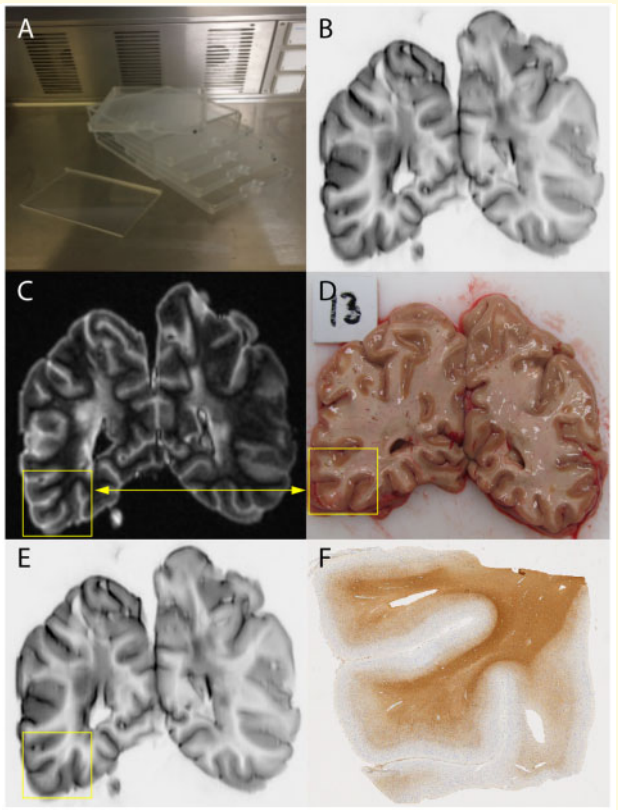

Figure 1 Overview of methods. (**A**) Five coronally cut brain slices were obtained as part of the Amsterdam MS Center rapid autopsy protocol, which were scanned simultaneously in a custom-made brain-slice holder. (**B**) Acquired pulse-sequences were 2D-T_1-weighted and 2D-PD/T_2-weighted, 3D-FLAIR, 3D-DIR and 2D-PSIR (shown). (**C** and **D**) Tissue samples are obtained based on a standardized protocol aided by MRI guided tissue dissection. (**E**) MRI scans were prospectively scored for cortical lesions, blinded to histopathology. (**F**) Subsequently, histopathological validation was performed using myelin-staining followed by a retrospective, unblinded, scoring for cortical lesions and sensitivity and specificity measures were calculated.

measures in different regions of interest, which were placed in cortical lesions ($n = 30$), normal-appearing grey matter ($n = 30$) and normal-appearing white matter ($n = 30$). Regions of interest were drawn separately for each patient in the subsample, once in the DIR sequence (3D, also used in FLAIR) and once in the PSIR sequence (2D, also used in T_1 and PD/T_2). Regions of interest were drawn using FSL 6.0.2 (FMRIB Software Library: http://fsl.fmrib.ox.ac.co.uk/fsl/fslwiki). Contrast ratio was defined as $(SI_1 - SI_2)/SI_2$, in which SI_1 denotes signal intensity of the lesion and SI_2 represents signal intensity of the normal-appearing grey matter or normal-appearing white matter. Outcomes of all selected patients were then averaged.

Statistical analysis

Histopathological lesion count was considered the gold standard. Sensitivity of the sequences for detecting cortical lesions was determined by dividing the number of detected lesions in either the prospective or the retrospective scoring by the number of lesions detected on histopathology, multiplied by 100%. Specificity of MRI sequences was calculated by dividing the total number of lesions counted on MRI by the number of histopathologically validated lesions × 100%. Comparisons between sequences and sequences to histopathology were made using a mixed model in SPSS 24.0 (SPSS, Chicago, Illinois), controlling for age, sex and post-mortem delay. Furthermore, η^2 was calculated to explore the added value of combining DIR and PSIR sequences in a regression model. Results from the pair-wise comparison were Bonferroni-corrected for multiple comparisons, after which P-values of $\leqslant 0.05$ were considered statistically significant. Wilcoxon signed-rank tests were performed to determine the relationship between cortical lesion surface area and visibility on DIR and PSIR. Prospective and retrospective inter- and intra-rater variability were expressed as an intraclass correlation coefficient for all sequences using a two-way random effects model with absolute agreement.

Data availability

Anonymized data, not published in the article, will be shared upon reasonable request from a qualified investigator.

Results

Based on MRI and tissue availability, coronally cut brain slices and MRI scans of 23 patients with progressive multiple sclerosis [mean post-mortem delay 4 h 30 min, standard deviation (SD) 1 h 3 min; five brain slices per patient, five sequences per patients, five images per sequence, 115 brain slices, 575 images in total] were included for prospective scoring. In the selected tissue samples for histopathology ($n = 93$), we identified a total of 224 cortical lesions. Of these lesions, seven were type I, 56 were type II, 141 were type III and 20 were type IV; 34 white matter lesions were detected.

Prospective lesion detection

The number of lesions that were counted and the corresponding sensitivity percentages are displayed in Table 3. After histopathological validation, DIR detected 3.25-fold more cortical lesions than FLAIR and PD/T_2, with 51 versus both 12 cortical lesions ($P < 0.001$), and 1.55-fold more than T_1, with 51 versus 20 cortical lesions ($P = 0.001$). PSIR detected 3.42-fold more cortical lesions than FLAIR and PD/T_2, with both 53 versus 12 cortical lesions ($P < 0.001$). PSIR detected 1.65-fold more cortical lesions than T_1, with 53 versus 20 cortical lesions ($P < 0.001$). PSIR did not detect more cortical lesions than DIR with 53 and 51, respectively (not significant). There were no differences in cortical lesion detection between T_1, PD/T_2 and FLAIR sequences. Few of the detected lesions showed signs of partial re/demyelination: three on T_1, two on PD/T_2, one on FLAIR, six on DIR and four on PSIR. Thirty-four white matter lesions were detected in the samples. There were no differences in white matter lesion detection between all included sequences.

Table 2 Sequence parameters

	2D-T_1	2D-PD/T_2	3D-FLAIR	3D-DIR	2D-PSIR
Repetition time, ms	400	3000	8000	8000	4000
Echo time, ms	7.95	16.2/113.2	127	126.1	13.1
Inversion time, ms	–	–	2346	3160/500	419
Field of view, mm	187 × 250	187 × 250	256 × 256	256 × 256	187 × 250
Flip angle	111°	111°	90°	90°	111°
Slice thickness, mm	5.0	5.0	1.2	1.2	5.0
Acquisition matrix	256 × 320	384 × 256	224 × 224	224 × 224	256 × 320
Resolution (reconstructed), mm	0.5 × 0.5 × 5.0	0.5 × 0.5 × 5.0	1 × 1 × 1.2	1 × 1 × 1.2	0.5 × 0.5 × 5.0
Acquisition time, min: s	3:27	3:18	5:40	8:52 (2 averages)	1:44

Table 3 Lesion count (sensitivity in %) MRI scoring

Histology		MRI rating									
Lesion type	*n*	T_1 pro (%)	T_1 retro (%)	PD/T_2 pro (%)	PD/T_2 retro (%)	FLAIR pro (%)	FLAIR retro (%)	DIR pro (%)	DIR retro (%)	PSIR pro (%)	PSIR retro (%)
Type I	7	2 (28.6)	3 (42.9)	–	2 (28.6)	1 (14.3)	3 (42.9)	4 (57.1)	5 (71.4)	3 (42.9)	6 (85.8)
Type II	56	4 (7.1)	8 (14.3)	2 (3.6)	5 (8.9)	1 (1.8)	7 (12.5)	7 (12.5)	12 (21.4)	7 (12.5)	17 (30.4)
Type III	141	13 (9.2)	46 (32.6)	9 (6.4)	34 (24.1)	6 (4.3)	30 (21.3)	33 (23.4)	63 (44.7)	32 (22.7)	84 (59.6)
Type IV	20	1 (5.0)	8 (40.0)	1 (5.0)	12 (60.0)	4 (20.0)	12 (60.0)	7 (35.0)	15 (75.0)	11 (55.0)	19 (95.0)
Type I-IV	224	20 (8.9)*	65 (29.0)	12 (5.4)*	53 (23.7)*	12 (5.4)*	52 (23.2)*	51 (22.8)	95 (42.4)	53 (23.7)	126 (56.3)**
WML	34	7 (20.6)	17 (50.0)	8 (23.5)	20 (58.9)	6 (17.6)	14 (41.2)	17 (50.0)	25 (73.5)	15 (44.1)	24 (70.6)
Total	258	27 (10.5)	82 (31.8)	20 (7.8)	73 (28.3)	18 (7.9)	66 (25.6)	68 (26.4)	120 (46.5)	68 (26.4)	150 (58.1)

pro = prospective (blinded) scoring; retro = retrospective (unblinded) scoring; WML = white matter lesion.
*Significant differences (i.e. $P < 0.001$) between DIR and other sequences;
**Significant difference between DIR and other sequences ($P < 0.05$).

Prospective inter-rater intraclass correlation coefficients were 0.840 for T_1, 0.656 for PD/T_2, 0.899 for FLAIR, 0.836 for DIR and 0.731 for PSIR. Prospective intra-rater intraclass correlation coefficients were 0.821 for T_1, 0.757 for PD/T_2, 0.738 for FLAIR, 0.896 for DIR and 0.827 for PSIR.

Twenty-three false positives (i.e. areas that were scored as cortical lesion but were found not to be cortical lesions on histopathology) were scored: 4/20 (20%) on T_1, 4/16 (25%) on PD/T_2, 3/15 (20%) on FLAIR, 5/56 (8.9%) on DIR and 7/60 (11.7%) on PSIR. Specificity was 80.0% for T_1, 75% for PD/T_2, 80% FLAIR, for DIR and PSIR this was 91.1% and 88.3%, respectively. Analysis of explained variance showed 5% increase of explained variance in the data when combining DIR and PSIR sequences.

Retrospective lesion detection

During the retrospective scoring, i.e. locating each histopathology-identified lesion on MRI, more cortical lesions were rated (for all sequences) than during the prospective scoring (Table 3). PSIR showed the largest retrospective increase in detection (237.7%) thereby retrospectively detecting more cortical lesions than DIR, with 126 versus 95 cortical lesions ($P = 0.014$). Retrospective detection rates between DIR and T_1 did not differ from each other ($P = 0.248$).

In regard to lesion type, PSIR tended to detect more type III lesions than DIR ($P = 0.056$). There were no significant differences in detection rates of specific lesion types. Retrospectively, four areas of partial re/demyelination were detected: four on T_1, two on T_2, three on FLAIR, seven on DIR and six on PSIR. Fifty-one cortical areas of partial re/demyelination were observed histopathologically, 36 of which were not visible on MRI on any of the included sequences. Retrospective inter-rater intraclass correlation coefficients were 0.711 for T_1, 0.615 for PD/T_2, 0.692 for FLAIR, 0.846 for DIR and 0.714 for PSIR. Retrospective intra-rater intraclass correlation coefficients were 0.750 for T_1, 0.714 for PD/T_2, 0.727 for FLAIR, 0.849 for DIR and 0.833 for PSIR.

Histopathological regions of interest were drawn on PLP sections for 117 cortical lesions that were visible on DIR and/or PSIR: five type I lesions, 15 type II lesions, 79 type III lesions and 18 type IV lesions. Average surface area of cortical lesions (all types) was 12.09 mm^2 (SD 18.66) for cortical lesions visible on DIR during the prospective scoring, and 11.74 mm^2 (SD 18.67) during the retrospective scoring. For PSIR, the average surface area of cortical lesions visible during the prospective scoring was 12.06 mm^2 (SD 18.60) and 11.52 mm^2 (SD 18.71) during the retrospective scoring. Wilcoxon signed-ranks tests indicated that, during the prospective scoring, lesions with a larger surface area were scored more often on DIR ($Z = -9.08$, $P < 0.001$) and PSIR ($Z = -9.11$, $P < 0.001$). Also, during the retrospective scoring, lesions with a larger surface area were scored more often on DIR ($Z = -8.92$, $P < 0.001$) and PSIR ($Z = -7.68$, $P < 0.001$).

Table 4 Contrast ratios

	T_1	PD-T_2	FLAIR	DIR	PSIR
CR CL–GM	0.07 (0.06)	0.08 (0.07)	0.20 (0.17)	0.61 (0.50)	0.20 (0.17)
CR GM–WM	0.16 (0.09)	0.21 (0.16)	0.50 (0.43)	2.29 (1.68)	9.76 (10.07)

Data are presented as contrast ratio (±SD); contrast ratio is defined as $(SI_1 - SI_2)/SI_2$. CL–GM = cortical lesion to grey matter; GM–WM = grey matter to white matter.

Image contrasts

Contrast ratios of cortical lesions to surrounding normal appearing grey matter and grey matter to white matter for different MRI sequences are described in Table 4. DIR showed the highest contrast ratios for cortical lesions versus grey matter; PSIR and FLAIR were comparable; and T_1 and PD/T_2 showed the lowest contrast ratios for cortical lesions versus grey matter. Contrast between grey matter and white matter was highest in PSIR and DIR. It should be noted that all values have a large range, because regions of interest were selected in a single slice and selected regions of interest in cortical lesions and grey matter contain a small number of voxels. Cortical lesion regions of interest contained on average 18.9 voxels (4.7 mm^2) in 2D sequences and 8.6 voxels (8.6 mm^2) in 3D sequences. Grey matter regions of interest had an average area of 17.2 voxels (4.3 mm^2) in 2D and 6.9 voxels (6.9 mm^2) in 3D sequences. White matter regions of interest had an average area of 62.2 voxels (15.5 mm^2) in 2D and 17.3 voxels (17.3 mm^2) in 3D.

Discussion

We found that, for all cortical lesion types, DIR and PSIR detected significantly more cortical lesions than conventional (clinical) sequences. They did not outperform each other in the prospective, blinded assessment, but PSIR did achieve higher cortical lesion ratings in the retrospective scoring.

The presented findings are in line with former studies showing a higher sensitivity of DIR for cortical lesions than that of PD/T_2 and FLAIR sequences (Geurts *et al.*, 2005*b*; Seewann *et al.*, 2012; Kilsdonk *et al.*, 2016), and with studies showing similar sensitivities of DIR and PSIR (Rinaldi *et al.*, 2010). Also, the finding that DIR is highly pathologically specific, superseding T_1, PD/T_2 and FLAIR, is concordant with former literature (Geurts *et al.*, 2005*a*; Seewann *et al.*, 2012; Kilsdonk *et al.*, 2016). This does not hold for specificity of PSIR—also highly pathologically specific—as this has not been histopathologically validated in previous studies. The results do not confirm findings from a former study in which PSIR was found to detect more cortical lesions than DIR *a priori* (mean: 18.1 intracortical lesions on PSIR versus 5.9 on DIR) (Sethi *et al.*, 2012). In a follow-up to the latter study, a substantial number of intracortical lesions scored on DIR was reclassified *a posteriori* as mixed grey matter–white matter lesions or juxtacortical lesions on PSIR (Sethi *et al.*, 2013). The number of prospectively scored type II-IV lesions on DIR that needed reclassification upon a check with PSIR in our total MRI sample was 19, showing that we cannot confirm any significant reclassification effect. These differences can be explained in a few ways. Whereas the aforementioned study used 2D-DIR and 3D-PSIR imaging modalities, our work was based on 3D-DIR and 2D-PSIR. This may have affected the scoring results to some degree. Also, this difference may have been influenced by (subtle) differences in adherence to DIR and PSIR scoring criteria, as well as in image resolution (our DIR had a higher in-plane resolution). We were not able to precisely match the image resolutions in the current study to those of the aforementioned study by Sethi *et al.* (2012), as this was a retrospective post-mortem study. Moreover, DIR and PSIR detected similar numbers of cortical lesions in our prospective scoring, whereas PSIR tended to detect more cortical lesions than DIR during the retrospective scoring. This increase seems to be driven by an improved retrospective detection of subpial lesions using PSIR, which, using the current PSIR scoring criteria were largely marked as artefacts in the prospective scoring. This finding needs to be confirmed in subsequent studies, but may indicate that PSIR enables visualization of at least part of the subpial lesions present, which would be a clear advantage over DIR. The current PSIR cortical lesion scoring criteria (Sethi *et al.*, 2012) emphasize the definition and exclusion of artefacts, such as those due to vessels and CSF, but might be overly strict in the end, resulting in an underestimation of subpial lesions in the cortex, and hence, a lower than desired sensitivity.

Although DIR and PSIR detected the highest number of cortical lesions, in the prospective, blinded-to-histopathology analysis, they still missed many, even when the rater was later unblinded to histopathology (Table 3). The fact that so many (i.e. ~75%) cortical lesions were still missed could be explained by the contrast ratio (Table 4), which shows that the images were relatively noisy, leading to subtle contrast differences between cortical lesions and normal-appearing tissue. This holds especially for the DIR images, in which many of the retrospectively identified cortical lesions had been marked as noise or artefacts during the prospective scoring. DIR images are noisy as a result of the double inversion scheme and loss of myelin in the thinly myelinated subpial areas produces little contrast, leading to disappearance of subpial lesions. Concerning detection of different cortical lesion types, type II lesions are usually very small, which tends to obscure these lesions against the (noisy) background grey matter (Seewann *et al.*, 2011). Type III lesions are particularly hard to detect as they are situated in the outer, far less densely myelinated regions of the cortex, therefore providing only very subtle contrast differences upon demyelination (Bø *et al.*, 2003; Bö *et al.*, 2004; Pitt *et al.*, 2010; Barkhof and Geurts, 2015). Meanwhile, type III lesions are known to be the most abundant cortical lesion type; they can become large and are predominantly present in progressive disease forms, stressing the need for better detection of this lesion type by MRI (Bø *et al.*, 2003; Seewann *et al.*,

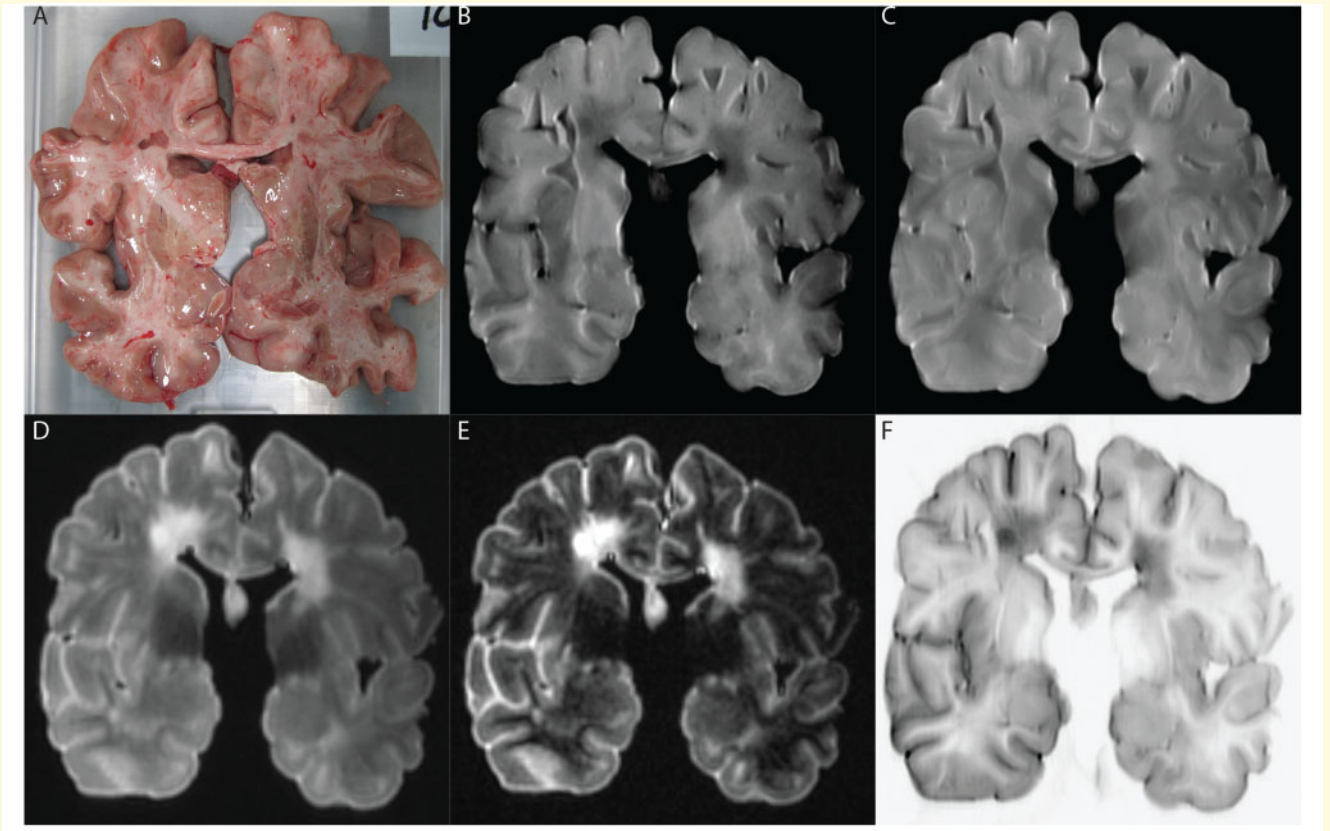

Figure 2 Representative example of MRI sequences evaluated for a specific patient. (**A**) Photo of scanned brain slice. (**B**) T_1-weighted MRI. (**C**) PD/T_2-weighted MRI. (**D**) FLAIR image. (**E**) DIR image. (**F**) PSIR image.

2011, 2012; Kilsdonk *et al.*, 2016; Trampel *et al.*, 2019). Whereas former studies stressed that cortical lesions are difficult to detect on MRI (Seewann *et al.*, 2011; Kilsdonk *et al.*, 2016), we confirm that larger cortical lesions that infiltrate into the deeper and more myelinated layers of the cortex (type IV) are usually easier to detect on MRI (Pitt *et al.*, 2010; Seewann *et al.*, 2011; Kilsdonk *et al.*, 2016). Some hyper/hypo-intensities that were scored as cortical lesion turned out not to be lesions upon histopathological analysis; these were scored as false positives. On many occasions, these false positives appeared to be areas of incomplete demyelination, or partial remyelination [as described by Strijbis *et al.* (2017)], or widened Virchow-Robin spaces instead. When solely possessing PLP-stained tissue sections, it is not possible to differentiate between partial demyelination and remyelination.

The number of detected cortical lesions on MRI could possibly be increased by obtaining higher image resolutions, but this would—for most sequences—include an increase in acquisition time that exceeds clinically acceptable terms. Higher resolution can be achieved using ultra-high field MRI, but this is not clinically feasible (Kilsdonk *et al.*, 2016). Clearly, increased cortical lesion detection is expected to help achieve an earlier diagnosis of patients (Filippi *et al.*, 2016, 2019). But, even using current tools, at clinical field strength, it is good to realise that a relatively low number of MRI-visible cortical lesions always correlates strongly with the total number of cortical lesions present (and largely unseen) (Seewann *et al.*, 2011). This means that, for clinical correlations, and if the visible and non-visible lesions are not in any way qualitatively different, the MRI-visible part of cortical lesions might just suffice. This was referred to as the tip-of-the-iceberg phenomenon (Seewann *et al.*, 2011). In addition, in terms of clinical understanding, it remains to be discussed whether visualization of cortical lesions should be preferred over, e.g. (cortical) atrophy measurement. Cortical atrophy measures have been repeatedly and strongly related to clinical disability, cognitive impairment and disease progression over time (Roosendaal *et al.*, 2011; Eijlers *et al.*, 2018; Dekker *et al.*, 2019). Furthermore, cortical atrophy measures are relatively easy to perform and are considered reliable, reproducible over different centres and predict subsequent cognitive decline (Steenwijk *et al.*, 2016; Eijlers *et al.*, 2018; Meijerman *et al.*, 2018). Cortical atrophy measures would therefore be appropriate to at least supplement, but perhaps altogether supplant the more arduous rating of cortical lesions.

A complicating factor for cortical lesion scoring in this study is the post-mortem setting. On post-mortem MRI, additional artefacts can be seen due to e.g. tissue-to-air

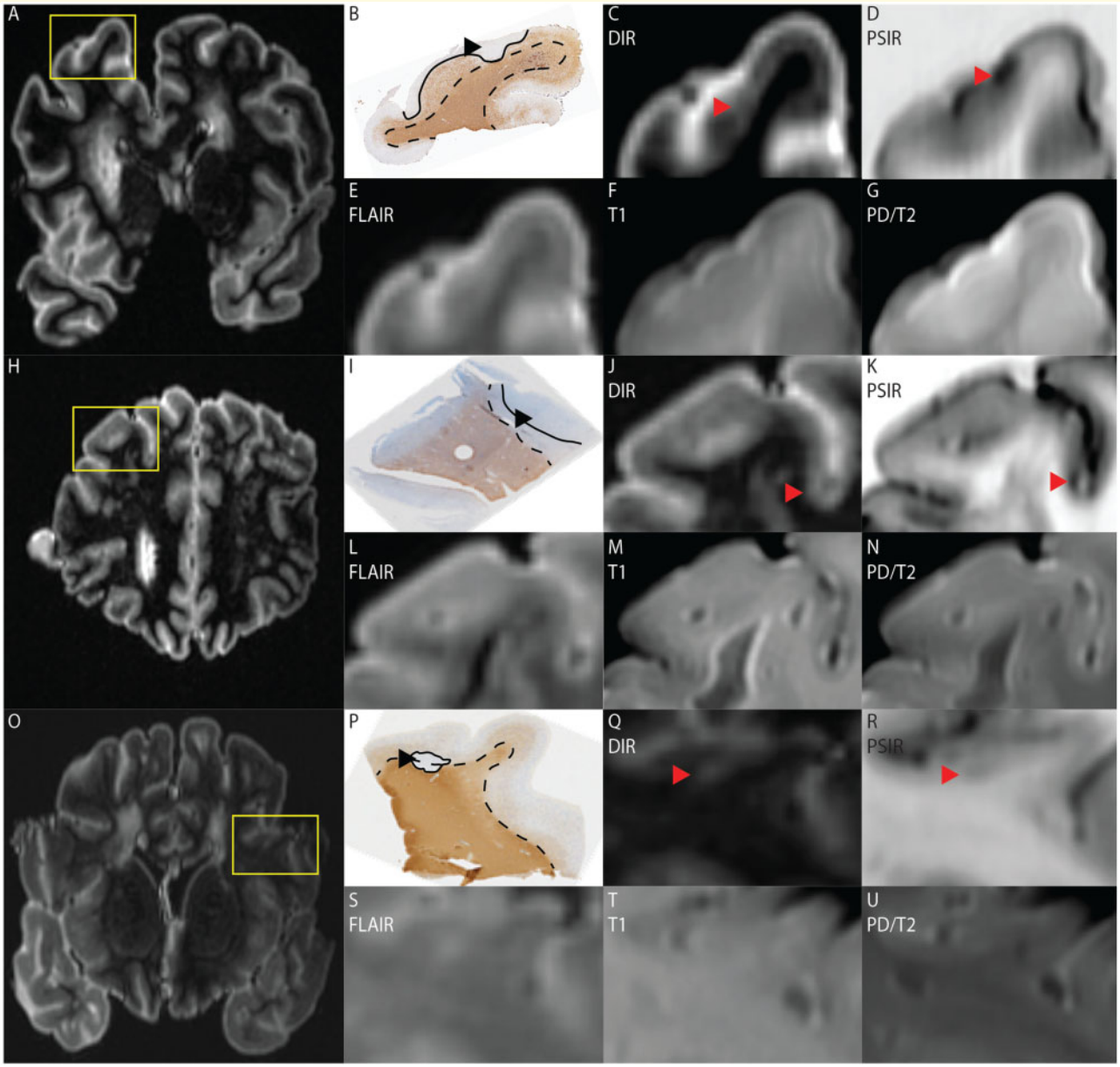

Figure 3 Three examples of cortical lesions visible on DIR and PSIR, but not on other sequences. (**A**) MRI overview of a DIR scanned brain slice from which a tissue block (yellow box) was processed. (**B**) Histopathological section of the processed tissue block stained for myelin, in which a type IV lesion is indicated by the black arrowhead and bordered by the black line; the dotted line indicates the cortex. (**C**–**G**) Excerpts of, respectively, DIR, PSIR, FLAIR, T_1 and PD/T_2 scans. The type IV lesion is indicated by the red arrowhead on DIR (**C**) and PSIR (**D**). (**H**) MRI overview of a DIR scanned brain slice from which a tissue block (yellow box) was processed. (**I**) Histopathological section of the processed tissue block stained for myelin, in which a type III lesion is indicated by the black arrowhead and bordered by the black line; the dotted line indicates the cortex. (**J**–**N**) excerpts of, respectively, DIR, PSIR, FLAIR, T_1 and PD/T_2 scans. The type III lesion is indicated by the red arrowhead on DIR (**J**) and PSIR (**K**). (**O**) MRI overview of a DIR scanned brain slice from which a tissue block (yellow box) was processed. (**P**) Histopathological section of the tissue block stained for myelin, in which a type I lesion is visible. The lesion is indicated by the black arrowhead and bordered by the black line; the dotted line indicates the cortex. (**Q**–**U**) Excerpts of, respectively, DIR, PSIR, FLAIR, T_1 and PD/T_2 scans. The type I lesion is indicated by the red arrowhead on DIR (**Q**) and PSIR (**R**).

transitions, blood and fluid in/around the sulci, and tissue thickness variations following slice cutting (Kilsdonk *et al.*, 2016). Owing to this point, an additional interpretational step may be required when comparing results from this study to existing *in vivo* studies, and differences in sensitivity and e.g. reclassification may at least in part be explained by a difference in settings. However, we have taken great care approaching the *in vivo* situation as nearly as possible: clinical statuses of the patients were taken into account and brain slice MRI was performed

with fresh tissue and with a post-mortem delay that was as short as possible.

Also, as this was a retrospective study, using a previously set MRI protocol, we could not add a 3D-T_1 (MPRAGE/FSPGR) sequence to our study. Interestingly, our results did indicate that T_1-based scoring may be highly valuable, as there was no retrospective difference between DIR and T_1 (while PD/T_2 and FLAIR did perform more poorly). In addition, there was a large difference between prospective and retrospective scorings for T_1, more so than for PD/T_2 and FLAIR. In the literature, a 3D-T_1 sequence has been reported to be of interest in cortical lesion detection before (Nelson *et al.*, 2008). However, the same authors later found the combination of DIR and PSIR to be superior to 3D-T_1 (Nelson *et al.*, 2014), which may reduce the gravity of our omission of a 3D-T_1 here. Furthermore, we have used both 2D and 3D sequences with parameters typically used in the clinical setting. Because of the differences in in-plane and through-plane resolution used in these protocols, measured signal intensities will suffer differently from variable partial volume effects, causing a large range in contrast ratio values, which cannot be directly compared between all sequences and should therefore be considered descriptive for now. In addition, since images were acquired with parallel imaging and geometric distortion correction, signal in the background was not uniformly distributed, such that we could not reliably determine signal-to-noise nor contrast-to-noise ratios. Last, in our material, a limited number of non-cortical (i.e. white matter) lesions were included; tissue blocks at autopsy were predominantly selected for presence of cortical grey matter, and we used whatever white matter pathology available within tissue blocks. An urgent matter in this perspective is the availability of DIR and PSIR worldwide, as these sequences are not current in standard multiple sclerosis radiological practice, and should be developed beyond their academic use.

In conclusion, the debate on which MRI sequence is most sensitive to cortical lesions in multiple sclerosis has long been inconclusive. The current study has furthered this debate by demonstrating, in a comprehensive and histologically verified dataset, using conventional and more advanced imaging sequences for cortical lesion detection, that DIR and PSIR clearly outperform other, more conventional (clinical) techniques such as T_1, PD/T_2 and FLAIR. However, they do not differ from each other in terms of the number and types of cortical lesions they detect *a priori*. It remains to be investigated whether the potential of PSIR for detecting subpial cortical lesions can be further optimized, given the difference between PSIR and DIR scores retrospectively. Learning programmes and amendment of scoring criteria for PSIR should now be considered. This study shows that, without the benefit of histopathological information at hand (in clinical practice), the use of DIR and PSIR (or a combination thereof) maximizes cortical lesion detection for clinical correlation studies. Adhering to published scoring criteria, moreover, ensures a high histopathological specificity.

Acknowledgements

The authors would like to thank the MS-MRI autopsy team of the Department of Anatomy and Neurosciences for their help in acquiring the data, the research technicians for staining the tissue and the Netherlands Brain Bank for their contribution in supplying the material.

Funding

This study was funded by the Department of Anatomy and Neurosciences of Amsterdam University Medical Centre, MS Center Amsterdam. Furthermore, F.B. is supported by the NIHR biomedical research centre at UCLH.

Competing interests

P.M.B. receives research support from the Dutch MS Research Foundation, grant number 19-1049. M.M.S. serves on the editorial board of *Frontiers of Neurology*, receives research support from the Dutch MS Research Foundation, grant number 13-820, and has received compensation for consulting services or speaker honoraria from Excemed, Genzyme and Biogen. F.B. serves as editorial board member of *Brain*, *European Radiology*, *Neurology*, *Multiple Sclerosis Journal* and *Radiology*. He has accepted consulting fees from Bayer-Schering Pharma, Biogen-IDEC, TEVA, Merck-Serono, Novartis, Roche, Jansen-Research, Genzyme-Sanofi, IXICO Ltd., GeNeuro, and Apitope Ltd., and speaker fees from Biogen, IDEC and IXICO, and has received grants from AMYPAD (IMI), EuroPOND (H2020), UK MS Society, Dutch MS Society, PICTURE (IMDI-NWO), NIHR UCLH Biomedical Research Centre (BRC), and ECTRIMS-MAGNIMS. J.J.G.G. is an editor of *Multiple Sclerosis Journal*, is a member of the board of the Netherlands Organization for Health Research and Innovation and has served as a consultant for Merck-Serono, Biogen, Novartis, Genzyme and Teva Pharmaceuticals. M.D.S., P.J.W.P. and L.E.J. report no competing interests.

References

Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. Curr Opin Neurol 2002; 15: 239–45.

Barkhof F, Geurts JJ. Laminar cortical damage in multiple sclerosis. Brain 2015; 138: 828–9.

Bö L, Geurts JJG, Ravid R, Barkhof F. Magnetic resonance imaging as a tool to examine the neuropathology of multiple sclerosis. Neuropathol Appl Neurobiol 2004; 30: 106–17.

Bø L, Vedeler CA, Nyland HI, Trapp BD, Mørk SJ. Subpial demyelination in the cerebral cortex of multiple sclerosis patients. J Neuropathol Exp Neurol 2003; 62: 723–32.

Calabrese M, Agosta F, Rinaldi F, Mattisi I, Grossi P, Favaretto A, et al. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. Arch Neurol 2009; 66: 1144–50.

Calabrese M, Filippi M, Gallo P. Cortical lesions in multiple sclerosis. Nat Rev Neurol 2010; 6: 438–44.

Compston A, Coles A. Multiple sclerosis. Lancet 2008; 372: 1502–17.

Dekker I, Eijlers AJC, Popescu V, Balk LJ, Vrenken H, Wattjes MP, et al. Predicting clinical progression in multiple sclerosis after 6 and 12 years. Eur J Neurol 2019; 26: 893–902.

Eijlers AJC, van Geest Q, Dekker I, Steenwijk MD, Meijer KA, Hulst HE, et al. Predicting cognitive decline in multiple sclerosis: a 5-year follow-up study. Brain 2018; 141: 2605–18.

Filippi M, Preziosa P, Banwell BL, Barkhof F, Ciccarelli O, De Stefano N, et al. Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. Brain 2019; 142: 1858–75.

Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, et al. MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. Lancet Neurol 2016; 15: 292–303.

Geurts JJG, Barkhof F. Grey matter pathology in multiple sclerosis. Lancet Neurol 2008; 7: 841–51.

Geurts JJG, Bø L, Pouwels PJW, Castelijns JA, Polman CH, Barkhof F. Cortical lesions in multiple sclerosis: combined postmortem MR imaging and histopathology. AJNR Am J Neuroradiol 2005a; 26: 572–7.

Geurts JJG, Calabrese M, Fisher E, Rudick RA. Measurement and clinical effect of grey matter pathology in multiple sclerosis. Lancet Neurol 2012; 11: 1082–92.

Geurts JJG, Pouwels PJW, Uitdehaag BMJ, Polman CH, Barkhof F, Castelijns JA. Intracortical lesions in multiple sclerosis: improved detection with 3D double inversion-recovery MR imaging. Radiology 2005b; 236: 254–60.

Geurts JJG, Roosendaal SD, Calabrese M, Ciccarelli O, Agosta F, Chard DT, et al. Consensus recommendations for MS cortical lesion detection scoring using double inversion recovery MRI. Neurology 2011; 76: 418–24.

Kidd D, Barkhof F, McConnell R, Algra PR, Allen IV, Revesz T. Cortical lesions in multiple sclerosis. Brain 1999; 122: 17.

Kilsdonk ID, Jonkman LE, Klaver R, van Veluw SJ, Zwanenburg JJ, Kuijer JP, et al. Increased cortical grey matter lesion detection in multiple sclerosis with 7 T MRI: a post-mortem verification study. Brain 2016; 139: 1472–81.

Klaver R, De Vries HE, Schenk GJ, Geurts JJ. Grey matter damage in multiple sclerosis: a pathology perspective. Prion 2013; 7: 66–75.

Kutzelnigg A, et al. Cortical demyelination and diffuse white matter injury in multiple sclerosis. Brain 2005; 128: 2705–12.

Meijerman A, Amiri H, Steenwijk MD, Jonker MA, van Schijndel RA, Cover KS, et al. Reproducibility of deep gray matter atrophy rate measurement in a large multicenter dataset. AJNR Am J Neuroradiol 2018; 39: 46–53.

Nelson F, Datta S, Garcia N, Rozario NL, Perez F, Cutter G, et al. Intracortical lesions by 3T magnetic resonance imaging and correlation with cognitive impairment in multiple sclerosis. Mult Scler 2011; 17: 1122–9.

Nelson F, Poonawalla A, Datta S, Wolinsky J, Narayana P. Is 3D MPRAGE better than the combination DIR/PSIR for cortical lesion detection at 3T MRI? Mult Scler Relat Disord 2014; 3: 253–7.

Nelson F, Poonawalla AH, Hou P, Wolinsky JS, Narayana PA. 3D MPRAGE improves classification of cortical lesions in multiple sclerosis. Mult Scler 2008; 14: 1214–9.

Pitt D, Pei W, Wohleb E, Jasne A, Zachariah CR, Rammohan K, et al. Imaging cortical lesions in multiple sclerosis with ultra-high-field magnetic resonance imaging. Arch Neurol 2010; 67: 812–8.

Popescu V, Klaver R, Voorn P, Galis-de Graaf Y, Knol DL, Twisk JW, et al. What drives MRI-measured cortical atrophy in multiple sclerosis? Mult Scler 2015; 21: 1280–90.

Rinaldi F, Calabrese M, Grossi P, Puthenparampil M, Perini P, Gallo P. Cortical lesions and cognitive impairment in multiple sclerosis. Neurol Sci 2010; 31 (Suppl 2): S235–7.

Roosendaal SD, Bendfeldt K, Vrenken H, Polman CH, Borgwardt S, Radue EW, et al. Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability. Mult Scler 2011; 17: 1098–106.

Roosendaal SD, Moraal B, Pouwels PJW, Vrenken H, Castelijns JA, Barkhof F, et al. Accumulation of cortical lesions in MS: relation with cognitive impairment. Mult Scler 2009; 15: 708– 14.

Seewann A, Kooi EJ, Roosendaal SD, Pouwels PJ, Wattjes MP, van der Valk P, et al. Postmortem verification of MS cortical lesion detection with 3D DIR. Neurology 2012; 78: 302–8.

Seewann A, Vrenken H, Kooi EJ, van der Valk P, Knol DL, Polman CH, et al. Imaging the tip of the iceberg: visualization of cortical lesions in multiple sclerosis. Mult Scler 2011; 17: 1202–10.

Sethi V, Muhlert N, Ron M, Golay X, Wheeler-Kingshott CA, Miller DH, et al. MS cortical lesions on DIR: not quite what they seem? PLoS One 2013; 8: e78879.

Sethi V, Yousry TA, Muhlert N, Ron M, Golay X, Wheeler-Kingshott C, et al. Improved detection of cortical MS lesions with phase-sensitive inversion recovery MRI. J Neurol Neurosurg Psychiatry 2012; 83: 877–82.

Steenwijk MD, Geurts JJ, Daams M, Tijms BM, Wink AM, Balk LJ, et al. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. Brain 2016; 139 (Pt 1): 115–26.

Strijbis EMM, Kooi EJ, van der Valk P, Geurts J. Cortical remyelination is heterogeneous in multiple sclerosis. J Neuropathol Exp Neurol 2017; 76: 390–401.

Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 2018; 17: 162–73.

Trampel R, Bazin PL, Pine K, Weiskopf N. In-vivo magnetic resonance imaging (MRI) of laminae in the human cortex. Neuroimage 2019; 197: 707–15.